

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
имени М. В. ЛОМОНОСОВА

ГЕОГРАФИЧЕСКИЙ ФАКУЛЬТЕТ

*На правах рукописи*

**Самсонов Тимофей Евгеньевич**

**Генерализация пространственных данных  
и ее картографические приложения**

1.6.20 — Геоинформатика, картография

ДИССЕРТАЦИЯ  
на соискание учёной степени  
доктора географических наук

**ТОМ 1**

Москва — 2025

# Содержание

## ТОМ 1

<b>Введение</b> . . . . .	<b>4</b>
<b>1 Современное состояние проблемы</b> . . . . .	<b>11</b>
1.1 Генерализация в картографии и геоинформатике . . . . .	11
1.2 Детальность пространственных данных . . . . .	37
1.3 Мультимасштабное картографирование . . . . .	57
1.4 Некартографические приложения генерализации . . . . .	69
1.5 Результаты и выводы . . . . .	76
<b>2 Адаптивная генерализация пространственных данных</b> . . . . .	<b>79</b>
2.1 Отбор точечных моделей населенных пунктов с учетом пространственной неравномерности их распределения . . . . .	79
2.2 Прореживание линейной модели автодорожной сети с сохранением связ- ности населенных пунктов . . . . .	91
2.3 Геометрическое упрощение множества линейных объектов с сохранением топологии . . . . .	101
2.4 Геометрическое упрощение линейных пространственных объектов разной формы . . . . .	112
2.5 Адаптивная генерализация цифровых моделей рельефа с сохранением структурных линий . . . . .	148
2.6 Агрегирование разнородных пространственных данных для построения баз данных характеристик городской среды . . . . .	169
2.7 Результаты и выводы . . . . .	195

## ТОМ 2

<b>3 Детальность пространственных данных</b> . . . . .	<b>4</b>
3.1 Плотность пространственных объектов на цифровых топографических кар- тах и ее связь с географическим положением . . . . .	4

3.2	Гранулярность цифровых моделей рельефа и ее связь с масштабом картографирования . . . . .	28
3.3	Контроль детальности линейных пространственных объектов при обработке разными алгоритмами геометрического упрощения . . . . .	56
3.4	Повышение детальности изображения при визуализации цифровых моделей рельефа с помощью дополнительных горизонталей . . . . .	75
3.5	Картографическая визуализация цифровых моделей рельефа с помощью штрихов и управление его детальностью . . . . .	99
3.6	Конфляция гидрографических данных и цифровых моделей рельефа разной детальности . . . . .	116
3.7	Взаимная координатная привязка цифровых геоизображений разной детальности . . . . .	154
3.8	Результаты и выводы . . . . .	170
<b>4</b>	<b>Мультимасштабное картографирование . . . . .</b>	<b>173</b>
4.1	Основные понятия и принципы мультимасштабного картографирования .	173
4.2	Вычисление параметров визуализации цифровых моделей дорожной сети и структуры населенных пунктов на мультимасштабных картах . . . . .	178
4.3	Мультимасштабное гипсометрическое картографирование на основе цифровых моделей рельефа . . . . .	188
4.4	Мультимасштабная визуализация цифровых моделей характеристик ветрового волнения . . . . .	206
4.5	Построение цифровых моделей полимагистралей при мультимасштабном картографировании транспортных сетей . . . . .	223
4.6	Визуализация пространственных данных способами картограмм и картодиаграмм на мультимасштабных картах . . . . .	233
4.7	Результаты и выводы . . . . .	249
	<b>Заключение . . . . .</b>	<b>251</b>
	Публикации по теме диссертации . . . . .	253
	<b>Список литературы . . . . .</b>	<b>262</b>
	<b>Приложение . . . . .</b>	<b>315</b>

# Введение

**Актуальность темы.** Картографическая генерализация представляет собой отбор, обобщение, выделение главных типических черт изображаемых объектов в соответствии с назначением, масштабом, содержанием карты, особенностями картографируемой территории и самого объекта [Салищев, 1976; Берлянт, 2014]. При использовании геоинформационных технологий карты составляются на основе *пространственных данных*, под которыми понимаются данные о пространственных объектах — *цифровых моделях* материальных или абстрактных объектов реального или виртуального мира с указанием их идентификаторов, координатных и атрибутивных данных [ГОСТ Р 57657-2017, 2018]. Таким образом, задача картографической генерализации решается путем *генерализации пространственных данных* — обобщения позиционных [координатных] и атрибутивных данных о пространственных объектах в ГИС в автоматическом или интерактивном режиме с использованием операторов генерализации, их наборов или последовательностей, часть из которых имеет соответствие в приемах и методах картографической генерализации [Геоинформатика, 1999].

Географические принципы обобщения карт, связанные с учетом закономерностей организации явлений в разных пространственных масштабах, трудно формализуемы, и в силу этого существующие подходы к генерализации пространственных данных реализуют их лишь частично. Требуется углубление географической содержательности методов генерализации, а также адаптация ее концептуального и технологического аппарата к оперированию цифровыми моделями географических объектов. Решение этой проблемы представляет собой важную и актуальную научную задачу в области геоинформатики и географической картографии.

**Объект исследования** — пространственные данные как совокупность цифровых моделей географических объектов.

**Предмет исследования** — методы генерализации пространственных данных для создания карт разных масштабов и мультимасштабных карт.

**Цель исследования** — разработка автоматизированных методов генерализации пространственных данных, обеспечивающих воспроизведение географических особенностей картографируемых объектов, представленных в исходных цифровых моделях, а также достижение необходимого уровня их детализации и создание мультимасштабных карт.

Для достижения цели потребовалось решить следующие **задачи**:

1. Выполнить анализ современного состояния теории, методов и технологий картографической генерализации; выделить особенности генерализации пространственных данных для целей картографирования, а также проблемы, обуславливающие их развитие в части адаптивных подходов, детальности и мультимасштабного картографирования.
2. Разработать методики адаптации процессов генерализации пространственных данных к распределению, структуре, форме и отношениям географических объектов, которые могут быть выявлены по их цифровым моделям.
3. Предложить подходы к оценке детальности пространственных данных, управлению детальностью при генерализации пространственных данных, а также интеграции пространственных данных различной детальности для целей картографирования.
4. Определить теоретические и методологические основания мультимасштабного картографирования (МК), обосновать принципы построения баз данных и преобразований способов изображения на мультимасштабных картах.
5. Выполнить реализацию разработанных подходов на разнообразных картографических примерах.

**Методология и методы исследования.** В основе исследования лежит цифровое моделирование географических объектов. Разработанные методики генерализации опираются на предположение о том, что исходные пространственные данные достоверно описывают моделируемые объекты, что позволяет оценивать качество их генерализации путем сравнения результатов с начальным состоянием. Выявление пространственных закономерностей организации географических объектов опирается на реализацию приемов картографического метода исследования посредством геоинформационных технологий и их применение к цифровым моделям. Для решения задач исследования выполнялась разработка и применение специализированного программного обеспечения, картографическая визуализация исходных и генерализованных пространственных данных, качественный и количественный (в том числе статистический) анализ полученных результатов.

**Основные положения**, выносимые на защиту:

1. *Генерализация пространственных данных для целей картографирования* опирается на принципы адаптивности, детальности и мультимасштабности. Применение данных принципов позволяет реализовать учет таких факторов генерализации как особенности картографируемых объектов, необходимый уровень детализации выходных данных (в том числе определяемый масштабом карты), а также интерактивность среды их визуализации при работе с цифровыми моделями географических объектов.

2. *Адаптивность* генерализации пространственных данных обеспечивает сохранение и воспроизведение пространственного распределения, пространственных отношений, характера формы и основных структурных элементов картографируемых объектов, а также соотношения их площадей, представленных в исходных цифровых моделях. Реализация принципа адаптивности основывается на включении в геоинформационные алгоритмы генерализации операций морфометрического, сетевого, оверлейного и статистического анализа, а также вспомогательных разбиений пространства, определяемых множеством генерализуемых объектов.
3. *Детальность* пространственных данных в геометрической части определяется характеристиками цифровых моделей географических объектов, которые монотонно изменяются в процессе их генерализации. Направленное воздействие на эти характеристики с помощью операторов генерализации позволяет достигать требуемой детальности картографического изображения, а моделирование этих характеристик дает количественные оценки детальности и возможность ее управляемого изменения. Согласование данных разной детальности осуществляется с помощью конфляции (совмещения) их соответствующих элементов, а их взаимная координатная привязка — с использованием методов компьютерного зрения.
4. *Мультимасштабность* картографического изображения обеспечивается интерактивным откликом процессов генерализации пространственных данных и преобразования графических средств и способов изображения на изменение масштаба карты. Анимации переходов между масштабными уровнями карты позволяют явным образом визуализировать геометрические и атрибутивные преобразования пространственных данных и повысить информативность мультимасштабных карт.

**Научная новизна** работы состоит в том, что впервые:

1. Сформулированы основные принципы генерализации пространственных данных для целей картографирования: адаптивность, детальность и мультимасштабность.
2. Разработано понятие адаптивной генерализации пространственных данных. Предложен и реализован ряд практических методов и алгоритмов генерализации, реализующих принципы адаптивности.
3. Получен комплекс методик оценки, управления и интеграции для пространственных данных и карт в контексте их детальности. Введено и исследовано понятие гранулярности цифровых моделей рельефа как ключевой геометрической характеристики их детальности.
4. Сформулированы основные положения мультимасштабного картографиро-

вания (МК) как нового направления картографии. Показана их практическая применимость в задачах общегеографического, физико-географического и социально-экономического картографирования.

5. Созданы новые уникальные наборы данных: глобальная мультимасштабная цифровая модель высот HYPSO для мелкомасштабного гипсометрического картографирования; база данных характеристик подстилающей поверхности и геометрии городских каньонов Московского региона.

**Практическая значимость.** Результаты работы применимы в широком спектре географических исследований, связанных с построением общегеографических и тематических карт. Предложенные методы позволяют снизить субъективный фактор при выполнении процедур генерализации и построения мультимасштабных карт, ускорить решение соответствующих задач. Разработанные автором программные средства размещены в открытом доступе в репозитории *GitHub* и могут быть использованы для решения практических задач обработки пространственных данных, а также в образовательных целях.

**Достоверность** полученных результатов обеспечивается:

- использованием численных и графоаналитических методов для оценки геометрической точности и географического правдоподобия результатов, получаемых каждым из разработанных методов генерализации;
- экспертными оценками отдельных результатов методом визуального сравнения с образцами аналогичных карт ручного составления, а также методом экспертного опроса;
- строгостью используемых математических методов;
- апробацией результатов на международных и российских научных мероприятиях, а также их публикацией в ведущих (в том числе, высокорейтинговых) журналах, соответствующих тематике диссертации.

**Апробация работы.** Основные результаты работы докладывались на ряде научных мероприятиях по теме диссертации: XXV-XXXI Международных картографических конференциях (Париж, 2011; Дрезден, 2013; Рио-де-Жанейро, 2015; Вашингтон, 2017; Токио, 2019; Флоренция, 2021; Кейптаун, 2023); 7-й и 8-й Международной конференции по геоинформатике GIScience (Коламбус, 2012; Вена, 2014); 19-м и 24-м Международном симпозиуме по автоматизированной картографии AutoCarto (Коламбус, 2012; Редландс, 2022); 34-м Международном географическом конгрессе (Стамбул, 2021); Региональной конференции Международного географического союза (Москва, 2015); 9-м, 11-м и 12-м совещании Комиссии по горной картографии Международной картографической ассоциации (Банфф, 2014; Хвар, 2018; Сноу-Ранч, 2023); 18-м и 19-м и 23-м совещании Комиссии по генерализации и множественным представлениям Международной картографической ассоциации (Дрезден, 2014; Хельсинки, 2016; Делфт, 2020); 3-м Меж-

дународном симпозиуме по геоморфометрии (Редландс, 2011); II Всероссийской научно-практической конференции с международным участием «Цифровая география» (Пермь, 2024), 1-й и 2-й Национальных картографических конференциях (Москва, 2016, 2018); III и V Всероссийской научно-практической конференции «Геодезия, картография, геоинформатика и кадастры. Наука и образование» (Санкт-Петербург, 2019, 2024); 18-й и 27-й международной конференции Интеркарто/ИнтерГИС (Смоленск, 2012; Апатиты, 2021); 1-й, 3-й и 11-й Всероссийской научно-практической конференции «Геоинформационное картографирование в регионах России» (Воронеж, 2009, 2013, 2020); XIV-XIX и XXII Международных конференциях пользователей Esri в России и СНГ (Голицыно, 2008, 2009, 2010; Поведники, 2011, 2012, 2013, 2016); XXXIII Пленуме Геоморфологической Комиссии РАН (Саратов, 2013); Летних школах Лаборатории дискретной и вычислительной геометрии имени Б. Н. Делоне ЯрГУ (Ярославль, 2012, 2013); I и II Международных форумах молодых ученых «Наука будущего — наука молодых» (Севастополь, 2015; Казань, 2016); V Международной конференции молодых ученых и специалистов памяти академика А. П. Карпинского (Санкт-Петербург, 2017); Всероссийской конференции «Ломоносовские чтения» (Москва, 2018); Международной научно-технической конференции «Пространственные данные — основа стратегического планирования, управления и развития» (Москва, 2019).

Результаты исследования были обсуждены на семинарах кафедр картографии и геоинформатики Санкт-Петербургского государственного университета (СПбГУ), Пермского государственного национального исследовательского университета (ПГНИУ), Сибирского государственного университета геосистем и технологий (СГУГиТ), кафедры картографии Московского государственного университета геодезии и картографии (МИИГАиК), кафедры геоинформатики Кубанского государственного университета (КубГУ).

На основе опыта многолетних исследований по тематике диссертационного исследования автором разработаны и внедрены в учебный процесс курсы «*Картографические базы данных*» (читался с 2014 по 2023 г.) и «*Генерализация пространственных данных*» (читается с 2024 г.) для студентов бакалавриата МГУ по направлению 05.03.03 «Картография и геоинформатика».

**Фактический материал и личный вклад автора.** В работе использованы открытые источники пространственных (базы пространственных данных, цифровые карты, цифровые модели рельефа, космические снимки) и статистических данных. Постановка проблемы, разработка методик и анализ результатов во всех исследованиях осуществлялись лично автором. Для 12 из 18 частных методик<sup>1</sup> (2/3 от общего объема) автором также выполнена их техническая (в т.ч. программная) реализация и непосредственно генерализация пространственных данных.

---

<sup>1</sup> во всех разделах кроме 2.1-2.4, 3.7, 4.2

**Публикации.** Соискатель имеет 33 опубликованных работы по теме диссертации, включая 23 статьи в рецензируемых научных изданиях, индексируемых в базе ядра Российского индекса научного цитирования «eLibrary Science Index». Получено 4 свидетельства о регистрации прав на программное обеспечение. В материалах российских и международных конференций опубликовано 20 работ. Участие соискателя во всех опубликованных работах является значительным, конкретный вклад обозначен в списке работ. В работах, где соискатель идет первым автором или имеет вклад более 50% (19 работ из 23 в изданиях, рекомендованных для защиты), его роль была определяющей при подготовке публикации, включая написание текста, составление графических материалов, предоставление статьи в редакцию и прохождение рецензирования. При подготовке диссертации использованы тексты всех 23 публикаций в изданиях, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования.

**Структура и объем работы.** Диссертация состоит из введения, четырех глав, заключения, списка литературы (617 наименований, из них 535 на иностранном языке) и приложения. Объем работы составляет 526 страниц, включая 198 иллюстраций и 32 таблицы. Количество томов диссертации — 2.

**Благодарности.** Автор выражает искреннюю благодарность сотрудникам кафедры картографии и геоинформатики д.г.н. В. И. Кравцовой, к.г.н. С. В. Чистову, к.г.н. С. М. Кошелю и д.т.н. М. В. Нырцову за обстоятельное и плодотворное обсуждение диссертации, внимательное ознакомление с ее текстом и авторефератом, конструктивные замечания. Автор глубоко признателен своим университетским учителям — к.г.н. С. М. Кошелю, под руководством которого он выполнил свои первые исследования и дипломную работу по автоматизации в картографии; д.г.н., профессорам А. М. Берлянту и И. К. Лурье, при поддержке и под руководством которых была подготовлена кандидатская диссертация автора, заложившая основы настоящего исследования.

Целый ряд задач в настоящей работе был решен совместно с к.ф.-м.н. О. П. Якимовой, которой автор благодарен за многолетнее плодотворное сотрудничество в исследованиях по геометрическим алгоритмам генерализации. Исследование по полимагистралам выполнено совместно с к.г.н. Е. А. Прохоровой. Важное место в диссертации занимают гидрометеорологические приложения генерализации пространственных данных, которые были осуществлены благодаря сотрудничеству с к.г.н. М. И. Варенцовым, к.г.н. П. И. Константиновым и к.ф.-м.н. С. А. Мысленковым. Автор также признателен студентам и аспирантам, совместно с которыми был получен ряд результатов: А. А. Шурьгиной, А. М. Семенковой (Кривошеиной), Н. Д. Юровой-Доронкиной, А. С. Подольскому, А. Р. Высоцкой (Халиуллиной), Д. О. Вальтеру, Э. А. Усмановой, А. Р. Мухаметшину и Г. А. Карташову.

Особую благодарность автор хочет выразить своей *Alma Mater* — географическому факультету МГУ и кафедре картографии геоинформатики, в дружеской атмосфере кото-

рых он сформировался как географ и ученый-исследователь. Вместе с тем, своим техническим уровнем автор во многом обязан опыту, полученному в бытность старшим экспертом по серверным геоинформационным системам в компаниях Дата+ и ESRI CIS в 2008-2011 гг.

Работа выполнена при поддержке грантов МК (793.2013.5, 4829.2016.5), РФФИ (12-05-31318-мол\_а, 14-05-00888-а, 15-07-06789-а, 18-07-01459), РГО (21/2016-Р) и РНФ (19-77-00071). Оригинал-макеты диссертации и автореферата подготовлены с помощью открытого и бесплатного программного обеспечения Quarto, Pandoc, Zotero, Typst, XeLaTeX, R, RStudio и Positron.

Наконец, самая сердечная благодарность адресована моим родным и близким за поддержку на всех жизненных этапах.

# 1 Современное состояние проблемы

В данной главе проведен анализ современного состояния исследований в области картографической генерализации. Показано, как развитие концептуальных (теоретических) моделей генерализации и автоматизации в картографии привело к возрастанию проблематики трех направлений, рассматриваемых в настоящей работе: адаптивной генерализации, детальности пространственных данных и мультимасштабного картографирования.

## 1.1 Генерализация в картографии и геоинформатике

**Картографическая генерализация**, согласно определению, принятому в школе географической картографии, представляет собой отбор, обобщение, выделение главных типических черт изображаемых объектов в соответствии с назначением, масштабом, содержанием карты, особенностями картографируемой территории и самого объекта [Салищев, 1976; Берлянт, 2014]. Американская картографическая школа акцентирует внимание на том, что генерализация — часть более широкого процесса **картографической абстракции**, который также включает в себя процессы *выбора* и *визуализации* [Robinson et al., 1995]. Под выбором понимается интеллектуальный процесс решения, какие классы объектов (например, рельеф, растительность, населенные пункты) необходимо отразить на карте в соответствии с ее назначением. Подчеркивается, что в процессе выбора не происходит какой-либо модификации исходной информации — по этой причине он вынесен за скобки процесса генерализации. В отечественной же картографической школе за это отвечает этап определения содержания и легенды в процессе проектирования и составления карты [Берлянт, 2014].

При использовании геоинформационных технологий карты составляются на основе **пространственных данных**, которые определяются как данные о пространственных объектах и их наборах [ГОСТ Р 52438-2005, 2005] или — более развернуто — как данные о пространственных объектах, включающие сведения об их форме, местоположении и свойствах, в том числе представленные с использованием координат [ГОСТ Р 57657-2017, 2018]. В свою очередь **пространственный объект** (англ. *feature*) определяется как *цифровая модель* материального или абстрактного объекта реального или виртуального мира с указанием его идентификатора, координатных и атрибутивных данных [ГОСТ Р 52438-

2005, 2005; ГОСТ Р 57657-2017, 2018], что соответствует определениям международных стандартов ISO и OGC. В более поздних отечественных стандартах [ГОСТ Р 70316-2022, 2022; ГОСТ Р 70846.4-2023, 2023], однако, используется формулировка в соответствии с Федеральным законом от 30 декабря 2015 г. № 431 «О геодезии, картографии и пространственных данных и о внесении изменений в отдельные законодательные акты Российской Федерации», которая определяет пространственные объекты как природные объекты, искусственные и иные объекты (в том числе здания, сооружения), местоположение которых может быть определено, а также естественные небесные тела. И.К.Лурье отмечает двоякость трактовки данного термина: как объекта реальности и как его цифрового представления [Лурье, 2016, с. 16]. Эту амбивалентность отмечают и авторы Толкового словаря терминов по геоинформатике [Геоинформатика, 1999, с. 72]. Однако в обоих случаях «цифровая» трактовка термина ставится на первое место. Этой позиции будем придерживаться и мы: в настоящей работе используется понимание пространственного объекта как цифровой модели. Для обозначения моделируемых объектов реальности в работе используется понятие *географический объект*.

Таким образом, геоинформационные методы картографирования манипулируют цифровыми моделями географических объектов, представленными в виде пространственных данных. Соответственно, задача картографической генерализации решается путем **генерализации пространственных данных** — обобщения позиционных [координатных] и атрибутивных данных о пространственных объектах в ГИС в автоматическом или интерактивном режиме с использованием операторов генерализации, их наборов или последовательностей, часть из которых имеет соответствие в приемах и методах картографической генерализации [Геоинформатика, 1999]. Можно обратить внимание на то, что в вышеприведенном определении картографической генерализации неявно присутствует некий переход между картографируемыми объектами и их изображением. Генерализация пространственных данных позволяет формализовать этот переход и наполнить его конкретным процедурным смыслом, который соответствует технологическому уровню развития геоинформатики. Вместе с тем, взгляды на сущность генерализации весьма разнятся, а выявление конкретных проблем ее развития требуют детального и предметного анализа современного состояния связанных с ней вопросов. Последующие параграфы настоящей главы посвящены такому анализу. Начать же следует с концептуальных моделей генерализации, которые определяют ее содержание.

### 1.1.1 Концептуальные модели генерализации

**Концептуальная модель** — это абстрактная модель, определяющая структуру исследуемого объекта (составные части и связи), свойства составных частей, причинно-следственные связи [ГОСТ Р 10032-2007, 2009]. С точки зрения концептуального модели-

рования генерализация может быть рассмотрена как процесс: а) имеющий некоторые предпосылки (факторы), б) рассматривающий входную географическую информацию под определенным углом зрения (имеющий особый «подход») и в) состоящий из множества атомарных операций, которые могут быть объединены в алгоритмы анализа и обработки этой информации. Задачи концептуальной модели генерализации — это формализация подхода в целом, а также систематизация факторов и операций.

Впервые задача картографической генерализации как научная проблема была сформулирована немецким картографом Максом Эккертом [Eckert, 1921]. Его труд, однако, носил чисто эмпирический характер и являлся, по сути, сборником удачных рецептов составления карт. Появление и развитие концептуальных моделей началось в середине XX века и изначально было обусловлено становлением картографии как полноценной академической дисциплины. Развитие картографии как науки требовало обобщения, систематизации и объективизации разнообразных проявлений и видов генерализации и символизации на картах, часто выполнявшихся на интуитивном уровне, на основе экспертного опыта. В 70-80-е годы с развитием вычислительной техники и увеличением объема накапливаемой информации к этому добавился второй серьезный фактор: необходимость выполнения генерализации и оформления карт на компьютере в автоматизированном режиме. Это означало, что данные процессы должны быть не только обобщены и четко систематизированы, но и формализованы в той степени, которая позволяет представить их в виде компьютерных алгоритмов.

Первенство в теоретическом обобщении различных проявлений генерализации, по всей видимости, принадлежит К.А.Салищеву, который еще в книге [Салищев, Сухов, Филиппов, 1947, с. 76-80] выделил такие виды генерализации, как:

- обобщение очертаний (формы) картографируемых объектов;
- обобщение их количественной характеристики;
- обобщение качественной характеристики;
- отбор картографируемых объектов;
- замена индивидуальных объектов (понятий) их собирательными обозначениями (видовыми и далее родовыми понятиями).

Там же автором произведено четкое выделение и описание влияния основных факторов генерализации, включая [Салищев, Сухов, Филиппов, 1947, с. 80-82]:

- назначение карты;
- масштаб;
- особенности картографируемой действительности.

Наконец, были отмечены различные условия, которые необходимы для научно обоснованного и практически целесообразного осуществления генерализации, такие как указание минимальных размеров условных знаков и контуров, определение опти-

мальной и предельной графической нагрузки, изучение картографируемых явлений, их взаимных связей и обусловленностей, выяснение требований к характеристике явлений и их изображению в зависимости от назначения карты, а также разработка методов графического изображения картографируемых явлений на основе изучения их характеристик и особенностей [Салищев, Сухов, Филиппов, 1947].

Важной предпосылкой осмысления и обобщения сущности генерализации стала систематизация визуальных форм представления информации на картах, а также концептуализация изображаемых явлений с точки зрения их локализации. Систематизацию способов изображения на тематических (тогда — специальных) картах Салищев предложил еще в первой редакции «Основ картоведения» [Салищев, 1939], включив в нее такие методы как:

1. Изолинии
2. Способ цветного фона (для качественных характеристик)
3. Внемасштабные знаки (для объектов, локализованных по пунктам)
4. Точечный способ
5. Ареалы
6. Линии движения
7. Картодиаграммы
8. Картограммы

В дальнейшем его идеи постоянно развивались. Например, в учебнике «Картография» [Салищев, 1966] впервые была введена классификация явлений по характеру размещения и передаваемым характеристикам, предполагающая возможную локализацию по пунктам, на линиях, по площадям, рассеянное и сплошное; при передаче характеристик состояния в определенный момент времени, перемещения (движения) и изменения во времени.

Наиболее же зрелый и целостный концептуальный подход к картографической генерализации можно найти в труде «Картоведение» [Салищев, 1976]. Его отличительными особенностями являются:

- высокий уровень абстракции (акцент на видах генерализации, а не атомарных операциях);
- концептуализация на уровне явлений, а не отдельных объектов;
- учет взаимного влияния способов изображения и генерализации;
- внимание к дополнительному процессу отбора подписей;
- анализ соотношения геометрической точности и содержательной верности.

Введенные в этой работе виды (стороны) генерализации используются в отечественной картографии и по сей день:

1. Отбор картографируемых явлений.
2. Геометрическая пространственная сторона генерализации.
3. Обобщение количественной характеристики.
4. Обобщение качественной характеристики.
5. Переход от простых объектов и понятий к более сложным.

А.М.Берлянт позднее использовал аналогичную классификацию, но явно разделил геометрическую сторону генерализации на несколько аспектов: обобщение очертаний, объединение контуров (выделов), смещение элементов изображения, а также утрирование, или показ объектов с преувеличением [Берлянт, 2014].

Начиная с 60-х гг. идеи концептуального представления генерализации начали появляться и за рубежом. Так, польский картограф Лех Ратайский в 1967 году предложил одну из первых формальных моделей генерализации, разделив процесс на 2 составляющие: качественную и количественную, подчеркивая тем самым различия между переходом на новый уровень абстракции и постепенным изменением содержания карты [Ratajski, 1967]. Под количественной генерализацией Ратайский понимал постепенное уменьшение объема содержания карты в соответствии с изменением масштаба; качественная же генерализация обозначала трансформацию элементарных символов в более абстрактные формы символьного представления. Жаком Бертэном эти типы генерализации описываются как концептуальная и структурная [Bertin, 2010]. Оба автора использовали понятие «точки генерализации», обозначая им момент, в котором требуется переход к новой форме картографического представления.

Особый интерес представляет работа американского картографа Джоэла Моррисон, который представил модель генерализации, основанную на формальной теории множеств, пытаясь придать этому процессу максимально строгое описание, подходящее для реализации средствами компьютера [Morrison, 1974]. Моррисон рассмотрел четыре составляющих генерализации: упрощение, классификацию, символизацию и индукцию. Для их формализации он ввел два класса функций вида  $A \rightarrow B$  :

- «один-к-одному» (*one-to-one*), при которой различные элементы в  $A$  имеют различные образы в  $B$ ;
- «на» (*onto*), при которой каждый элемент  $B$  имеет прообраз в  $A$ .

Наличие (1) или отсутствие (0) каждого из этих свойств определяет 4 возможных вида трансформаций, которые соответствуют обозначенным составляющим генерализации: 11 — упрощение, 00 — индукция, 10 — символизация и 01 — классификация, где первая цифра соответствует функции «один-к-одному», а вторая — функции «на». Важно, что при объяснении практического применения своей модели Моррисон оперировал представлением картографических слоев в виде оцифрованных последовательностей координат или изображений, что позволяет назвать его концептуальную модель

первой, опирающейся на обработку пространственных данных.

Повышенный интерес к разработке концептуальных моделей генерализации возник во второй половине 1980-х гг, что по всей видимости, было связано с накопленным багажом алгоритмов автоматизированной генерализации, а также резким увеличением доступности персональных компьютеров. Первой моделью этой эпохи можно назвать концептуальную модель Никерсона-Фримена, разработанную по аналогии с подходом экспертных систем [Nickerson, Freeman, 1986]. Они выделили 5 последовательно выполняемых этапов:

1. Модификация объектов — удаление, упрощение, комбинирование и изменение типа.
2. Масштабирование символов.
3. Смещение объектов, размещение символов.
4. Уменьшение масштаба.
5. Размещение подписей.

Любопытной особенностью этой модели было то, что авторы предлагали не уменьшать масштаб карты на первых трех этапах. Вместо этого на третьем этапе размещались увеличенные в размере символы, которые потом уже уменьшались на 4 этапе вместе с масштабом карты. В дальнейшем Никерсон опубликовал реализацию этой модели применительно к линейным объектам [Nickerson, 1988].

В свою очередь, Брассель и Вайбель разработали модель генерализации, основанную на библиотеке процессов [Brassel, Weibel, 1988]. Как и в модели Никерсона-Фримена, процесс генерализации в этой модели разбит на этапы, однако их фокус и содержание совсем иные:

1. *Распознавание структуры* — идентификация генерализуемых пространственных объектов и их пространственных отношений в контексте факторов генерализации, таких как результирующий масштаб, назначение карты и т.д.
2. *Распознавание процессов* — выбор операторов генерализации и их параметров, наиболее подходящих с учетом выявленной структуры пространственных объектов, а также факторов генерализации.
3. *Моделирование процессов* — выстраивание логической цепочки генерализации с использованием функций, извлеченных из библиотеки процессов в соответствии с предыдущим этапом.
4. *Выполнение процесса* — запуск модели применительно ко входной базе пространственных данных и с результатом в виде выходной базы данных.
5. *Отображение данных* — конвертация выходной базы пространственных данных в картографическое изображение.

Данная модель важна в контексте настоящей работы сразу по нескольким причи-

нам. Во-первых, она, хоть и предназначена прежде всего для картографических целей, оперирует явным образом пространственными данными. Во-вторых, в ней явным образом заложена *адаптация* процесса генерализации к особенностям картографируемых объектов и их пространственных отношений. В третьих, используется понятие «моделирование процесса» для представления сложной цепочки операторов генерализации, которой в дальнейшем получило существенно развитие. Наконец, Брассель и Вайбель явным образом допускают некартографические приложения своей модели, разделяя возможность *статистической* и *картографической* генерализации. Под статистической генерализацией ими понимался процесс фильтрации данных, преследующий сжатие данных и их статистический анализ.

Следующая в хронологическом порядке и одновременно самая сложная и комплексная концептуальная модель генерализации была предложена МакМастером и Ши [McMaster, Shea, 1988]. Если в моделях Моррисона, Никерсона-Фримена и Брасселя-Вайбеля отчетливо прослеживался формально-технический уклон, то в этой новой модели авторы постарались уделить равное внимание как технологическим аспектам генерализации, так и ее картографическим факторам. МакМастер и Ши разделили процесс генерализации на 3 аспекта:

1. *Философские причины* (зачем). Состоят из теоретических, прикладных и вычислительных аспектов. К теоретическим авторы отнесли такие факторы как уменьшение сложности, поддержание пространственной и атрибутивной точности, эстетического качества и логической иерархии, а также последовательное применение правил генерализации. Прикладными являются аспекты, связанные с назначением карты, ее масштабом, а также читаемостью. Наконец, к вычислительным элементам авторы отнесли эффективность алгоритмов, максимальное уменьшение объема данных и минимизацию требований к их хранению.
2. *Картометрическая оценка* (где). Данный аспект раскладывается на три элемента: геометрические условия, пространственные метрики, а также контроль трансформации. К первому элементу относятся различные операторы обнаружения геометрических проблем типа наложений объектов и их неразличимости. Пространственные метрики направлены на выявление паттернов и закономерностей в структуре генерализуемых объектов за счет оценки плотности, распределения, длины, извилистости, формы, расстояний и прочих характеристик. Контроль трансформации же осуществляется посредством выбора операторов генерализации, алгоритмов генерализации (реализующих операторы), а также параметров алгоритмов.
3. *Пространственная и атрибутивная трансформация* (как). Этот аспект состоит из двух элементов, обозначенных в его названии. К пространственным трансформациям авторы относят упрощение, сглаживание, агрегирование (точек), амаль-

гамирование (полигонов), слияние (линий), коллапс (уменьшение размерности: полигоны в точки и т.п.), прореживание (для множеств), преувеличение, усиление (символа, а не геометрии) и смещение. К атрибутивным — классификацию и, как ни странно, символизацию. Отнесение последнего к атрибутивным преобразованиям авторы мотивируют тем, что это, по сути, отображение атрибутов объектов на графические переменные.

МакМастер и Ши не вдаются в подробности технической реализации их модели, акцентируя внимание на необходимости теоретического обобщения всевозможных аспектов генерализации в контексте цифрового моделирования.

Наконец, последней моделью в череде разработок 80-х годов стала модель МакМастера-Монмонье, разработанная специально для генерализации в растровом режиме. Авторы разделили все операции на 4 категории: структурная, численная, численная категоризация и категоризация [McMaster, Monmonier, 1989]:

1. *Структурная генерализация* отвечает за изменение геометрии растра (передискретизацию).
2. *Численная генерализация* подразумевает преобразование числовых данных растра, приводящее к появлению также числовых данных. Примером числовой генерализации является фильтрация растра.
3. *Численная категоризация* приводит к преобразованию непрерывных растров в категориальные, по сути представляя собой реклассификацию растра.
4. *Категоризация* представляет собой перевод растра из одной классификации в другую — например, при укрупнении классов для отображения в более мелких масштабах картографирования.

Данная модель также не ограничена применением в картографических целях и может быть успешно адаптирована применительно к обработке растровых данных любого типа, в том числе данных дистанционного зондирования.

Таким образом, к началу 90-х годов XX века был накоплен уже солидный багаж в области теоретического обобщения картографической генерализации. Наметившийся в 80-е годы переход в сторону компьютеризации породил технологические концепции, которые хоть и были ориентированы на решение картографических задач, были уже полноценно ориентированы на обработку *пространственных данных*, т.е. цифровых моделей пространственных объектов.

Как было сказано ранее, появление технологических концепций во многом было стимулировано развитием автоматизации в картографии, и в частности накопленным багажом компьютерных алгоритмов генерализации. Проанализируем разработки в этой области, чтобы выявить современное состояние и имеющиеся проблемы. Для анализа выбраны некоторые задачи генерализации, часто встречающиеся при составлении карт: отбор точечных объектов, прореживание дорожных сетей, геометрическое упрощение

линий, генерализация цифровых моделей рельефа и агрегирование данных. Прочие задачи генерализации рассмотрены кратко в отдельном параграфе.

### 1.1.2 Отбор населенных пунктов<sup>1</sup>

Отбор объектов на картах должен осуществляться с учетом географического рисунка их размещения. В малообжитых районах важно сохранить даже незначительные по статусу объекты, в густонаселенных, наоборот, должен быть произведен интенсивный отбор с сохранением наиболее значимых объектов [Заруцкая, Сваткова, 1982]. Одновременно с этим должны быть сохранены и переданы территориальные контрасты. Наиболее разработаны методики отбора населенных пунктов и их количественное обоснование [Бородин, 1948; Сухов, 1957; Салищев, 1972; Бугаевский, Подольская, 2004].

С точки зрения пространственных данных эта задача решается путем применения алгоритмов генерализации точечных множеств. Одни из первых работ в области автоматизации этого процесса опубликованы в работах [Иванов, 1964; Kadmon, 1972], в которых населенные пункты ранжируются по множеству критериев, включающих административные экономические и культурные характеристики, транспортную значимость, близость к ближайшим населенным пунктам более высокого ранга, а также изолированность. Здесь видим явную адаптацию к свойствам населенных пунктов. Ли [Li, 2006] выделяет две группы алгоритмов генерализации точечных множеств: выборочное исключение и структурное упрощение. Первая группа включает алгоритмы прореживания населенных пунктов, коэффициента распределения, гравитационного моделирования, сегментации множества, редукции по квадрантам, предложенные в [Langran, Poiker, 1986], а также алгоритм растущей окружности [van Kreveld, van Oosterom, Snoeyink, 1995]. В алгоритме прореживания вокруг населенных пунктов строятся окружности радиусом, обратно пропорциональным их значимости и населенные пункты последовательно удаляются, если их окружности захватывают более важные населенные пункты. В алгоритме растущей окружности зависимость радиуса от важности аналогичная, но он последовательно увеличивается и населенный пункт исключается, если его окружность охватила окружность более важного населенного пункта.

Вторая группа рассматривает точки не индивидуально, а как множество, удаляя из него элементы на основе разнообразных характеристик. Наиболее зарекомендовавший себя подход из данной группы основан на применении диаграммы Вороного. Диаграмма Вороного заданного множества точек представляет такое разбиение плоскости, при котором каждая область этого разбиения (*полигон Вороного*) образует геометрическое место точек, более близких к одному из элементов множества, чем к любому другому. Точ-

<sup>1</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov, Krivosheina, 2012; Самсонов, Кривошеина, 2015а]

ки с малой площадью полигонов Вороного располагаются плотно — это означает, что их отбор должен производиться более интенсивно. По этому принципу работают алгоритмы, предложенные в работах [Ai, Liu, 2002; Yan, Weibel, 2008]. При этом, кроме площади полигонов при определении весов населенных пунктов учитывается также их численность населения и административный статус.

Отдельного внимания заслуживает алгоритм генерализации точечных объектов, основанный на построении квадратомиического дерева, и предложенный в работах [Bereuter, Weibel, 2013; Bereuter, Weibel, 2017]. В данной группе методов дерево строится таким образом, что каждая ячейка содержит одну точку исходных данных. Далее в соответствии с масштабом определяется минимально возможный размер ячейки, при котором не будет происходить наложение точек. Ячейки размером меньше заданного агрегируются в более крупные, и среди их точек выбирается либо наиболее важная по совокупности признаков (отбор), либо вычисляется средневзвешенная точка (кластеризация).

В работе [Gröbe, Burghardt, 2021] выполнено сравнение трех алгоритмов отбора точек: на основе отбора по регулярной сетке, на основе функциональной важности, а также методом «дискретной изоляции», в которой ранг точки определяется расстоянием до ближайшей точки с более высоким весом. Авторы отмечают, что при отборе по регулярной сетке нельзя учитывать пространственное распределение точек внутри ячеек, а результат зависит от ее конфигурации: формы ячеек, их размера и смещения. В то же время, алгоритм, основанный на функциональной важности, показал себя достаточно трудоемким и требующим перерасчета весов для каждого масштаба визуализации. Метод дискретной изоляции показал себя наиболее оптимальным образом.

Весовые коэффициенты населенных пунктов могут быть вычислены не только аналитически, но также и с применением машинного обучения, путем сопоставления исходных и вручную генерализованных карт относительно различных метрик, вычисленных для населенных пунктов [Karsznia, Weibel, 2018; Lisiewicz, Karsznia, 2021]. В данных публикациях используется метод деревьев решений, который анализирует такие пространственные факторы, как тип поселения, его административный статус, численность населения, а также соотношение функций поселения (культурная, промышленная, финансовая и т.д.) дополнительно к площади ячеек диаграммы Вороного, ассоциированных с населенными пунктами и расстояниям от них до ближайших соседей. В работе [Xiao et al., 2024] использован подход к отбору точек с применением глубокого обучения и графовых сверточных нейросетей, который опирается на триангуляции точек (она автоматически определяет близость и соседство) и расчет различных параметров. Обучение также происходит на примере вручную отобранных населенных пунктов.

Отметим, что большинство алгоритмов работает по принципу исключения, то есть вместо отбора наиболее значимых точек происходит удаление незначимых. Здесь и далее, чтобы не вводить путаницу, будем понимать под отбором процесс, в результате ко-

того на карте остаются наиболее значимые объекты. Каким образом он реализуется — выборкой или исключением, относится уже к алгоритму.

Алгоритмы выборочного исключения вряд ли могут считаться полноценным средством отбора точечных объектов (в т.ч. населенных пунктов), поскольку они работают локально в окрестности каждой точки, не учитывая географический контекст. В алгоритмах структурного упрощения географический контекст учитывается путем сортировки всего множества точек по отнесенным к ним площадям. Более того, путем введения эффекта «заморозки», при котором удаление каждой точки временно запрещает удаление ее соседей, удается избежать значительного прореживания густонаселенных районов и равномерно распределяет активность отбора точек по всей площади карты [Yan, Weibel, 2008]. Однако в статье показано использование данного алгоритма на абстрактном множестве точек. Применительно к населенным пунктам, при отборе которых играют роль особые географические критерии, данный алгоритм показывает недостаточный уровень контроля над отбором в разных по плотности населения регионах. Это приводит к слабо контролируемому балансу территориальных контрастов. В то же время, в области генерализации имеется положительный опыт региональной дифференциации допусков генерализации применительно к автоматизированной обработке гидрографической сети [Buttenfield, Stanislawski, Brewer, 2013]. При этом стратегия генерализации заключается во введении различных параметров и допусков для районов, отличающихся по географическим условиям.

Отечественный опыт масштабной и региональной дифференциации при автоматизированном отборе населенных пунктов изложен в работе [Krylov et al., 2013], где установлена полезная зависимость между плотностью населенных пунктов и масштабом, а также произведена и апробирована на практике классификация регионов России по плотности населенных пунктов. Недостатком предложенного подхода является отсутствие учета реального рисунка размещения населенных пунктов, что, с одной стороны, выразилось в отборе населенных пунктов по субъектам, а не по их естественным группировкам (в крупных субъектах плотность населения сама по себе неравномерна, а сгустки расселения могут простираться на несколько соседних субъектов), а, во-вторых, в том, что сам процесс отбора основан на атрибутивных признаках и не учитывает локальную плотность расселения (расстояния между соседними пунктами). Сам отбор производится не в автоматическом, а в автоматизированном интерактивном режиме.

### 1.1.3 Прореживание дорожных сетей<sup>2</sup>

Главные задачи отбора дорог — отражение степени развитости дорожной сети в разных районах картографируемой территории, подчеркивание магистральных направлений дорог и связей между населенными пунктами, показ местных особенностей дорог, особенно в слабообжитых районах [Сухов, 1945; Заруцкая, Сваткова, 1982]. При этом для учета территориальных контрастов выполняется дифференциация области картографирования как по густоте дорог, так и по густоте связываемых ими населенных пунктов [Бугаевский, Подольская, 2005]. Ряд исследований был посвящен изучению процентного изменения количества символов дорог при уменьшении масштаба карты. В статье [Li, Choi, 2002] для масштабов от 1 : 1 000 до 1 : 200 000 анализируется связь этих изменений с тематическими атрибутами, в качестве которых взяты следующие: класс, длина, ширина, количество проезжих частей, количество полос движения и связность. Аналогичное исследование касательно изменения графической нагрузки в мелких масштабах картографирования представлено в работе [Krylov *et al.*, 2013].

Задачу отбора дорог с точки зрения пространственных данных необходимо рассматривать как прореживание сетевой модели. Дорожная сеть, в отличие от гидрографии, не имеет однонаправленного характера и содержит множество кольцевых, звездчатых, сетчатых структур [Heinzle, Anders, 2007], что не позволяет выделить ранги, эквивалентные порядкам притоков. Помимо этого, элементы дорожной сети несут семантические и геометрические данные, которые также необходимо учитывать при генерализации. Для обработки такого рода данных применяется теория графов [Mackanness, Beard, 1993]. С использованием различных метрик может быть определена важность отдельных узлов и ребер графа, что служит основанием при отборе сегментов дорожной сети [Li, Choi, 2002; Han *et al.*, 2020]. Удаление ребер графа может основываться на принципе сохранения связности оставшихся ребер и максимизации суммы их весов, зависящих от их важности [Chimani, Van Dijk, Haunert, 2014], а также сохранении основных топологических свойств графа [Pung *et al.*, 2022]. Основные топологические свойства конфигурации транспортных сетей — пространственная связность и изолированность, взаимное расположение элементов, отношения соседства, цикличность и разветвленность линейных элементов сети [Heinzle, Anders, 2007]. В крупных масштабах, где улично-дорожная сеть выражается явно на картах населенных пунктов, выполняется обогащение пространственных данных путем выявления рисунка сети дифференциация критериев отбора дорог для городских и сельских территорий [Touya, 2010], явный учет границ населенных пунктов и выходов основных магистралей через них [Lyu *et al.*, 2022].

<sup>2</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov, Krivosheina, 2012; Самсонов, Кривошеина, 2015b]

В работе [Thomson, Brooks, 2007] предложен алгоритм, основанный на принципе «штрихов» (strokes), который заключается в объединении ребер графа, стыкующихся под малым углом и таким образом наиболее вероятных для прокладки маршрута. В сети оставляются «штрихи», объединяющие дороги высоких классов. Недостатком данного подхода является невозможность учета неравномерной густоты дорожной сети, что было исправлено авторами алгоритма [Liu, Zhan, Ai, 2010], в котором густота учитывается путем построения диаграммы Вороного для ребер графа (аналогично рассмотренному ранее методу отбора населенных пунктов). В работе [Chen et al., 2009] предложен «площадной» подход к исключению ребер графа дорожной сети, основанный на построении сетки кварталов по дорогам и исключении ребер в тех кварталах, которые имеют размер ниже заданного порога (т.е. высокую плотность дорожной сети). В дальнейшем получила развитие комбинация штрихового и площадного подходов в комбинированные алгоритмы [Li, Zhou, 2012; Benz, Weibel, 2014].

Описанные подходы применимы главным образом для крупномасштабного и среднемасштабного картографирования. Для отбора дорог в целях мелкомасштабного картографирования в работе [Weiss, Weibel, 2014] предложено использовать показатель центральности по промежуточности, который отражает долю от общего числа всех маршрутов в сети, которая проходит через данное ребро дорожного графа. Поскольку мелкомасштабное отображение дорог связано с очень сильной степенью отбора, соответствующим образом повышается и разнообразие критериев, которые заложены в этот процесс. Несмотря на то, что общие принципы мелкомасштабного отбора давно исследованы и задокументированы [Сухов, 1945; Заруцкая, Сваткова, 1982], в каждом конкретном случае экспертиза картографа могла допускать индивидуальный подход к объектам.

Для определения таких индивидуальностей, выходящих за рамки общих инструкций отбора, так же как и в случае с отбором точечных объектов, может быть использовано машинное обучение [Karsznia, Wereszczyńska, Weibel, 2022; Karsznia et al., 2024]. В этих работах показано, что реализация алгоритмов исключительно по руководствам и наставлениям приводит к менее качественным результатам, чем подход, основанный на машинном обучении, в рамках которого модель отбора (дерево решений) обучается на сопоставлении эталонных вручную генерализованных картах, используя в качестве предикторов различные семантические характеристики дорог, полученные в результате обогащения базы пространственных данных. Вычисление этих характеристик может выполняться не явным образом стандартными операциями ГИС, а средствами моделей глубокого обучения, таких как графовые сверточные нейросети, которые способны самостоятельно извлекать информацию, позволяющую предсказывать отбор тех или иных элементов дорожной сети [Zheng et al., 2021; Tang et al., 2024].

В целом, следует отметить, что вопросы мелкомасштабной генерализации дорожных сетей на основе пространственных данных получают значительно меньше внимания, чем аналогичные задачи в крупных и средних масштабах. Однако во всех работах

проявляется необходимость адаптации алгоритмов отбора к особенностям организации дорожной сети.

#### 1.1.4 Геометрическое упрощение линейных объектов<sup>3</sup>

Геометрическое упрощение линий — несомненно, самая проработанная область генерализации в картографии и геоинформатике. Традиционные алгоритмы геометрического упрощения линий, применяющиеся для этих целей, характеризуются следующими особенностями:

- *глобальность* — линия подвергается обработке целиком, без разделения на сегменты;
- *точечность* — фактическим объектом анализа являются вершины линии, а не сама линия; процесс генерализации предваряется определением значимости каждой вершины;
- *консервативность* — новые вершины в линию не добавляются, существующие не смещаются, происходит только удаление исходных вершин; множество вершин упрощенной линии является подмножеством множества исходных вершин;
- *нетопологичность* — каждая линия подвергается геометрическому упрощению без учета ее топологических отношений с другими объектами: точками, линиями, полигонами, как внутри текущего слоя, так и по отношению к объектам других слоев.

К методам, обладающим данными свойствами, следует отнести алгоритмы, предложенные в классических работах Дугласа-Пейкера [Douglas, Peucker, 1973], Розенфельда-Джонстона [Rosenfeld, Johnston, 1973] и Висвалингам-Уайатта [Visvalingam, Whyatt, 1993]. В первом алгоритме значимость каждой точки определяется по ее расстоянию от стягивающей хорды, во втором — углом поворота (кривизной) линии в данной точке, в третьем — по площади треугольника, который образуется предыдущей, текущей и последующей вершиной. Существует также целый ряд других подходов аналогичного типа [Li, 2006]. Перечисленные алгоритмы обладают недостатками. Глобальность и точечность исключают возможности анализа кривой и определения ее основных изгибов, консервативность ограничивает возможности сдвигов, преувеличений и согласования геометрии нескольких линий, а нетопологичность исключает внутренний механизм подобного согласования. Традиционные алгоритмы не могут обеспечить требуемого качества картографической генерализации, поэтому их называют алгоритмами редуцирования точек, подчеркивая механистическую природу заложенных в них геометрических операций [Li, 2006].

<sup>3</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Самсонов *et al.*, 2014; Samsonov, Yakimova, 2017]

Недостатки перечисленных алгоритмов преодолеваются. Так, в алгоритме Ли-Оупеншоу [Li, Openshaw, 1992] использован «естественный принцип», суть которого заключается в удалении деталей, размер которых менее видимой величины. Для этого на исходное множество линий накладывается сетка с заданным шагом и удаляются все изгибы, находящиеся внутри ячеек. Наиболее продуманным и приближенным к географической постановке задачи является алгоритм Ванга-Мюллера [Wang, Müller, 1998], известный также как «Bend Simplify». В основе его работы лежит сегментация линий на отдельные изгибы, которые подвергаются удалению, слиянию и преувеличению. Подход Ванга-Мюллера успешно разрешает первые три недостатка традиционных алгоритмов, однако в него не заложено сохранение топологии, что затрудняет генерализацию множества линий, требуя постобработки. В то же время, методы манипуляции геометрией объектов с сохранением топологии разработаны в вычислительной геометрии [Dey et al., 1999]. Аналогичные методы применяются в геоинформационном картографировании для повышения качества работы алгоритмов редуцирования точек [de Berg, van Kreveld, Schirra, 1998]. Однако общей проблемой существующих в данной области методов является работа с одним слоем данных, в то время как для обеспечения достоверности картографирования алгоритм геометрического упрощения линий должен поддерживать учет топологических отношений между множеством слоев карты.

Важным аспектом генерализации является оценка результатов. Распространенным подходом к численной оценке близости исходной линии и генерализованной является Хаусдорфово расстояние [Hangouët, 1997] между множествами их узлов. Для каждого узла одной линии находится расстояние до ближайшего узла другой линии. Аналогичным образом процедура выполняется для второй линии. Из всех полученных расстояний (в обе стороны) выбирается максимальное. Хаусдорфово расстояние дает лишь показатель близости линий. В то же время, известно, что при оценке генерализации не меньшее внимание уделяется географическому правдоподобию линии, т. е. ее морфологическому соответствию реальному объекту. Берлянт, Мусин, Собчук [1998] исследовали поведение таких величин, как извилистость, удельная кривизна и длина линии при уменьшении масштаба на примере генерализации. В качестве меры извилистости они использовали простой коэффициент, выражающий число изгибов на единицу длины линии. Анализ карт различных масштабов привел их к выводу о том, что кривизна линий уменьшается при переходе к более мелкому масштабу, в то время как длина и извилистость не имеют масштабной зависимости.

Один из крупных специалистов по алгоритмам генерализации, Роберт Вайбель отметил, что многие алгоритмы, разработанные для обобщения линий, пренебрегают значительными ограничениями, которые хорошо известны в картографической практике [Weibel, 1996]. Следовательно, результаты могут быть использованы для незначительного упрощения. Среди ограничений, необходимых для использования в упрощении линий, Вайбелем были отмечены необходимость *минимизировать искажения формы (мет-*

рическое) и сохранять исходный характер линии (*гештальт*). Первое касается сохранения длины и угловатости линии с минимальным пространственным смещением. Второе требует сохранения характера линии, отражающего признаки процесса, посредством которого она была сформирована.

Характер и структуру формы линии можно оценить из анализа координат в разных (и расширяющихся) пространственных контекстах, включая отдельные вершины, изгибы и отрезки. Обычно характеристики, полученные в отдельных вершинах, такие как угол поворота или эффективная площадь, используются в качестве критериев в алгоритмах обобщения линий, например для идентификации характерных точек [Dutton, 1999]. В работе [Nie, Huang, 2011] показано, что характерные ключевые точки могут быть обнаружены с использованием универсального алгоритма, который основан на вычислении кривизны линии в локальном и глобальном контексте [Chen He, Yung, 2008]. Ключевые точки используются для разделения линии на сегменты, а затем для обобщения каждого сегмента с использованием алгоритма Ли-Оупеншоу. Морфологические свойства линий между ключевыми точками при этом не рассматриваются.

Характеристики, вычисленные на уровне точек, дают основу для описания формы последовательности точек и последующей сегментации линии на однородные сегменты [Buttenfield, 1991]. Промежуточный уровень между точечным и сегментным анализом занимает анализ *изгибов*, введенный в статье [Wang, Müller, 1998]. Ванг и Мюллер формализовали термин «изгиб» как отрезок линии с постоянным знаком угла поворота (перегиба) и предложили различные операции генерализации изгибов, включая *исключение* (для небольших изгибов), *комбинацию* (для близких аналогичных изгибов) и *преувеличение* (для изолированных изгибов). Форма индивидуального изгиба описывается различными численными характеристиками, такими как размер (площадь) и компактность. Хотя методика Ванга и Мюллера, ориентированная на изгибы, очень близка к ручной методологии, она не превратилась в новые алгоритмы обобщения, основанные на ее концепциях, и подход, основанный на сегментации, доминирует в исследованиях по упрощению геометрических линий до сих пор.

Баттенфилд [Buttenfield, 1987] была одной из первых, кто подчеркнул необходимость анализа формы линии для выбора параметров алгоритмов генерализации. В своей работе она подчеркнула, что величина допуска при генерализации линий должна иметь возможность меняться для сохранения реалистичности линии (подразумевая под реалистичностью сочетание точности и узнаваемости). С этой целью она ввела иерархические метрики геометрии, которые неявно учитывают масштабную-зависимую природу картографических деталей и предложила цифровой метод идентификации типов линий на основе этих метрик. Эти метрики были положены в основу выбора величин допуска.

В работах [Plazanet, Affholder, Fritsch, 1995; Plazanet, Bigolin, Ruas, 1998] был предложен метод сегментации линий на фрагменты с различными характеристиками формы. Сегментация выполнялась с целью последующего выбора подходящих инструментов ге-

нерализации и их параметров. Авторы применяли иерархический процесс сегментации, в котором линия рекурсивно расщепляется на сегменты до тех пор, пока все они не станут однородными по своим морфологическим характеристикам. Для каждого сегмента можно вычислить несколько показателей, что позволяет провести их дальнейшую классификацию. Наконец, для каждого класса применяются собственные процедурные решения генерализации [Plazanet, Bigolin, Ruas, 1998]. Редким примером комбинации анализа изгибов и сегментации является работа [Ai et al., 2017], в которой сначала выделяются критические точки линии, разбивающие ее на сегменты, а затем последовательности изгибов внутри каждого сегмента огибаются по средней линии, либо правой/левой огибающей в зависимости от соотношения площадей изгибов по правую и левую сторону линии.

Мюстьер [Mustiere, 2005] предложил локальный адаптивный подход к обобщению дорог, в котором различные алгоритмы могут последовательно применяться к каждой части дороги. Выбор подходящего алгоритма определяется с использованием контролируемых методов машинного обучения, основанных на определенном наборе мер. В работе [Garcia Balboa, Ariza López, 2007] посредством искусственных нейронных сетей была выведена классификация дорог, базирующаяся на метриках извилистости линий и ориентированная на выполнение процедур генерализации. В схожем ключе выполнена и работа по сегментации береговых линий [Saga, 1995].

Парк и Ю [Park, Yu, 2011] разработали гибридный подход к геометрическому упрощению линий. Предложенный подход к сегментации линий, так же как и предыдущие, основан на вычислении количественных характеристик линий. Авторы проанализировали эффективность существующих алгоритмов упрощения, обращая внимание на то, какой алгоритм приводит к более точному представлению линии по метрикам отклонения от исходной. Для тестирования на разных геометрических линиях были выбраны алгоритмы Дугласа-Пейкера [Douglas, Peucker, 1973], алгоритм «sleeve-fitting» [Zhao, Saalfeld, 1997] и алгоритм на основе поворотной функции [Rangayyan et al., 2008]. Для каждой линии выбирается алгоритм, обеспечивающий наименьшую позиционную ошибку среди всех алгоритмов, то есть наименьшее отклонение от исходной линии. Метод был испытан на примере зданий, дорог и рек с топографических карт. В этом подходе есть несколько важных вопросов, которые следует отметить для следующего анализа:

- Для выбора наилучшего алгоритма используется только позиционная точность; характер линии в расчет не принимается. Это означает, что технология не может гарантировать картографическое качество результатов.
- Кластеризация используется для получения данных обучения, а количество классов должно быть выбрано пользователем. Хорошо известно, что кластеры не всегда четко идентифицируются; это может привести к случайному попаданию сегментов линии в неподходящие классы.

- Выбранные примеры в основном однородны по своей природе (например, здания, дороги и реки) и разделены на разные слои, что делает сегментацию более простой, чем когда линии разной формы смешиваются в одном наборе данных, как в случае административных границ.

Все упомянутые методы основаны на анализе количественной информации о форме линии и ее изменчивости. Однако общий характер линии, который относится к гештальт-характеристикам согласно типологии Вайбеля [Weibel, 1996], является скорее качественной, нежели количественной характеристикой, и может быть выражен в таких интуитивных понятиях, как *прямоугольный*, *гладкий*, *остроугольный*, *схематичный* и т.д. Гештальт-характеристики могут быть использованы двояко при генерализации. Во-первых, они могут работать и как предварительное знание о характере линии, и как ограничение во время генерализации. В этом случае алгоритм должен упрощать геометрию объекта определенным образом, который позволяет избавиться от несущественных деталей и одновременно сохранить характер линии. Этот подход наиболее разработан в обобщении зданий, где в качестве гештальт-характеристики используется ортогональность формы [Staufenbiel, 1973]. Во-вторых, характеристика формы может использоваться только как ограничение, без предварительного знания характера линии. Этот вариант обычно используется в алгоритмах *схематизации* линий, которые преобразуют исходную форму в некоторую упрощенную карикатуру, подчиняющуюся математическим правилам аппроксимации формы, например ортогональными сегментами [Meulemans, van Renssen, Speckmann, 2010], дугами окружностей [van Goethem et al., 2015] или с возможностью выбора допустимых азимутов линий [Buchin et al., 2016].

Одной из ранних работ по геометрическому упрощению зданий была диссертация [Staufenbiel, 1973], в которой были предложены общие принципы разработки соответствующих алгоритмов. Зестер [Sester, 2005] разработала подход к генерализации зданий, основанный на применении метода наименьших квадратов для ортогонализации формы и сохранения ортогональности. В работе [Shan, Sampath, 2013] разработан алгоритм построения здания ортогональной формы по данным лазерного сканирования. Аналогичная задача, но для оцифровки наборов данных растровых построений была решена в статье [Gribov, Bodansky, 2006]. Существует также семейство рекурсивных подходов к генерализации зданий [Bayer, 2009], а также семейство методов, основанных на применении операторов математической морфологии [Cámara, López, 2000; Damen, van Kreveld, Spaan, 2008].

Оптимизационный метод топологического упрощения контуров зданий был предложен в работе [Hauert, Wolff, 2010a]. Их достаточно сложная методика основана на выборе подмножества ребер многоугольника, представляющего здание. Затем эти ребра соединяются так называемыми *шорт-катами* (short-cuts), которые строятся путем продолжения или обрезки отобранных ребер. Решение получается путем нахождения ми-

нимального количества выходных ребер, которое удовлетворяет заданному пороговому значению расстояния (отсюда — оптимизация). Алгоритм упрощения полигональных разбиений, представленный [Buchin, Meulemans, Speckmann, 2011; Buchin *et al.*, 2016], основанный на процедуре *перемещения ребра*, также позволяет учитывать различные ограничения при генерализации.

В большинстве рассмотренных алгоритмов для всех объектов применяется глобальная стратегия обработки. В то же время, унифицированный подход не представляется подходящим для наборов данных с гетерогенным характером линий, таких как случай административных границ, представленный выше. Необходимость применения разных обобщающих подходов в разных пространственных контекстах всячески подчеркивается современными исследователями [Touya, Duchêne, Ruas, 2010; Buttenfield, Stanislawski, Brewer, 2013], что подразумевает необходимость разработки адаптивного к форме подхода к упрощению картографических линий. Такой подход должен быть нацелен на сохранение общего характера линии и требует разработки алгоритмов распознавания характеристик линии для выполнения сегментации. Если такая *качественная* сегментация выполнена, то каждая часть линии может быть далее обобщена с помощью наиболее подходящего алгоритма, либо сегментирована далее на фрагменты, которые являются однородными в терминах *количественных* метрик [Buttenfield, 1987; Plazanet, Bigolin, Ruas, 1998; Mustiere, 2005]. Высокий потенциал в повышении адаптивности геометрического упрощения к особенностям линейных объектов имеют появившиеся в последнее время подходы, основанные на глубоком обучении [Du *et al.*, 2022; Yu, Chen, 2022; Yang *et al.*, 2022; Knura, 2023; Yan, Yang, 2024], однако их применимость к разнообразным типам географических линий для картографирования в разных масштабах требует дальнейших исследований.

### 1.1.5 Генерализация цифровых моделей рельефа<sup>4</sup>

К обобщенному изображению рельефа предъявляются следующие требования [Заруцкая, 1958]:

- сохранение на карте форм (положительных и отрицательных), имеющих размеры больше установленной величины или типичных для изображаемого ландшафта;
- сохранение географического подобия изображаемых форм путем выделения характерных особенностей их рисунка;
- получение наглядности в изображении обобщенных форм, т.е. создание впечат-

<sup>4</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov, 2011; Лурье, Самсонов, 2011; Самсонов, 2011]

ления их выпуклости, объемности;

- сохранение определенной степени точности местоположения и высоты крупных форм.

При обобщении изображения рельефа средствами цифровых технологий необходимо использовать методы генерализации цифровых моделей рельефа. Вайбелем были сформулированы основные требования к таким методам [Weibel, 1987]:

- Максимальная степень автоматизации.
- Эффективность в широком диапазоне масштабов.
- Адаптируемость к характеристикам рельефа.
- Оперирование непосредственно ЦМР.
- Возможность анализа результатов.
- Возможность сдвигов и преувеличений форм, основанных на распознавании основных орографических элементов и отдельных форм рельефа (для значительной генерализации).

Эти требования наряду с географическими принципами могут быть учтены при разработке, анализе и применении методов генерализации ЦМР. Анализ современного состояния исследований по генерализации сеточных (растровых) ЦМР позволяет выделить следующие группы методов:

- передискретизация и агрегирование ячеек;
- глобальная фильтрация;
- структурная генерализация;
- адаптивная фильтрация и интерполяция;
- спектральная декомпозиция на разночастотные составляющие.

Проанализируем эти методы, обратив внимание на их преимущества и недостатки.

**Передискретизация** позволяет изменить разрешение модели путем интерполяции и может быть использована как непосредственный инструмент генерализации [Li, Ho, 2004]. Однако в процессе передискретизации не учитывается структурность рельефа, при этом модель обычно получается излишне детальной по отношению к разрешению (если она предварительно не генерализована). Основное назначение передискретизации — приведение разрешения модели в соответствии с ее подробностью. Близким к передискретизации методом генерализации является агрегирование ячеек — их объединение. Если коэффициент агрегирования равен 2, объединяются 4 соседних ячейки. Если 3, то объединяются 9 ячеек. Значение «укрупненной» ячейки получается осреднением значений исходных ячеек внутри плавающего окна.

Суть **глобальной фильтрации** заключается в последовательном проходе всех ячеек ЦМР с помощью «плавающего окна», внутри которого вычисляется некая величина

на основе попадающих в его пределы значений высоты. Результат присваивается ячейке, попадающей в центр плавающего окна, представляющего собой обычно квадратную, прямоугольную или круговую область. Для целей генерализации используются сглаживающие фильтры, которые осредняют значения ячеек [Loon, 1977]. Наиболее популярны такие сглаживающие фильтры, как гауссова фильтрация, вычисление среднего и вычисление медианы. Недостатком фильтрации является выполаживание всех без исключения форм рельефа. Как отмечает Вайбель [Weibel, 1987], метод фильтрации годится только для незначительного обобщения рельефа, когда изменения масштаба картографирования несут существенны и требуется лишь приглушить мелкие неровности. В большинстве случаев требуется более качественная генерализация, которая может быть осуществлена с привлечением информации о структурных линиях рельефа.

Идея привлечения **структурных линий** к генерализации ЦМР появилась после того, как методы фильтрации не оправдали себя. Первые опыты по использованию структурных линий были проведены еще в 80-е годы [Wu, 1981; Weibel, 1987; Yoeli, 1990]. По цифровой модели автоматизированно или вручную выделяется подробная сеть структурных линий (хребты и тальвеги) с известными высотами узловых точек. Затем эта сеть генерализуется путем геометрического упрощения, сдвигов и удаления мелких отрогов и тальвегов. Наконец, на основе генерализованной сети структурных линий восстанавливается поверхность одним из методов интерполяции. Основную сложность представляет автоматизированное выделение структурных линий. Если с тальвегами проблема более-менее решена [O'Callaghan, Mark, 1984], то с хребтами ситуация сложнее [Band, 1986; Lindsay, Seibert, 2013; Гарцман, Шекман, Ли, 2016]. В случае же рельефа с нечетко выраженными структурными линиями (например, холмисто-западинный), ситуация фактически заходит в тупик.

Вайбелем [Weibel, 1987] был предложен метод адаптивной генерализации, который позволил обойти проблему нечеткости структурных линий. Для слаборасчлененного рельефа и при незначительном уменьшении масштаба использовалась фильтрация модели. В других случаях — генерализация на основе структурных линий с последующей интерполяцией между ними. Выбор между методами осуществлялся на основе экспертной оценки или автоматически на основе морфометрической характеристики рельефа. Однако такой подход не столько устраняет недостатки фильтрации и структурной генерализации, сколько объединяет их. Более сбалансированный подход используется в методах адаптивной фильтрации и интерполяции.

**Адаптивные методы** используют структурные линии и морфометрические коэффициенты для локальной адаптации фильтра или локальной интерполяции значений. Теоретически такой подход позволяет определить, какие формы рельефа надо сгладить (удалить) фильтром или интерполяцией, какой конкретно тип фильтра необходимо использовать (среднее, максимум, отклонение и т.п.) и где обработку не надо делать вообще.

Закшек и Подобникар предложили комбинировать структурные линии, характерные точки и саму модель в качестве входных данных для интерполяции [Zakšek, Podobnikar, 2005]. При этом структурные линии определялись с помощью детектора границ на основе перепада яркости по аналитической отмывке, а характерные точки (пики и впадины) — на основе разности высот между исходной и сглаженной ЦМР. Исходная ЦМР подвергалась фильтрации и передискретизации, затем ее ячейки в качестве точечных данных участвовали в моделировании вместе с «несглаженными» ячейками структурных линий и характерных высот. Предложенный подход, однако, дает достаточно невыразительные результаты генерализации ЦМР. Йорданом [Jordan, 2007] предлагается метод, основанный на удалении долин определенного порядка. Для этого сначала выделяется и классифицируется по Стралеру сеть водотоков. Для каждого водотока строится его бассейн. Точки, лежащие на водоразделах (границах бассейнов) используются в качестве узлов локальной триангуляции. Выбирая иерархический уровень водотока, можно на основе триангуляции интерполировать новые значения внутри бассейна и таким образом «заполнить» долину, т.е. удалить ее. Этот подход эффективно решает проблему заполнения эрозионных форм, не выражающихся в масштабе, но не решает проблему генерализации целиком.

Аи и Ли [Ai, Li, 2010] также используют выделение тальвегов и бассейнов. Процесс разбивается на 3 этапа: генерализация сети водотоков; расширение границ крупных долин путем поглощения малых, соответствующих удаленным водотокам; и сглаживание модели в областях поглощенных долин. В работе [Fan, Yang, Hu, 2007] предлагается четырехстадийный метод генерализации, использующий три вида фильтров: фильтр низкой частоты (low-pass), пороговый и сглаживающий (под которым авторы на самом деле понимают локальную реинтерполяцию ЦМР по высотам водораздельных линий аналогично методу [Jordan, 2007]). Для локальной адаптации фильтров используются морфометрические коэффициенты — кривизна и уклон поверхности. Метод, предложенный Леонович и Йенни [Leonowicz, Jenny, Hurni, 2009], использует разные фильтры для долин и хребтов, что позволяет преувеличивать крупные как положительные, так и отрицательные формы рельефа путем поглощения мелких форм. Для этих целей используется взвешенная сумма двух версий ЦМР, обработанной фильтрами нижней четверти (преувеличены отрицательные формы) и верхней четверти (преувеличены положительные формы), где веса определяются расстоянием до ближайшего тальвега.

**Спектральный анализ** дает возможность представить поверхность в виде суммы разночастотных составляющих. Традиционным методом спектральной декомпозиции является преобразование Фурье [Clarke, 1988; Пузаченко, Онуфреня, Алещенко, 2002]. Отбросив высокочастотные составляющие, можно получить генерализованное представление поверхности. Более современным методом спектральной декомпозиции является двухмерное вейвлет-преобразование [Daubechies, 1992], которое в компьютерной графике нашло применение в оптимизации скорости отображения поверхностей [Gross,

Gatti, Stadt, 1995; Bonneau, 1998] и их мультимасштабном анализе [Lounsbury, DeRose, Warren, 1997]. Исследования показали, что оно может быть использовано и как инструмент генерализации рельефа [Wu, 2000; Kalbermatten *et al.*, 2009]. Представляется также перспективным использование для генерализации ЦМР спектрально-аналитического метода, предложенного И.В.Флоринским и А.Н.Панкратовым [Florinsky, Pankratov, 2016], который комбинирует в себе возможности аппроксимации поля высот, удаления шумов, генерализации, и ориентирован прежде всего не геоморфометрические задачи.

Алгоритмы упрощения **триангуляционных моделей** можно разделить на 2 группы: алгоритмы детализации («сверху вниз») и алгоритмы опустошения («снизу вверх») [Lee, 1991; Pedrini, 2001; Скворцов, 2002]. И в том и в другом случае в качестве граничного условия выступает максимальное отклонение по высоте относительно исходной модели, либо максимальное количество треугольников. Алгоритмы детализации работают по принципу «сверху вниз» [Fowler, Little, 1979]. Исходная поверхность аппроксимируется минимальным количеством треугольников, необходимым, чтобы покрыть все точки. Далее в триангуляцию встраиваются новые узлы, пока не будет достигнуто требуемое разрешение. Алгоритмы опустошения работают в «прямом» порядке («снизу вверх»), т.е. берется исходная триангуляция и из нее удаляются элементы, удовлетворяющие допуску упрощения. Стратегия детализации, как правило, дает более точные результаты [Скворцов, 2002]. Однако могут быть использованы различные методы оптимизации алгоритмов опустошения. Л. Коббельтом было показано, что при коллапсе ребра или треугольника место размещения нового узла на месте удаленного элемента большой роли не играет. С точки зрения сохранения морфологии поверхности гораздо важнее сам факт того, удален элемент или нет, а на это влияет метрика, используемая для вычисления отклонения [Kobbelt, Campagna, Seidel, 1998]. Исследования в этой области направлены в основном на: а) разработку оптимальных метрик для вычисления отклонения и б) учет топографической значимости удаляемых элементов.

В работе [Fowler, Little, 1979] предложен метод генерализации сеточной модели с преобразованием ее в триангуляционную. Основные узлы триангуляции отбираются путем фильтрации исходной ЦМР с помощью скользящего окна  $2 \times 2$ , которая выделяет точки перегиба поверхности (вершины, понижения, тальвеги и т.д.). После того как получен каркас из основных узлов, дальнейшие действия идут по вышеописанной схеме. Таким образом, этот алгоритм более географичен и учитывает основные структурные элементы поверхности. Чен и Гевара предлагают похожий алгоритм, который основан на выделении важнейших точек [Chen, Guevara, 1987]. Важность точки интерпретируется как ее вклад в формирование локальной формы поверхности и вычисляется как разница между ее действительной высотой и высотой, полученной на основе интерполяции 8-ми ее соседних точек. Далее точки сортируются по степени важности и отбираются на основе минимальной величины важности либо на основе требуемого количества. Полученный набор ключевых узлов используется как основа триангуляции. Педри-

ни [Pedrini, 2001] предлагает сначала вычислять среднеквадратическое отклонение высот вокруг каждой точки в пределах плавающего окна  $3 \times 3$ , и использовать его в качестве критерия «важности» точки. Этот критерий используется в паре с ошибкой аппроксимации поверхности — расстоянием от узла до покрывающего треугольника.

Шрёдер и Россбах в качестве критерия значимости точки предлагают использовать расчлененность поверхности [Schröder, Roßbach, 1994]. Вершина удаляется из триангуляции только, если она не вносит весомого вклада в форму поверхности. В работах [Gueziec, Hummel, 1995; Pedrini, 2001] одним из критериев возможности коллапса ребра является незначительное изменение ориентации нормалей к изменившимся треугольникам. Таким образом, используется не высотная, а угловая метрика, характеризующая изменение экспозиции и уклона поверхности в результате генерализации. Алгоритм с выделением топографических элементов поверхности излагается в работе [Li, Wang, Yang, 2008]. В качестве критерия наличия важного элемента (вершина, хребет, тальвег и т.п.) выступает *средняя кривизна* поверхности, для вычисления которой она локально аппроксимируется параболоидом. Одно из важнейших свойств триангуляции — возможность встраивания в ее структуру линий и полигонов. Триангуляция с ограничениями использует сегменты заданных линейных и полигональных объектов в качестве ребер триангуляции [Preparata, Shamos, 1985]. Это позволяет фиксировать в структуре модели линии водотоков, хребтов, водоразделов, а также плоские участки, например озера [Buys, Messerschmidt, Botha, 1991].

В работе [Fei, Huang, He, 2008] предложен и разработан **трехмерный вариант алгоритма Дугласа-Пейкера**, который может быть использован для генерализации любых ЦМР, как сеточных, так триангуляционных и изолинейных. Идея трехмерного алгоритма аналогична оригиналу [Douglas, Peucker, 1973], где берется начальная точка линии (она становится «якорной»), соединяется прямой с конечной точкой (она становится «плавающей»), после чего среди промежуточных точек выбирается та, что наиболее удалена по перпендикуляру от получившейся прямой. Если расстояние от нее до прямой больше порога генерализации, точка становится плавающей, якорная остается на своем месте, и процесс повторяется. Если меньше — то все точки между якорной и плавающей отбрасываются, плавающая точка становится якорной, а в качестве новой плавающей точки опять выбирается последняя точка линии. Если плавающая точка совпала с якорной, то алгоритм завершил свою работу.

Существующие подходы к **генерализации горизонталей (изолинейных ЦМР)** можно разделить на 2 большие группы: те, которые оперируют непосредственно линиями, и те, которые производят обобщение косвенно — путем генерализации растровой или триангуляционной ЦМР на основе горизонталей [Chen, 1987; Weibel, 1992]. Генерализация горизонталей сопряжена с множеством проблем. Основная из них заключается в том, что алгоритм должен оперировать не каждой линией по отдельности, а совокупностью горизонталей в пределах участков, обеспечивая таким образом генерализацию

целых форм [Wu, 1981; Peng, Pilouk, Tempfli, 1996]. Попытки использовать глобальные алгоритмы удаления узлов наподобие Дугласа-Пейкера приводят к рассогласованности горизонталей. Наилучшие результаты генерализации достигаются на основе предварительно выделенных структурных линий [Fei, 1993]. Однако выделение структурных линий по горизонталям [Tang, 1992; Fei, 1993; Ai, 2007] существенно сложнее, чем по растровым или триангуляционным ЦМР [O'Callaghan, Mark, 1984], поскольку требует выявления топологических связей, соседства горизонталей.

В работе [Li, Sui, 2000] предложен подход, который использует для генерализации изолиний алгоритм Ли-Оупеншоу [Li, Openshaw, 1992], основанный на «естественном» принципе: удаляются все изгибы линии меньше видимой величины (авторы экспериментально установили, что она равна 0.6–0.7 мм в масштабе). Для этого на исходные линии накладывается растровая сетка с шагом видимости и участок каждой линии внутри квадрата заменяется одним узлом. К преимуществам этого метода можно отнести, что он исключает как самопересечения линий, так и взаимные пересечения соседних горизонталей. К недостаткам — то, что он является глобальным и не учитывает структурные линии и формы рельефа. В работе [Zhang et al., 2007] для контроля топологии (пересечений) горизонталей между ними дополнительно строится триангуляция Делоне, после чего они упрощаются с помощью алгоритма Дугласа-Пейкера.

Еще одна сложность при генерализации горизонталей, на которую редко обращается внимание — то, что новое сечение может быть не кратным предыдущему (например, было 10 м, а стало 25 м). Это требует интерполяции промежуточных горизонталей. В работах [Peled, Loon, Bossler, 1989; Li, Sui, 2000] можно найти варианты решения этой задачи, в том числе на основе триангуляции Делоне с ограничениями между горизонталями и последующей трассировки промежуточных высотных уровней. Однако эти алгоритмы работают хорошо далеко не во всех случаях, что опять же связано с множеством вариантов взаимного расположения горизонталей. Также при генерализации горизонталей сложно оценить высотную точность получившейся модели. Если в случае растровой или триангуляционной модели достаточно посчитать отклонения по высоте в каждой точке, то с горизонталями так не получится: линии при сглаживании смещаются, и для того, чтобы определить отклонение относительно исходной модели в каждом узле, потребуется интерполяция между горизонталями. Чаще оценивается плановая точность расположения горизонталей — их отклонение от оригинала. Разработаны специальные алгоритмы обобщения с заданной плановой точностью [Gökgöz, 2005; Cetinkaya et al., 2006].

Стоит также отметить, что для представления рельефа часто используются мультимасштабные модели [de Floriani, Marzano, Puppo, 1996], которые можно отнести к иерархическим и пирамидальным.

*Иерархические модели* реализуются путем рекурсивной детализации области определения, которая на начальном этапе покрывается минимальным набором граней (треугольников, квадратов). Далее каждая из этих граней детализируется вставкой новых уз-

лов, пока полученная аппроксимация модели не будет отклоняться от исходного набора точек менее чем на заданную величину. Поскольку новые элементы создаются внутри существующих, последовательность разбиения может быть представлена в виде иерархического дерева — отсюда и название данного типа модели. Среди иерархических моделей можно выделить основанные на квадродереве [Chen, Tobler, 1986; Von Herzen, Barr, 1987], кватернарной триангуляции [Gomez, Guzman, 1979], тернарной триангуляции [De Floriani *et al.*, 1984]. Адаптивная иерархическая триангуляция, предложенная в работе [Scarlatos, 1990; Scarlatos, Pavlidis, 1992] обеспечивает явную мультимасштабность и в среднем хорошую форму треугольников, не требуя при этом создания новых узлов [De Floriani, Puppo, 1992]. Основная идея этого метода — использовать триангуляцию Делоне, и каждый раз при добавлении нового узла перестраивать ее целиком, а не соединять новую точку ребрами с углами треугольника, внутри которого она находится. Разновидность иерархических моделей составляют сфероидические иерархические модели [Bernardin *et al.*, 2011]. Для сфероидических моделей необходимо использовать специальные методы расчета морфометрических характеристик [Florinsky, 1998; Florinsky, 2017]. Такие модели удобны для исследований континентального и планетарного уровня, выделения линеаментов и планетарных структур [Флоринский, 2009].

*Пирамидальные модели* отличаются от иерархических тем, что каждый уровень детализации строится глобально и независимо от предыдущего. Он строится на основе полного множества точек, а не собирается по кусочкам, получаемым разбиением граней более низкого уровня. По сути, это послойное представление цифровой модели в виде нескольких уровней детализации. Пирамида Делоне — стандартный вариант пирамидальной ЦМР, каждый уровень детализации которой представляет собой триангуляцию Делоне исходного множества точек с заданной ошибкой аппроксимации [De Floriani, 1989].

Можно также выделить следующие особенности рассмотренных методов построения мультимасштабных ЦМР:

1. *Формальность.* Критерием генерализации является ошибка аппроксимации по высоте относительно исходного множества узлов. В то же время, как было установлено ранее, существуют продвинутое методы генерализации как сеточных, так и триангуляционных ЦМР, учитывающие структурные линии и расчлененность. Необходима интеграция этих методов в алгоритмы построения мультимасштабных ЦМР. Ошибка по высоте не может быть единственным критерием генерализации.
2. *Использование исходных узлов.* Практически все алгоритмы, за исключением адаптивной триангуляции, генерируют уровни детализации как подмножества исходных узлов с теми же высотами и координатами. Такой подход исключает возможность сдвигов и преувеличений форм рельефа. В то же время эти операции неотъ-

емлемы в географической генерализации, без них изображение рельефа может быть невыразительным и морфологически неправдоподобным.

### 1.1.6 Прочие аспекты генерализации<sup>5</sup>

Рассмотренные задачи отбора точечных объектов, прореживания дорожных сетей, геометрического упрощения линий и генерализации цифровых моделей составляют лишь некоторую долю общей проблематики исследований в области генерализации. Среди прочих направлений следует выделить специализированные методы *генерализации полигональных разбиений*, которые используются для картограмм и карт качественного фона [de Berg, van Kreveld, Schirra, 1998; Haunert, Wolff, 2010b; Buchin, Meulemans, Speckmann, 2011; Buchin *et al.*, 2016; Meijers, Savino, Van Oosterom, 2016], в том числе с использованием специальных структур данных, таких как tGAP [Van Oosterom, 2005; Meijers *et al.*, 2020]. Также большой интерес привлекает задача *регионализации* — кластеризации пространственных объектов с пространственными ограничениями, при которой полученные кластеры должны быть пространственно связными [Openshaw, 1977; Assunção *et al.*, 2006; Guo, 2008; Miele, Picard, Dray, 2014; Chavent *et al.*, 2018].

Комплексность процесса генерализации, попытки его полностью автоматизировать, привели к разработке *моделей генерализации*, основанных на различных правилах и ограничениях, позволяющих выполнять обобщение карт как целостных цифровых продуктов [Buttenfield, McMaster, 1991; Müller, Lagrange, Weibel, 1995; Mackaness, Ruas, 2007a; Burghardt, Duchene, Mackaness, 2014]. Вместе с тем, появились полезные обобщения существующих алгоритмов генерализации [Li, 2006; Yan, 2019], исследованы особенности генерализации аэрокосмических изображений [Кравцова, 2000], сформулированы основные положения генерализации геоизображений [Берлянт, 2006].

## 1.2 Детальность пространственных данных

### 1.2.1 Уровень детализации и детальность<sup>6</sup>

Концепция *уровня детализации* (УД) была введена в оборот в картографии в рамках направления, посвященного базам данных с множественными представлениями —

<sup>5</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov, Konstantinov, 2014; Samsonov, Konstantinov, Varentsov, 2015; Samsonov, Yakimova, 2020]

<sup>6</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov, Yakimova, 2020; Samsonov, 2022; Samsonov *et al.*, 2023]

*multiple representation databases* [Jones, Abraham, 1986; Jones *et al.*, 1996; Kilpeläinen, 2000; Sarjakoski, Sarjakoski, 2007]. Понятие уровня детализации кажется интуитивно очевидным, но его трудно определить точным и универсальным способом, который был бы принят широким научным сообществом и использован для построения соответствующих формул для оценки УД. Интерпретации данного понятия варьируются от общих и расплывчатых до подробных и включающих развернутый перечень характеристик. Менг и Форберг [Meng, Forberg, 2007] определяют УД как точку в масштабно-пространственном континууме, которая соответствует определенной степени обобщения. Лемменс [Lemmens, 2011] приравнивает УД к совокупности разрешения и количества пространственных, временных и семантических деталей. В работе [Ruas, Bianchin, 2002] уровень детализации базы пространственных данных определяется как комбинация концептуальной схемы данных, семантического разрешения, геометрического разрешения, геометрической точности и детализации. Нет общего согласия относительно того, что следует считать УД, но исследователи едины в своем понимании того, что этот сложный термин не может быть сведен к одной упрощенной мере.

При решении проблемы УД возникает множество вычислительных проблем, среди которых две представляют особый интерес для нашего исследования: как оценить УД некоторого объекта/набора данных и как уменьшить УД объекта/набора данных на заданную величину. Оба вопроса требуют формализованного определения УД и числовых метрик для оценки его абсолютного значения или изменения.

Исторически формализация УД представляет большой интерес в области трехмерной географической информации, особенно в области многомасштабных 3D-моделей городов и зданий, где континуум детализации может быть дискретизирован в последовательность уровней, каждый из которых определяется набором архитектурных элементов здания [Kolbe, Gröger, Plümer, 2005; Meng, Forberg, 2007; Biljecki *et al.*, 2014]. Пространственные данные для топографического картографирования также получают на нескольких УД, которые составляют мультимасштабную топографическую базу данных [Jones, Abraham, 1986; Kilpeläinen, 2000]. При этом официальные стандарты составления топографических карт — это, по сути, определения уровней детализации, которые соответствуют определенному масштабу карты [Руководство, 1978; Руководство, 1980; Руководство, 1985]. Они предписывают множество правил выделения и обобщения признаков (пространственных объектов), а также точность их отображения.

Поскольку правила выведения УД могут быть довольно сложными, различия и несоответствия в УД можно определить с помощью методов машинного обучения. Тойя [Touya, Brando-Escobar, 2013] разработал многокритериальный подход для оценки УД пространственных объектов и обнаружения неоднородностей в базе данных OpenStreetMap на основе сравнения этих данных с эталонной информацией, полученной с национального геопортала Франции. Тип объекта и несколько геометрических свойств используются в процедуре принятия решений. Позже этот метод был дополнен

эмпирическим подходом оценки уровня детализации, основанным на инвертированном радикальном законе [Touya, Reimer, 2015], а также набором процедур для выполнения гармонизации детальности представления объектов на крупномасштабных картах [Touya, Baley, 2017]. В стремлении найти универсальный и эффективный подход к оценке УД разрабатываются и методы растрового анализа. В частности, метод разрешения деталей Ченга и др. [Cheng *et al.*, 2017] основан на расчете растрового перекрытия линий. Аналогичный подход может быть применен для описания различимости отдельных пространственных объектов [Cheng, Liu, Zhang, 2021].

Понятие УД также можно проследить в литературе по картографической генерализации (или генерализации пространственных данных). В частности, формализованные представления на основе УД широко распространены в моделировании поверхностей (особенно на основе TIN), где точность полученного УД определяется простыми метрическими критериями, такими как вертикальная ошибка [de Floriani, Marzano, Puppo, 1996]. Аналогичные критерии, обычно выраженные в терминах расстояний, площадей или плотности точек, используются при геометрическом упрощении линий [Douglas, Peucker, 1973; Li, Openshaw, 1992; Visvalingam, Whyatt, 1993], полигонов [Haunert, Wolff, 2010a; Buchin *et al.*, 2016] или при выборе точек [Töpfer, Pillewizer, 1966]. Поскольку УД сам по себе является сложным понятием, его изменение при обобщении, скорее всего, должно быть выражено как комбинация нескольких характеристик.

Несмотря на устойчивый интерес к вопросам детализации в географической информационной науке, подобные исследования остаются довольно редкими. Помимо этого, во всех перечисленных исследованиях, связанных с уровнями детализации, прослеживается единая логика: сначала определяются характеристики, совокупность значений которых идентифицирует уровень детализации, а затем конкретные уровни детализации задаются путем фиксации этих значений на заданных уровнях. Совокупность (вектор) этих значений в настоящей работе называется **детальностью**<sup>7</sup>. Таким образом, существует некий континуум детальности пространственных данных, а уровень детализации — это некая точка в этом континууме.

Детальность геометрии характеризует подробность геометрического описания пространственных объектов. Методы оценки детальности разнятся в зависимости от используемых моделей пространственных данных. Наиболее распространены векторная модель [Herring, 2010] и модель покрытий [OGC 07-011r2, 2024] (к которой, в частности, относятся растровые и векторные данные), используемые для представления дискретных объектов и геополей соответственно. Рассмотрим текущее состояние исследований по данному направлению для линейных векторных объектов и цифровых моделей рельефа.

---

<sup>7</sup>Можно было бы также использовать термин «детализация», однако лингвистически он скорее характеризует процесс увеличения подробности, нежели состояние.

### 1.2.2 Детальность линейных объектов<sup>8</sup>

Оценка детальности линейных объектов и ее изменения требует вывода перечня геометрических характеристик, которые бы эффективно описывали форму линии и были бы чувствительны к генерализации. Одно из ранних систематических исследований в этой области было проведено в работе [McMaster, 1986]. Автор проанализировал 30 метрик и выявил шесть из них, которые могут быть использованы для оценки изменения детальности без статистической избыточности, включая процентное изменение числа точек, процентное изменение стандартного отклонения числа точек на дюйм, процентное изменение угловатости, общее векторное смещение на дюйм, общее смещение площади на дюйм и процентное изменение числа криволинейных сегментов (изгибов). Ранее Дженкс [Jenks, 1981] с помощью группы респондентов выявил, что мера векторного (линейного) смещения наиболее тесно коррелирует с перцептивными (визуальными) различиями между исходной и геометрически упрощенной линией. Он также отметил, что во многих случаях респонденты испытывали проблемы с упорядочением линий в соответствии с последовательностью обобщений.

Расчет геометрических характеристик имеет важное значение на всех этапах генерализации, включая настройку допусков упрощения, оценку состояния линии во время обобщения и оценку результата. МакМастер [McMaster, 1987] предложил модель выбора величины допуска для процедуры геометрического упрощения. В число параметров модели он включил степень уменьшения масштаба, технологию вывода и характер использования карты. Он выделил четыре качественных градации обобщения: воспринимается как одно и то же / узнаваемое, воспринимается как различное / узнаваемое (2 градации) и воспринимается как различное (неузнаваемое). В этом исследовании Макмастер подчеркивает, что *«к сожалению, картографы, которые ищут в картографической литературе рекомендации по выбору параметров упрощения, будут разочарованы: мало или вовсе нет работ, посвященных этим проблемам»*. И хотя Бирд [Beard, 1991] позже сместила акцент на *ограничения (constraints)*, что несколько облегчило проблему, это не привело к ее полному решению, так как для удовлетворения ограничений необходимо найти оптимальные значения параметров обобщения.

Баттенфилд [Buttenfield, 1987; Buttenfield, 1991] обратила внимание на то, что линии часто неоднородны по своему характеру, и поэтому для разных частей линии могут потребоваться различные допуски упрощения. Она разработала основанный на оценке точности подход для иерархического разбиения линий на однородные сегменты, что дает потенциальную возможность упрощать эти сегменты независимо друг о друга. Данная идея была позже реализована в ряде работ, где линии подвергались классификации (целиком) или сегментации (на фрагменты) для настройки допусков упрощения и/или

<sup>8</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov, Yakimova, 2020]

совместного применения нескольких алгоритмов упрощения [Plazanet, Affholder, Fritsch, 1995; Mustiere, 2005; Mitropoulos, Nakos, 2011; Park, Yu, 2011; Samsonov, Yakimova, 2017].

Выбор оптимальных допусков генерализации требует оценки не только характера исходной линии, но и интенсивности изменений детальности, вносимых каждым алгоритмом при вариации значения допуска. Верегин [Veregin, 2000] количественно оценил позиционную ошибку алгоритмов упрощения и показал, что существует возможность достижения желаемой позиционной точности при упрощении. При этом было выявлено, что график эмпирической кривой кумулятивного отклонения линии от исходной имеет форму логистической функции для умеренных упрощений и степенной функции для высоких значений допуска. Его исследование было сфокусировано на объектах гидрографической сети и дорог, но рассматривался в нем только один алгоритм (Дугласа-Пейкера). Ши и Чунг [Shi, Cheung, 2006] попытались связать геометрическую точность и время обработки с количеством вершин в линиях, полученных разными алгоритмами геометрического упрощения. Это позволило сопоставить алгоритмы по эффективности в плане соотношения точности и объема информации. В работе [Visvalingam, 2016] представлен подробный разбор поведения алгоритма Висвалингам-Уайатта при упрощении линий. Автор использовала различные метрики и эвристики, чтобы формализовать управление алгоритмом для получения предсказуемых результатов. Сонг и Миао [Song, Miao, 2016] разработали подход к оценке алгоритмов упрощения линий, учитывающий графические параметры отображения, масштаб карты и способность человеческого глаза различать пиксели. Последовательность упрощений, полученная каждым алгоритмом при вариации значения параметра, позволила выявить значения параметров, при которых алгоритмы дают близкое по расстоянию отклонение результирующей линии от исходной.

Величина допуска при геометрическом упрощении обычно выражается в терминах критических значений геометрических дескрипторов, таких как расстояние, площадь и т.д. Эти дескрипторы, однако, сами по себе не отражают интенсивность изменения детальности линий и не одинаковы для разных алгоритмов. Поэтому их оказывается сложно подбирать вручную. Небольшое количество исследований посвящено поиску простых и интуитивно понятных метрик, которые бы характеризовали желаемое уменьшение детализации линии и были бы применимы к разным алгоритмам. Жу и Джонс [Zhou, Jones, 2003] ввели меру *степени генерализации* (*DoG — Degree of Generalization*), которая позволяла определить, насколько интенсивное упрощение будет выполнено при переходе к другому уровню детализации базы данных в случае использования алгоритма взвешенной эффективной площади. Другой способ приведения алгоритмов к «общему знаменателю» был предложен в работе [Foerster, Stoter, Lemmens, 2007]. Они заменили допуски нескольких алгоритмов на единый *коэффициент сжатия количества точек*. Однако, как отмечают сами авторы, этой упрощенной метрики недостаточно для обработки всех картографических аспектов обобщения. Эти картографические аспекты требуют удовлетворения многочисленных ограничений при упрощении линий, включая метрические, топологи-

ческие, семантические и гештальт (линейный символ) [Weibel, 1996]. Стоит отметить, что понятия уровня детализации и степени генерализации рассматривают одну и ту же проблему с разных точек зрения. УД — это свойство статическое, которое характеризует состояние карты или базы пространственных данных с точки зрения ее детальности. DoG — это свойство динамическое, которое больше ориентировано на эффект обобщения и связь между начальным и результирующим состоянием карты или базы пространственных данных. Поэтому оба понятия дополняют друг друга и могут быть использованы для характеристики процесса обобщения и его результата.

Представленный анализ демонстрирует широкий спектр исследований, касающихся определения и оценки уровня детализации, характеристики формы линий, настройки алгоритмов генерализации и оценки результатов их работы. Однако существующие подходы, которые соотносят некую частную метрику (расстояние, степень сжатия и т. д.) со значением допуска алгоритма, не раскрывают возможный способ объединения этих метрик в интегральную метрику детальности. Кроме того, отсутствуют работы, в которых бы исследовался характер изменения значений метрик при упрощении разных по форме линий. В то же время, решение данных задач имеет определяющее значение для выбора наилучшего алгоритма генерализации (который должен основываться на сравнении упрощенных разными алгоритмами копий линии с сопоставимой детализацией), а также для совместного применения различных алгоритмов к разным объектам (которое должно приводить к сопоставимому изменению их детализации).

### 1.2.3 Детальность цифровых моделей рельефа<sup>9</sup>

#### 1.2.3.1 Геометрия ЦМР

Цифровые модели рельефа (ЦМР) широко используются в современной картографии. Помимо вычисления полезных топографических величин, ЦМР часто просто визуализируют — для представления рельефа на карте. Поскольку карты создаются в широком диапазоне масштабов, необходима соответствующая генерализация ЦМР, иначе представление рельефа будет перегружено излишними деталями. За последние десятилетия разрешение и охват данных о рельефе значительно увеличились. Современные глобальные ЦМР, такие как MERIT DEM [Yamazaki *et al.*, 2017] и NASADEM [NASADEM, 2020], обеспечивают глобальное или полуглобальное покрытие суши с разрешением 30 — 90 м. ЦМР, полученные с помощью лидарной и оптической аэрофотосъемки, имеют меньший охват, но могут иметь разрешение до первых сантиметров [Guillaume *et al.*, 2021]. Поскольку более высокое разрешение часто считается показателем более высокого качества ЦМР, нередко мелкомасштабные карты, созданные на основе таких детальных источников, не

<sup>9</sup> При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov, 2022]

имеют генерализации. Однако отсутствие генерализации снижает читаемость карты и портит визуальное впечатление.

Понимание того, что такое крупный, средний и мелкий масштаб, различается между стандартами, принятыми национальными картографическими агентствами, академическими институтами и областями применения. В данной работе используется классификация Имгофа [Imhof, 1982], разработанная специально для картографического изображения рельефа. По наблюдениям Имгофа, масштабы менее 1 : 100 000 характеризуются неизбежной необходимостью обобщения. Упрощение поверхности, представленной горизонталями, распространяется в большей или меньшей степени на все изображение карты, а не только на отдельные места. Такие масштабы считаются мелкими. Поэтому в данном исследовании интерес представляет применение ЦМР для картографирования в масштабах 1 : 200 000 и мельче.

Как было отмечено ранее, в концепции [Ruas, Bianchin, 2002], детальность пространственных данных характеризуется в том числе такой величиной как *гранулярность*. В то время как схема данных, пространственное разрешение и точность отражают потенциальную способность базы данных представлять реальный мир, гранулярность определяет фактический размер пространственных объектов и элементов их формы (таких как изгибы), которые будут показаны на карте. В случае с рельефом такими объектами являются формы рельефа — холмы, долины и т. д. и подчиненные им элементы.

Генерализация рельефа — хорошо разработанная тема еще со времен традиционной ручной картографии. Полезные рекомендации можно найти в специализированных монографиях [Заруцкая, 1958; Imhof, 1982] и официальных руководствах по созданию топографических карт, изданных национальными картографическими агентствами [Руководство, 1980; Cartographic generalization, 1987]. Хотя эти публикации достаточно полно описывают качественные принципы упрощения поверхности, они несколько скупы на числовое описание желаемой зернистости рельефа. Например, линейный размер поверхностного разрыва, такого как скала, овраг или разлом, для отображения на российской топографической карте масштаба 1 : 200 000 должен составлять не менее 3—10 мм в зависимости от типа объекта [Руководство, 1980]. Рекомендуемый минимальный размер цветных символов территории составляет 4 мм<sup>2</sup> согласно [Cartographic generalization, 1987]. Тем не менее, рекомендации по минимальному или типичному размеру рельефа найти практически невозможно, и обычно говорят, что этот размер должен отражать «особенности поверхности, характерные для картографического масштаба» [Заруцкая, 1958], что мало пригодно для реализации компьютерных алгоритмов.

Представляется, однако, что гранулярность ЦМР должна быть сопоставима с общими рекомендациями по изображению самых мелких объектов, которые на карте составляют около первых миллиметров по линейным размерам. Кроме того, воспринимать объекты с экрана компьютера обычно сложнее, чем с печатного носителя, поэтому этот порог может быть сознательно увеличен. Однако для практического применения необ-

ходим автоматизированный метод.

Качество цифровых моделей рельефа — часто обсуждаемая тема в научных исследованиях. Понимание качества, а также возможные меры качества и их предпочтительные значения не являются универсальными и зависят от задач пользователя [Polidori, El Hage, 2020]. Мелкомасштабное картографирование на основе ЦМР — специфическая задача, которая приоритезирует степень обобщение над другими аспектами качества [Imhof, 1982]. В то же время, согласно опросам, пользователи обычно не хотят уделять много времени оценке влияния качества ЦМР на результаты их анализа [Wechsler, 2003]. Эти условия означают, что необходимы различные методы оценки качества ЦМР, причем эти методы должны быть максимально автоматизированы.

Точность — один из наиболее часто обсуждаемых аспектов качества ЦМР [Florinsky, 1998; Höhle, Höhle, 2009; Mesa-Mingorance, Ariza-López, 2020]. В частности, выявлены различия в точности в зависимости от расчлененности рельефа [Zhou, Liu, Sun, 2006; Erdoğan, 2010]. По этой причине сегментация ЦМР на морфологически однородные зоны имеет смысл для прогнозирования ожидаемой точности анализа. Однако, как отмечается в работе [Mesa-Mingorance, Ariza-López, 2020], такая сегментация обычно представляет собой лишь стратификацию (равнинные, горные районы и т. д.), что означает, что отдельные формы рельефа не анализируются. По мнению авторов публикации [Polidori, El Hage, 2020], точность важна для двух сторон качества ЦМР: качество высот (абсолютная или относительная точность) и качество формы и топологии (связанное с точностью производных ЦМР, таких как уклон, экспозиция склона, кривизна и т. д.). Те же авторы объясняют разрешение ЦМР как способность ЦМР различать объекты, в то время как термин «размер ячейки» имеет другое значение и может не иметь отношения к фактическому разрешению ЦМР. Небольшой размер ячейки — необходимое условие для обеспечения высокого разрешения [Polidori, El Hage, 2020]. Хотя авторы утверждают, что разрешение часто используется неправильно (также взаимозаменяемо с размером ячейки), проблема в том, что оба термина представляют *возможность*, но не *фактический* размер форм рельефа, представленных в ЦМР, что, вероятно, является одной из причин, почему их трудно различить. Гут с соавторами [Guth et al., 2021] рассуждает более технически, определяя разрешение как «горизонтальные размеры наименьшей детали, обнаруживаемой сенсором и изменяемой после процедуры построения сетки, обычно в метрах». Однако такое определение разрешения, ориентированное на дистанционное зондирование, слишком специфично для ЦМР, полученных из несенсорных источников (например, топографических карт) или обработанных с помощью фильтрации и других процедур генерализации.

В то время как численные меры точности ЦМР необходимы для оценки распространения ошибок при анализе рельефа, они гораздо менее информативны при оценке пригодности ЦМР для создания карт. Причина заключается в том, что эти меры не могут справиться с качеством представления рельефа на карте, то есть описать морфологию и

размер изображаемых объектов рельефа. Чем меньше масштаб, тем менее важна точность, и тем важнее морфологическое правдоподобие представления рельефа [Imhof, 1982]. Подобникар [Podobnikar, 2008] высказался за использование визуальной аналитики для оценки качества ЦМР. Он предложил несколько приемов, основанных на визуализации результатов пространственных аналитических и статистических операций, а также непространственных визуализаций, облегчающих понимание погрешностей и ошибок, присутствующих в ЦМР. В частности, он предложил рассматривать характерные точки, линии и области и искать в них ложные закономерности. Представляется, что если такие характерные элементы являются наиболее отличительными чертами поверхности, то свойства, основанные на их распределении, могут привести нас к мерам качества ЦМР, эффективно характеризующим пригодность ЦМР для целей картографирования.

Выявление характерных особенностей поверхности - хорошо разработанная тема в геоморфометрии и гидрологическом анализе ЦМР. Пейкер и Дуглас [Peucker, Douglas, 1975] описывают так называемые *специфические точки поверхности*, понимая под ними те, которые предоставляют больше информации о поверхности, чем просто их координаты. Традиционно специфическими считаются точки, расположенные на вершинах, котловинах, перевалах, хребтах, тальвегах и разломах. Такие точки могут быть определены на основе соотношений между высотами в движущемся окне  $3 \times 3$  [Peucker, Douglas, 1975; Band, 1986]. Подмножество таких точек, расположенных вдоль тальвегов и водоразделов, также можно определить с помощью анализа направлений стока [O'Callaghan, Mark, 1984; Jenson, Domingue, 1988]. На первый взгляд, плотность точек, характерных для поверхности, является потенциальным косвенным показателем для оценки среднего размера форм рельефа (чем больше точек, тем меньше формы рельефа). Однако такая мера не является надежной, поскольку данная плотность точек может быть вызвана как формами рельефа схожего размера, так и одинаковым количеством форм рельефа, значительно отличающихся по размеру. Поэтому необходимо знать площадь, занимаемую каждой формой рельефа, или ее линейные размеры.

Предложено множество методов разделения поверхности на морфологически или гидрологически обоснованные площадные элементы. Они могут быть основаны на классификации производных поверхности [Wood, 1996; Schmidt, Hewitt, 2004; Iwahashi, Pike, 2007], сегментации [Drăguț, Blaschke, 2006; Romstad, Etzelmüller, 2012], направлении стока [O'Callaghan, Mark, 1984; Jenson, Domingue, 1988] или фокальном анализе [Jasiewicz, Stepinski, 2013]. Недостатком всех этих методов является то, что каждый из них должен быть параметризован, а это значит, что размер полученных элементов рельефа, таких как холмы, склоны, долины, седловины, водоразделы или тальвеги, зависит от исходных значений параметров, таких как максимальная кривизна, расстояние поиска или минимальная площадь водосбора. Более того, только сегментация и выделение водоразделов являются по-настоящему объектно-ориентированными, в то время как другие методы

просто относят ячейки ЦМР к одному из типов в классификации. Эти неоднозначности характеризуют существующие в настоящее время подходы к выделению форм рельефа и их элементов как недостаточно надежные для оценки размеров форм рельефа, представленных на поверхности ЦМР.

Можно взглянуть на проблему со стороны картографической генерализации, целью которой является получение представления рельефа с оптимальной гранулярностью. На сегодняшний день разработано множество алгоритмов автоматической генерализации ЦМР (фактически, упрощения поверхности) [Guilbert, Gaffuri, Jenny, 2014], которые были рассмотрены ранее в Разделе 1.1.5. Кроме того, Raposo [2020] показал, что фильтр переменного размера может быть использован для эффективного обобщения поверхности на основе локальной меры энтропии. Однако несмотря на обилие существующих методов генерализации ЦМР, удивительно мало исследований, показывающих, как можно оценить гранулярность упрощенной ЦМР для картографирования в определенном масштабе. В то же время каждый алгоритм генерализации ЦМР управляется своим набором параметров, таких как количество итераций сглаживания, или порог кривизны для выделения хребтов, или длина кратчайшего водотока для сохранения. Оптимальное значение параметра обычно подбирается методом проб и ошибок, пока ЦМР не достигнет нужного уровня детализации, который оценивается визуально. Этот процесс невозможно автоматизировать, пока не будет разработана эффективная мера детальности ЦМР.

### 1.2.3.2 Визуализация ЦМР<sup>10</sup>

Детальность картографического изображения определяется не только детальностью отображаемых пространственных данных, но также и параметрами их визуализации. Наиболее сложно этот процесс устроен для цифровых моделей рельефа и других геополей, при визуализации которых в ряде сценариев происходит не просто отображение ячеек, а преобразование модели данных. В частности, горизонтали (изолинии) и штрихи представляют собой векторные объекты, извлеченные из цифровой модели рельефа с определенными параметрами, которые определяют плотность их размещения. Рассмотрим современное состояние разработок в этом направлении.

Изолинии отображают двумерные скалярные поля, такие как высота земной поверхности или другие типы непрерывных данных. Применительно к изображению рельефа изолинии обычно называют *горизонталями*. При выборе интервала сечения картографам приходится идти на компромисс. При большом интервале карта содержит небольшое количество горизонталей, и конфликты с другими элементами карты сводятся к минимуму. Однако редкие горизонтали не могут выявить рельеф почти плоских участков или детали мелких объектов, таких как террасы, седловины, холмы, впадины

<sup>10</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov, 2014; Samsonov et al., 2019]

или дамбы. Большой интервал также приводит к большому заложению — горизонтальному расстоянию между горизонталями, что затрудняет визуальную интерпретацию изображения, восстановление форм рельефа на его основе. При малом интервале горизонтали показывают мелкие детали, но на крутых участках могут стать слишком плотными, вплоть до того что полностью покрывают склон без зазоров. Кроме того, плотно расположенные горизонтали чаще всего графически мешают другим элементам карты [Imhof, 1982; Keates, 1989]. Использование дополнительных горизонталей — элегантное решение этой дилеммы. Дополнительные горизонтали — это короткие участки горизонталей с тем же сечением, что и обычные горизонтали, но с базовым уровнем, смещенным на половину сечения. Они размещаются там, где небольшие, но важные особенности рельефа не отображаются обычными горизонталями, и на плоских участках, где заложение между основными горизонталями велико. Дополнительные горизонтали обычно изображаются штриховыми (русская школа) или пунктирными (швейцарская школа) линиями, чтобы отличить их от сплошных горизонталей основного сечения.

Линии дополнительных горизонталей исторически используются во многих сериях национальных топографических карт [Collier, Forrest, Pearson, 2003], а учебники по картографии неизменно рекомендуют использовать их для карт местности [Imhof, 1982; Keates, 1989; Robinson *et al.*, 1995; Верещака, 2002; Kenneth Field, 2018; Slocum *et al.*, 2022]. Поэтому удивительно, что до сих пор ни один цифровой метод автоматического нанесения дополнительных горизонталей не обсуждался в литературе и не был доступен в программном обеспечении ГИС. По этой причине дополнительные горизонтали стали редкостью на картах, где используются горизонтали, полученные на основе цифровых моделей рельефа. Например, недавно обновленная серия карт Швейцарского федерального бюро топографии Swisstopo больше не включает дополнительные горизонтали [Geländedarstellung, 2008].

Поскольку горизонтали в основном используются для визуализации рельефа, картографы сосредоточились на формализации и документировании принципов проектирования горизонталей на топографических картах. В начале истории цифровой картографии было разработано множество алгоритмов построения горизонталей, и Тоблер [Tobler, 1965] сообщил о существовании нескольких десятков приложений для построения горизонталей к середине 60-х годов. Горизонталы извлекаются либо из двумерных скалярных полей, таких как цифровые модели рельефа в формате сетки, либо из множества точек. Для обычных скалярных полей обычно используется алгоритм шагающих квадратов (*marching squares*), который является двумерной версией алгоритма шагающих кубов для извлечения изоповерхностей из трехмерных скалярных полей [Lorensen, Cline, 1987].

В литературе не описано ни одного цифрового метода создания дополнительных горизонталей. Единственный задокументированный метод размещения дополнительных горизонталей упоминается в работе [Jaara, Lecordix, 2011], где рассматривается цифровая

генерация горизонталей для серии национальных карт Франции. Они выявляют регионы, где необходимы дополнительные горизонталы, но не приводят ни описания метода, ни наглядных примеров. Аналогичная ситуация с работой [Kettunen, Koski, Oksanen, 2017] на территорию Финляндии, где обосновывается необходимость отрисовки дополнительных горизонталей, но не представлено конкретное алгоритмическое решение для этих целей.

Альтернативный подход к размещению дополнительных горизонталей заключается в уменьшении интервала сечения и удалении или перемещении горизонталей, которые слишком плотно расположены на крутых участках. Так, например, Ordnance Survey и канадское национальное картографическое агентство удаляют горизонталы, чтобы избежать соприкосновения или перекрытия горизонталей на крутых участках [Keates, 1989; Mackaness, Steven, 2006]. Другой пример — серия карт 1 : 25 000 швейцарского национального картографического агентства Swisstopo, где обычные горизонталы удалены из районов со скальным рельефом, поскольку в противном случае горизонталы сливались бы в этих, как правило, очень крутых районах [Jenny et al., 2014]. Смещение горизонталей необходимо для того, чтобы избежать слияния горизонталей на более мелких масштабах [Imhof, 1982]. Гаффури [Gaffuri, 2009] представил автоматизированный метод увеличения расстояния между плотными горизонталями на крутых участках. Он деформирует горизонталы с помощью алгоритма, основанного на ограничениях, который сохраняет общую форму поверхности.

**Штрихи** представляют собой особый способ изображения рельефа, который может иметь преимущества перед другими техниками, такими как горизонталы и отмывка. В частности, штрихи обладают довольно выразительной особенностью представлять дифференциальные характеристики поверхности, прослеживая направление линий тока и подчеркивая их конвергентное или дивергентное поведение. Размещение штрихов вдоль горизонталей и вариация их ширины позволяют выявить форму и кривизну поверхности. Как и другие способы изображения, штрихи могут выполнять рисовку рельефа с разной детальностью: на это влияет и шаг сечения рельефа горизонталями, и шаг расстановки штрихов.

Изначально введенные Леманом [Lehmann, 1799], топографические штрихи оставались основным способом изображения рельефа на картах в первой половине XIX века. Один из известнейших шедевров «эпохи штрихов» — карта Дюфура. Имгоф сформулировал 5 основных принципов построения штрихов [Imhof, 1982]:

1. Штрихи представляют собой плотные упаковки небольших сегментов линий тока. Их направление везде соответствует уклону поверхности. Вследствие этого они пересекают горизонталы под прямым углом.
2. Штрихи располагаются горизонтальными<sup>11</sup> рядами.

<sup>11</sup>Имгоф действительно использует термин «horizontal rows» (в англоязычном издании [Imhof, 1982]), по всей

3. Длина каждого штриха соответствует заложению между горизонталями.
4. Светлота штриховки варьируется по принципу «чем круче, тем темнее».
5. На каждом квадратном сантиметре карты должно быть расположено одинаковое количество штрихов. Это количество должно быть адаптировано для конкретной карты.

В крупных и средних масштабах обычно использовались два типа штрихов — крутизны и теневые, а в мелких — смешанная техника штриховки общего характера. В случае штрихов крутизны более толстые линии проводились на более крутых склонах, а при теневом подходе более толстые линии покрывали менее освещенные участки.

Первый алгоритм для автоматического построения штрихов был предложен Йоэли [Yoeli, 1985]. В качестве исходных данных он использовал набор горизонталей и структурных линий (талъвегов). Штрихи строились между двумя последовательными уровнями горизонталей как прямые отрезки, перпендикулярные более высокому уровню. Структурные линии использовались для обрезки штрихов на дне долин и предотвращения их пересечения. Они также использовались в качестве ограничителей на участках с малыми углами наклона. Можно отметить некоторые особенности полученных результатов, которые открывают перспективы дальнейших исследований:

1. Штрихи — это *прямые* линии. Прямая линия может быть перпендикулярна только одной горизонтали (и только в начальной или конечной точке), а ее направление не везде следует направлению максимального уклона. Чтобы выполнить первое правило Имгофа, штрих должен быть изогнутой линией тока, которая следует [анти]градиенту поверхности.
2. Локальная изменчивость формы горизонтали в значительной степени влияет на направление штрихов, поскольку они строятся перпендикулярно горизонталям. В ряде мест это нарушает схожесть направлений соседних штрихов, теряется их взаимная когерентность.
3. Расхождение штрихов на выпуклых склонах не учитывается: дополнительные штрихи в зонах дивергенции не вставляются, в результате чего образуются широкие пустые пространства.
4. Штрихи строятся только вниз по склону. Это означает, что вершины положительных форм рельефа штрихами не заполняются и имеют платообразное изображение.

В работе [Regnauld, Mackaness, Hart, 2002] представлен алгоритм построения штрихов для археологического картографирования антропогенных форм рельефа. Авторы использовали набор данных *DNF (Digital National Framework)* с линейными объектами

---

видимости понимая под этим расположение штрихов между горизонталями или расположение начальных (или конечных) точек штрихов на одной горизонтали.

«бровка искусственного склона» и «тыловой шов искусственного склона» для построения штрихов между ними перпендикулярно линиям бровки. Алгоритм итеративно выравнивает направление штрихов, сравнивая ориентации трех последовательных линий, что решает проблему №2 алгоритма Йоэли. Также была реализована вставка дополнительных штрихов между исходными в случае их расхождения, что решает проблему 3 алгоритма Йоэли. Однако полученные линии были по-прежнему прямыми и строились только в направлении спуска.

Следующий шаг в автоматизации способа штрихов был пройден благодаря использованию методов *нефотореалистичного рендеринга (НФР)*. Методы создания изображений, имитирующих рукописные эскизы рельефа были предложены в работах [Dowson, 1994; Lesage, Visvalingam, 2002; Mower, 2009] с использованием цифровой модели рельефа (ЦМР) и алгоритмов анализа изображений. В исследованиях [Hurni, Dahinden, Hutzler, 2001; Dahinden, Hurni, 2007] предложили методы автоматизированного изображения скал на топографических картах. Обширный обзор применения НФР в картографии можно найти в работе [Kennelly, Kimerling, 2006].

НФР-алгоритмы для визуализации векторных полей с помощью равномерно распределенных линий тока предложены в [Turk, Banks, 1996; Jobard, Lefer, 1997]. Равномерная плотность изображения была обеспечена путем контроля расстояния между соседними линиями и сужения их концов. При этом плотность размещения линий тока, а также разрешение подстилающей модели геополя определяет детальность изображения. Такой подход хорошо вписывается в первое правило построения штрихов, поскольку использует кривые линии тока. Схожая техника была применена в работе [Buchin et al., 2004] для генерации перспективных видов местности. Они использовали регулярное разбиение ЦМР сеткой квадратов и прослеживали линию тока из каждого квадрата вниз и вверх по склону, пока она не достигнет соседней линии тока или пока градиент не станет слишком мал. После этого моделируется светотень в зависимости от положения источника света с помощью двух техник: путем отбор линий тока (чем ярче освещен склон, тем сильнее отбор) и путем изменения ширины и светлоты линий (линии тока на более ярко освещенных склонах тоньше и светлее). Штрихи крутизны при этом были реализованы путем изменения их толщины в зависимости от угла наклона поверхности.

Обратим внимание на то, что рассмотренные техники в стиле НФР не используют горизонталы в качестве опорных линий для размещения штрихов. Следует отметить, что изображения, полученные в [Buchin et al., 2004], сложнее интерпретировать, чем изображения Йоэли [Yoeli, 1985]. Причина в том, что изображение плохо структурировано. Регулярно-сеточное или случайное размещение начальных точек штрихов не позволяет визуально объединить их в группы, огибающие формы рельефа. При распределении же начальных точек вдоль горизонталей визуальная связь между соседними штрихами в одном ряду становятся гораздо сильнее. Штрихи в этом случае располагаются в соответствии с формой поверхности, а их изображение дает не только локальную, но и зональ-

ную характеристику поверхности по периметрам форм рельефа.

В этом контексте техника Кеннелли и Кимерлинга по построению штрихов путем их размещения по регулярной сетке также не является морфологически обоснованной [Kennelly, Kimerling, 2000]. Нельзя не упомянуть, что Кеннелли также предложил использовать ориентированные полутона (oriented halftones) при выполнении отмывки рельефа чтобы подчеркнуть экспозиционные различия склонов [Kennelly, Kimerling, 2004]. Каждый полутон представлял собой штриховку параллельными линиями, а ее направление соответствовало экспозиции склона. В итоге получилась техника, родственная штрихам, хотя и достаточно сильно отличающаяся от них по принципам построения.

Наш анализ показывает, что во всех ранее разработанных алгоритмах выполняется либо правило 1 Имгофа, либо правила 2-3, но не все одновременно. То есть, штрихи либо располагаются между горизонталями, но они прямые, либо это кривые линии тока, но при этом не ограниченные горизонталями. Помимо этого, не исследован вопрос управления детальностью штрихового изображения путем соответствующей параметризации алгоритма их построения.

## 1.2.4 Интеграция данных разной детальности<sup>12</sup>

### 1.2.4.1 Конфляция

**Конфляция** — это процесс объединения (слияния) информации из двух источников данных в один источник, при котором по возможности устраняются несоответствия в пространственном положении объектов [de Smith, Goodchild, Longley, 2018]. Заалфельд [Saalfeld, 1988] впервые выделил конфляцию в особый метод ГИС и установил ее математические и технологические основы. По сути, конфляция предполагает идентификацию соответствующих объектов и точек в двух наборах данных с использованием пространственных и семантических атрибутов, после чего выполняется пространственная корректировка объектов или перенос атрибутов между ними. Одной из наиболее широко используемых техник пространственной трансформации является так называемый *rubbersheeting*, или *метод резинового листа*, в ходе которого объекты сначала триангулируются, а затем вершины триангуляции локально смещаются в соответствии со связями смещения, которые соединяют похожие точки в двух наборах данных [White, Griffin, 1985].

Конфляция — непростой процесс, поскольку идентификация соответствующих объектов всегда в той или иной степени неопределенна. Более того, значительные сложно-

<sup>12</sup> При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov, 2020b; Мухаметшин, Самсонов, Лурье, 2024]

сти вносит гетерогенность совмещаемых наборов данных, которая может выражаться в их *разной детальности*. Стандартным подходом является использование минимаксной стратегии, при которой соответствующие объекты идентифицируются с использованием минимального или максимального значения некоторой меры сходства. В частности, Станиславски, Нельсон и Хаманн [Stanislawski, Nelson, Hamann, 2002] использовали максимальную площадь перекрытия и минимальное расстояние между двумя гидрографическими наборами данных. Кобб с соавторами [Cobb *et al.*, 1998] подчеркнули важность использования атрибутивной информации для повышения надежности совмещения, а в работе [Walter, Fritsch, 1999] было показано, что определение наиболее вероятных пар объектов для совмещения может быть основано на статистических метриках. Задача сопоставления объектов часто формулируется в терминах оптимизации, которая ищет наилучшее решение, максимизирующее общее сходство между всеми сопоставленными парами [Li, Goodchild, 2011] или минимизирующее меру общего расхождения между наборами данных [Lei, Lei, 2019]. Поскольку интеграция выполняется в основном по линейным объектам, особый интерес в последних исследованиях представляют специальные меры сходства, такие как расстояние Фреше, которые учитывают непрерывность линий [Gheibi, Maheshwari, Sack, 2016; Chambers *et al.*, 2020].

Методы конфляции в своем большинстве ориентированы на данные о дорожных сетях [Chen, Knoblock, Shahabi, 2006; Mustière, Devogele, 2008] и привязку данных мобильного/GPS-трекинга к дорожным сетям [Lou *et al.*, 2009; Bergman, Oksanen, 2016]. Национальные картографические агентства указывают конфляцию как одну из технологий, включенных в процессы генерализации карт [Duchêne *et al.*, 2014]. В частности, производственный процесс создания набора пространственных данных OS VectorMap™ District Национальной картографической службы Великобритании Ordnance Survey GB включает в себя конфляцию железных дорог и надписей [Revell, Regnauld, Bulbrooke, 2011]. Хотя большинство методов конфляции сосредоточено на векторных данных, эта процедура релевантна и для геополей, дискретизированных в виде растровых или триангуляционных покрытий. В частности, исследование [Kyriakidis, Shortridge, Goodchild, 1999] позволило получить геостатистический подход к конфляции цифровых моделей рельефа, где совмещенный результат представляется как реализация стохастического пространственного процесса.

Увеличение разнообразия источников пространственных данных в последние десятилетия повысило значение конфляции как одной из ключевых технологий в обработке пространственных данных [Adams *et al.*, 2010]. Среди наиболее влиятельных технологических факторов — широкое распространение *краудсорсинга* (упрощенно говоря, совместного решения задач группой волонтеров), который существенно расширил диапазон и разброс качества пространственных данных [Goodchild, Li, 2012]. Как следствие, методы конфляции распространились на различные типы пространственных данных и стали востребованы даже для неточных, схематичных пространственных

наборов данных [Wallgrün, Wolter, Richter, 2010]. Базы пространственных данных, такие как OpenStreetMap [OpenStreetMap, 2025], часто обновляются с использованием разнородных и неточных источников, и устранение топологических несоответствий в перекрывающихся покрытиях является необходимым для обеспечения согласованности данных [Borges, Laender, Davis, 1999]. В настоящее время инструменты конфликации доступны в виде специализированных модулей для ГИС-платформ общего назначения, таких как ArcGIS [Lee, Yang, Ahmed, 2014] и QGIS [Dalang, 2025], а также в виде самостоятельных решений с открытым исходным кодом [Hootenanny, 2025] и проприетарных [L3Harris Technologies, 2020].

Рельеф и гидрография часто анализируются и картографируются вместе из-за их генетической связи. Естественная гидрографическая сеть следует за местами вдоль отрицательных форм рельефа, которые имеют высокое значение аккумуляции стока и составляют так называемую *дренажную сеть* Jenson, Domingue [1988]. Когда гидрографические данные и данные о рельефе поступают из разных источников, эта связь часто нарушается. Причин этому может быть множество, начиная от несоответствия уровней детализации и методов сбора данных, и заканчивая неточной информацией о геодезических системах отсчета. Но следствие всегда проявляется одинаково: гидрографические линии не совпадают с дренажной сетью, неявно представленной в данных о рельефе. Эта проблема обычно решается с помощью специализированных техник, таких как «выжигание» водотоков [Lindsay, 2016a]. Эти техники изменяют дренажную сеть на ЦМР таким образом, чтобы она соответствовала конфигурации реальной гидрографической сети. Однако измененная таким образом ЦМР все равно непригодна для картографических целей, поскольку несоответствие между гидрографическими линиями и поверхностью ЦМР фактически остается неисправленным.

Для решения этой проблемы необходимо использовать специализированный метод совмещения, который выравнивает данные о рельефе и гидрографии. Выравнивание необходимо не только при составлении карт, но и при генерализации ЦМР. В частности, обобщенная ЦМР должна быть пространственно согласована с обобщенной векторной гидрографией. До настоящего времени не было разработано алгоритмов для решения этой задачи.

Проблеме несоответствия между векторными гидрографическими данными и автоматически выделенной дренажной сетью уделяется большое внимание. Она особенно важна в задачах, где ЦМР используется для обогащения существующих гидрографических наборов данных [Stanislawski, Battenfield, Doumbouya, 2015; Stanislawski et al., 2018]. Для обеспечения соответствия распределения направлений стока на ЦМР реальной гидрографии разработано множество методик так называемого *выжигания водотоков* (*stream burning*). Стандартный подход к выжиганию заключается в снижении значений ячеек ЦМР под гидрографическими объектами таким образом, чтобы они аккумулялировали большую часть потока [Saunders, 1999; Soille, Vogt, Colombo, 2003; Getirana et

*al.*, 2009]. Ву с соавторами [Wu *et al.*, 2019] модифицировали процедуру приоритетного затопления [Wang, Liu, 2006], которая заполняет искусственные понижения методом поиска пути наименьшей стоимости, для реализации алгоритма выделения дренажной сети вдоль существующих гидрографических объектов. Линдси [Lindsay, 2016a] разработал сложную процедуру выжигания водотоков, которая позволяет получить растр аккумуляции стока (площади водосбора) с дренажной сетью, согласованной с реальными гидрографическими данными. Хотя его метод эффективен для гидрологического анализа, он не корректирует поверхность, закодированную в исходной ЦМР, и поэтому не подходит для решения проблемы с картографической точки зрения.

Картографически ориентированный подход к автоматизированному извлечению гидрографически скорректированных горизонталей был представлен в статье [Arundel, Thiem, Constance, 2018]. Их метод объединяет растровые ЦМР National Elevation Dataset (NED) и независимо созданные гидрографические данные National Hydrography Dataset (NHD). Это достигается путем реинтерполяции центров ячеек ЦМР в точки рельефа со сглаживанием и принудительным дренированием вдоль особенностей NHD с помощью программы ANUDEM [Hutchinson, 1989]. Хотя метод показал себя эффективным для получения горизонталей, согласованных с гидрографическими данными, он все еще использует прямой подход для вжигания водотоков на поверхности. Поэтому он, как правило, неэффективен при плохой согласованности данных о рельефе и гидрографии [Arundel, Thiem, Constance, 2018, с. 54].

В работе [Yadav, Hatfield, 2018] рассмотрен альтернативный случай, когда ЦМР считается более качественной, чем векторная гидрографическая сеть. Они предложили алгоритм DEM-Stream-Conflation (DSC), который идентифицирует водотоки, соответствующие реальным гидрографическим объектам, автоматически прослеживая их от начальной точки каждой гидрографической линии. Однако если данные не выровнены, начальная точка (источник) векторного гидрографического объекта может попасть в неправильный водосбор. Метод также не учитывает случай, когда трассировка соответствующего водотока невозможна из-за несоответствия направлений стока на ЦМР.

Алгоритмы генерализации ЦМР с ограничениями по дренажу опираются на дренажную сеть, полученную из ЦМР с помощью автоматизированных методов, таких как маскировка или трассировка раstra аккумуляции стока (площади водосбора) [O'Callaghan, Mark, 1984]. Однако такой подход не гарантирует, что извлеченные водотоки представляют особенности обобщенной гидрографической сети. В качестве решения такого рода проблем Гаффури [Gaffuri, 2007] разработал общую модель GAEL для одновременного обобщения векторных объектов и геополей с использованием агентно-ориентированного подхода. Для частного случая векторной гидрографии и поля высот его подход основан на определении пространственных отношений между гидрографией и элементами триангулированной ЦМР в исходном наборе данных. Если в процессе генерализации река смещается, то эти изменения могут быть перенесены

на поле высот путем соответствующего смещения координат треугольников. Однако это возможно только в том случае, если данные о рельефе и гидрографии изначально выровнены друг относительно друга, что является идеализированным случаем.

Проведенный обзор показывает, что разработано множество методов устранения несоответствий между ЦМР и векторной гидрографией. Однако существующие подходы не мотивированы качеством визуализации и поэтому не предполагают пространственного совмещения ЦМР и гидрографических линий. В то же время пространственные данные, совместно используемые на картах, часто поступают из разных источников. Существующие бесплатные базы пространственных данных, такие как Natural Earth [Patterson, Kelso, et al., 2020] и OpenStreetMap [OpenStreetMap, 2025], не содержат готовых слоев для отображения высот земной поверхности. Недостающие данные о высоте можно легко извлечь из глобальных ЦМР, таких как GMTED 2010 [Danielson, Gesch, 2011] или GEBCO [GEBCO, 2019], но полученный гипсометрический слой не будет согласован с гидрографическими объектами из базы данных в силу разного их происхождения. В конечном итоге это приведет к географически некорректным и эстетически неудовлетворительным картографическим результатам. Следовательно, необходимо разработать специализированный метод совмещения ЦМР с опорными объектами гидрографии. Этот метод должен быть максимально автоматизирован и встраиваться в процессы картографической генерализации.

#### 1.2.4.2 Координатная привязка

В практике картографа часто возникает необходимость осуществления *координатной привязки геоизображений* — сканированных бумажных карт и атласов, космических снимков с отсутствующими данными о привязке, карт в системах координат с неизвестными параметрами и т. д. Под геоизображением понимается любая пространственно-временная, масштабная, генерализованная модель земных (планетных) объектов или процессов, представленная в графической образной форме [Берлянт, 2006]. В настоящем исследовании рассматриваются картографические и снимковые геоизображения.

Координатная привязка геоизображения может осуществляться по опорным точкам, и в этом случае она требует задания их местоположения в системе координат с известными параметрами. Часто в качестве таких точек служат характерные элементы географических объектов — углы зданий, пресечения дорог, места слияния речных каналов и т. д. Для определения их координат может использоваться второе геоизображение, обладающее координатной привязкой. В этом случае возникает задача поиска идентичных точек на исходном (не привязанном) и эталонном (привязанном) геоизображении. На основе найденных точек формируется модель *взаимной привязки геоизображений*, которая определяет параметры трансформации непривязанного изображения в известную систему координат. И если для снимков и топографических карт существуют методы ав-

томатического поиска соответствующих точек [Li, Briggs, 2012; Cléri, Pierrot-Deseilligny, Vallet, 2014; Heitzler *et al.*, 2018; Luft, Schiewe, 2021], то для карт произвольного содержания и масштаба такая работа выполняется вручную, что занимает большое количество времени и может подвергаться ошибкам, обусловленным человеческим фактором. В связи с этим существует необходимость в разработке методики автоматизированного поиска опорных точек на двух произвольно заданных геоизображениях, покрывающих общую территорию.

На сегодняшний день уже были разработаны некоторые подходы в автоматизированной привязке геоизображений. Так, например, в исследовании [Cléri, Pierrot-Deseilligny, Vallet, 2014] описана многомасштабная привязка старых аналоговых снимков к современным цифровым по линейным объектам. Авторы работы [Wolter, Blank, Henrich, 2017] рассматривали автоматизированную географическую привязку растровых изображений речных сетей, применяя методику, схожую с приведенной в статье [Li, Briggs, 2012], но с отличием в использовании дорог и спутниковых снимков вместо речных сетей и карт. В исследовании [Howe *et al.*, 2019] была предложена модель деформируемых частей для автоматической географической привязки исторических карт. Это приближенно напоминает методику автоматического распознавания слов на исторических картах, где источником слов послужил географический справочник США [Weinman, 2013]. В работе [Weinman, 2017] автор дополнил ее возможностью использования информации о текстовых метках с учетом их стиля и географического положения. Ещё одна работа с историческими картами была проведена в статье [Yilmaz, Gullu, 2012], где для привязки исторической карты Кипра к современной использовались двумерное аффинное преобразование координат и искусственная нейронная сеть с обратным распространением. В публикациях [Heitzler *et al.*, 2018; Burt *et al.*, 2020] представлены методики координатной привязки отсканированных карт по картографической сетке. В статье [Luft, 2020] автор предложил методику привязки с использованием геокодирования. В еще одном исследовании этого же автора [Luft, Schiewe, 2021] рассматривался другой подход к географической привязке — с помощью картографического содержания карт, а именно — водных объектов.

Каждое из перечисленных исследований вносит вклад в развитие исследуемой проблематики, однако у них есть общий недостаток — отсутствие универсальности разработанных моделей. Все алгоритмы работают либо с определенными географическими объектами (реки, дороги, водоемы), либо с топонимами и картографическими сетками. В практике может возникнуть ситуация, когда специалисту потребуется осуществить привязку геоизображения, на котором нет нужного типа объекта или координатной сетки. Более того, весьма распространена ситуация, когда привязываемое геоизображение отличается своей детальностью от эталонного. На сегодняшний день решения данной проблемы предложено не было. Следовательно, наиболее актуальной задачей в этом направлении является разработка универсального алгоритма, который будет способен работать

со всем геоизображением в целом, а не с отдельными его элементами.

## 1.3 Мультимасштабное картографирование

### 1.3.1 Развитие концепции мультимасштабности<sup>13</sup>

В 1983 году Американским космическим агентством NASA была поставлена задача определить основные направления исследований в области пространственных данных. Результатом работы стала публикация [Marble, 1984], в которой среди прочих направлений была обозначена необходимость создания мультимасштабных представлений объектов (*multiscale representation*), т.е. формирования наборов данных, в которых одни и те же объекты были бы представлены несколько раз с различной степенью генерализации. Таким образом, можно сказать, что история мультимасштабного картографирования ведет свой отсчет с 1984 года.

Во второй половине 80-х гг. начали появляться работы, развивающие тему мультимасштабных представлений [Jones, Abraham, 1986; Abler, 1987; Rhind, 1988]. Основной упор в них делался на многоуровневой генерализации с последовательным формированием наборов данных разной детализации. Были разработаны методы сопряженной генерализации объектов из разных слоев и так называемой *rule-based*-генерализации на основе наборов правил [Brassel, Weibel, 1988]. Интерес к данной тематике возрастал, инициировались международные исследовательские проекты.

Одним из наиболее крупных стал проект *Multiple Representation*, запущенный в 1989 г. Национальным центром географической информации и анализа США (NCGIA). В процессе работы специалистами исследовательской группы были предложены различные модели автоматизированной генерализации [Buttenfield, McMaster, 1991; McMaster, Shea, 1992], а результаты были изложены в финальном отчете [Buttenfield, 1993]. Эта работа стала первым серьезным шагом в развитии мультимасштабного картографирования и послужила толчком к появлению учебных пособий и монографий, связанных с многоуровневой генерализацией и созданием мультимасштабных представлений, в частности [Müller, Lagrange, Weibel, 1995].

Задача создания мультимасштабных представлений оказалась тесно переплетена с технологиями баз данных, что привело к появлению такого понятия как база данных с множественными представлениями [Jones et al., 1996; Kilpelainen, 1997]. Термин «база данных с множественными представлениями» относится структуре базы данных, в которой несколько представлений одной и той же географической сущности или явления представлены в виде отдельных, связанных между собой объектов [Sarjakoski, Sarjakoski,

<sup>13</sup> При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Самсонов, 2009а; Самсонов, 2011]

2007]. Хранение организуется в виде нескольких уровней детализации, а иерархические связи могут носить как двунаправленный [Kilpelainen, 1997], так и однонаправленный характер [Harrie, 1998]. Концепция базы данных с множественными представлениями естественным образом связана с генерализацией модели.

Существует как минимум 3 подхода к наименованию таких баз:

1. База данных с множественными представлениями (multiple representation database)
2. База данных переменного разрешения (multiresolution database)
3. Мультимасштабная база данных (multiscale database)

Каждый из этих терминов характеризует базу со своей стороны. Множественность представлений подчеркивает наличие нескольких сущностей в БД, сопоставленных одному реальному объекту. Переменность разрешения говорит о наличии нескольких уровней детализации с различным разрешением. Наконец, мультимасштабность свидетельствует о возможности получать картографические представления в разных масштабах, в том числе и мультимасштабные карты. По всей видимости, нельзя отдать однозначного предпочтения какому-либо из этих терминов, однако в литературе наибольшее распространение получил вариант «база данных с множественными представлениями». Введем для таких баз данных следующее обозначение: *МБД*.

Необходимость применения МБД связана с тем, что алгоритмы автоматической генерализации до сих пор не удовлетворяют требованиям ни по качеству и географической достоверности результатов, ни по затратам времени на выполнение [Wang, Chen, Jian, 2003], и, следовательно, в интерактивной среде неприемлемы. Таким образом, обеспечить масштабирование одним набором базовых данных невозможно [Лурье, Самсонов, 2010]. Переключение между уровнями детализации БД позволяет имитировать интерактивную генерализацию, которая теоретически может осуществляться на основе одного детального набора данных.

Методы создания мультимасштабных БД интенсивно развиваются начиная с 1990-х годов. Они охватывают такие разносторонние аспекты, как разработка оптимальной структуры баз данных для МК [Jones *et al.*, 1996]; организация иерархических связей между представлениями объектов на различных уровнях [Yin, 2005], каскадное обновление данных [Hae-Kyong, Jung-Wook, Ki-Joune, 2004] и т.д. В отечественной литературе рассматриваются принципы построения объектно-ориентированных мультимасштабных геоинформационных систем в задачах экологического мониторинга, а также вопросы интеграции картографических данных в мультимасштабных ГИС на примере геоморфологических карт [Говоров, Хорев, 1997] и мультимасштабных общегеографических карт для интернета [Самсонов, 2009b].

Различают уровень детализации карты и уровень детализации базы данных. Один уровень детализации карты может использовать данные из нескольких уровней детали-

зации БД и наоборот. Когда речь идет об уровне детализации, всегда подразумевается наличие нескольких уровней, образующих иерархию. Уровень детализации диалектически связан с понятием масштаба, представляя его как бы «вывернутым наизнанку». Поэтому часто в терминах эти понятия подменяются друг другом, как, например, в случае с мультимасштабным представлением.

*Экземпляр представления (representation instance)* — это запись в базе данных, описывающая состояние объекта на определенном уровне детализации: его геометрию и семантику (значения атрибутивных полей). Один и тот же объект имеет единицу представления в каждом уровне детализации, включенном в его жизненный цикл.

*Жизненный цикл (объекта или единицы представления)* — совокупность уровней детализации, в пределах которых существует объект или его единица представления. Соответственно, различают жизненный цикл объекта и жизненный цикл единицы представления. Жизненный цикл объекта является объединением жизненных циклов его единиц представления.

Основные преимущества, представляемые МБД [Sarjakoski, Sarjakoski, 2007]:

- Гибкость поддержания: автоматическая трансляция обновлений с детального уровня на все остальные.
- Минимизация избыточности хранения: каждый объект представлен одной сущностью.
- Возможность автоматической проверки БД на предмет целостности с использованием связей.
- Повышение скорости получения данных.
- Мультимасштабный анализ географической информации.
- Возможность получать ориентированные на конкретные приложения производные картографические продукты.

При создании МБД возможно два подхода: аккумуляция независимо полученных наборов данных с организацией связей между ними либо получение мелкомасштабных уровней детализации путем генерализации модели. Для хранения геометрии множественных представлений возможно 2 подхода: множественные геометрические атрибуты у одного объекта [Jones et al., 1996] либо несколько представлений с одним геометрическим столбцом [Harrie, Hellström, 1999].

Килпелайнен в своих работах [Kilpeläinen, 1997; Kilpeläinen, 2000] подчеркивает разницу между картографическими базами данных (КБД) и МБД. МБД формируются в среде, управляемой алгоритмами генерализации модели, объекты в них связаны четко определенными топологическими правилами, у них нет необходимости конкурировать за графическое пространство. В свою очередь, КБД предназначены для визуализационных задач, и их формирование требует привлечения множества операций, таких как утри-

рование, смещение, сглаживание, символизация. Структурная единица верхнего уровня в МБД — уровень детализации [Kilpeläinen, 2000]. Каждый уровень оптимизирован для отображения в определенном диапазоне масштабов карты. Уровни делятся на слои. Обратный подход позволяет каждый слой внутри себя представить в виде нескольких уровней детализации [Vangenot, 2004]. Такой вариант более компактен, однако технологически пока недостаточно разработан и отсутствует в ГИС-пакетах. В то же время существуют специальные иерархические и пирамидальные типы данных, которые позволяют хранить ЦМР с переменным разрешением и детализацией [de Floriani, Marzano, Puppo, 1996].

В системах управления базами данных широкое распространение получили так называемые кубы данных, которые позволяют получать осмысленные абстракции тех данных, которые содержатся в реляционных таблицах, представляя их в агрегированном виде. Использование кубов данных позволяет осуществлять мультимасштабную визуализацию реляционных БД, в которых данные организованы в виде множества различных иерархий. Поскольку процесс визуализации происходит в многомерной среде, он происходит не линейно, а вдоль некоего графа, или иерархического дерева (zoom graph), позволяя пользователю менять переменные, участвующие в визуализации, как независимо друг от друга, так и в паре. Каждая вершина графа — это проекция куба данных, а ребро — это масштабный переход.

Одной из пионерных разработок в области мультимасштабной визуализации реляционных БД стала программная среда Polaris, разработанная в Стэнфордском университете [Stolte, Tang, Hanrahan, 2003]. В рамках эксперимента были реализованы следующие фиксированные сценарии навигации по базе данных:

- Стековые графики с возможностью детализации по оси Y и уменьшения грануляции по оси X
- Картограммы с переходом на более дробную сетку АТД
- Точечные графики с детализацией по отраслям
- Матрицы с регулируемой грануляцией по осям X и Y

Для формального описания графов масштабирования была разработана нотация. Система поддерживала только линейное масштабирование вдоль каждой оси и не предусматривала возможность ветвления графа (схема снежинки — snowflake scheme). Транспонирование осей также не предусматривалось.

В мультимасштабной визуализации реляционных БД участвует два типа абстракции: абстракция данных и визуальная абстракция [Stolte, Tang, Hanrahan, 2003]. Абстракция данных изменяет исходные данные, например путем выборки или агрегирования, перед их отображением. В свою очередь, визуальная абстракция применяет различные способы изображения и графические средства для морфинга изображения при его масштабировании. Прослеживается очевидная аналогия с генерализацией модели и карто-

графической генерализацией [Sarjakoski, Sarjakoski, 2007]. Под абстракцией данных понимается их осмысленное обобщение. В свою очередь, выбор абстракции соответствует определенной проекции куба данных. Определяющую роль в этом процессе играет отбор релевантных измерений и уровней детализации по этим измерениям. Разработанные методы далее легли в основу системы бизнес-аналитики Tableau [Tableau Software, 2025].

Подавляющее число исследований в области мультимасштабного картографирования на начальных этапах его развития было связано с генерализацией и созданием баз данных. Меньше внимания уделялось непосредственно картографированию. В 2000-е годы появились работы, посвященные интерактивной генерализации при веб-картографировании и созданию «карт по запросу» — *map-on-demand* [Torun, Köbber, Lemmens, 2000; Cecconi, Weibel, Barrault, 2002], особенностям дизайна мультимасштабных карт для мобильных устройств [Harrie, Sarjakoski, Lehto, 2002; Hampe, Sester, Harrie, 2004] и систем обеспечения автомобильной навигации [Li, Ho, 2004]. В большинстве исследований можно проследить унифицированный, двухстадийный подход к мультимасштабному картографированию [Frye, Eicher, 2003; Frye, 2006; Brewer, Buttenfield, 2007; Brewer, Buttenfield, 2009]:

1. Построение мультимасштабной базы данных, обеспечивающей картографирование в разных масштабах данными необходимой детализации. На этой стадии происходит ресурсоемкая генерализация данных.
2. Составление карты на основе подготовленной БД. На этой стадии происходит оформление данных и их интерактивная (насколько это возможно) генерализация.

В ряде работ была обозначена тесная взаимозависимость между оформлением карты и уровнем детализации данных. В частности, в работе [Bedard, Bernier, 2002] разработана концепция «вьюелов» (англ. VUEL — View Element), которая объединяет геометрическую, семантическую и семиотическую составляющую картографического объекта, т.е. геометрию его представления, классификационное положение и символику в зависимости от масштаба.

Методика интерактивного веб-картографирования, предложенная в работе [Cecconi, Weibel, Barrault, 2002], опирается на пределы применимости уровней детализации БД для использования в разных масштабах. Авторы помещают выбор условных обозначений в начало процесса генерализации, так как символика находится во взаимозависимости с необходимой степенью генерализации. Это согласуется с классическим принципом, который говорит о том, что генерализация начинается с легенды [Салищев, 1990].

Брюэр и Баттенфилд обращают внимание на то, что выбору символики в зависимости от масштаба просмотра уделяется недостаточно внимания [Brewer, Buttenfield, 2009].

В этой статье также анализируется «чувствительность» уровней детализации БД к масштабу отображения — для этого была разработана специальная диаграмма, получившая название ScaleMaster. Она показывает, на каких масштабах карты меняется уровень детализации данных, когда производится отбор объектов и изменение их символики.

Наконец, мультимасштабное картографирование нашло применение в серьезных проектах национального и глобального уровня. В качестве примера можно привести электронную версию топографических карт США [Brewer, Akella, 2008] и национальный атлас Швейцарии [Oberholzer, Hurni, 2000]. Экспериментально установлено, что разные объекты неодинаково чувствительны к изменению масштаба. Антропогенные объекты — дороги, здания, линии электропередач и т.д. — менее чувствительны к масштабу, поскольку часто имеют простой геометрический рисунок, построенный по математическим законам. Природные объекты имеют более сложный рисунок и, соответственно, более чувствительны к изменению масштаба. Для рельефа и гидрографии уменьшение масштаба в 2 раза уже требует существенной генерализации данных [Brewer, Buttenfield, 2007]. При этом методы генерализации ЦМР «на лету», как уже отмечалось, недостаточно разработаны, и требуется их имитация переключением уровней детализации.

В последние годы фокус исследователей стал смещаться в сторону тематического мультимасштабного картографирования [Roth *et al.*, 2019; Raposo, 2020; Becony t  *et al.*, 2022], однако исследования в этой области носят предварительный характер и нуждаются в теоретическом и методологическом обосновании.

## 1.3.2 Масштабные уровни и их преобразования

### 1.3.2.1 Общегеографические карты<sup>14</sup>

Разработка мультимасштабных карт основывается на применении ряда операций, обеспечивающих логичное и последовательное изменение картографического изображения при смене масштаба. Данные операции можно разделить на четыре группы, соответствующие изменению содержания, геометрии, символики и подписей на карте [Roth, Brewer, Stryker, 2011]. Изменение содержания и геометрии на карте входят в собственно классическое определение генерализации [Берлянт, 2014], автоматизация этих методов наиболее разработана и представлена в обзорных публикациях, в частности [Stanislawski *et al.*, 2014]. В то же время показано, что оперирование символикой объектов при создании мультимасштабных карт выполняет важнейшую функцию и позволяет не только заменить или отложить в ряде случаев применение генерализации [Brewer, Buttenfield,

<sup>14</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Самсонов, Подольский, 2014]

2009], но также придать ее результатам визуально логическую последовательность, отражающую смену единиц картографирования [Самсонов, Подольский, Юрова, 2013].

Предоставление мультимасштабных карт онлайн в режиме «по запросу», или «on-demand», когда внешний вид карты (состав объектов, детализация, символика и т. д.) определяется требованиями пользователя [Sarjakoski, Sarjakoski, 2007], предполагает полную автоматизацию всех четырех составляющих генерализации. Следует отметить, что несмотря на актуальность данного вопроса, количество опубликованных работ, посвященных автоматизации преобразований символики, а также связанных с ними геометрических преобразований, невелико. Так, например, в статье [Choi, Hwang, 2009] предложен подход, в котором для отображения каждого слоя предлагается использовать базу знаний, которая определяет символику объекта на трех уровнях детализации. Переключение между разными символами происходит автоматически при достижении критических масштабов. Однако определение самих символов для каждого уровня детализации и самих критических масштабов производится пользователем вручную, что не позволяет говорить о полноценной автоматизации этих процессов.

Выбор и преобразование способов изображения — сложный многофакторный процесс, автоматизация которого требует многостадийного подхода. В первую очередь, необходимо описание процессов, представляющее собой совокупность общих правил преобразования способов изображения, рекомендаций по изменению символов на карте при уменьшении масштаба. К примерам таких правил можно отнести логичность и последовательность изменения графических переменных, их циклическое заимствование в рамках объектов одной иерархии, переход от аналитических способов изображения к комплексным и синтетическим, уменьшение информационной и графической нагрузки [Самсонов, Подольский, Юрова, 2013]. Описание правил логически дополняется системой операторов изменения содержания, геометрии, символики и подписей [Roth, Brewer, Stryker, 2011]. При наличии разработанных правил может быть предложена формализация выбора и преобразования способов изображения, которая уже более конкретным образом предписывает, какие именно операторы следует применять в соответствии с условиями конкретной ситуации. Конечным результатом формализации является алгоритм или набор алгоритмов. Наконец, успешная реализация алгоритмов в программном виде и их тестирование с использованием реальных данных позволяет говорить об автоматизации выбора и преобразования способов изображения.

Одним из важных направлений последних лет в области автоматизации проектирования масштабных уровней стали работы по трансляции оформления из одного масштаба в другой, а также генерации промежуточных представлений с использованием нейронных сетей [Dumont, Touya, Duchêne, 2018; Dumont, Touya, Duchêne, 2020; Christophe *et al.*, 2022], исследования по коммуникативным проблемам использования мультимасштабных карт [Touya, Gruget, Muehlenhaus, 2023; Touya, Potié, Mackaness, 2023].

### 1.3.2.2 Рельеф и другие геополья<sup>15</sup>

Отправной точкой для исследований по мультимасштабным представлениям рельефа послужили работы по генерализации цифровых моделей (ЦМР), рассмотренные ранее в диссертации. Помимо собственно алгоритмов генерализации, хорошим подспорьем в развитии мультимасштабных методов послужили разработки из области вычислительной геометрии, применяющиеся в средах виртуального моделирования, компьютерных играх и т.д. Трехмерные объекты в подобных задачах представлены, как и рельеф ЗП, в виде поверхностей. Для них также необходимо формировать уровни детализации, которые бы позволили осуществлять максимально эффективную визуализацию (рендеринг) сцены [Lindström *et al.*, 1996; Duchaineau *et al.*, 1997; Cheng, 2000]. Наиболее быстрые из этих методов всегда опираются на заранее рассчитанное мультимасштабное представление поверхности в базе данных.

В то же время, как показывает изучение литературы, вопросам мультимасштабного картографирования рельефа не уделено достаточно внимания. Исследования, так или иначе затрагивающие визуализацию мультимасштабных моделей рельефа, ориентированы на их тематическую обработку. Так, например, целый ряд работ посвящен анализу масштабных изменений морфометрических характеристик рельефа, таких как расчлененность, кривизна, углы наклона и т.д. [Wood, 1996; Gallant, Hutchinson, 1997; Пузаченко, Онуфреня, Алещенко, 2002; Drăguț, Blaschke, 2006; Arrell, Carver, 2009; Grohmann, Smith, Riccomini, 2011]. Эванс, а также Динеш и Ахмад Фадзил автоматизированно выделяют границы форм рельефа по цифровым моделям различного разрешения, далее анализируя размеры выделенных объектов в каждом масштабе [Evans, 2003; Dinesh, Ahmad Fadzil, 2007]. Драгут и Блашке представили метод классификации форм рельефа различного ранга на основе анализа ЦМР [Drăguț, Blaschke, 2006]. Мультимасштабный анализ орографических барьеров по ЦМР позволил установить степень и характер их влияния на перераспределение осадков по территории [Garvert, Smull, Mass, 2007]. Тьен Тай и Дайа Сагар используют мультимасштабные ЦМР для выделения геофизических сетей [Tay, Sagar, Chuah, 2005]. Пузаченко, а также Пэйн применяют мультимасштабный анализ для изучения рельефа как фактора ландшафтной дифференциации [Пузаченко, 1997; Пузаченко, Онуфреня, Алещенко, 2002; Pain, 2008].

Все эти исследования не являются собственно картографическими и используют изображение лишь для иллюстрации тематических показателей в разных масштабах, а мультимасштабные ЦМР — как источник данных для анализа. Специфика подготовки моделей для использования на мультимасштабных картах, приемы их визуализации различными способами изображения практически нигде не рассматриваются. За исключе-

<sup>15</sup> При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov, 2011; Самсонов, 2011; Myslenkov *et al.*, 2023]

нием работы Брюэр и Баттенфилд, где горизонталы изучаются как один из общегеографических слоев на предмет чувствительности к изменению масштаба и границ применимости уровней детализации [Brewer, Buttenfield, 2009], разработкам собственно по мультимасштабному картографированию не уделяется должного внимания.

Результаты исследований гидрометеорологических полей обычно представляются в виде статических или анимированных карт. В некоторых случаях для визуализации данных требуется специальное программное обеспечение. Хорошим альтернативным примером является приложение APPMAR 1.0, которое помогает визуализировать и анализировать данные о ветре и волнах из баз данных Национального управления океанических и атмосферных исследований (NOAA) [Rivillas-Ospina *et al.*, 2022]. Однако разрешение волнового реанализа NOAA достаточно грубое, а пользователю необходимо знать язык программирования Python.

Электронные веб-атласы более удобны для пользователей, так как предоставляют возможность выбора необходимых параметров, района и масштаба карты. Есть пример такого атласа для побережья Австралии [Hemer *et al.*, 2017] и для Черного моря [Bingölbali *et al.*, 2020]. Атлас глобальной энергии волн, основанный на грубом ретроспективном прогнозе (разрешение  $1,25^\circ \times 1^\circ$ ), представлен в работе [Arinaga, Cheung, 2012]. Ирландский морской атлас (<https://atlas.marine.ie>) — это региональный проект с данными о ветре и волнах высокого разрешения. Несколько веб-сервисов предоставляют данные о распределении энергии ветра и волн с грубым пространственным разрешением глобальном Мировом океане или в локальных регионах [Larsén *et al.*, 2022]. Атлас возобновляемых источников энергии для территории России (<http://gisre.ru/>) позволяет пользователю оценить потенциал солнечной или ветровой энергии [Gridasov *et al.*, 2011]. Особенностью перечисленных ресурсов является отсутствие полноценной мультимасштабности, т.е. генерализации картографического изображения в соответствии с масштабом визуализации.

### 1.3.2.3 Транспортная сеть: полимагистрали<sup>16</sup>

Транспортная сеть, как и многие пространственные системы, обладает сложной внутренней организацией, которая характеризуется наличием дорог разного типа и класса, объединенных в топологическую структуру и вместе обеспечивающих сложную систему транспортных связей и потоков [Тархов, 2005]. Широкий диапазон географического охвата и неравнозначность элементов транспортной сети говорит о том, что ей, как и многим географическим явлениям, свойственна полимасштабность организации [Трейвиш, 2007]. Отсюда следует, что аспекты картографирования транспортных сетей, связанные с масштабом, следует трактовать не только узко — в разрезе изменения

<sup>16</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Самсонов, Прохорова, 2019]

масштаба и уровня детализации карты, но и в контексте отображаемых на ней объектов, которые могут иметь различный пространственный (расстояния и зоны обслуживания), семантический (транспортная нагрузка, интенсивность сообщения) и временной (продолжительность, цикличность перевозок) масштаб.

Масштаб транспортного потока может формироваться не только высокой интенсивностью сообщения по одной магистрали, но также сочетанием сонаправленных близко идущих путей сообщения разного типа, образующих линейно вытянутые кластеры. Для обозначения таких сочетаний отечественным географом П.М.Поляном введено понятие полимагистрали [Полян, 1988]. Выделение полимагистралей представляет собой один из важных методов генерализации транспортных сетей на тематических картах.

Известно, что географический рисунок транспортной сети формируется под воздействием ряда факторов, отражая как экономико-, так и физико-географические условия [Kohl, 1841]. В свою очередь, его особенности существенно влияют на экономические связи и во многом определяют уровень хозяйственного развития территории [Баранский, 1960]. В соответствии с этим при генерализации изображения транспортных сетей на картах принимаются во внимание такие важные функции путей сообщения, как обеспечение связности населенных пунктов, перемещение грузов и населения, доступность центров обслуживания [Евтеев, 1999]. Введение полимагистралей на тематических картах позволяет выделить комплексный характер транспортного сообщения на ключевых направлениях перевозок. В то же время, понятие полимагистрали, основанное на оценке географической близости путей сообщения, может реализовываться в широком спектре пространственных масштабов, поскольку результат выделения полимагистралей зависит от расстояния, используемого в качестве порогового значения близости.

Строгое обоснование выбора тех или иных алгоритмов генерализации является трудно разрешимой задачей [Флоринский, 1991], тем не менее существует ряд зарекомендовавших себя подходов, связанных с геометрическим обобщением рисунка транспортной сети. С.А. Тархов [Тархов, 2005] определил ряд критериев отбора социально-экономических элементов на картах, и среди них один из самых значимых — принцип «географической типичности объектов», который, с одной стороны, весьма отличен от формальных подходов к отбору объектов, а с другой — может учитывать их геометрическую сущность в пределах одного типа территории. Отбор элементов транспортных сетей осуществляется по разным принципам в зависимости от масштаба картографирования. В крупных масштабах особое внимание уделяется отбору наиболее значимых дорог, выделению сквозных проездов и городских магистралей [Thompson, Soni, Weatherill, 1999; Touya, 2010].

При переходе к мелким масштабам картографирования нюансы транспортной сети отходят на второй план и основными принципами отбора элементов становятся сохранение связности населенных пунктов и топологической структуры сети [Weiss, Weibel, 2014]. Слияние дорог часто применяется при крупномасштабном картографировании,

когда полосы движения преобразуются в коридоры движения и далее в осевые линии магистралей. Для этого могут быть использованы как алгоритмы конфляции (совмещения) данных [Saalfeld, 1988; Walter, Fritsch, 1999], так и методы, основанные на построении медианной оси [Nickerson, 1988].

В работах по картографированию транспортных сетей важную роль играют не только методы генерализации, но и различные абстракции, возникающие при описании полимасштабной структуры транспортной системы на высоких уровнях ее организации. К таким абстракциям относятся полимагистрали и транспортные коридоры. С одной стороны, полимагистрали объединяют несколько путей сообщения, образуя пространственно более крупный объект, покрывающий некоторую полосу, с другой — аккумулируют в себе пропускную способность и типы входящих в их состав транспортных путей, что наделяет их более крупным масштабом характеристик — как пространственных, так и содержательных (семантических). Под транспортным коридором понимается определенное направление (маршрут) массовых перевозок пассажиров и грузов между центрами социально-экономической активности, которое обслуживается несколькими видами транспорта, способными обеспечить высокие скорости перевозки [European Yearbook 1994, 1996]. Транспортные коридоры вводятся, как правило, в сверхмелких масштабах картографирования при необходимости сделать акцент на общегосударственный и международный аспект работы транспорта.

В своей монографии П.М.Полян [Полян, 2014] отмечал, что важнейшей проблемой в изучении и картографировании полимагистралей является определение критериев и пороговых расстояний для их выделения. Эта проблема отмечается также в работах [Воробьев, 1973; Weiss, Weibel, 2014]. Очевидно, что постановка задачи выделения полимагистралей и транспортных коридоров с геометрической точки зрения родственна процедуре слияния, упомянутой ранее, однако к настоящему моменту методы и технологии, решающие эту проблему в приложении к мультимасштабному картографированию транспортных сетей, не разработаны.

#### 1.3.2.4 Картограммы и картодиаграммы<sup>17</sup>

Социально-экономические карты часто основываются на наборах статистических данных, которые положены на сетку административно-территориального деления, образующую партномическую иерархию [Tversky, 1989], или привязаны к объектам точечной локализации. Для отображения этих данных чаще всего используются картодиаграммы (для абсолютных показателей) или картограммы (для относительных показателей), для точечных данных большую популярность приобрело кластерное отображение с группировкой точек по пространственной близости, которое по сути представляет собой

<sup>17</sup> При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Самсонов, Юрова, 2014]

традиционный точечный способ. Процесс формализации выбора способов изображения при отображении тематических данных рассмотрен в работе [Иванов, Булыгина, 2012], однако нас интересуют более низкий уровень формализации, при котором происходит определение параметров уже выбранного способа изображения. Кроме того, требуется учесть эффекты возникающие при изменении масштаба.

Тема мультимасштабной визуализации статистики активно обсуждается в бизнес-аналитике, где различные диаграммы и графики используются для исследования статистических агрегатов больших баз данных, получаемых средствами *OLAP* (*online analytical processing*). В работах [Stolte, Tang, Hanrahan, 2003; Wang *et al.*, 2010] рассмотрены программные среды визуализации данных для систем онлайн-процессинга. Эти методы реализованы в таких системах, как *Dundas Dashboard* [Dundas Data Visualization, 2009] и *Tableau* [Tableau Software, 2025], обладающих также простейшими функциями картографирования.

К недостаткам данных средств социально-экономического картографирования можно отнести независимость выбора способов изображения от конфигурации территориальных единиц, а также неучет масштабных эффектов, связанных с изменением иерархического уровня картографирования и статистического распределения картографируемой величины при смене уровня детализации карты. В то же время эти аспекты чрезвычайно важны в системах картографирования по запросу, где пользователю предлагается уже готовое решение, которое должно быть оптимальным.

Если сами способы картографического изображения к настоящему моменту образовали достаточно устоявшуюся систему [Берлянт, 2014], то возможности их трансформации в интерактивной среде разработаны лишь фрагментарно. Исследовательский анализ пространственных данных опирается на многовариантность визуального представления географической информации, в том числе в разных масштабах картографирования и с применением различных преобразований данных и способов изображения [Kraak, Ormeling, 2021; Slocum *et al.*, 2022]. При этом исследования в области визуализации данных обеспечили возможность программирования анимационных переходов между графическими представлениями информации, которые способствуют пониманию того, как трансформируются данные при смене метода визуализации [Heer, Robertson, 2007; Ruchikachorn, Mueller, 2015; Kim, Heer, 2021]. Недавняя работа [Lee *et al.*, 2023] показала, что такие плавные переходы возможны между пространственными (значки, картограммы) и непространственными (диаграмма рассеяния, столбчатая диаграмма) представлениями. При этом в картографии область применения анимаций традиционно ограничена показом пространственно-временной динамики явлений [DiBiase *et al.*, 1992; Берлянт, Ушакова, 2000; Medyńska-Gulij, Forrest, Cybulski, 2021]. Применение же анимаций для визуализации переходов между способами картографического изображения, в частности, картограмм и картодиаграмм, остается недостаточно изученным, в том числе в мультимасштабной среде.

## 1.4 Некартографические приложения генерализации

Генерализация пространственных данных используется не только для картографических приложений. Методологический приём генерализации находит применение в самых разных сферах наук о Земле применительно к обработке пространственных данных. В целом можно разделить такие приложения на несколько крупных групп:

1. **Обработка** — генерализация «сырых» пространственных данных для того чтобы сделать их содержательно интерпретируемыми.
2. **Агрегирование** — статистическое обобщение информации по заданной сетке территориальных единиц.
3. **Моделирование** — разработка и реализация математических моделей для разных уровней детализации (в том числе путем переноса данных с одного уровня на другой).
4. **Масштабирование или апскейлинг** (англ. *upscaling*) — обобщение территориально разрозненных высокодетальных пространственных данных на территории со сплошным и более крупным охватом.

Рассмотрим кратко некоторые примеры из данных областей применения, не претендуя на полноту обзора, но ставя цель показать разнообразие некартографических задач в области обработки пространственных данных, в которых используется прием генерализации.

### 1.4.1 Обработка

Обработка «сырых» пространственных данных, полученных по результатам измерений и наблюдений, в ряде случаев требует их генерализации для того, чтобы из данных можно было извлечь полезную информацию (интерпретировать). В особенности эта процедура оказывается востребована при обработке *данных дистанционного зондирования (ДДЗ)*: снимков и облаков точек, полученных методом лазерной локации (лидаром). Проблема этих данных часто заключается в избыточном количестве информации и/или наличии шумов, которые убираются с помощью специализированных методов фильтрации и классификации, которые явно соотносятся с такими процедурами генерализации, как отбор, сглаживание, обобщение качественных и количественных характеристик. Помимо этого, такие данные могут быть комбинированы с альтернативными источниками информации, имеющими другую детальность, что также требует согласования.

Первое, о чем следует сказать, что данные дистанционного зондирования, и в частности космические снимки, сами по себе обладают эффектом генерализации, поскольку имеют ограниченную разрешающую способность. В частности пиксельная структура растровых снимков способна выявлять объекты лишь определенного порядка [Кравцова, 2000]. Классификация растровых изображений согласно концептуальной

модели [McMaster, Monmonier, 1989] относится к количественно-категориальной генерализации, при которой количественные переменные (в случае снимка — яркость в отдельных каналах) переходят в качественные или порядковые (в случае снимка — идентификаторы классов). С этой точки зрения любые методы классификации и сегментации, применяемые для обработки снимков и других изображений [Lillesand, Kiefer, Chipman, 2015; Gonzalez, Woods, 2017; Sabins, Ellis, 2020; Chuvieco, 2023], можно отнести к методам генерализации пространственных данных. Однако принцип генерализации проникает существенно глубже в обработку данных ДДЗ. Результаты множества независимых классификаций могут быть объединены в один результат с помощью метода т.к. стековой генерализации, который имеет обобщающий эффект и повышает надежность и устойчивость результата [Wolpert, 1992]. Так, например, в работе [Healey et al., 2018] метод стековой генерализации использован для выявления изменений лесного покрова путем комбинирования результатов 8 разных алгоритмов методом случайного леса (*random forest*), что позволило уменьшить ошибки классификации в среднем в 2 раза по сравнению с применением отдельно взятых алгоритмов.

В статье [Becker-Reshef et al., 2010] предложена регрессионная модель предсказания урожайности озимых с эффектом генерализации, основанная на комбинации снимков MODIS (AVHRR) и детальной статистики по урожайности. В процессе обработки данных авторы столкнулись с проблемой существенного шума данных AVHRR, который удалось минимизировать методом *Maximum Value Composite (MVC)*, выбирающим пиксели максимального качества из каждого временного фрейма [Holben, 1986]. В работе [Chen et al., 2014] генерализация использована как подход к обобщению принципа слияния (*fusion*) изображений применительно к комбинации гиперспектральных снимков низкого пространственного разрешения и более детальных многозональных снимков. А в исследовании [Gafurov et al., 2023] генерализующая способность нейросетей типа LSTM (Long Short-Term Memory) использована для применения модели распознавания сельскохозяйственных земель, обученной на одной территории (Канада), для дешифрирования аналогичных земель на другой территории (Европейская территория России).

Следует также упомянуть о работах, в которых приемы, весьма близкие к картографической генерализации, используются для повышения качества визуального представления космических снимков и результатов их классификации. Так, в статье [Li, Shen, Dai, 2024] предложен оригинальный способ повышения наглядности мультимасштабного представления ДДЗ, при котором возникает проблема неразличимости мелких островов и их групп. В предложенном подходе острова сначала отделяются от моря методом К-средних, затем агрегируются путем суперпиксельной сегментации, а после этого проливы между объединенными островами заполняются методом текстурного синтеза. Другим примером применения генерализации при обработке ДДЗ является работа [Strand et al., 2024], в которой предложен подход к генерализации классов наземного покрова (*land cover*), выделенных по ДДЗ (в работе использованы снимки *Sentinel 2*). Схема дешифри-

рования классов наземного покрова разделяется на отдельные бинарные маски классов, каждый из которых подвергается геометрической генерализации с использованием морфологических фильтров [Serra, Vincent, 1992], а затем результаты агрегируются в единое растровое покрытие, обладающее более высокой степенью обобщения и наглядностью, чем изначальный результат классификации.

Другим источником пространственных данных, остро нуждающимся в генерализации, являются данные, получаемые с помощью устройств лазерной локации. Лидарные облака точек нуждаются в фильтрации, классификации, восстановлении по ним поверхностей и границ объектов [Dong, Chen, 2018; Shan, Toth, 2018; Thenkabail, 2024], которые по сути также представляют собой результаты генерализации исходных данных, имеющих точечную локализацию. Так, в одной из ранних работ [Filin *et al.*, 2007] рассмотрена проблема восстановления земной поверхности с явно представленными дорогами по данным лазерного сканирования с использованием различных приемов генерализации. Особенно актуальной для лидарных данных является проблема упрощения облака точек, т.е. исключения из нее нерелевантных, чему посвящено большое количество работ [Song, Feng, 2009; Zhang *et al.*, 2018; Ji *et al.*, 2019; Fang, Fan, 2021; Hu *et al.*, 2024]. Однако для упрощения облаков точек может быть использован не только отбор, но также и агрегирование. Так, в исследовании [Petras, Newcomb, Mitasova, 2017] агрегирование лидарных данных по трехмерным (воксельным) ячейкам позволило количественно охарактеризовать фрагментарность растительного покрова, а классификация вокселей по их расположению относительно внутренней и внешней области леса с последующей фильтрацией дала еще более тонкую оценку пространственной структуры растительности.

Аналогично космическими снимками, для моделей восстановления характеристик зондируемых объектов по лидарным данным актуальна проблема генерализации, т.е. применения модели обученной на ограниченной выборке (как территориально, так и по видам зондируемых объектов), на «незнакомых» ей данных [Kim, Im, 2025]. В частности, такая проблема исследована применительно к инвентаризации характеристик лесного покрова [Bouvier *et al.*, 2015; Knapp *et al.*, 2020], выделения различных типов поверхностей методами классификации и сегментации [Weidner, Walton, Kromer, 2020; Kim *et al.*, 2023; Eskandar, 2024], навигации с использованием компьютерного зрения [Zhu, Hayashibe, 2023; Zhi *et al.*, 2025].

#### 1.4.2 Агрегирование

Проблема *агрегирования* информации сопровождает пространственный анализ в тех случаях, когда множество объектов объединяются в один, либо когда в пределах крупных географических объектов (единиц административно-территориального деления, физико-географических границ и т.д.) или ячеек регулярной сетки собирается статистика по более мелким вложенным объектам. В случае когда у исследователя есть

выбор агрегирующих объектов, возникает так называемая проблема модифицируемой территориальной единицы — MAUP [Openshaw, 1984], смысл которой заключается в том, что результаты вычисляемых статистик существенно зависят от расчетной сетки, и возникает задача выбора оптимальной формы и размеров ячеек [Fotheringham, Wong, 1991; Jelinski, Wu, 1996; Zhang, Kukadia, 2005; Wang, Di, 2020]. Агрегирование, наряду с интерполяцией, является очень важным звеном процесса генерации пространственно-временной информации [Scheider *et al.*, 2016; Krueger *et al.*, 2021], в рамках которого сырые пространственные данные могут проходить через сложные цепочки трансформаций. При этом оказывается важным различие интенсивных (не зависящих от опоры измерений) и экстенсивных (зависящих от опоры измерений) пространственных переменных [Stasch *et al.*, 2014; Scheider, Ballatore, 2018]. Значения первых (например, плотности населения) должны осредняться при агрегировании, в то время как значения вторых (например, численности населения) — суммироваться.

Одним из современных и актуальных примеров использования агрегирования является задача подготовки регулярно-сеточных баз данных характеристик подстилающей поверхности для информационного обеспечения моделей городской погоды и климата [Masson *et al.*, 2020]. Уже сейчас в этой области активно используются региональные модели атмосферы с шагом сетки 1 км и менее, позволяющие учитывать мозаичность городских ландшафтов и их влияние на атмосферу [Wouters *et al.*, 2016]. Для описания взаимодействия атмосферы с застройкой на таком масштабе применяют так называемые параметризации урбанизированной поверхности [Garuma, 2018]. При этом предполагается, что ячейка расчетной сетки модели или ее часть занята квази-однородной городской средой, характеризующейся ограниченным набором параметров. Успешность моделирования метеорологического режима городов во многом определяется точностью задания входных параметров городской среды. Задаче их определения посвящено большое число исследований, однако выбор оптимального метода по-прежнему носит дискуссионный характер [Masson *et al.*, 2020; Lipson *et al.*, 2022; Cheng *et al.*, 2024]. Дополнительной проблемой является то, что данные для подготовки баз данных таких параметров необходимо собирать из различных источников, что неизбежно приводит к проблеме их согласования для корректного агрегирования соотношения разных видов поверхностей. Универсального решения этой проблемы к настоящему моменту не получено.

### 1.4.3 Моделирование

В задачах моделирования генерализация может быть полезна как инструмент обобщения натуральных наблюдений или модельных данных для выявления генеральных закономерностей. Наиболее ярко это проявляется в различных направлениях геофизики и гидрометеорологии. Так, например, в работе [Dickinson, Gavin, 2011] предложен спо-

соб введения параметрических моделей сейсмических возмущений, откалиброванных по натурным данным, для решения анализа сейсмической опасности при недостаточной плотности пунктов измерений. Эффект генерализации отмечается в гидрологическом анализе и моделировании при использовании данных, покрывающих большие водосборы — это позволяет выявлять генеральные закономерности соотношения осадков и стока, но вместе с тем утрачивается глубина анализа, типичная для исследований на малых водосборах [Gupta *et al.*, 2014].

Результатами таких исследований могут быть модели, которые объясняют поведение гидрологической системы в целом, в противоположность детальным моделям прогноза стока. Так, например, в работе [Wang, Tang, 2014] предложена модель водного баланса, которая обобщает (генерализует) ряд других моделей, объясняющих соотношение осадков, стока и испарения на водосборах исходя из принципов предложенных М.И.Будыко [Будыко, 1956]. Вместе с тем, отмечается проблема масштабирования региональных моделей стока на более крупные водосборы, которая на современном этапе эффективно решается посредством искусственных нейронных сетей типа LSTM, способных как учитывать локальные соотношения осадков и стока, так и их генеральные закономерности [Kratzert *et al.*, 2019].

Также весьма показательные примеры двунаправленного процесса генерализации — детализации в гидрометеорологическом моделировании. В частности, в исследовании [Ford *et al.*, 2022] разработана модель прогноза поверхностного стока, в которой сначала на основе данных о расходах выделяются пространственные кластеры со схожим потенциалом стока (генерализация), а далее для каждого кластера калибруется гибридная модель стока, опирающаяся как на физические принципы, так и на модель машинного обучения для прогнозирования стока уже с километровым разрешением (детализация). Аналогичный принцип использован в статье [Bolton, Zanna, 2019], где проблема недостаточного разрешения океанологических данных (таких как направление и скорость движения водных масс) решается путем обучения сверточной нейронной сети на исходных и генерализованных (с помощью сглаживающей фильтрации) данных модели высокого разрешения. Это позволяет далее предсказывать с ее помощью подсеточные вихревые процессы (имеющие масштаб меньше размера ячейки).

Аналогично рассмотренным ранее проблемам, для моделей глубокого обучения в геофизике тоже актуальна проблема генерализации — переносимости модели на новые объекты, на которых она не обучалась, что часто является одним из критериев оценки качества модели [Sonnewald *et al.*, 2021]. В качестве примера можно привести исследования по разработке моделей обнаружения океанических вихрей [Duo, Wang, Wang, 2019], прогноза высоты волн [Jörges, Berkenbrink, Stumpe, 2021], оценке суммарной облачности [Krinitskiy *et al.*, 2021], восстановлению геополей по данным сейсморазведки [Zhang *et al.*, 2020].

#### 1.4.4 Масштабирование (апскейлинг)

Термином *масштабирование* или *апскейлинг* (от англ. upscaling) обычно обозначают экстраполяцию данных, полученных в отдельно взятых локациях, на более крупные территории, вплоть до глобального охвата [Harvey, 2000]. Очевидно, что в данном процессе имеет место обобщение информации, а значит он может рассматриваться как еще одна область использования генерализации пространственных данных как методологического приема. При этом обобщенные данные обычно соотносятся с уровнями организации явлений, которые выделяются в соответствии с принципом полимасштабности [Трейвиш, 2007; Хорошев, 2016].

Примеры масштабирования в географии весьма многочисленны и разнообразны. Внимания заслуживают исследования растительного покрова, в частности работы по масштабированию листового индекса (leaf area index) путем моделирования зависимости его измеренных в натуре значений и характеристик данных дистанционного зондирования [Williams et al., 2008; Tripathi et al., 2014]. Если измерения возможно провести на очень ограниченных территориях, то мозаики снимков высокого и среднего разрешения могут покрывать огромные площади, что и дает возможность обобщать результаты локальных измерений в региональных и глобальных масштабах. Аналогичный подход используется и при оценки биомассы различных видов [Clifton-Brown et al., 2017; Villoslada et al., 2024], а в работе [Pang et al., 2023] акцент сделан на масштабировании уже сезонных паттернов листового индекса и биомассы путем совместного использования полевых данных и космических снимков.

Отдельным направлением можно назвать оценку потоков влаги и углекислого газа в атмосферу. Так, в работе [Ryu, 2024] исследована проблема масштабирования потоков CO<sub>2</sub> в атмосферу путем совместного применения натуральных наблюдений и данных дистанционного зондирования, предложена концептуальная схема масштабирования в глобальном охвате. В исследовании [Li et al., 2018] для целей масштабирования данных эвапотранспирации выполнено сравнение 6 разных моделей, включая осреднения, симуляции, геостатистический подход и 3 разных метода машинного обучения. По итогам работы авторы рекомендуют использовать ансамбль взвешенного осреднения и регрессионного кригинга для гомогенных территорий, и случайного леса для территорий с гетерогенным наземным покровом.

Достаточно близкими являются проблемы масштабирования почвенных данных, для которых требуется генерализация закономерностей выявленных на точках измерений и их экстраполяции по обширным территориям. В качестве примера можно привести исследования по масштабированию данных о влажности почв [Crow et al., 2012; Zhang et al., 2017; Brown et al., 2023] и запасах органического углерода [Martin et al., 2014; Peng et al., 2015; Ladd et al., 2022], в каждом из которых для целей масштабирования строятся модели зависимости между измерениями и характеристиками поверхности, извле-

каемыми из внешних источников: аэро- и космических снимков, лидарных данных, цифровых моделей рельефа.

Важную роль масштабирование играет в геоморфологических исследованиях. Так, в работе [Davy, Crave, 2000] локальные зависимости интенсивности эрозионных процессов от крутизны склонов и поверхностного стока использованы для обоснования макроскопических закономерностей в масштабах крупнейших водосборов. Принципиально другой диапазон масштабов эрозии — от миллиметровых к сантиметровым — исследован в работе [Yuan, Kennedy, Ierodiaconou, 2022] применительно к береговым процессам. В исследовании [Zhang, Lin, 2020] выполнена геоморфологическая регионализация территории в районе Пекин-Тяньцзинь, которая опирается на генерализацию ЦМР путем приведения к более низкому разрешению методом HASM (High Accuracy Surface Modeling) [Yue *et al.*, 2015].

Достаточно остро проблема масштабирования стоит и в геологическом моделировании, где нет возможности сплошного покрытия всего объема измерениями. По этой причине различные подходы к масштабированию геологических данных используются в седиментологии [Liu *et al.*, 2002; Hodge, Hoey, 2012], гидрогеологии [Noëtinger, Artus, Zargar, 2005; Quental, Almeida, Simões, 2012], моделировании пористых сред и потоков веществ [Rasaei, Sahimi, 2009; Milad *et al.*, 2020], анализе структурных и линеаментных данных [Carmichael, Ailleres, 2016].

В гидрологическом анализе масштабирование используется для обобщения информации о локальных направлениях стока на континентальные масштабы. Такая задача весьма востребована в моделях Земной системы, которые воспроизводят состояние атмосферы, гидросферы и других сфер в глобальном масштабе для целей анализа и прогноза их состояния. Подходящие для этих целей растровые модели направлений стока получают путем генерализации моделей D8 [O'Callaghan, Mark, 1984] более высокого разрешения (по сравнению с целевым разрешением сетки). Для этого используют специализированные алгоритмы, такие как: COTAT и COTAT+ Paz, Collischonn, Lopes Da Silveira [2006], DMM [Olivera *et al.*, 2002], DRT Wu *et al.* [2012] и IHU [Eilander *et al.*, 2021].

Смежная область метеорологии и климатологии также нуждается в методах масштабирования своих данных. Так, например, в работе [Muthusamy *et al.*, 2017] предложен подход к геостатистическому масштабированию данных осадкомеров для дальнейшего использования в гидрологическом моделировании. Особое внимание в последние десятилетия уделяется городской климатологии, в рамках которой также предпринимаются попытки обобщить климатические процессы и закономерности, наблюдаемые в крупных городах, на глобальный масштаб и выработать практики их возможного регулирования и адаптации к ним [Creutzig *et al.*, 2019; Creutzig *et al.*, 2020].

Наконец, весьма оживленная дискуссия относительно возможностей масштабирования информации традиционно идет в социально-экономической географии. Так, в работе [Zen *et al.*, 2019] предложена методика агрегирования информации об экосистемных

услугах по единицам административно-территориального деления, которая позволила выявить кластеры различных видов сервисов. Исследование [Hemerijckx *et al.*, 2020] посвящено масштабированию информации характеристик выборочных городских домохозяйств, таких как количество детей, доход, доступность источников питьевой воды и т.д., на масштаб всего города с применением данных дистанционного зондирования. Большое разнообразие подходов к масштабированию обсуждается в области моделирования взаимодействия человека и окружающей среды [Dressler *et al.*, 2022; Friis *et al.*, 2023; Arlati, 2024].

## 1.5 Результаты и выводы

В результате проведенного обзора была получена картина современного состояния исследований в области генерализации пространственных данных, выявлены ее картографические и геоинформационные аспекты, в том числе связанные с оценкой уровня детализации данных, особенностями их визуализации в интерактивной компьютерной среде с возможностью изменения масштаба изображения.

Основные выводы по Главе 1 заключаются в следующем:

1. Развитие методик генерализации пространственных данных неразрывно связано с усилением их географической содержательности посредством явного учета особенностей картографируемых объектов в алгоритмах генерализации, оперирующих цифровыми моделями объектов. Этот вывод лежит в основе *принципа адаптивности*, который определяет необходимость учета пространственной организации картографируемых объектов. Данный принцип далее раскрывается в Главе 2.
2. Цифровая трансформация картографии и ее переориентация на геоинформационные методы составления карт, оперирующие пространственными данными, привели к поиску адекватной альтернативы понятию масштаба. Такой альтернативой стало понятие уровня детализации. Этот вывод лежит в основе *принципа детальности*, который определяет необходимость оценки подробности цифровых моделей географических объектов не масштабом, а количественными характеристиками их геометрии и атрибутов. Данный принцип далее раскрывается в Главе 3.
3. Возможность интерактивного взаимодействия с картами на цифровых устройствах, а также запрос исследователей на необходимость представления географических объектов на разных уровнях их организации естественным образом породили требование к возможности адаптации картографического изображения к изменению масштаба карты. Этот вывод позволяет сформулировать *принцип мультимасштабности*, постулирующий необходимость интерактивной гене-

рализации пространственных данных в соответствии с масштабом их визуализации в цифровой среде. Данный принцип раскрывается в Главе 4.

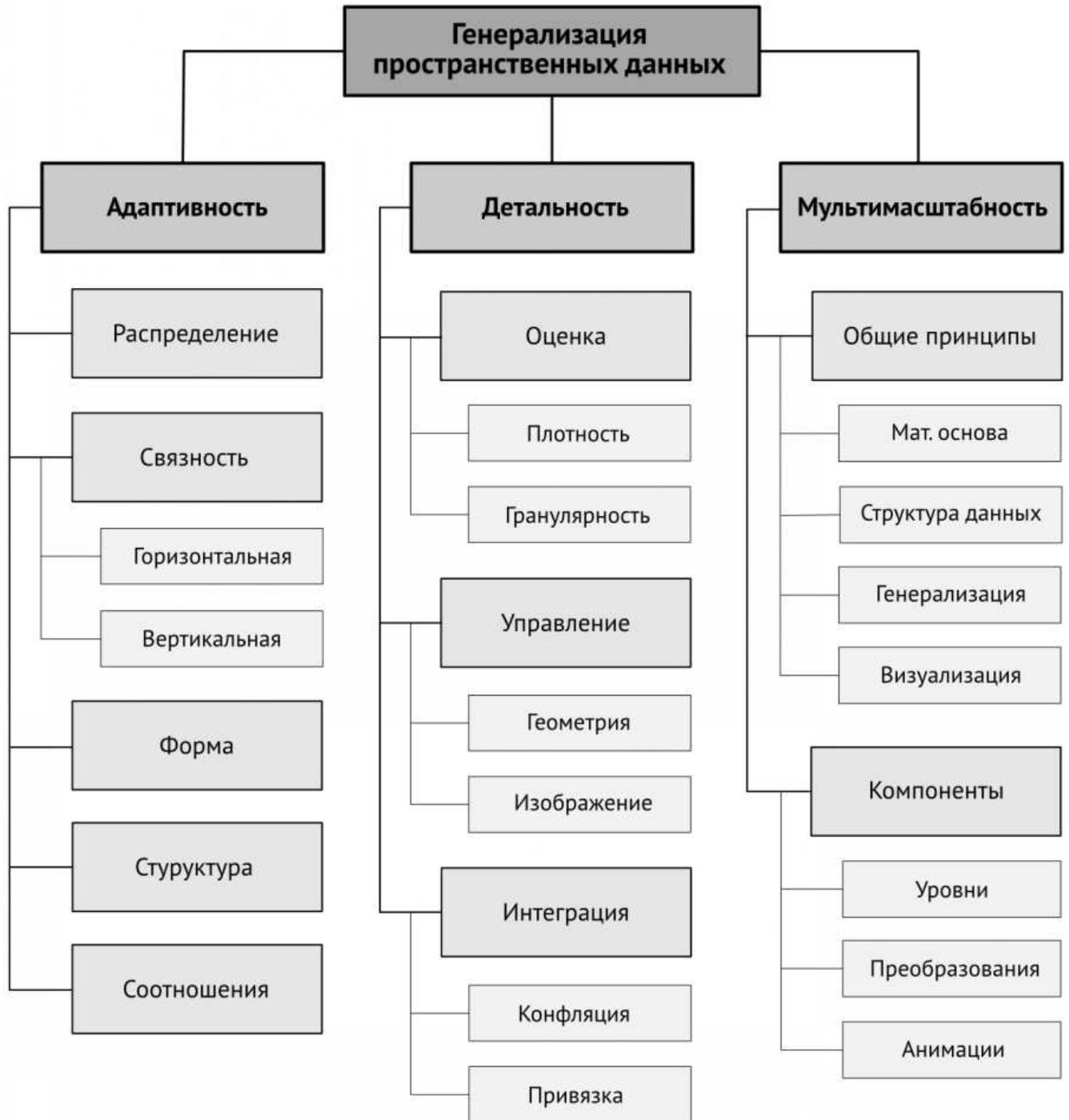
4. Генерализация как методологический прием используется не только в картографических задачах, но также и для обработки «сырых» пространственных данных, их агрегирования по сетке территориальных единиц, а также в задачах пространственного моделирования и масштабирования (апскейлинга).

Обобщение выводов 1-3 позволяет сформулировать первое защищаемое положение настоящей диссертации:

#### **Защищаемое положение №1**

*Генерализация пространственных данных для целей картографирования опирается на принципы адаптивности, детальности и мультимасштабности. Применение данных принципов позволяет реализовать учет таких факторов генерализации как особенности картографируемых объектов, необходимый уровень детализации выходных данных (в том числе определяемый масштабом карты), а также интерактивность среды их визуализации при работе с цифровыми моделями географических объектов.*

Дальнейшая структура исследования обусловлена этими выводами и упорядочена в соответствии с Рисунком 1.1. Адаптивность генерализации пространственных данных рассматривается на примерах учета распределения, связности, формы, структуры и соотношений географических объектов, которые можно выявить по их цифровым моделям. Детальность пространственных данных исследована с точки зрения ее оценки, управления ею, а также интеграции пространственных данных различной детальности. Наконец, мультимасштабность в генерализации пространственных данных раскрывается через общие принципы мультимасштабного картографирования, а также реализацию ее компонент на примерах мультимасштабных карт различного содержания и тематики.



**Рисунок 1.1** – Принципы генерализации пространственных данных для целей картографирования и различные аспекты их реализации

## 2 Адаптивная генерализация пространственных данных

В главе рассмотрен принцип *адаптивности*, определяющий необходимость учета пространственной организации картографируемых объектов. Реализующая его группа методов генерализации учитывает форму объектов, структуру и плотность их пространственного распределения, характер пространственных отношений между ними. Распознавание и воспроизведение этих особенностей обеспечивает реализацию географических принципов картографической генерализации и таким образом повышает достоверность получаемых пространственных данных и карт.

### 2.1 Отбор точечных моделей населенных пунктов с учетом пространственной неравномерности их распределения<sup>1</sup>

Одним из первых факторов, влияющих на сложность генерализации карт, является **пространственное распределение** объектов. В частности, проблема *неравномерности* распределения весьма характерна для населенных пунктов, плотность которых выше в исторически более освоенных регионах.

В настоящем разделе диссертации представлены методика и алгоритм отбора населенных пунктов для мелких масштабов картографирования, позволяющие ввести территориальную дифференциацию интенсивности отбора, а также производить его автоматическим путем [Samsonov, Krivosheina, 2012; Самсонов, Кривошеина, 2015a]. Результаты отбора оценены численно и графически.

Разработанная методика состоит из нескольких шагов, которые последовательно рассмотрены в настоящем разделе:

1. Деление территории на области (классы), обладающие разной плотностью размещения населенных пунктов.
2. Ранжирование и вычисление весов точек.
3. Итеративный отбор точек по выделенным классам плотности размещения.

---

<sup>1</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Самсонов, Кривошеина, 2015a]

#### 4. Численная, графическая и визуальная оценка результатов отбора.

Методика нацелена на мелкомасштабную генерализацию, в качестве исходных данных использована цифровая топографическая карта России масштаба 1 : 1 000 000, предоставленная компанией Дата+ по состоянию на конец 2013 г.

### 2.1.1 Деление территории

Деление территории на области, однородные по густоте размещения населенных пунктов, используется для того, чтобы обеспечить раздельное управление интенсивностью отбора в малообжитых и освоенных районах. В разработанной методике деление опирается на построение *диаграммы Вороного*. Согласно определению, диаграмма Вороного конечного множества точек  $P$  на плоскости представляет такое разбиение плоскости, при котором каждая область этого разбиения образует множество точек, более близких к одному из элементов множества  $P$ , чем к любому другому элементу множества [Voronoi, 1908]. Пример диаграммы Вороного и обозначение ее элементов представлены на Рисунке 2.1. Как видно, ячейки диаграммы представляют собой своеобразные зоны влияния исходных точек  $P$  (обозначенных серым цветом), при этом площадь ячейки тем меньше, чем выше локальная плотность размещения точек. Именно эта особенность используется в разработанной методике.

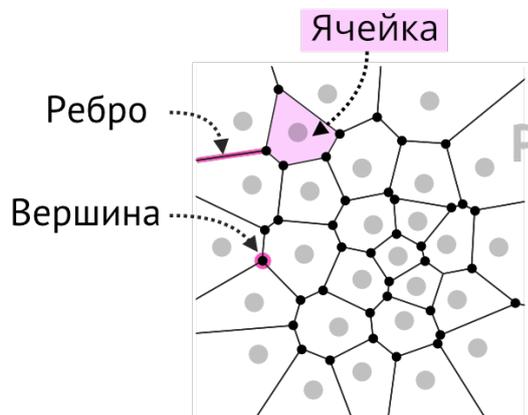


Рисунок 2.1 – Диаграмма Вороного и ее элементы

Предложенная последовательность действий при районировании выглядит следующим образом:

1. Построение диаграммы Вороного для всего множества точек.
2. Расчет площадей полученных полигонов.
3. Классификация полигонов по значениям площадей полученных полигонов.
4. Объединение полигонов Вороного одного класса.

Приемлемый результат в случае населенных пунктов России дает классификация методом *естественных интервалов*, в котором классы основаны на естественном группи-

ровании данных. Границы классов определяются таким образом, чтобы сгруппировать схожие значения (минимизировать дисперсию внутри каждого класса) и максимально увеличить различия между классам [Jenks, 1967].

В результате выполненного деления методом естественных интервалов получены 4 класса густоты населенных пунктов, передающих различия в заселенности страны. На первый (самый густозаселенный) класс приходится более 90% всех населенных пунктов, на второй — 7%, на третий — 1%, на четвертый менее 0.5. Выбранное количество классов было подобрано эмпирически путем визуальной оценки рисунка размещения населенных пунктов на карте масштаба 1 : 1 000 000.

Районирование территории как результат объединения полигонов Вороного показано на Рисунке 2.2, а увеличенный фрагмент одного из регионов — на Рисунке 2.3.

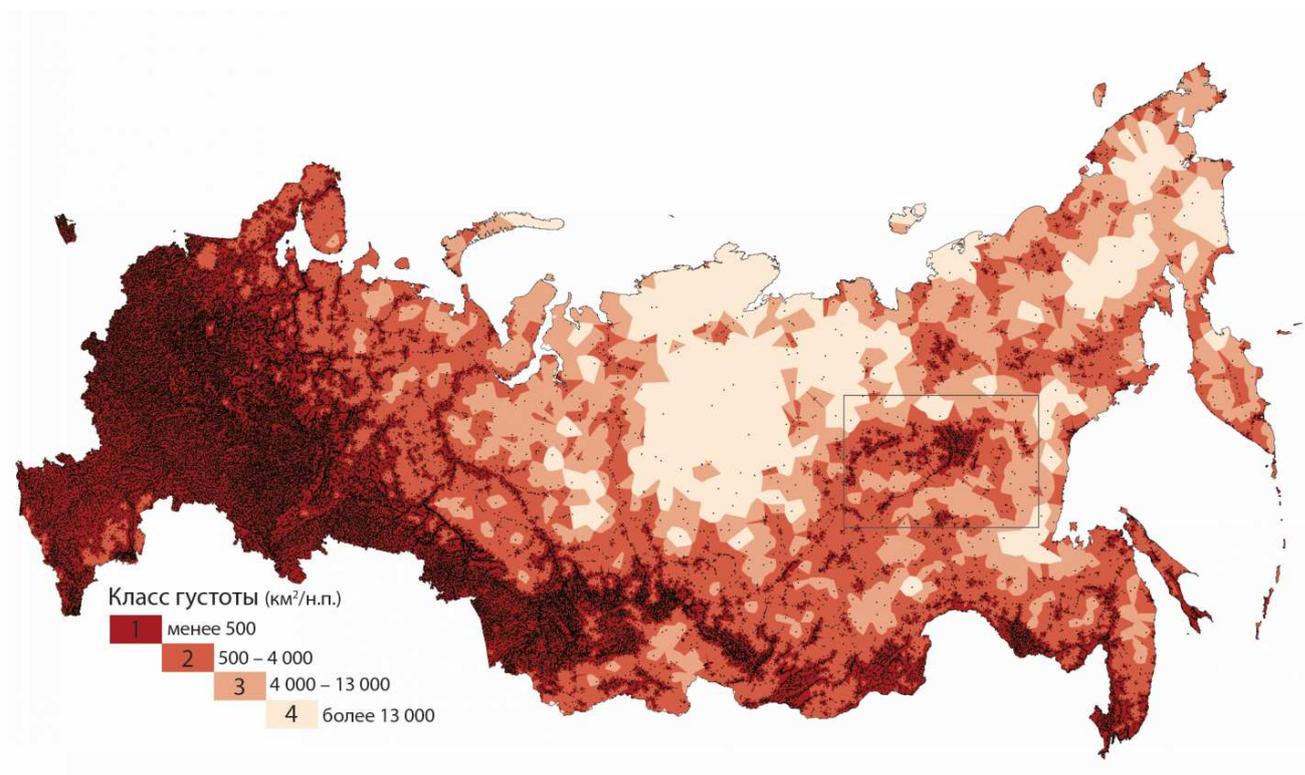
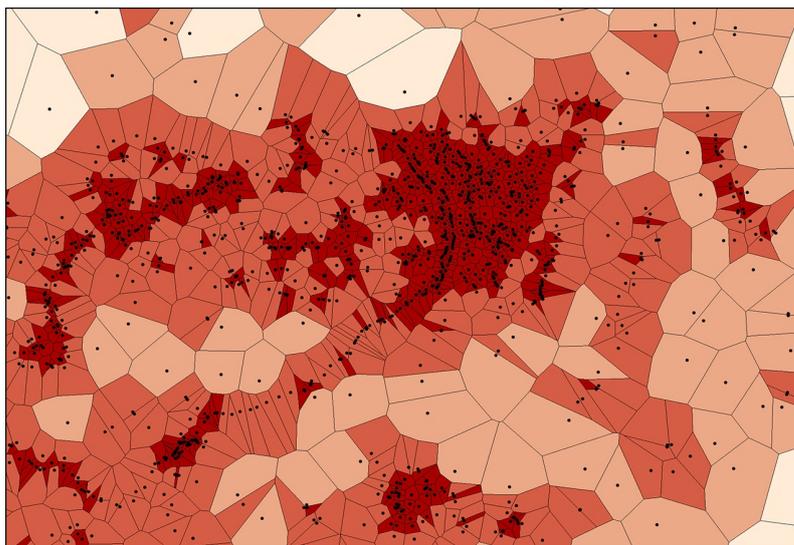


Рисунок 2.2 – Деление территории России по густоте населенных пунктов (данные на 2013 г.)

### 2.1.2 Ранжирование и вычисление весов точек

Для ранжирования населенных пунктов были использованы их административный статус и численность населения. По административному статусу были выделены следующие пять классов (переменная  $W_C$ ):

1. Города федерального значения (5);
2. Центры федеральных округов (4);
3. Центры субъектов РФ (3);



**Рисунок 2.3** – Фрагмент диаграммы Вороного, соответствующий выделенному прямоугольнику на Рисунке 2.2

4. Центры муниципальных районов и городских округов (2);
5. Прочие населенные пункты (1).

По численности населения выделены также пять классов (переменная  $W_P$ ):

1. Более 1 млн чел. (5);
2. 500 000 – 1 млн чел. (4);
3. 100 000 – 500 000 чел. (3);
4. 10 000 – 100 000 чел. (2);
5. Менее 10 000 чел (1).

Итоговый вес точки, учитывающий площадь полигона Вороного  $A$ , численность населения и административный статус населенного пункта, вычисляется по формуле

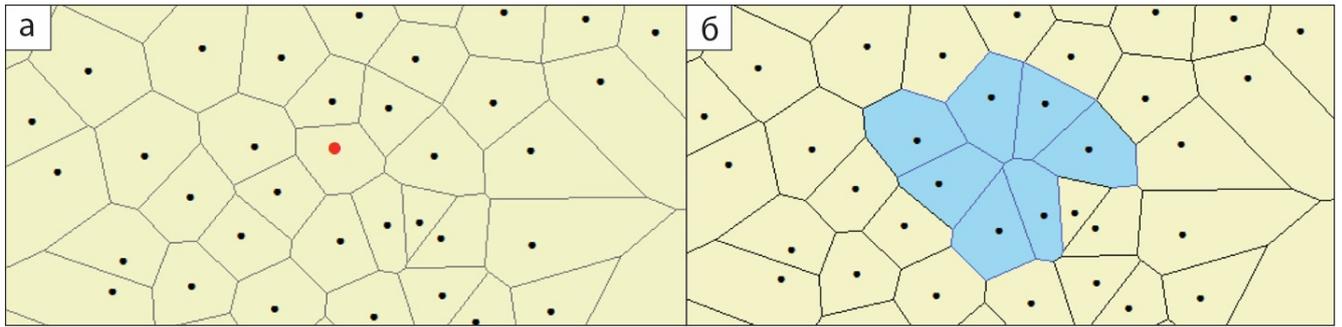
$$W = A \times W_C \times W_P, \quad (2.1)$$

Основным фактором является площадь, а численность населения и административный статус «взвешивают» ее, увеличивая вплоть до 25 раз.

### 2.1.3 Алгоритм отбора точек

Отбор населенных пунктов для итогового масштаба основан на последовательном исключении точек с минимальным весом. При этом происходит локальное перестроение диаграммы Вороного. Этот процесс проиллюстрирован на Рисунке 2.4.

Заметим, что прямолинейное следование этой стратегии без дополнительных условий приведет к тому, что алгоритм начнет удалять точки только в областях с наибольшей их густотой. Эффект «заморозки» алгоритмически не очень удобен, поскольку он пред-



**Рисунок 2.4** – Перестроение диаграммы Вороного при удалении точки. а – удаляемая точка выделена красным цветом; б – полигоны Вороного соседних точек покрывают освободившуюся площадь

полагает проверку наличия доступных для удаления точек (если все точки оказались «замороженными», то этот параметр сбрасывается). Вместо этого предлагается параметр задержки  $S$  (*Suspension*). Данный параметр изначально для всех точек имеет значение 0. Точкам, полигоны Вороного которых являются смежными с полигоном удаленной на предыдущей итерации точки, к значению параметра  $S$  прибавляется 1 (на Рисунке 2.4б эти полигоны выделены голубым цветом).

После того как построена диаграмма Вороного и рассчитаны исходные веса, алгоритм представляет собой следующую последовательность операций:

1. Сортировка точек по убыванию значений  $S$  (первичная) и  $W$  (вторичная).
2. Удаление последней в отсортированном списке точки.
3. Локальное перестроение диаграммы Вороного для точек, смежных с удаленной точкой.
4. Увеличение параметра  $S$  на 1 для смежных точек.
5. Пересчет  $W$  для смежных точек с учетом обновленной площади.

Перечисленные шаги повторяются до тех пор, пока не будет исключен заданный процент точек  $P$ . Таким образом, точки, для которых удалены соседи, оказываются вверху списка и не могут быть удалены при последующей итерации. После того как все точки получают параметр  $S = 1$ , они становятся равноправными, и удаление в наиболее плотных областях снова возможно. Через несколько итераций точки, у которых удалено  $n$ -е количество соседних, имеют параметр  $S = n$ . В итоге множество точек, получаемое на выходе работы алгоритма, обогащается значением числа удаленных соседей вокруг каждой точки, что можно использовать при построении карт для оценки интенсивности генерализации в разных регионах.

Данным алгоритмом последовательно обрабатывается каждый полигон каждого класса густоты. Если задано удалить 50% точек, то при обработке первого в списке полигона из него будет удалена половина точек. Таким же образом будут обработаны все остальные полигоны. Отметим, что генерализация по классам, когда все точки одного

класса рассматриваются как одно множество, приводит к некорректным результатам, т. к. точки будут исключаться из полигонов с несколько большей плотностью, а более разреженные полигоны того же класса будут менее затронуты процедурой отбора.

#### 2.1.4 Методика отбора с использованием алгоритма

Программная реализация алгоритма отбора точек была выполнена на языке *Python*. Апробация производилась на примере данных цифровой топографической карты масштаба 1 : 1 000 000, предоставленных компанией Дата+ (актуальность на 2013 г.).

При выполнении значительной генерализации (в 5—10 раз) наиболее контролируемый процесс отбора точек достигается не при единоразовом запуске алгоритма с заданным процентом (например 90%), а путем последовательной генерализации в несколько итераций, на каждой из которых удаляется 50% населенных пунктов. В частности, при выполнении экспериментальных работ для масштабного перехода от 1 : 1 000 000 к 1 : 8 000 000 генерализация была проведена в пять последовательных стадий.

Многостадийный отбор позволяет реализовать гибкий процесс выбора точек, в котором возможна регулировка соотношения плотности между разными классами плотности в результирующем масштабе. Так, например, при выполнении генерализации для масштаба 1 : 8 000 000 обнаружилось, что одинаковые параметры алгоритма при обработке полигонов разных классов приводят к чрезмерному прореживанию слабозаселенных регионов. Чтобы уменьшить этот эффект, первые три стадии были проведены для всех населенных пунктов, четвертая была выполнена без отбора в четвертом классе, и пятая — без отбора в третьем и четвертом классе плотности. Таким образом, классы плотности исключаются из процесса отбора каскадным образом.

Помимо этого, многостадийный отбор позволяет учесть гранулярность первого класса плотности, который часто представлен множеством близкорасположенных небольших полигонов с малым количеством входящих точек, при этом полигоны разделены областями 2-го класса. Подобная ситуация типична, например, для агломераций, образующих центральные места. В результате генерализации в каждом таком полигоне с высокой вероятностью остается (а иногда может и изначально быть) по одной точке. Если параметр  $P$  более 50%, то такие точки будут удалены и все полигоны окажутся пустыми. Анализ этой ситуации привел к выводу, что в алгоритм должен быть включен запрет на удаление единственной точки в полигоне 1 класса.

После введения данного условия поведение алгоритма изменилось таким образом, что в классах наивысшей плотности за счет гранулярности удаляется меньше точек, чем задано. Для устранения этого эффекта после первой итерации генерализации (с удалением 50% точек) первые два класса плотности были объединены для последующих итераций.

Для формализации технологического процесса ведем следующие функции:

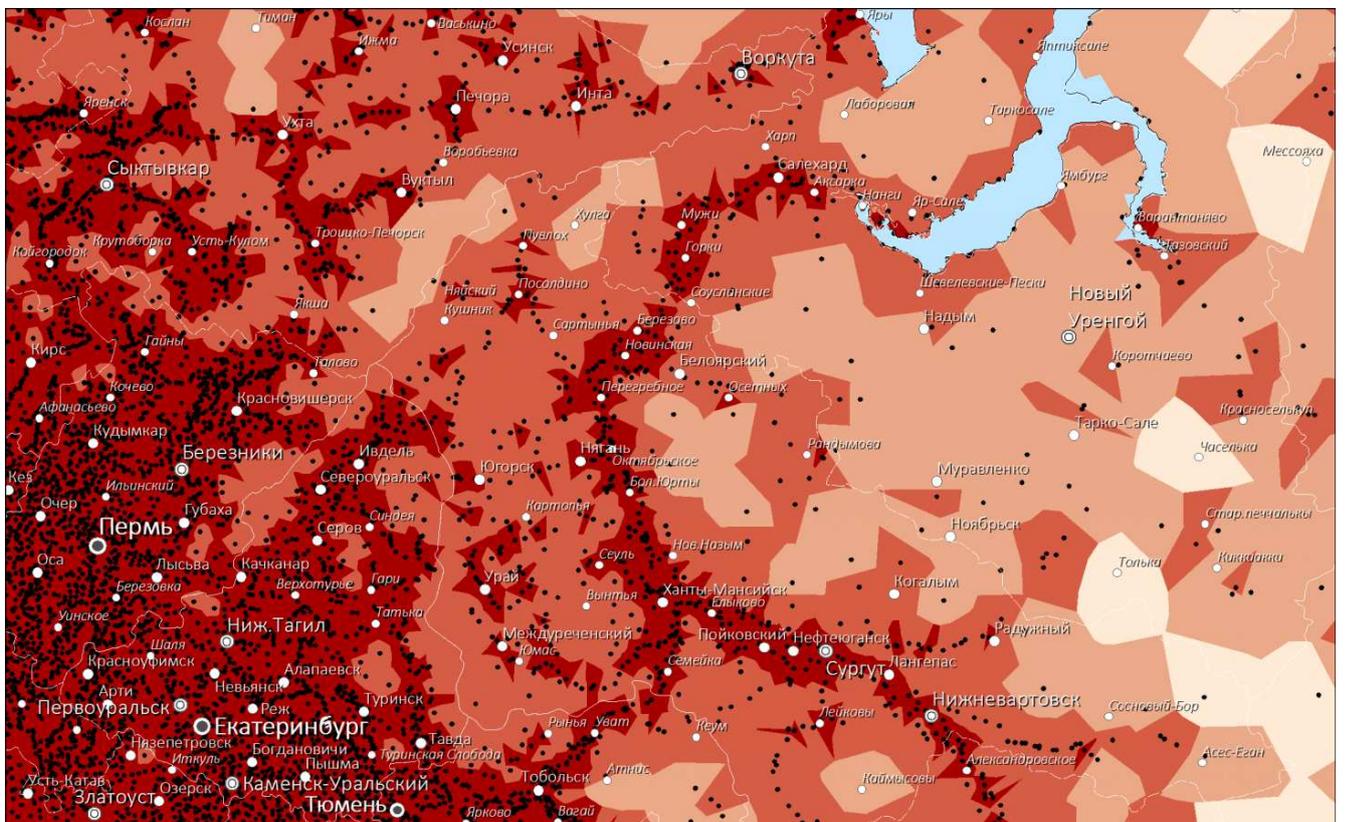
1.  $E_{i=1}^n(i, P)$  — исключение  $P\%$  точек в полигонах  $i$ -го класса от 1 до  $n$ ;

2.  $U(i, j)$  – объединение  $i$ -го и  $j$ -го класса.

В терминах данных функций общий технологический процесс, который был использован для отбора населенных пунктов при масштабном переходе с 1 : 1 000 000 до 1 : 8 000 000, выглядел следующим образом:

$$E_{i=1}^n(i, 50) \rightarrow U(1, 2) \rightarrow E_{i=1}^n(i, 50) \rightarrow E_{i=1}^n(i, 50) \rightarrow E_{i=1}^{n-1}(i, 50) \rightarrow E_{i=1}^{n-2}(i, 50) \quad (2.2)$$

В результате работы алгоритма в целом сохраняется относительная густота размещения населенных пунктов по территории, основной отбор производится в самых густозаселенных регионах и сохраняется достаточное их количество в слабозаселенных (Рисунок 2.5).



**Рисунок 2.5** – Результаты отбора (север Западной Сибири и Урала). Черными точками показаны исходные населенные пункты

Разработанный технологический процесс не претендует на универсальность и был применен конкретно для отбора населенных пунктов в заданном диапазоне масштабов. Однако собственно алгоритм исключения точек по классам густоты остается неизменным. Задачей дальнейших исследований является автоматизация выбора числа классов густоты и их границ, последовательности операций исключения точек, объединения и исключения классов в зависимости от степени уменьшения. Стоит также отметить, что исключение классов может быть заменено на разные значения  $P$ .

### 2.1.5 Численная и графическая оценка результатов

Существует ряд исследований, рассматривающих связь между количеством объектов на картах разных масштабов. Одни из первых исследований по изменению пространственного отображения при смене масштаба были отражены в статье [Töpfer, Pillewizer, 1966]. Топфер сформулировал так называемые «принципы отбора» и «основной закон» для выражения отношения между масштабом карты и количеством отображаемых на ней объектов:

$$N_G = N_S \times \sqrt{\frac{S_S}{S_G}} \quad (2.3)$$

где  $N_G$ ,  $N_S$  — число объектов в целевом и исходном масштабе, а  $S_G$  и  $S_S$  — знаменатели целевого и исходного масштабов.

В данную формулу может быть введен ряд коэффициентов, учитывающих, например, назначение карты ( $C_P$ ), форму картографических знаков ( $C_F$ ), которые в свою очередь зависят от масштаба. В этом случае формула примет следующий вид:

$$N_G = C_P \times C_F \times N_S \times \sqrt{\frac{S_S}{S_G}} \quad (2.4)$$

Поскольку значения  $C_P$  и  $C_F$  являются константами, а список потенциальных факторов ими не исчерпывается (т.е. формула не является универсальной), Ли [Li, 2006] предложил это выражение упростить и вынести все факторы в единственный степенной параметр  $x$ :

$$N_G = N_S \times \sqrt{\left(\frac{S_S}{S_G}\right)^x} \quad (2.5)$$

Параметр  $x$  определяет степень отбора при генерализации и принимает следующие характерные значения: 0 — нет отбора, интервал (0, 4) — увеличение густоты размещения объектов, 4 — сохранение густоты, > 4 — уменьшение густоты.

Данная модификация закона была применена для оценки результатов генерализации населенных пунктов. Имея исходный и результирующие масштабы, а также количество населенных пунктов до и после генерализации по классам густоты, был определен параметр  $x$  для каждого класса (Таблица 2.1).

На основе полученных значений параметра  $x$  можно оценить, насколько изменилась густота размещения населенных пунктов по территории по выделенным классам густоты. В первом классе густота после генерализации несколько уменьшилась, в остальных густота увеличилась, причем наибольшее ее увеличение наблюдается в последнем классе. Это объясняется стремлением показать в самых слабозаселенных регионах максимальное количество населенных пунктов и меньшим процентом их отбора. В среднем же по всей территории России густота размещения населенных пунктов сохранена

Таблица 2.1 – Оценка степени отбора

Класс	Значение параметра $x$
1	4.2
2	3
3	2
4	1.2
Среднее	3.8

с небольшим ее увеличением.

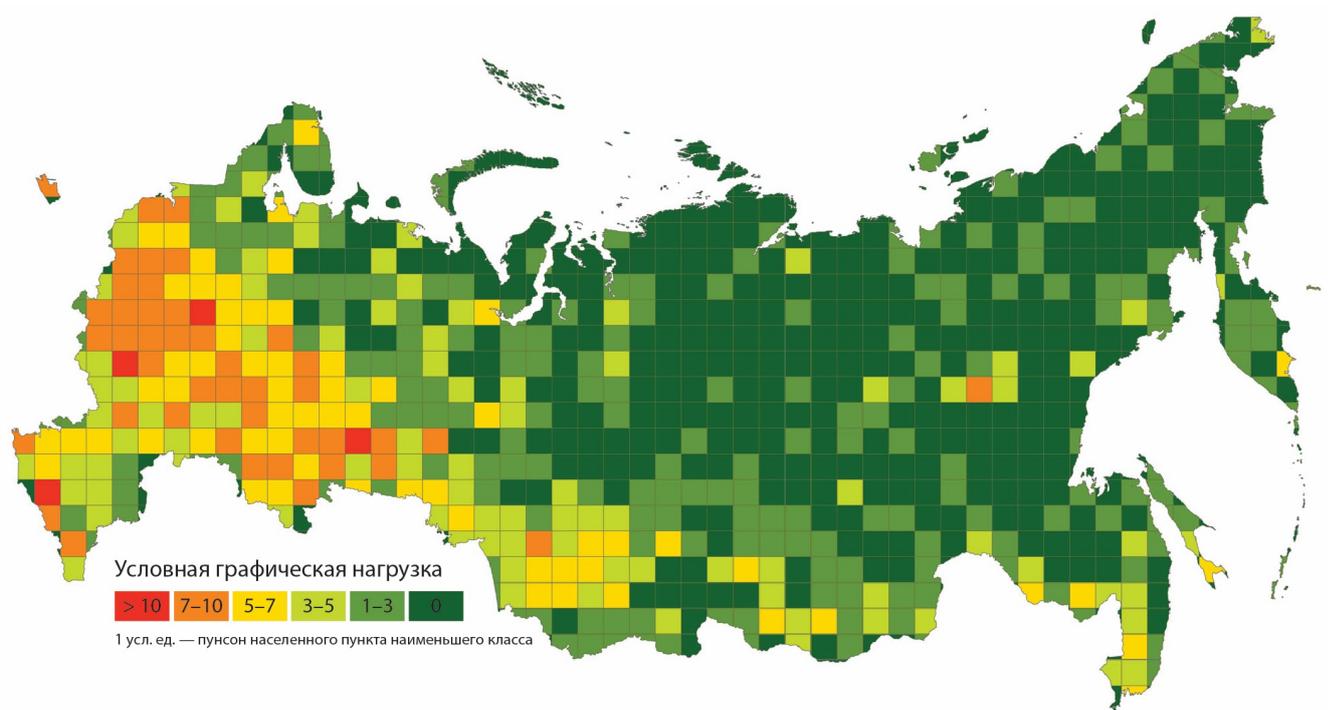
Эту оценку можно назвать географической, поскольку она характеризует изменение густоты по естественным регионам. В топографическом картографировании важную роль играет также техническая оценка, которая связана с графической нагрузкой на единицу площади карты. Для сравнения нагрузки до и после генерализации на исходные и полученные в результате генерализации данные была наложена сетка с шагом 2 см в масштабе карты. Для каждой ячейки было подсчитано количество попадающих в нее населенных пунктов в равновеликой конической проекции. Количество населенных пунктов  $i$ -го класса населения умножалось на  $i$ , что примерно соответствует соотношению видимой площади пунсонов разных классов, которые в дальнейшем использовались для картографирования. Поскольку пунсоны и классификация при переходе к целевому масштабу в нашем случае остаются неизменными, то полученные по квадратам значения можно использовать как условную аппроксимацию графической нагрузки.

Далее на основе классификации были построены карты условной графической нагрузки населенных пунктов на единицу площади. Для классификации был выбран метод равных интервалов, разбивающий диапазон значений атрибута на поддиапазоны равного размера. Для удобства сравнения нескольких карт границы интервалов и их цвета были приняты одинаковыми. Результирующие картограммы представлены на Рисунке 2.6 и Рисунке 2.7.

В исходных данных масштаба 1 : 1 000 000 на 4 см<sup>2</sup> на карте (400 км<sup>2</sup> на местности) в густозаселенных районах в среднем приходится 7 – 9 населенных пунктов. После генерализации в густозаселенных районах на 4 см<sup>2</sup> карты (40 000 км<sup>2</sup> на местности) приходится в среднем 4 – 5 населенных пунктов. Однако условная графическая нагрузка сохраняется в силу того, что сохраненные населенные пункты имеют высокий класс и их пунсоны занимают больше места на карте. В регионах средней густоты размещения населенных пунктов их густота до и после генерализации составляет 3 – 4 на 4 см<sup>2</sup> карты, в малообжитых – менее 2. Относительный порядок графической нагрузки по классам сохранен при генерализации, но контрасты несколько уменьшены в силу необходимости поддерживать присутствие населенных пунктов на территориях с рассеянным заселением. В целом, полученный при генерализации результат по формальным признакам можно считать географически правдоподобным.



**Рисунок 2.6** — Условная графическая нагрузка на карте масштаба 1 : 1 000 000



**Рисунок 2.7** — Условная графическая нагрузка на карте масштаба 1 : 8 000 000

### 2.1.6 Визуальная оценка результатов

Рассмотрим фрагменты составленной карты, иллюстрирующие различные пространственные рисунки населенных пунктов, сохраненных в ходе генерализации разработанным алгоритмом.

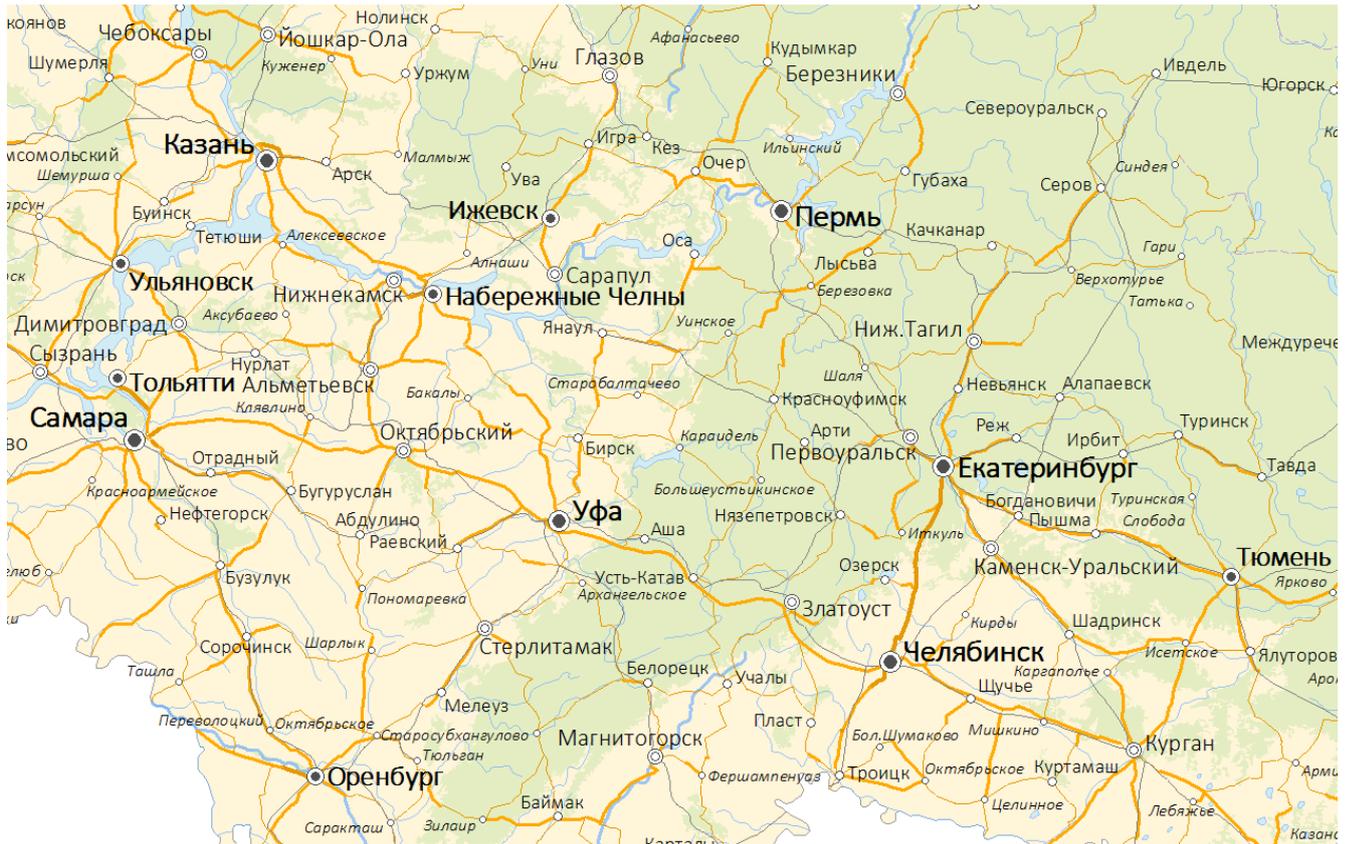
Для крупных рек России, таких как Обь, Енисей, Лена, характерна особая разновидность кластерной группировки — ленточное размещение населенных пунктов. Густота размещения населенных пунктов в этих регионах, особенно на севере, низкая. Несмотря на то, что речная сеть не учитывается в разработанном алгоритме, данный пространственный рисунок при генерализации сохранен (Рисунок 2.8). Для второго региона — Поволжье и Южный Урал характерно регулярное размещение населенных пунктов, которое также сохранено (Рисунок 2.9).



**Рисунок 2.8** — Фрагмент карты масштаба 1 : 8 000 000 (нижнее течение Оби)

Полученные пространственные рисунки размещения населенных пунктов достаточно близки к результатам ручной генерализации [Россия, 2012]. Следует также обратить внимание на то, что в приведенных примерах дорожная сеть отличается высокой степенью согласованности с размещением населенных пунктов. Данный результат был получен также автоматизированно с помощью алгоритма, представленного в статье [Самсонов, Кривошеина, 2015b].

Поскольку отбор населенных пунктов происходил автоматически и управлялся



**Рисунок 2.9** – Фрагмент карты масштаба 1 : 8 000 000 (Южный Урал и Поволжье)

лишь косвенными параметрами (процент удаляемых точек, последовательность стадий отбора и объединения классов), полученная карта содержит также и ряд ошибок. Так, например, в Пермском крае не сохранился крупный город Соликамск, а в ЯНАО отсутствует населенный пункт Лабитнанги, являющийся конечной станцией железной дороги, а в Подмосковье не сохранились вообще все населенные пункты в пределах трассы А108, что связано с чрезвычайно высокой плотностью их размещения. Дальнейшее совершенствование алгоритма и технологического процесса должно позволить более точно контролировать отбор в особо густых районах, а также вводить дополнительные критерии для сохранения населенных пунктов, в частности, их роль в качестве транспортных и промышленных узлов, историко-культурных центров.

### 2.1.7 Итоги

В данном разделе диссертации предложена методика автоматизированного отбора точечных моделей населенных пунктов, которая реализует принцип *адаптивности генерализации к пространственному распределению объектов*. В основе методики лежит алгоритм последовательного исключения точек населенных пунктов в соответствии с их весом, учитывающим три параметра: административный статус, численность населения и относительная локальная плотность, выраженная через площадь полигона Вороного, отнесенного к точке. Новизна методики заключается в действиях, которые позволяют

сохранить географический рисунок размещения точек и включает:

1. Деление территории области (классы) по густоте размещения точек населенных пунктов и выполнение алгоритма отдельно по классам различной густоты.
2. Многостадийную обработку, позволяющую учитывать гранулярность классов и дифференцировать интенсивность отбора в разных классах путем исключения низких классов на заключительных стадиях генерализации.

Стоит отметить, что в качестве альтернативы исключения низких классов, для каждого класса может быть использован разный процент сохраняемых населенных пунктов. Дальнейшее развитие методики видится в автоматизации выбора числа классов густоты и их границ, в автоматизации выбора последовательности и числа процедур исключения точек, объединения и исключения классов при обработке данных, а также совершенствовании отбора в особо плотных областях и введении дополнительных весовых коэффициентов, отражающих промышленную, транспортную и историко-культурную значимость населенных пунктов.

Результаты могут быть использованы при автоматизации составления мелкомасштабных общегеографических карт, а также при мультимасштабном картографировании в рамках процедур подготовки уровней детализации для разных масштабов.

## 2.2 Прореживание линейной модели автодорожной сети с сохранением связности населенных пунктов<sup>2</sup>

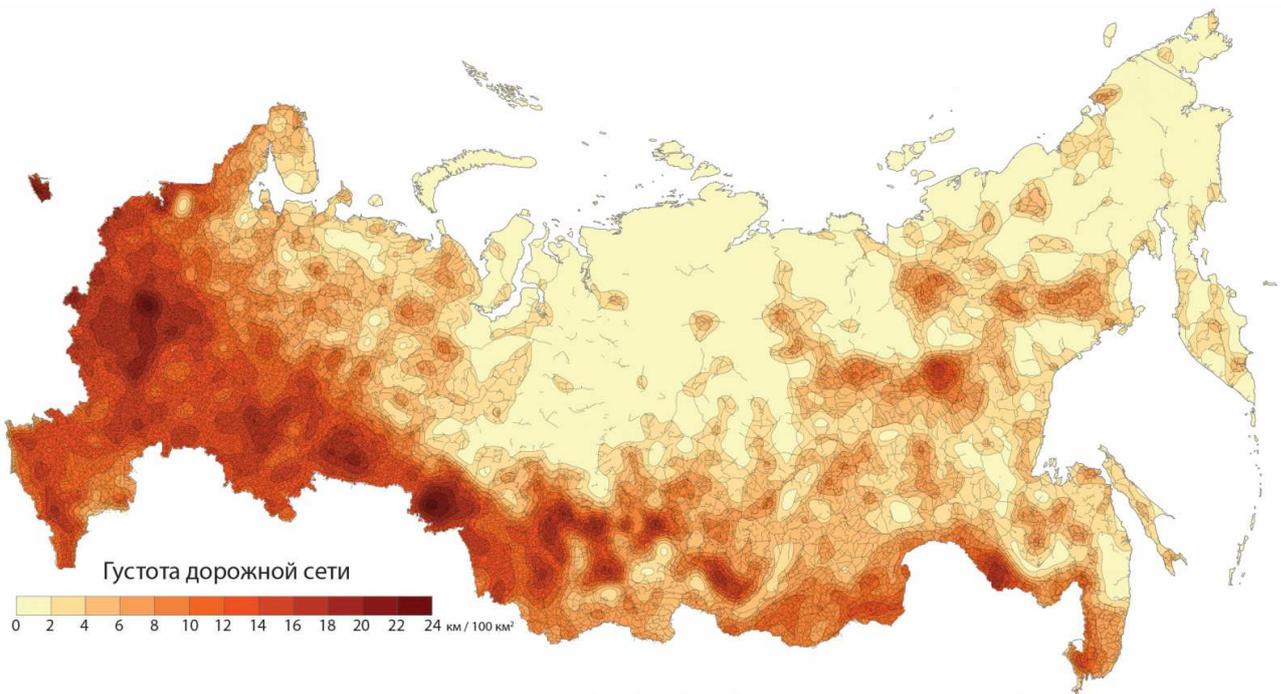
**Горизонтальная связность** возникает между географическими объектами, находящимися в разных местоположениях. Частным случаем такой связности является транспортная связность населенных пунктов, отображение которой является важнейшим условием при мелкомасштабной генерализации карт. Для реализации этого требования был разработан оригинальный метод прореживания (отбора сегментов) линейной модели дорожной сети, который основан на построении кратчайших маршрутов между точками населенных пунктов, предварительно отобранными для целевого масштаба картографирования [Samsonov, Krivosheina, 2012; Самсонов, Кривошеина, 2015b]. Предметом интереса настоящего раздела диссертации является конкретно сеть автомобильных дорог. Здесь и далее, употребляя термин «дорожная сеть», будем подразумевать сеть автомобильных дорог.

---

<sup>2</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Самсонов, Кривошеина, 2015b]

### 2.2.1 Деление территории по густоте дорожной сети

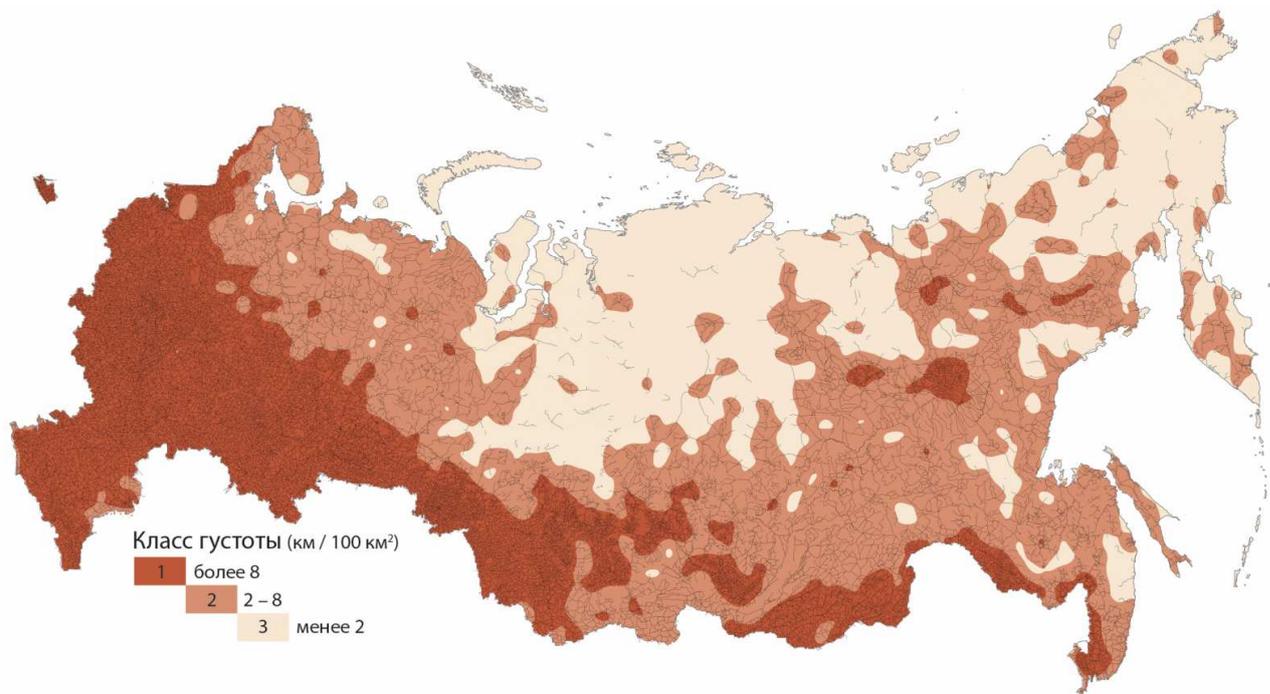
Аналогично рассмотренному в предыдущем параграфе подходу к отбору точек населенных пунктов, для прореживания дорожной сети целесообразно выполнить деление территории на классы плотности, чтобы в дальнейшем можно было управлять интенсивностью отбора независимо в каждом из них. Оптимальным методом оценки густоты линий, представляющих дорожную сеть, является растровый анализ на основе плавающего окна. На исходное множество линий накладывается регулярная растровая сетка с разрешением  $R$ , которое зависит от среднего расстояния до ближайшей линии в исходном наборе данных. Далее в плавающем окне (круглом или квадратном) относительно каждой ячейки растра подсчитывается суммарная длина линий, которая делится на площадь плавающего окна  $S$ . Результирующее значение характеризует локальную густоту дорожной сети в данной точке, а полученное поле распределения дает представление об изменении густоты по территории. Поверхность, построенная таким образом по данным карты масштаба 1 : 1 000 000, представлена на Рисунке 2.10. В целом, можно отметить, что распределение густоты дорожной сети имеет высокую степень корреляции с плотностью размещения населенных пунктов [Самсонов, Кривошеина, 2015а].



**Рисунок 2.10** – Густота дорожной сети на карте масштаба 1 : 1 000 000 (состояние на 2013 г.)

Затем на основе классификации полученной поверхности распределения густоты транспортной сети следует выделить  $N$  основных классов густоты дорог. Экспериментальным путем было установлено, что оптимальное количество классов равняется 3, а их выделение произведено визуально по карте густоты в изолиниях. Отличие от числа классов густоты размещения населенных пунктов, предложенного в предыдущем разде-

ле и равного 4, обусловлено отличием распределения густоты дорожной сети. Дальнейшие исследования должны позволить определить оптимальное число и границы классов на основе численного анализа данных. Выделенные классы густоты дорожной сети приведены на Рисунке 2.11. По результатам классификации каждому населенному пункту присваивается номер класса густоты дорожной сети, которому он принадлежит.



**Рисунок 2.11** – Деление территории России по классам густоты дорожной сети (состояние на 2013 г.)

### 2.2.2 Построение графа дорожной сети

Отбор дорог производится в зависимости от их значимости и густоты размещения. Так, например, сегменты длинных маршрутов являются более значимыми, нежели те, которые используются только для местного передвижения.

В целях расстановки приоритетов движения по магистралям различного уровня следует присвоить им соответствующие коэффициенты сопротивления: чем выше коэффициент, тем труднее проехать по дороге. Коэффициенты присваиваются согласно классификации, в качестве которой для генерализации в масштабном диапазоне 1 : 1 000 000 — 1 : 8 000 000 было использовано следующее разделение:

1. Автомагистрали (автострады), автодороги с усовершенствованным покрытием, автодороги с покрытием.
2. Улучшенные грунтовые дороги, зимники.
3. Проселочные грунтовые дороги и др.

Далее на основе исходных линий строится взвешенный неориентированный граф

дорожной сети, в котором вес каждого ребра графа вычисляется путем умножения его длины на коэффициент сопротивления, по умолчанию равный номеру класса. При необходимости могут быть выбраны более мягкие масштабирующие коэффициенты — например, 1, 1.2 и 1.5 для классов 1, 2 и 3 соответственно, однако такая экспериментальная работа не проводилась.

### 2.2.3 Отбор сегментов дорожной сети

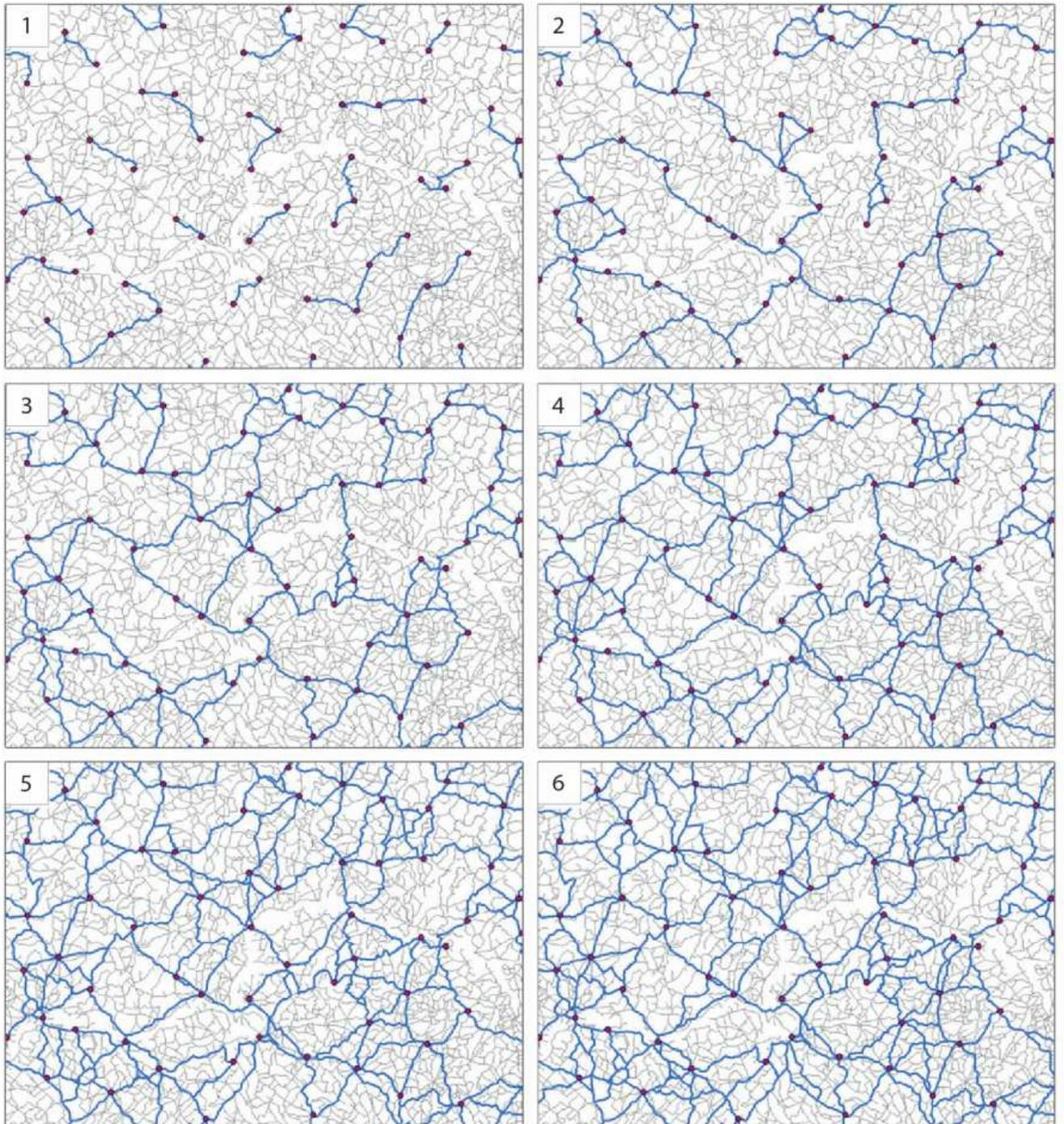
В основе предлагаемого принципа отбора сегментов сети лежит естественное предположение о том, что следует сохранить только те дороги, которые необходимы для передвижения между населенными пунктами по кратчайшему пути (с учетом установленной иерархии дорог). При этом множество населенных пунктов должно быть уже генерализовано для результирующего масштаба.

От каждого населенного пункта строятся маршруты до ближайших  $N$  населенных пунктов. Данная задача относится к типу *Closest Facility* (ближайший пункт обслуживания). При этом число  $N$  зависит от класса густоты дорожной сети, которому принадлежит анализируемый населенный пункт. Сами линии не сегментируются на классы и участвуют в анализе в полном наборе (то есть допускается, что маршрут может быть построен из одного класса плотности в другой). Этот подход обеспечивает стыковку на границе классов. Основной вопрос заключается в выборе числа  $N$ . Для его разрешения необходимо проанализировать как меняется количество отобранных при генерализации дорог при модификации этого числа.

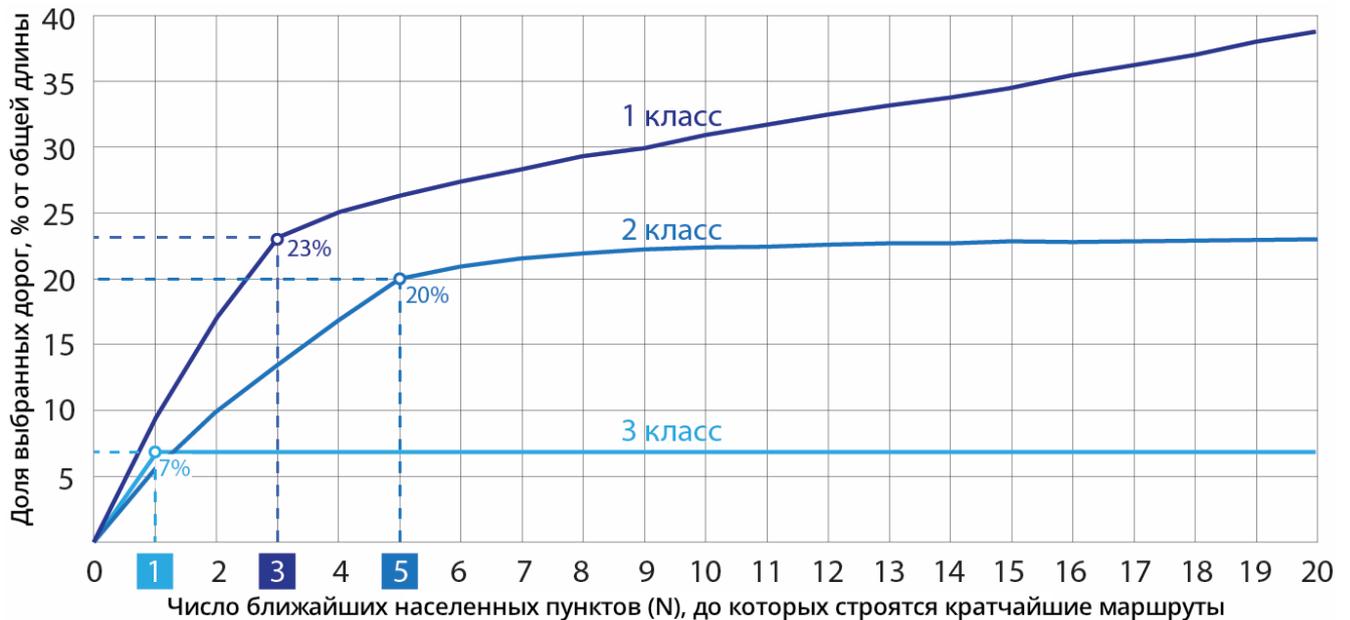
На Рисунке 2.12 показана динамика отбора дорог в регионе 1 класса густоты при изменении количества населенных пунктов, к которым строятся кратчайшие маршруты. При анализе данных иллюстраций видно, что сеть приобретает более-менее связный вид при  $N = 3$ , а далее начинается ее сгущение сегментами, позволяющими сократить путь до дальних населенных пунктов.

Однако анализ двумерного распределения достаточно сложен и неоднозначен. Для принятия решения об оптимальном числе требуется снизить размерность. Для этого можно исследовать функцию зависимости между  $N$  и процентной долей длины сегментов дорожной сети, задействованных в маршрутах (от суммарной длины). Примеры графиков для трех выделенных классов представлены на Рисунке 2.13.

В данном случае для областей с одинаковой плотностью графики имеют схожую точку перегиба, которая может служить оптимальным параметром для генерализации. Эта точка характеризуется спадом роста процента сохраняемых дорог и максимальной выпуклостью графика, что соответствует резкому уменьшению производной, которая легко вычисляется через значения на графике. Таким образом, число  $N$  может быть найдено автоматизированным путем.



**Рисунок 2.12** – Итеративный отбор дорог путем построения кратчайших маршрутов до ближайших населенных пунктов. Номер в углу – количество маршрутов  $N$



**Рисунок 2.13** – Зависимость доли длины отобранных дорог от количества маршрутов

**Таблица 2.2** – Оценка степени отбора

Класс плотности	Исходная суммарная длина дорожной сети, км	Суммарная длина дорожной сети после генерализации, км	Доля сохраненных в ходе генерализации дорог, %
1	500895	101062	20.2
2	389999	85219	21.9
3	23708	5858	24.7
Всего	914601	192140	21

## 2.2.4 Численная и графическая оценка результатов генерализации

Реализация предложенной методики была выполнена средствами модуля сетевого анализа *Network Analyst* для ГИС-пакета *ArcGIS for Desktop 10.1*. После выполнения отбора было произведено геометрическое упрощение линий с помощью алгоритма *Bend Simplify* [Wang, Müller, 1998] с установленным минимальным размером изгиба равным 10 км.

Для проведения анализа были рассчитаны данные по суммарной протяженности дорог разного класса густоты (Таблица 2.2):

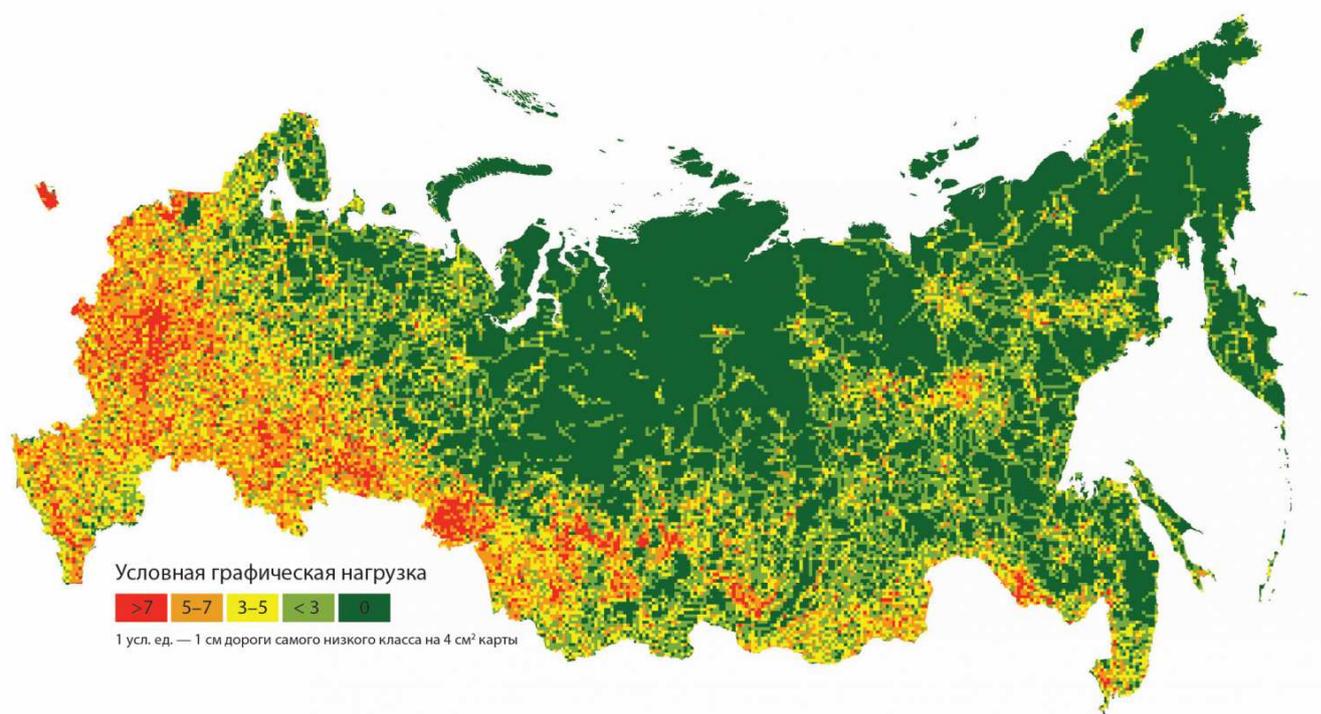
Далее для оценки связи между количеством объектов до и после генерализации аналогично методике отбора населенных пунктов был применен обобщенный закон Топфера [Töpfer, Pillewizer, 1966; Li, 2006] (Уравнение 2.5). Заметим, что при оценке генерализации линий правильней учитывать не количество объектов, а именно их суммарную длину [Brewer et al., 2013], которая и была использована вместо значений  $N_G$  и  $N_S$ .

Имея исходный и результирующий масштабы, а также суммарную протяженность дорожной сети до и после генерализации по классам плотности, был определен параметр  $x$  для каждого класса. На основе него можно оценить, насколько изменилась густота дорог по выделенным классам плотности. В первых двух классах плотности этот

параметр приблизительно равен 1.7, в третьем — 1.5. Таким образом, можно сделать вывод, что в целом густота дорожной сети увеличилась, причем в третьем классе несколько больше чем в первых двух. Это объясняется стремлением показать в самых малоосвоенных регионах связи между всеми населенными пунктами.

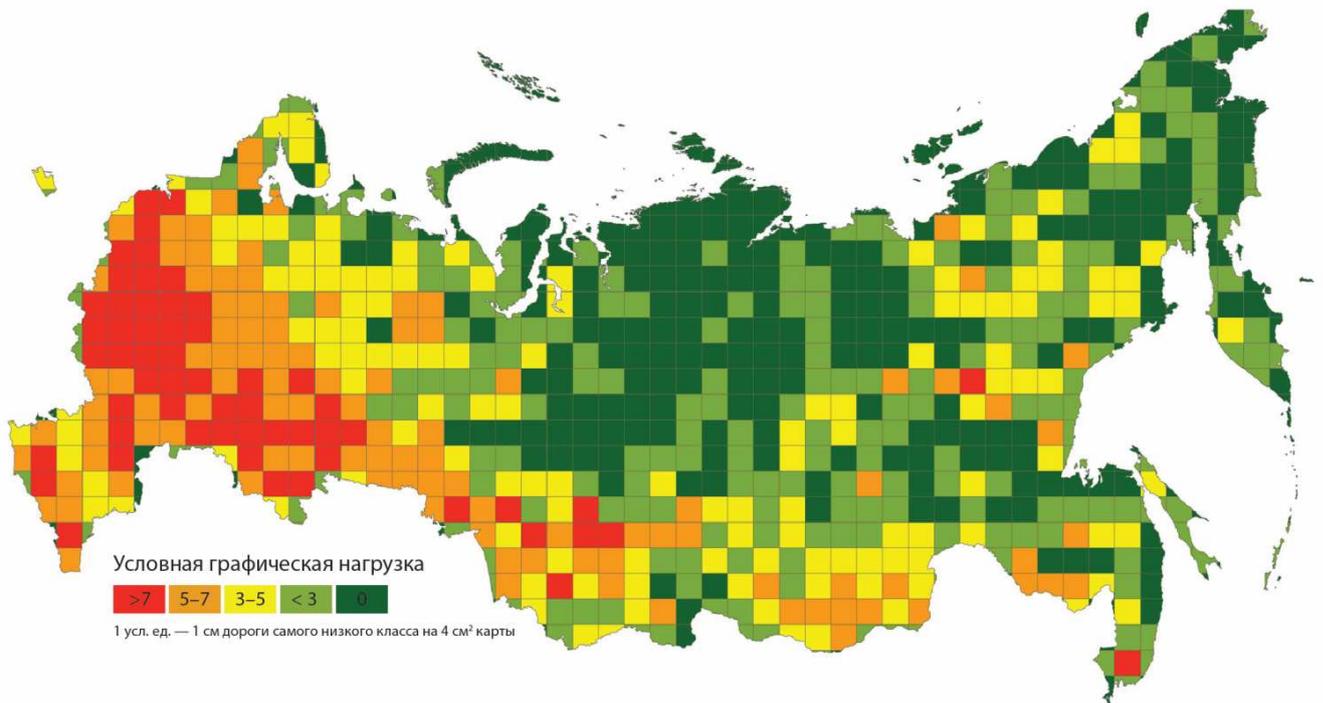
Для сравнения графической нагрузки объектов на единицу площади до и после генерализации была построена сетка с шагом 2 см в масштабе карты (как и для населенных пунктов). Для каждой ячейки была подсчитана суммарная протяженность попадающих в нее дорог. Длина каждого  $i$ -го иерархического класса дорог была умножена на весовой коэффициент равный  $i$ , чтобы учесть площадь, занимаемую знаком на карте. Далее на основе классификации были построены карты условной графической нагрузки дорожной сети на единицу площади карты (предполагается, что классификация дорог и сами условные знаки остались неизменными).

В исходных данных масштаба 1 : 1 000 000 на 4 см<sup>2</sup> на карте (400 км<sup>2</sup> на местности) в густозаселенных районах в среднем приходится 7 — 15 см автомобильных дорог (Рисунок 2.14), а в малоосвоенных сохраняются практически все дороги, соединяющие населенные пункты.



**Рисунок 2.14** — Условная графическая нагрузка дорог на карте масштаба 1 : 1 000 000

При сравнении двух картограмм графической нагрузки, можно сделать вывод, что в целом рисунок ее пространственной дифференциации сохраняется. Однако максимальные значения густоты дорожной сети после генерализации выше, чем в исходных данных.



**Рисунок 2.15** — Условная графическая нагрузка дорог на карте масштаба 1 : 8 000 000

### 2.2.5 Визуальная оценка результатов генерализации

Для визуальной оценки качества генерализации наиболее интересно рассмотреть регионы с кластерным и равномерным размещением населенных пунктов, которое в свою очередь определяет и морфологию транспортной сети.

На Рисунке 2.16 представлены фрагменты карты масштаба 1 : 8 000 000 на территории юга Восточной Сибири, а также Центральной части Европейской России, включая столичный регион. Анализ данных фрагментов позволяет утверждать, что алгоритм позволяет сохранить связность населенных пунктов, используемых в качестве начального условия, вне зависимости от типа размещения населенных пунктов.

При этом, благодаря сильной связи размещения населенных пунктов и транспортных сетей в мелких масштабах, удастся также сохранить и характерные пространственные рисунки, образы территории. Так, например, на Рисунке 2.17 хорошо прослеживается радиальный рисунок дорожной сети, расходящейся от Москвы в разных направлениях. Сравнение результатов генерализации с рукописной картой масштаба 1 : 8 000 000 [Россия, 2012] показывает, что комбинация предложенных алгоритмов дает схожий с ручным отбором результат.

Следует обратить внимание также на то, что в результаты отбора можно легко включить те сегменты дорожной сети, которые ведут не к населенным пунктам, а к промышленным (месторождения), рекреационным объектам и любым точкам интереса. Для этого достаточно эти точки добавить к населенным пунктам. При необходимости такие осо-

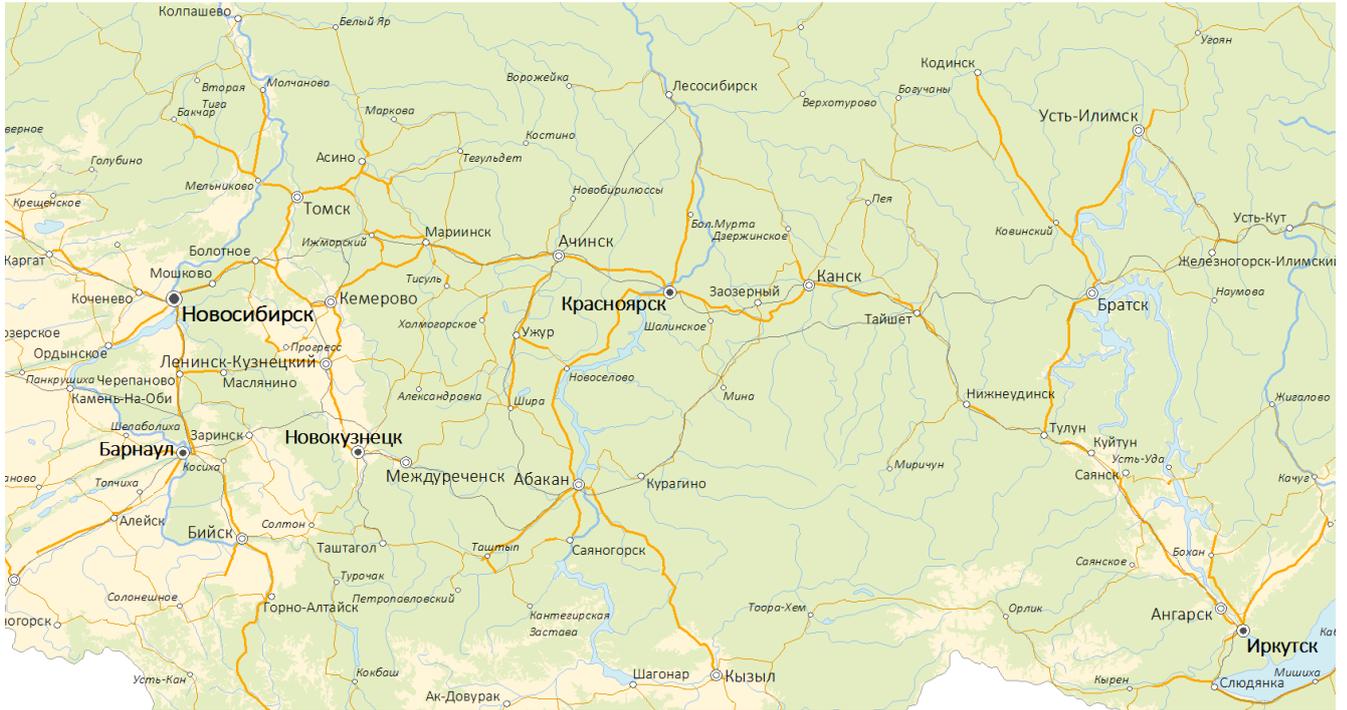


Рисунок 2.16 – Фрагмент карты масштаба 1 : 8 000 000 (юг Восточной Сибири)

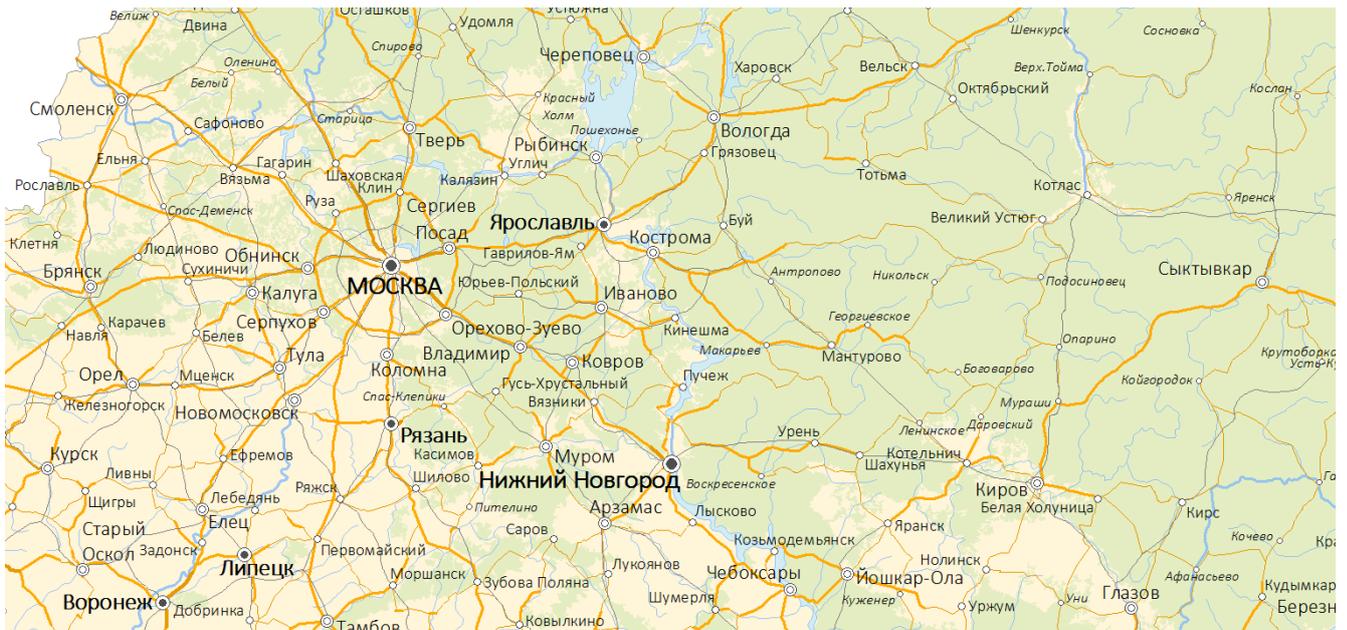
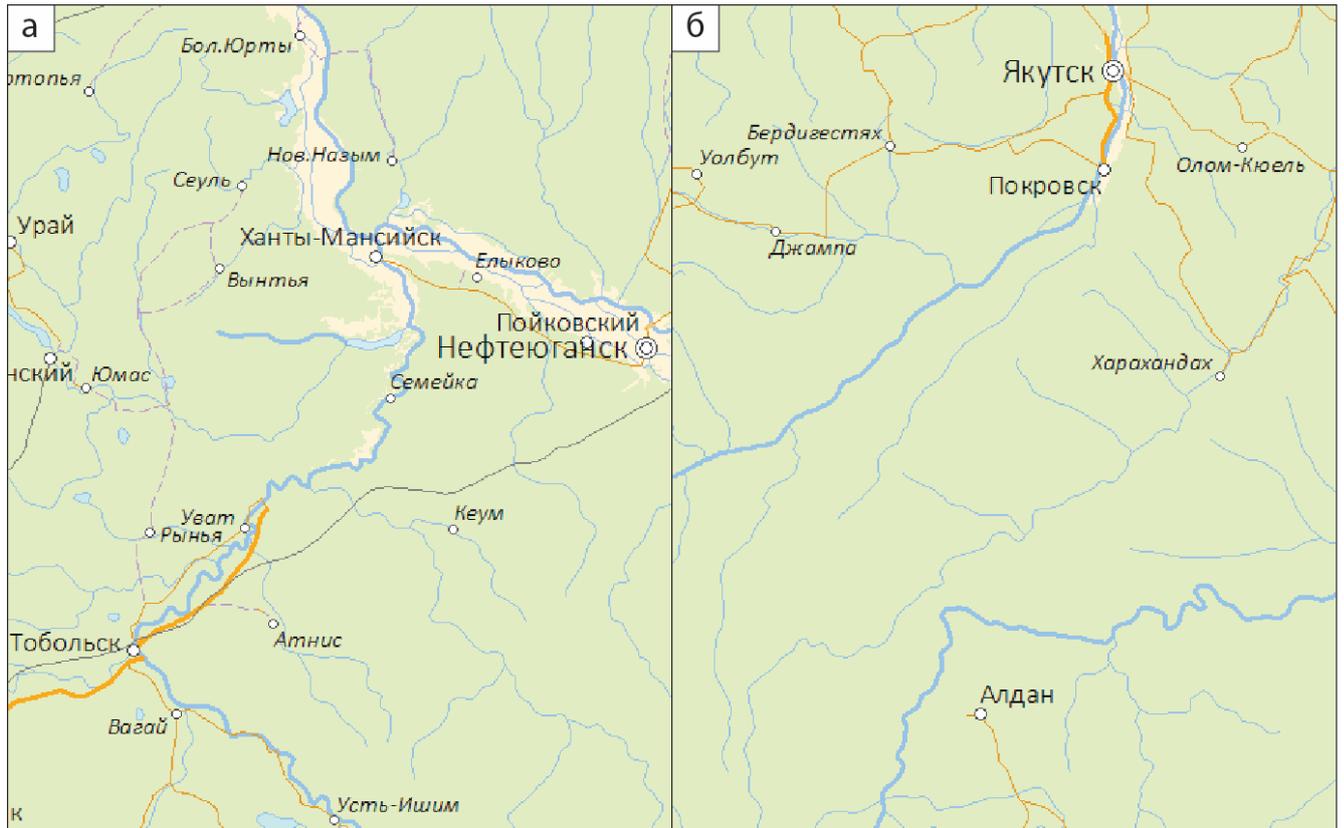


Рисунок 2.17 – Фрагмент карты масштаба 1 : 8 000 000 (центр Европейской части России)

бые точки могут быть установлены оператором вручную.



**Рисунок 2.18** – Неверно исключенные при отборе автодороги. а) Тобольск – Нефтеюганск; б) Алдан – Якутск

Данные ошибки частично обусловлены географической асимметрией построения маршрутов, которая обусловлена конфигурацией населенных пунктов. Так, например, если к северу от населенного пункта сосредоточено большое число соседей, то будут отобраны только магистрали ведущие на север. Для устранения этого недостатка в будущем предполагается интегрировать подход, предложенный в работе [Schmid, Janatzek, 2013] по мотивам публикации автора [Samsonov, Krivosheina, 2012] и основанный на вычислении маршрутов до соседей по диаграмме Вороного.

## 2.2.6 Итоги

В главе предложен новый подход к автоматизированному прореживанию (отбору сегментов) линейной модели дорожной сети для мелких масштабов картографирования, реализующий принцип *адаптивности к горизонтальной связности* населенных пунктов. Основной параметр отбора — число маршрутов — дифференцируется по классам различной густоты дорожной сети на основе численного анализа зависимости процентной доли длины сохраняемых дорог от количества маршрутов.

В результате генерализации разработанным алгоритмом сохраняется соотношение графической нагрузки дорожной сети между разными классами, при этом основной от-

бор производится в самых густозаселенных регионах и сохраняется достаточное количество дорог в малоосвоенных. В силу выраженной связи морфологической структуры транспортной сети и конфигурации населенных пунктов с помощью предложенного алгоритма удастся также сохранить характерные рисунки дорог в мелких масштабах. На основе визуального анализа полученной карты масштаба 1 : 8 000 000 выявлены ошибки в работе алгоритма и предложено направление их разрешения с использованием анализа соседства по диаграмме Вороного.

Дальнейшее совершенствование сопряженной методики генерализации населенных пунктов и дорожной сети видится в усилении обратной связи, при которой также и отбор населенных пунктов будет зависеть от их транспортной значимости, что может быть введено на основе сетевого и топологического анализа.

## 2.3 Геометрическое упрощение множества линейных объектов с сохранением топологии<sup>3</sup>

*Вертикальная связность*, в отличие от горизонтальной, возникает в случае, когда пространственные объекты или их части совпадают по местоположению. Так, например, административные границы пересекаются объектами гидрографии и путями сообщений (автомобильными и железными дорогами). В настоящем разделе диссертации представлен новый алгоритм геометрического упрощения объектов, позволяющий с использованием графов генерализовать *множество* линейных объектов, в том числе из разных слоев, с сохранением топологии [Самсонов *et al.*, 2014]. Теория графов показала себя эффективной при генерализации картографических данных [Mackness, Beard, 1993]. Предложенный алгоритм также является неконсервативным, что повышает его возможности с точки зрения согласования линий. Еще одним важным аспектом работы является подход к оценке результатов, который основан на совместном численном описании критерия близости (Модифицированное расстояние Хаусдорфа) и морфологичности (коэффициент относительной извилистости). Ниже рассмотрены основные операции, применяемые в процессе упрощения линий, качественной и количественной оценке результатов.

### 2.3.1 Построение графа и операция стягивания ребра

Линейные объекты представляются в виде графа. Для этого узлы линий преобразуются в вершины графа, а соединяющие их отрезки — в ребра графа. При этом каждая точка пересечения линий представляется в виде одной вершины результирующего гра-

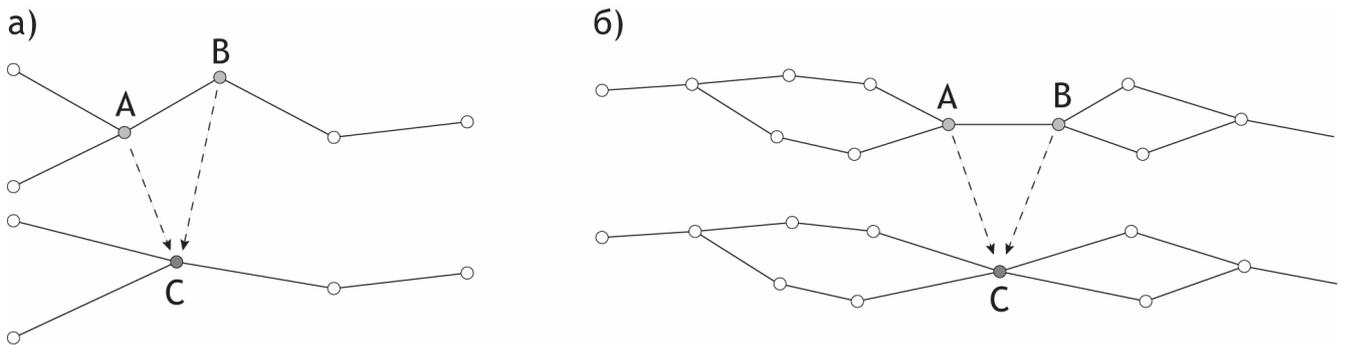
<sup>3</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Самсонов *et al.*, 2014]

фа (вне зависимости была ли эта точка представлена в явном виде в исходном наборе линий).

Основной операцией, применяемой для упрощения объектов, является операция стягивания ребра в точку, в результате которой ребро удаляется, а вершины на его концах заменяются на одну новую вершину. На примере, представленном на Рисунке 2.19 таким ребром является  $AB$ , а результирующей вершиной —  $C$ . При этом вершина  $C$  заменяет вершины  $A$  и  $B$  для всех ребер графа, инцидентных вершинам  $A$  и  $B$ . При стягивании возможны случаи, когда топология системы остается прежней (Рисунок 2.19а) и когда она изменяется (Рисунок 2.19б). Во втором случае стягивание ребра приводит к тому, что два цикла в графе (типичных для разветвленного русла реки) получили одну общую точку. Индикатором топологической ошибки является следующее правило:

$$\text{deg}(C) \neq \text{deg}(A) \wedge \text{deg}(C) \neq \text{deg}(B), \quad (2.6)$$

где  $\text{deg}$  — степень вершины (число инцидентных ей ребер), а  $\wedge$  — логическое И. Таким образом, ошибка возникает, когда степень изменилась для *обеих* вершин.



**Рисунок 2.19** — Стягивание ребра  $AB$  в вершину  $C$ : а) с сохранением топологии; б) с изменением топологии

Положение новой вершины графа при стягивании ребра может быть выбрано различными способами. В основу метода, применяемого в настоящей работе, взят алгоритм, изложенный в работе [Dey et al., 1999], который разработан для упрощения трехмерных поверхностей и применяется в компьютерной графике. Каждой вершине графа ставится в соответствие матрица коэффициентов, которая строится по следующему правилу.

Пусть число ребер, инцидентных некоторой вершине  $A$ , равно  $N$ . Для каждого  $i$ -го ребра, инцидентного вершине, рассматривается уравнение прямой, на которой лежит ребро:  $a_i x + b_i y + c_i = 0, i = 1, \dots, N$ . Предполагается, что уравнения нормализованы, т. е.  $a_i^2 + b_i^2 = 1$  и, таким образом, расстояние от точки  $X$  с координатами  $(x, y)$  до  $i$ -й прямой равно  $|a_i x + b_i y + c_i|$ . Если вершина  $A$  в результате стягивания переходит в точку  $X$ , то можно вычислить расстояния от этой точки до прямых, на которых лежат исходные ребра, и использовать их при оценке цены стягивания. Для этого рассмотрим

квадратичную функцию вида:

$$F(x, y) = \sum_{i=1}^N (a_i x + b_i y + c_i)^2 = q_{11}x^2 + 2q_{12}xy + q_{22}y^2 + 2q_1x + 2q_2y + q_0, \quad (2.7)$$

где суммирование производится по всем ребрам, инцидентным данной вершине. Эту квадратичную функцию будем называть ошибкой вершины, а матрицу этой квадратичной функции

$$Q_A = \begin{pmatrix} q_{11} & q_{12} & q_1 \\ q_{12} & q_{22} & q_2 \\ q_1 & q_2 & q_0 \end{pmatrix} \quad (2.8)$$

– матрицей коэффициентов для вершины  $A$ .

Ребру  $AB$  поставим в соответствие матрицу  $Q_{AB} = Q_A + Q_B$ , где  $Q_A$  и  $Q_B$  – матрицы коэффициентов для вершин  $A$  и  $B$  соответственно. При стягивании ребра  $AB$  положение новой точки  $X = (x, y, 1)^T$  определяется как решение задачи минимизации квадратичной функции с матрицей  $Q_{AB}$ :

$$XQ_{AB}X^T \rightarrow \min. \quad (2.9)$$

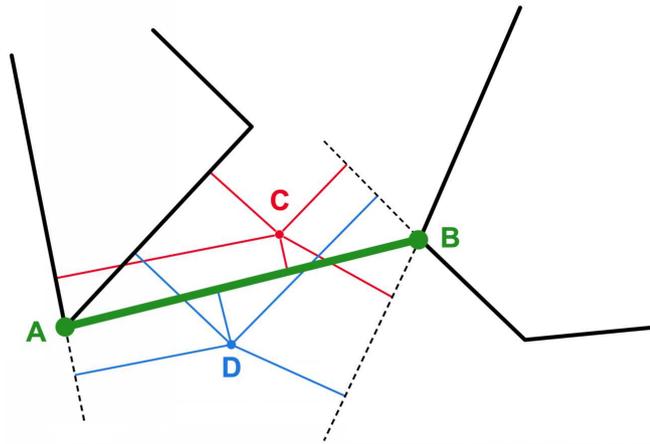
минимальную величину будем считать мерой погрешности, вносимой в систему при стягивании ребра  $AB$  и введения вместо него точки  $X$ , и называть ценой стягивания ребра  $AB$ .

Решение задачи минимизации квадратичной формы сводится к отысканию решения системы двух линейных уравнений с двумя неизвестными. В случае, если решение системы не определено однозначно, в качестве новой вершины  $X$  выбирается та из вершин  $A$  или  $B$ , на которой выражение  $XQ_{AB}X^T$  принимает меньшее значение, и ценой стягивания ребра  $AB$  считаем соответствующее значение.

По сути, положение новой вершины графа при стягивании ребра определяется путем минимизации суммы квадратов расстояний от нее до прямых вида  $ax + by + c$ , проходящих через ребра, инцидентные двум стягиваемым вершинам. Этот принцип проиллюстрирован на Рисунке 2.20, где сравнивается положение двух альтернативных вершин  $C$  и  $D$ , которые могут быть результатом стягивания ребра  $AB$ . В данном случае будет выбрана вершина  $C$ , поскольку сумма квадратов расстояний от нее (красные отрезки) меньше, чем сумма квадратов расстояний от вершины  $D$  (голубые отрезки).

### 2.3.2 Принцип сохранения топологии

Математическим аналогом требования сохранения относительного расположения объектов предлагается считать топологическую эквивалентность исходного и получен-



**Рисунок 2.20** – Альтернативные положения результирующей вершины ( $C$  и  $D$ ) при стягивании ребра  $AB$

ного графов. Графы называются топологически эквивалентными, если существует гомеоморфизм (непрерывное обратимое отображение плоскости в себя, обратное к которому тоже непрерывно), при котором один граф переходит в другой. Преобразовывая граф, будем стягивать только те ребра, стягивание которых приводит к топологически эквивалентному графу. Общие условия сохранения топологической эквивалентности при стягивании ребер в симплициальных комплексах могут быть достаточно сложны [Dey et al., 1999]. Для случая планарного графа имеем очень простой вид: стягивание ребра  $AB$  в точку приводит к топологически эквивалентному графу тогда и только тогда когда по меньшей мере одна из вершин  $A$ ,  $B$  принадлежит в точности двум ребрам и вершины  $A$  и  $B$  не имеют в графе общих соседей. Выполнение этого условия так же обеспечивается условием Уравнения 2.6.

### 2.3.3 Алгоритм геометрического упрощения

Алгоритм геометрического упрощения состоит из следующих этапов:

1. *Объединение графов.* Графы, построенные для отдельных слоев, объединяются в один граф, при этом для ребер и вершин сохраняется информация об их принадлежности соответствующим объектам (допускается, что элемент графа принадлежит объектам из разных слоев). В точках пересечения ребер разных слоев добавляются новые вершины графа.
2. *Добавление весов в вершинах изгибов (опционально).* Для сохранения наиболее важных элементов ломаных производится анализ изгибов. Если точка является вершиной изгиба, соответствующая ей матрица коэффициентов умножается на константу, характеризующую значимость данного изгиба. Это приводит к увеличению цены стягивания ребер, инцидентных данной вершине.
3. *Сортировка.* Для всех вершин графа вычисляются матрицы коэффициентов. Для

каждого ребра вычисляется цена стягивания. Ребра включаются в список и упорядочиваются по возрастанию цены стягивания.

4. *Упрощение.* Производится упрощение структуры путем последовательного стягивания ребер в порядке возрастания цены. Перед применением операции стягивания, производится проверка условия сохранения топологической эквивалентности системы. Процесс продолжается, пока количество вершин графа не снизится до заданного числа процентов от исходного количества вершин. Этот процент может быть связан с масштабом, в котором предполагается представлять данные.
5. *Сглаживание.* Кривые сглаживаются В-сплайнами 2-й степени для устранения мелких неровностей.
6. *Разделение на слои,* содержащие генерализованные версии данных исходных слоев объекта.

При стягивании ребра, требуется пересчитать цены стягивания ребер, инцидентных новой вершине и обновить список ребер. В силу локальности этот пересчет не увеличивает трудоемкость алгоритма.

### 2.3.4 Критерии оценки качества генерализации

Для оценки близости исходной и генерализованной линии, т. е. геометрической точности результата, было адаптировано применительно к графам *Модифицированное расстояние Хаусдорфа* [Dubuisson, Jain, 1994]. Вторым показателем — *коэффициент относительной извилистости* — является морфологическим критерием, отражающим географическое правдоподобие результата.

Модифицированное расстояние Хаусдорфа (МНД), введенное в [Dubuisson, Jain, 1994], определяется следующим образом. Пусть даны два множества точек  $\mathcal{A} = \{a_1, \dots, a_{N_a}\}$  и  $\mathcal{B} = \{b_1, \dots, b_{N_b}\}$ . Тогда среднее расстояние между  $\mathcal{A}$  и  $\mathcal{B}$  может быть вычислено как  $\bar{d}(\mathcal{A}, \mathcal{B}) = \frac{1}{N_a} \sum_{a \in \mathcal{A}} d(a, \mathcal{B})$ , где  $d(a, \mathcal{B}) = \min_{b \in \mathcal{B}} \|a - b\|$ . Аналогично, обратное расстояние между  $\mathcal{B}$  и  $\mathcal{A}$  вычисляется как  $\bar{d}(\mathcal{B}, \mathcal{A}) = \frac{1}{N_b} \sum_{b \in \mathcal{B}} d(b, \mathcal{A})$ , где  $d(b, \mathcal{A}) = \min_{a \in \mathcal{A}} \|b - a\|$ .

Имея прямое и обратное расстояния между  $\mathcal{A}$  и  $\mathcal{B}$ , модифицированное расстояние Хаусдорфа МНД вычисляется как:

$$d_M(\mathcal{A}, \mathcal{B}) = \max \left[ \bar{d}(\mathcal{A}, \mathcal{B}), \bar{d}(\mathcal{B}, \mathcal{A}) \right]. \quad (2.10)$$

По сути, МНД есть максимальное из средних расстояний от  $\mathcal{A}$  к  $\mathcal{B}$  и от  $\mathcal{B}$  к  $\mathcal{A}$ . Чем меньше значение МНД, тем интегрально ближе  $\mathcal{A}$  и  $\mathcal{B}$  друг к другу. Введение этой метрики вдохновлено классическим расстоянием Хаусдорфа (в котором  $d(\mathcal{A}, \mathcal{B}) = \max_{a \in \mathcal{A}} d(a, \mathcal{B})$ ), которое, как видно из определения, чувствительно к

точкам-выбросам, поскольку использует максимальное расстояние вместо среднего<sup>4</sup>.

В нашем случае объектами сопоставления являются не множества точек, а графы, соответствующие тематическим слоям. Пусть граф  $G = (V, E)$  представляет совокупность слоев до геометрического упрощения, а  $G' = (V', E')$  — после. Обозначим индексом  $i$  некий подграф (подмножество вершин и ребер), который соответствует одному из слоев, участвующих в генерализации. В этом случае имеем для каждого слоя переход  $G_i \rightarrow G'_i$ . Тогда модифицированным расстоянием Хаусдорфа между графами будем считать следующую величину:

$$d_M(G_i, G'_i) = \max \left[ \bar{d}(V, E'), \bar{d}(V', E) \right]. \quad (2.11)$$

В качестве интегральной меры соответствия используется среднее из значений  $d_M$ , вычисленных по всем  $i$ . Например, если в генерализации участвует одновременно 2 слоя, то будет вычислено 2 значения  $d_M$  и взято их среднее арифметическое.

Коэффициент извилистости – мера извилистости объекта, вычисляемая как отношение длины линии к длине отрезка, соединяющего ее концы [Берлянт, 1988]. Пусть линия  $L$  состоит из  $n$  узлов, соединенных ребрами. Тогда коэффициент ее извилистости

$$K = \frac{1}{l_{1,n}} \sum_{i=1}^{n-1} l_{i,i+1}, \quad (2.12)$$

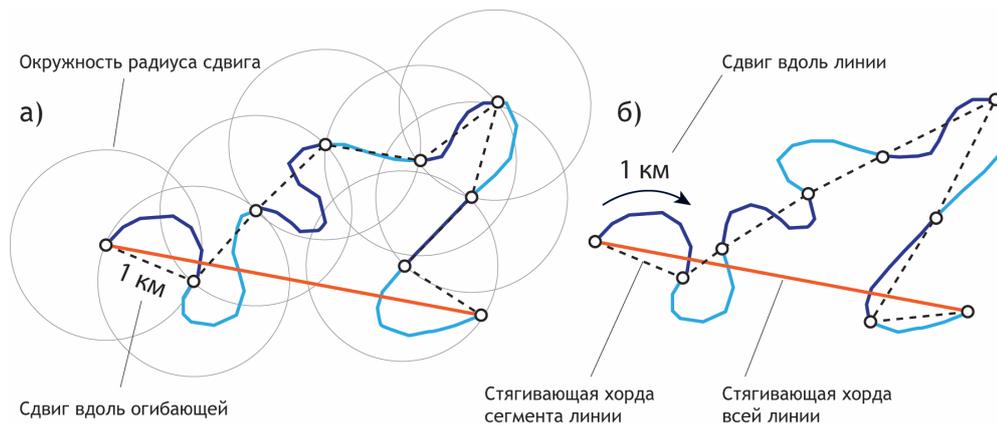
где  $l_{i,j}$  – Евклидово расстояние между  $i$ -м и  $j$ -м узлом линии.

Коэффициент извилистости зависит от конфигурации сглаживающей, описывающей общую траекторию линии. В предельном случае, когда имеют дело с замкнутой фигурой, величина  $K$  не определена, так как длина стягивающей хорды равна нулю. Чтобы избежать этого, можно вычислить отношение длины линии не к стягивающей, а к сглаживающей, т. е. вычислить относительную извилистость. Для этого можно сегментировать линию одним из двух способов, представленных на Рисунке 2.21: либо сдвигаться на одинаковое значение стягивающей хорды (Рисунок 2.21а), либо сдвигаться вдоль самой линии с равным шагом (Рисунок 2.21б). Второй подход вычислительно более эффективен, так как он не требует поиска пересечения с окружностью, а также поиска следующей по ходу линии точки (пересечений может быть несколько в случае извилистой линии). Коэффициент относительной извилистости, вычисляемый при таком подходе, имеет следующий вид:

$$\tilde{K} = \frac{1}{m} \sum_{j=1}^m K_j, \quad (2.13)$$

где  $K_j$  – коэффициент извилистости для  $k$ -го сегмента линии;  $m$  – число сегментов.

<sup>4</sup>Более подробно различные варианты расстояния Хаусдорфа рассмотрены далее в Разделе 3.6.1.3.



**Рисунок 2.21** — Подходы к вычислению коэффициента относительной извилистости: а) с фиксированной длиной стягивающей хорды; б) с фиксированной длиной сегмента линии

### 2.3.5 Апробация алгоритма

Разработанный алгоритм был реализован в виде компьютерной программы на языке C#. Для сравнения результатов использовались алгоритмы Ванга-Мюллера [Wang, Müller, 1998] и Ли-Оупеншоу [Li, Openshaw, 1992] как наиболее приближенные к географической постановке задачи. Исходные данные для экспериментальной работы (цифровая топографическая карта России масштаба 1:1 000 000) предоставлены компанией Дата+. Для сопоставления результатов использовался следующий подход. Сначала производилась генерализация разработанным методом с заданным процентом оставляемых точек. Затем параметры других алгоритмов подбирались таким образом, чтобы число точек было примерно тем же самым (с погрешностью 5%). Оценка при равном числе точек позволяет судить о том, насколько эффективно они используются алгоритмом для представления исходной формы линии. После генерализации оценивалась извилистость при разном шаге квантования, а также обобщенное Хаусдорфово расстояние.

Оценка извилистости (Таблица 2.3) произведена на примере рек, поскольку для гидрографии этот показатель наиболее важен. Сравнить результаты можно на Рисунке 2.22.

Как видно из Таблицы 2.3, значение коэффициента относительной извилистости зависит от шага квантования. Однако важно не абсолютное значение коэффициента, а его сохранение при уменьшении масштаба. При сравнении значений для исходных и генерализованных линий использовать одинаковый шаг квантования некорректно, так как длина линии на карте уменьшается. Для масштабной компенсации шага квантования можно использовать одну из существующим моделей изменения числа объектов при уменьшении масштаба, известную как закон Топфера, согласно которой число объектов снижается не линейно, а по закону квадратного корня [Töpfer, Pillewizer, 1966]. При распространении этой модели на линейные объекты в работе [Brewer et al., 2013] вместо числа объектов использована суммарная длина линий. Данная формализация может быть также применена к числу изгибов на единицу длины линии или к шагу квантования, с

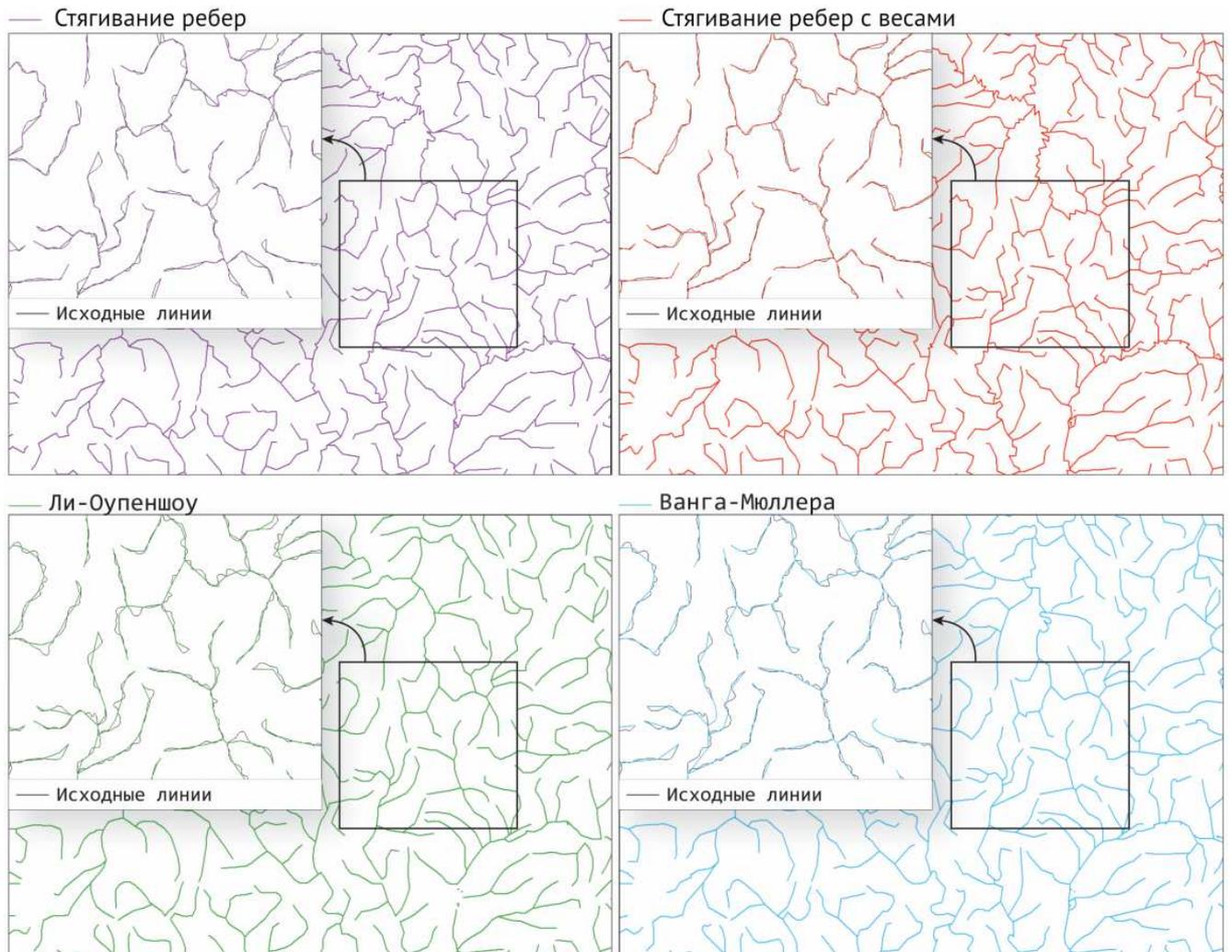
**Таблица 2.3** – Значения коэффициента относительной извилистости для слоя рек при обработке линий разными алгоритмами генерализации (оставлено 20% точек)

Шаг квантования, м	Исходный объект	Метод квадра- тической ошибки	Метод квadratic- ошибки с весами	Метод Ли-Оупеншоу	Метод Ванга- Мюллера
5000	1.125	1.041	1.049	1.022	1.025
7000	1.152	1.071	1.081	1.033	1.046
9000	1.169	1.096	1.105	1.046	1.061
11000	1.185	1.111	1.123	1.057	1.074
13000	1.198	1.128	1.14	1.067	1.087
15000	1.211	1.141	1.152	1.077	1.097
17000	1.22	1.153	1.164	1.084	1.106
19000	1.231	1.162	1.172	1.09	1.113
21000	1.238	1.17	1.183	1.096	1.121
23000	1.244	1.177	1.19	1.102	1.126

помощью которого подсчитывается извилистость. Таким образом, если шаг квантования в исходном масштабе составил  $D$ , то при уменьшении масштаба в 5 раз следует использовать шаг, равный  $D\sqrt{5}$ . Предложенный алгоритм дает наиболее близкие к исходным значения извилистости, использование повышенного приоритета вершин изгибов — наилучшие результаты.

При визуальной оценке результатов, изображенных на Рисунке 2.22 видно, что при одинаковом числе точек разработанный подход дает существенно более детальное представление линии. С одной стороны, это говорит о его морфологической эффективности, которая позволяет отразить характер интенсивного меандрирования даже в мелком масштабе. С другой стороны, линии полученные алгоритмом Ли-Оупеншоу, эстетически более привлекательны, для них нехарактерно излишнее сближение и угловатость. Метод Ванга-Мюллера дает наиболее выразительный и сбалансированный рисунок линий, что говорит об эффективности сегментации линии.

Оценка близости исходных и генерализованных линий по Модифицированному расстоянию Хаусдорфа представлена в Таблице 2.4. Для каждой комбинации слоев приведено исходное, результирующее число точек, а также среднее из расстояний, вычисленных между исходными и результирующими линиями. При оценке использовался только подход с расстановкой весов как более эффективный.

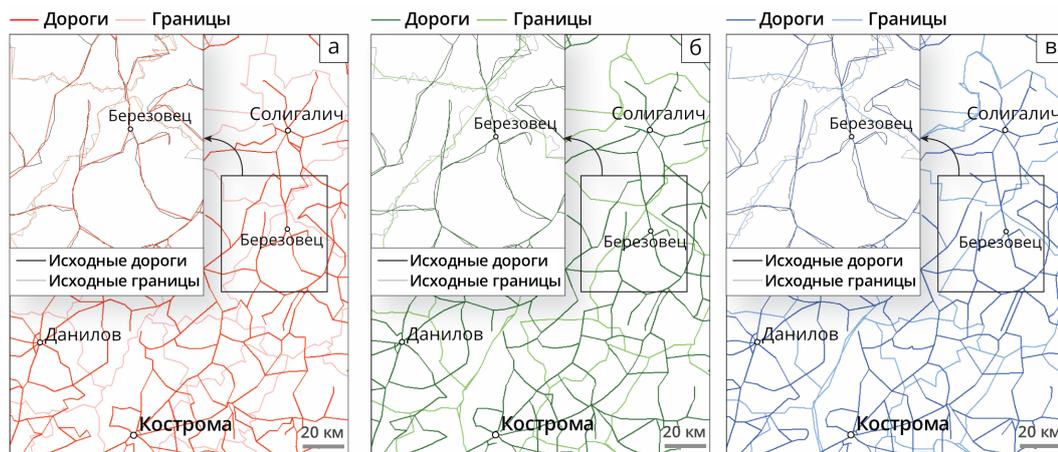


**Рисунок 2.22** – Результат генерализации слоя рек разными методами при уменьшении числа точек в 5 раз. На врезках приведены увеличенные фрагменты для сравнения с исходными линиями

**Таблица 2.4** – Модифицированное расстояние Хаусдорфа (MHD) между исходными и генерализованными объектами при использовании различных методов

Объекты	Исходное количество точек	Квадрики		Ли-Оупеншоу		Ванг-Мюллер	
		Количество точек	MHD, м	Количество точек	MHD, м	Количество точек	MHD, м
Реки	52550	10510	837	10693	1082	10576	1172
Границы муниципальных образований	14267	4855	854	4903	884	4859	841
Границы муниципальных образований, реки, береговые линии	82177	15460	969	15479	1177	15506	1072
Границы муниципальных образований и дорожная сеть	32292	11948	773	11942	672	11955	803
Границы областей и дорожная сеть	19270	832	780	8401	1014	8331	905

Результаты генерализации на примере графа, состоящего из границ областей и дорожной сети, представлены на Рисунке 2.23. Анализ полученных значений позволяет утверждать, что при сопоставимом числе узлов предлагаемый алгоритм дает линии, в среднем более точно соответствующие общей траектории исходных линий. В трех случаях из пяти этот показатель наилучший, и ни в одном случае алгоритм не показал худшего результата из трех. Результаты генерализации позволяют также оценить преимущества использования топологической структуры данных при генерализации. В частности, границы соседних районов при использовании алгоритмов Ли-Оупеншоу и Ванга-Мюллера в ряде мест не согласованы.



**Рисунок 2.23** — Результаты генерализации слоев дорожной сети и административных границ при уменьшении числа точек в 5 раз: (а) стягивание ребер; (б) алгоритм Ли-Оупеншоу; (в) алгоритм Ванга-Мюллера. На врезках приведены увеличенные фрагменты для сравнения с исходными линиями.

### 2.3.6 Итоги

В данном разделе диссертации предложен оригинальный подход к генерализации множеств линейных пространственных объектов, который *адаптируется к вертикальной связности* между ними путем сохранения их топологических отношений. В сравнении с предыдущими работами элементами его новизны являются:

- принцип геометрического упрощения — стягивание ребра на основе среднеквадратической ошибки вместо удаления узлов линии; процедура приводит к созданию новых узлов (аналогичным свойством обладает алгоритм Ли-Оупеншоу);
- Модифицированное расстояние Хаусдорфа для оценки геометрической точности результатов применительно к графовой модели;
- совместное использование Модифицированного расстояния Хаусдорфа и коэффициента относительной извилистости для численной оценки результатов генерализации с точки зрения баланса геометрической точности и географического правдоподобия;

- оценка результатов генерализации на уровне графа (множества объектов), а не отдельных объектов или слоев объектов.

Результаты экспериментов показывают, что предложенный подход дает наилучшие численные показатели генерализации в сравнении с известными методами Ли-Оупеншоу и Ванга-Мюллера. Разработанное программное обеспечение имеет свидетельство о регистрации [Якимова *et al.*, 2013].

В качестве перспектив дальнейших исследований следует назвать внедрение морфологического анализа линий, а также разработку подхода, осуществляющего коллапс циклов (разветвлений) в осевые линии. Наконец, требуется всестороннее тестирование в приложении к объектам разных типов, при разных уровнях генерализации, а также в мультимасштабном картографировании.

## 2.4 Геометрическое упрощение линейных пространственных объектов разной формы<sup>5</sup>

**Форма** географических объектов является их важнейшим свойством, на основе которого во многих случаях можно судить об их происхождении. Воспроизведение характера формы в мелких масштабах — еще один критериев географичности результатов картографической генерализации. В рамках диссертационного исследования был впервые предложен и реализован подход к геометрическому упрощению гетерогенных линий, содержащих одновременно участки несхематичного, схематичного и ортогонального типа [Samsonov, Yakimova, 2017].

Объектом анализа часто является множество гетерогенных линейных объектов, то есть объектов, имеющих разное происхождение, что отражается и на их внешних морфологических признаках. Административные границы представляют распространенный пример такого рода множества линий. Иногда эти линии следуют направлениям меридианов и параллелей, образуя ортогональный рисунок. На других участках они состоят из длинных прямых отрезков, которые уже не подчиняются направлениям сторон света, но имеют по-прежнему схематичный, визуально простой рукотворный характер. А с третьей стороны, линии границ часто совпадают с природными рубежами, такими как реки и горные хребты, что делает их характер менее угловатым, но более насыщенным изгибами, фрактальным (бесконечно детализируемым, но не обязательно самоподобным!), что выдает их природный генезис. Хорошо известным примером являются административные границы республики Коми и Архангельской области, изображенные на Рисунке 2.24.

Эти различия в форме линии можно концептуализировать в виде куба формы на

<sup>5</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov, Yakimova, 2017]



**Рисунок 2.24** — Административные границы со смешением линий различной формы. Муниципальные образования Архангельской области и Республики Коми

Рисунке 2.25, в котором три измерения соответствуют схематизму, гладкости и регулярности. *Схематичные линии* просты по форме и не включают сложную иерархию изгибов разных размеров, что характерно для несхематичных линий. *Гладкие линии* воспринимаются как имеющие постепенное изменение касательных направлений. Они контрастируют с угловатыми линиями, в которых острые углы четко воспринимаются из-за резких поворотов касательной. Наконец, *регулярность* означает ритмичное повторение некоторого геометрического паттерна вдоль протяжения линии. Например, повторение прямых углов приводит к ортогональной регулярности. Несхематичные линии также могут быть регулярными — например, реки с повторяющимся рисунком/формой меандров. Симметрию можно рассматривать также как тип регулярности [Hauert, 2012].

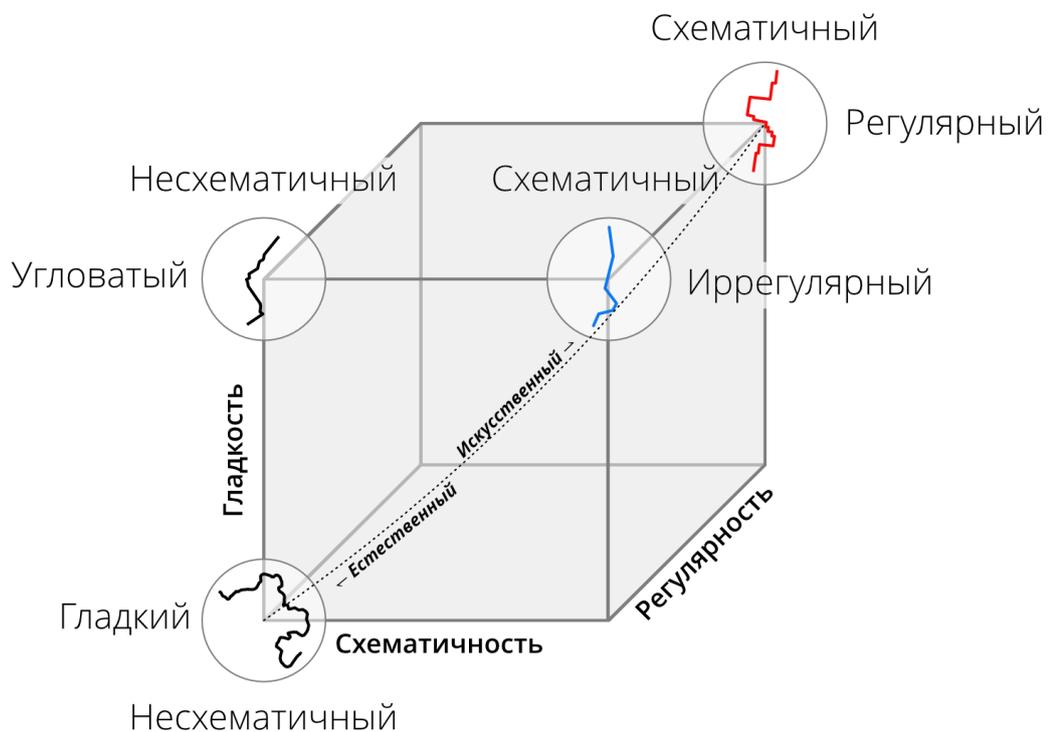
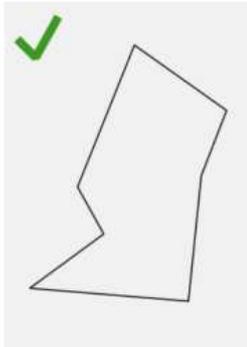
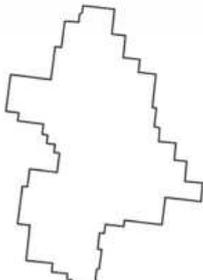
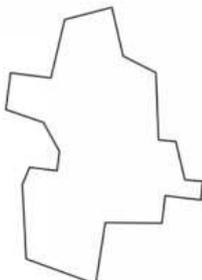
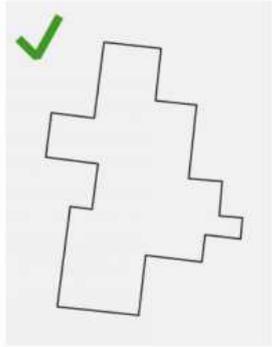


Рисунок 2.25 — Куб формы линий

Если в одном наборе данных объединены разные формы линий, ни один стандартный алгоритм из упомянутых выше, не будет эффективен глобально для геометрического упрощения. Эта проблема проиллюстрирована на Рисунке 2.26, который содержит примеры административных границ Республики Коми. Как было замечено, ортогональный рисунок может быть обобщен только с использованием некоторого специализированного алгоритма, такого как *Simplify Building* из программного пакета *ArcGIS*. Алгоритм, разработанный в [Douglas, Peucker, 1973], сохраняет угловатость схематического сегмента, но искажает углы в ортогональном сегменте. Он также разрушает характер гладких несхематичных линий, делая их острыми и схематичными. Наконец, алгоритм Ли-Оупеншоу [Li, Openshaw, 1992] делает схематичные сегменты гладкими, что также является ошибкой обобщения. Однако он хорошо работает для гладких несхематичных

сегментов.

<b>Алгоритм геометрического упрощения</b>				
<b>Тип линии</b>	<b>Источник</b>	Douglas-Peucker	Li-Openshaw	Simplify Building
Гладкий иррегулярный несхематичный				
Угловатый иррегулярный схематичный				
Угловатый регулярный схематичный (Ортогональный)				

**Рисунок 2.26** – Результаты обработки линий разной формы разными алгоритмами геометрического упрощения

В этом разделе диссертации фокус исследования сосредоточен на развитии качественной сегментации линий. Показано, как этот подход может помочь в сохранении формы линий в одном гетерогенном наборе данных при геометрическом упрощении в мелких масштабах.

### 2.4.1 Методика

Разработана модель геометрического упрощения линий, состоящих из сегментов трех типов: несхематичных, схематичных и ортогональных (частный случай схематичных). Технологический процесс модели состоит из трех этапов: предварительная обра-

ботка, обработка и постобработка. Этап предварительной обработки включает операции, такие как фильтрация, сегментация и ортогонализация. Обработка включает в себя итеративное геометрическое упрощение сегментов, в котором каждый из них обрабатывается с использованием выбранного алгоритма; затем сегменты малой длины объединяются с соседями, и упрощение выполняется снова. Постобработка реализует регуляризацию результатов, аналогичную предварительной обработке, но выполняется на упрощенной геометрии. Все эти операции описаны более подробно ниже.

Прежде чем перейти к алгоритмам и методам, необходимо пояснить, что ожидается увидеть при поиске вышеупомянутых типов линий. Имея ломаную линию (далее — просто *линию*), состоящую из отрезков (без дуг и кривых), определим сегмент (или участок) линии как *схематичный*, если он состоит из отрезков, которые достаточно длинны для их отдельного восприятия. Формально для этого нужно установить допуск длины  $D$ . Тогда каждый отрезок  $\mathbf{e}_i = \overrightarrow{p_i, p_{i+1}}$  с длиной  $l_i = \|\mathbf{e}_i\| \geq D$  будет классифицироваться как принадлежащий схематичному сегменту. В противном случае он классифицируется как несхематичный.

Предпочтительно определять  $D$  в миллиметрах на экране, поскольку это позволяет абстрагироваться от разрешения исходных данных, проводить экспериментальную работу, изменяя значение  $D$  и масштаб визуализации. В этом случае количество распознанных схематичных сегментов будет зависеть от интуитивно понятных параметров.

Дифференциация *угловатых* и *плавных* линий зависит от угла поворота линии в каждой вершине  $\delta_i = \angle(\mathbf{e}_{i-1}, \mathbf{e}_i) = \arccos \frac{\mathbf{e}_{i-1} \cdot \mathbf{e}_i}{\|\mathbf{e}_{i-1}\| \cdot \|\mathbf{e}_i\|}$ . Можно установить некую константу  $\delta$  в качестве граничного угла между плавными и угловатыми вершинами. Следовательно, если  $|\delta_i| \leq \delta$ , то вершина считается гладкой; в противном случае считается угловатой. Следует обратить внимание на то, что схематичность определяется через отрезки, тогда как анализ гладкости производится на уровне вершин.

Наконец, для распознавания линий *ортогонального* типа как частного случая регулярного угловатого схематичного типа нам нужно найти вершины, которые одновременно удовлетворяют следующим двум условиям:

- Каждое ребро при вершине  $p_i$  должно быть схематичным:  $l_{i-1} \geq S; l_i \geq S$ , где  $S$  — допуск длины для ортогональных схематичных сегментов, которые могут отличаться от  $D$ . Это условие гарантирует, что оба отрезка, составляющие угол, достаточно длинны, так что угол при вершине можно визуально оценить.
- Угол при вершине должен быть приблизительно равен  $90^\circ$ . Некоторые углы на карте воспринимаются как прямые или почти прямые, хотя на самом деле они немного больше или меньше  $90^\circ$ , что приводит к отклонению  $\alpha_i = |90^\circ - \delta_i|$ . Отрезки часто не образуют в точности ортогональную последовательность. Ошибки оцифровки, точность координат, а также искажения, вносимые проекцией, также будут влиять на значения угла. Таким образом, также необходим некоторый

угловой допуск  $A$ , который позволяет рассматривать угол как прямой, если его значение находится в окрестности  $90^\circ \pm A$ . Это приводит ко второму условию ортогонального угла,  $\alpha_i \leq A$ .

На этом этапе исследование не фокусировалось на разделении плавных и угловатых линий и было принято, что все несхематичные линии являются плавными, а все схематичные — угловатыми. Это предположение представляется приемлемым по двум причинам. Во-первых, несхематичные линии состоят из коротких отрезков в соответствии с введенным определением, и отклонение линии в точке недостаточно различимо. Во-вторых, метод ориентирован на мелкомасштабную генерализацию, где плавные схематичные линии редки по сравнению с крупномасштабным картографированием.

#### 2.4.1.1 Предобработка

Цель этапа предварительной обработки состоит в том, чтобы разбить линию на геометрически однородные участки, которые могут быть эффективно обработаны с использованием алгоритмов упрощения, сохраняющих их характер. Для достижения этой цели предлагается несколько процедур, таких как *фильтрация*, *сегментация* и *выпрямление углов*, которые позволяют удалить избыточные вершины, разбить линию на сегменты и выровнять ортогональные сегменты для применения специализированного алгоритма упрощения. Пример сегментации представлен на Рисунке 2.27. Исходный объект изображен тонкой черной линией, а выделенные сегменты показаны в цвете. Будем ссылаться на этот рисунок в дальнейших объяснениях.

#### Фильтрация

Переменная плотность точек вдоль линии, а также потенциальная возможность наличия шума в координатах приводят к необходимости удаления избыточных вершин в целях очистки линии для эффективной сегментации. Целью является оптимизация представления линии с помощью процедуры фильтрации, которая приводит к форме, которая визуально повторяет исходную, но представлена меньшим количеством вершин. Процедуру можно выполнить одним из алгоритмов редуцирования точек, предназначенных для геометрического упрощения [Li, 2006]. Параметр алгоритма должен быть установлен малым, поскольку необходимо удалить только лишние точки, которые не влияют на общую форму линии.

Был применен простой последовательный алгоритм редуцирования точек, основанный на перпендикулярном расстоянии, аналогичный алгоритму *sleeve-fitting*<sup>6</sup> [Zhao, Saalfeld, 1997]. Алгоритм управляется одним параметром  $d$ , который является самым коротким расстоянием от текущей точки до хорды, стягивающей две опорные точки аналогично итеративному алгоритму *Дугласа-Пейкера*. Однако, в отличие от алгоритма

<sup>6</sup>Дословно — *подгонка рукава*

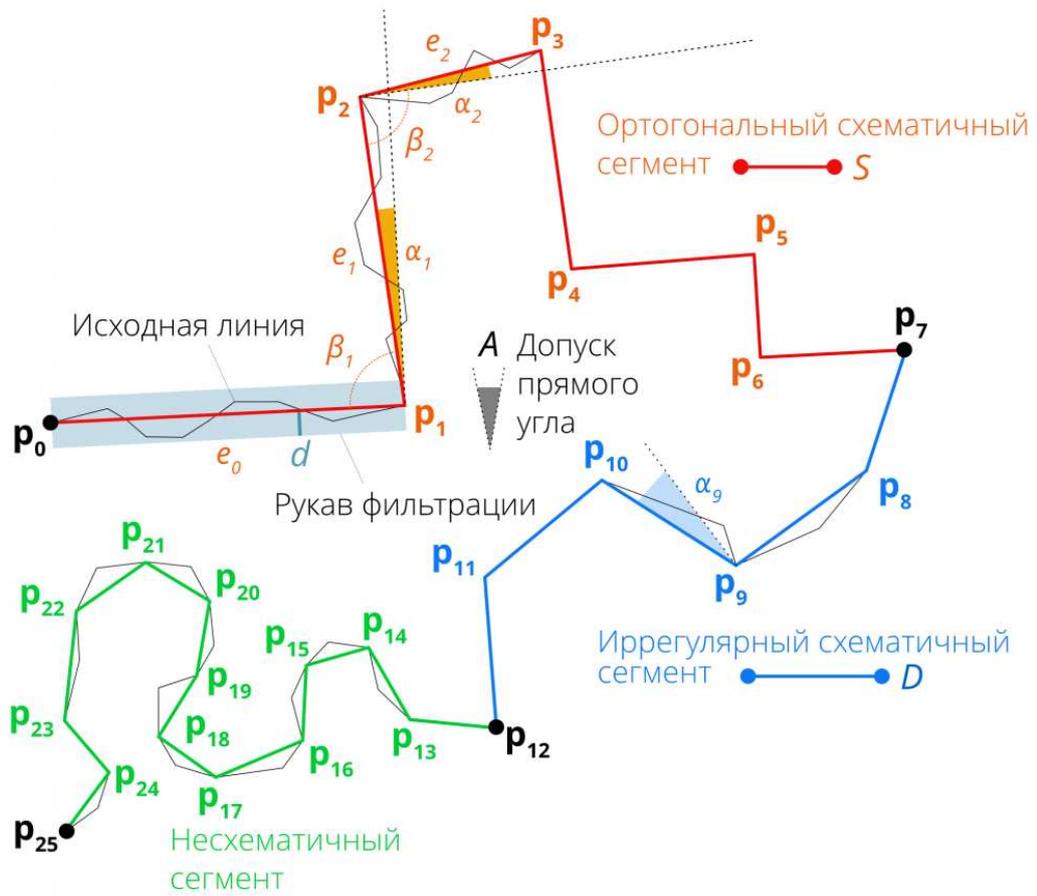


Рисунок 2.27 – Сегментация линий

Дугласа-Пейкера, алгоритм *sleeve-fitting* является последовательным с точки зрения классификации [Li, 2006]. Дополнительно контролируется топологическую согласованность во время фильтрации, проверяя, пересекается ли вновь созданный сегмент с любой существующей линией в наборе данных. Аналогичная стратегия была предложена ранее в работе [Pallero, 2013].

В нашем методе  $d$  указывается в миллиметрах на экране и привязывается к масштабу предварительной обработки (исходному масштабу). Когда алгоритм выполняется, параметр преобразуется в расстояния на местности, что позволяет привести значение допуска в соответствие с системой координат вершин линий.

Мы будем отмечать точки, которые были оставлены после фильтрации как  $p_i$ , где  $i$  — новый точечный индекс, пересчитанный для отфильтрованной линии. Пример результирующего сегмента представлен на Рисунке 2.27, где прямоугольник «рукава» покрывает все начальные точки между  $p_0$  и  $p_1$ , которые остаются после фильтрации. Эти точки находятся внутри прямоугольника шириной  $2d$  с главной осью  $p_0p_1$ . Затем следующий рукав покрывает все начальные точки между  $p_1$  и  $p_2$  и т.д. Процесс фильтрации может быть концептуализирован как последовательная подгонка таких рукавов к исходной форме линии.

## Сегментация

Сегментация включает идентификацию ортогональных, схематичных и несхематичных сегментов.

Идентификация ортогональных сегментов предполагает обнаружение и связывание соседних прямых углов в линии. Углы вычисляются между каждой парой смежных отрезков линии, длина которых больше, чем  $S$ . Пусть  $s_i = \min(l_{i-1}, l_i)$  — длина самого короткого отрезка при угле. Если  $\alpha_i \leq A$  и  $s_i = S$ , то угол отмечается как прямой.

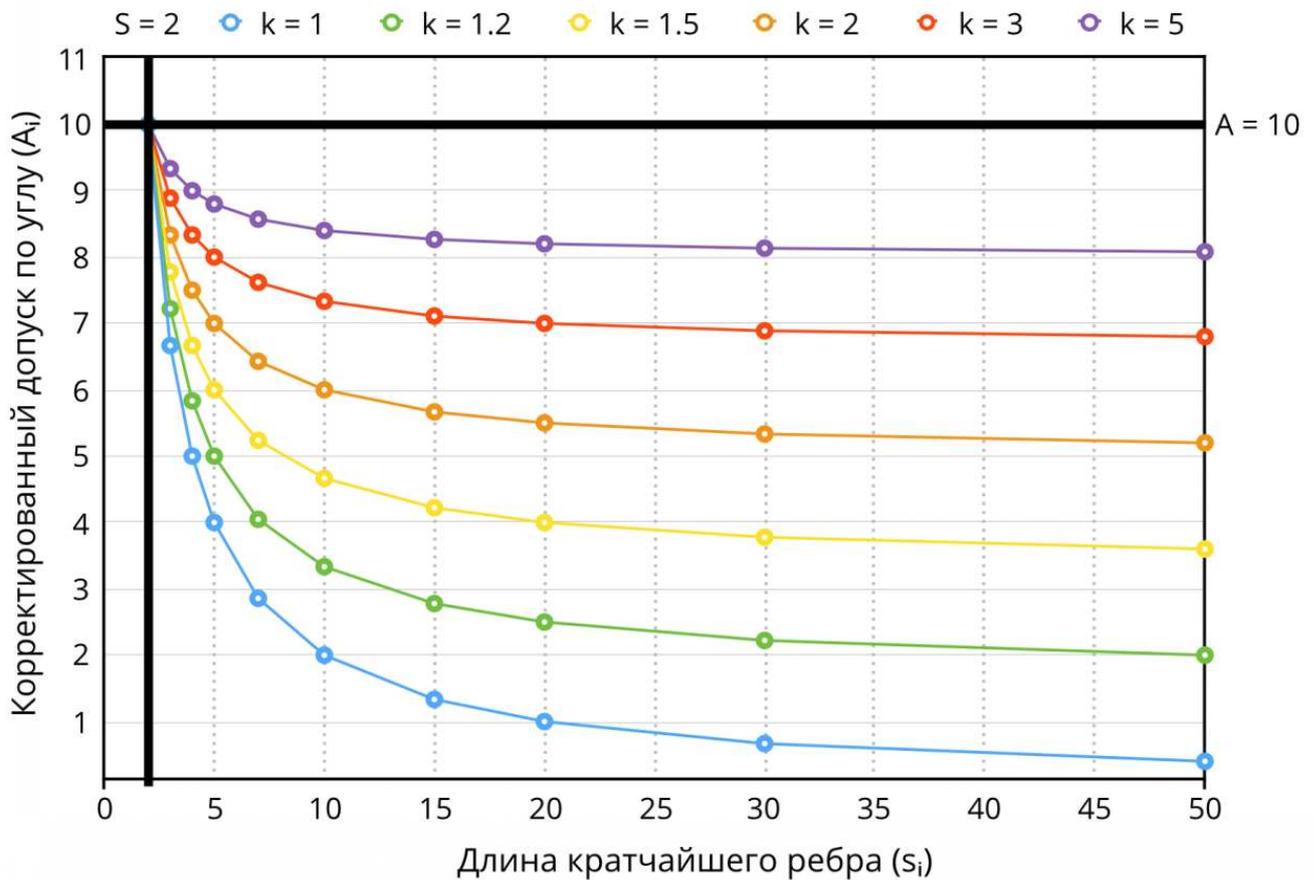
Далее можно сделать предположение, что  $A$  необходимо скорректировать для случаев, когда  $s_i > S$ . Причиной этой коррекции является то, что ортогонализация углов с длинными ребрами приводит к значительному смещению вершин, что в конечном итоге может привести к искажению формы и топологическим ошибкам. Следовательно, значение  $A$  должно быть малым для углов, образованных длинными отрезками, и наоборот.

Скорректированный допуск по углу  $A_i$  вычисляется по следующей формуле:

$$A_i = A \left( 1 - \frac{s_i - S}{ks_i} \right), \quad (2.14)$$

где  $k \geq 1$  — параметр ингибции. Чем больше значение  $k$ , тем медленнее уменьшается угол отклонения пропорционально увеличению стороны самого короткого угла. Это преобразование показано на Рисунке 2.28, с  $S = 2$  мм и  $A = 10^\circ$ . Когда  $k = 1.0$ , тогда соотношение прямо пропорционально: если ребро увеличивается в 5 раз, то угловой

допуск уменьшается в 5 раз (от  $10^\circ$  до  $2^\circ$ ). Однако при  $k = 2.0$  уменьшение допуска по углу будет всего в 1.67 раза (от  $10^\circ$  до  $6^\circ$ ).



**Рисунок 2.28** – Отношения между скорректированным угловым допуском и длиной самой короткой стороны угла

Включение этого параметра позволяет рассматривать углы с относительно длинными сторонами для ортогонализации и проводить эксперименты по выбору оптимального значения  $k$ , которые будут описаны ниже.

После выделения ортогональных отрезков все неортогональные отрезки, длина которых больше  $D$ , классифицируются как схематичные. Затем все остальные ребра классифицируются как несхематичные. Ребра всех трех типов объединяются в сегменты для дальнейшей обработки.

Наконец, вернемся к примеру на Рисунке 2.27, чтобы оценить все этапы сегментации линии. Во-первых, начальная линия фильтруется с использованием алгоритма *sleeve-fitting*, в результате чего остается множество точек  $P = \{p_i\}$ . Во-вторых, сегмент  $p_0 \dots p_7$  классифицируется как ортогональный по заданным угловым ( $A$ ) и линейным ( $S$ ) допускам пороговых значений. В-третьих, сегмент  $p_7 \dots p_{12}$  классифицируется как иррегулярный схематичный в соответствии с линейным порогом  $D$ . Наконец, остальные точки  $p_{12} \dots p_{25}$  классифицируются как принадлежащие несхематичному сегменту.

Заключительным этапом предварительной обработки является выпрямление углов ортогональных сегментов.

### Выпрямление углов

Один из основных принципов картографической генерализации гласит, что нечеткие пространственные отношения должны быть преобразованы в четкие; в частности, это приводит к трансформации почти перпендикулярных линий в перпендикулярные линии [Touya *et al.*, 2014].

Выпрямление углов выполняется в соответствии со следующей процедурой для каждого ортогонального сегмента:

1. Вычислить общие длины нечетных и четных ребер. Наибольшая сумма соответствует основной ориентации сегмента.
2. Для каждого ребра в основной ориентации вычислить угол сближения  $\gamma_i$  между ребром и направлением  $y$ .
3. Получить средневзвешенное значение  $\gamma_i$  в качестве дескриптора основной ориентации:

$$\gamma_{av} = \frac{\sum_i l_i \gamma_i}{\sum_i l_i}. \quad (2.15)$$

4. Развернуть каждое ребро основного направления вокруг его средней точки, чтобы его направление было строго равно  $\gamma_{av}$ . Если ребро содержит конечную точку сегмента, он вращается вокруг конечной точки. Таким образом, вращение не влияет на геометрию смежного сегмента линии.
5. Реконструировать каждое ребро второстепенного направления, построив линию, проходящую через его среднюю точку и перпендикулярную смежным сегментам основного направления. Конечные точки нового сегмента совмещены с окончными точками смежных сегментов основного направления.

После того как определена процедура выпрямления углов, можно произвести оценку параметров сегментации. Для этого был выбран и пропущен через всю цепочку предварительной обработки один фрагмент данных. Результирующим точкам были присвоены следующие значения в соответствии с типами сегментов, примыкающих к узлу:

- *Ортогональные* — вершина с обоими ребрами  $s_i \geq S$  и  $\alpha_i \leq A$ .
- *Два схематичных* — вершина с  $s_i \geq S$  и  $\alpha_i > A$ .
- *Одно схематичное* — только один соседний край длиннее  $S$ , тогда как другой меньше  $S$ .
- *Несхематичные* — вершина с обоими ребрами  $s_i < S$ .

Рисунок 2.29а показывает влияние отклонения допуска по углу  $A$  на обнаружение прямых углов. Очевидно, чем больше допуск, тем больше число обнаруженных прямых

углов. Точка в юго-восточном углу сегмента является точкой стыковки двух сегментов, поэтому не помечается как ортогональная.

Рисунок 2.29б иллюстрирует влияние допуска длины  $S$  на результаты выделения прямых углов. В данном случае из меньшей длины не следует большее число прямых углов. В качестве примера можно сравнить случаи  $S = 1.0$  и  $S = 2.0$  мм. Угол в черной точке при  $S = 1.0$  не был распознан как ортогональный из-за формулы коррекции допуска по углу, которая делает его слишком жестким для ребер малой длины.

Рисунок 2.29в иллюстрирует влияние параметра ингибиции  $k$  на обнаружение прямых углов. Здесь большее ингибирование, как ожидается, приводит к увеличению числа обнаруженных прямых углов.

Комбинация параметров, которая приводит к наибольшему числу ортогональных углов, в данном случае следующая:  $A = 15.0^\circ$ ,  $S = 2.0$  мм и  $k = 3.0$ . Эта комбинация будет использована позже в разделе результатов во время оценки сегментации линий. Эти значения могут быть не оптимальными для других наборов данных и должны быть соответствующим образом выбраны (см. параграф «Обсуждение»).

### 2.4.1.2 Обработка

#### Геометрическое упрощение ортогональных сегментов

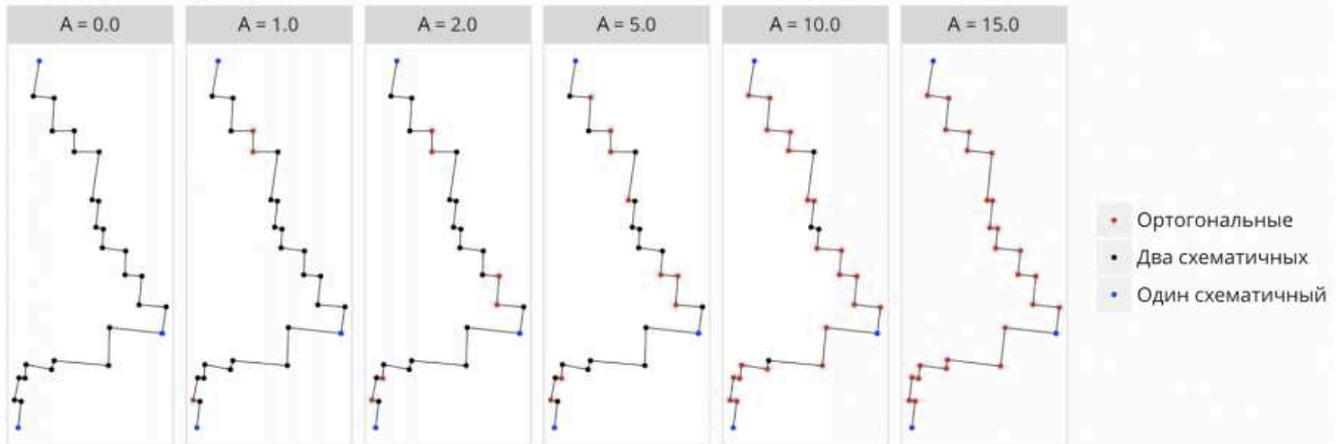
При выполнении геометрического упрощения пользователь должен установить целевой масштаб, который в  $R$  меньше масштаба исходного. Затем необходимо удалить ребра, меньшие  $S'$  в целевом масштабе. Следовательно, результат зависит от степени уменьшения масштаба, выбранной пользователем.

Сохранение характера ортогональной линии означает сохранение прямых углов в ее рисунке. В целях формализации возможных процедур геометрического упрощения целесообразно ввести ограниченное число возможных ортогональных конфигураций. Предлагается ввести  $Z$ -подобные,  $U$ -подобные, краевые и короткие конфигурации, как показано на Рисунке 2.30. Упрощение ортогонального сегмента основано на процедуре замены ребра, в которой удаленное ребро и его соседи заменяются новой конфигурацией. Предлагается использовать замены трех типов: медиана, срезка и диагональ.

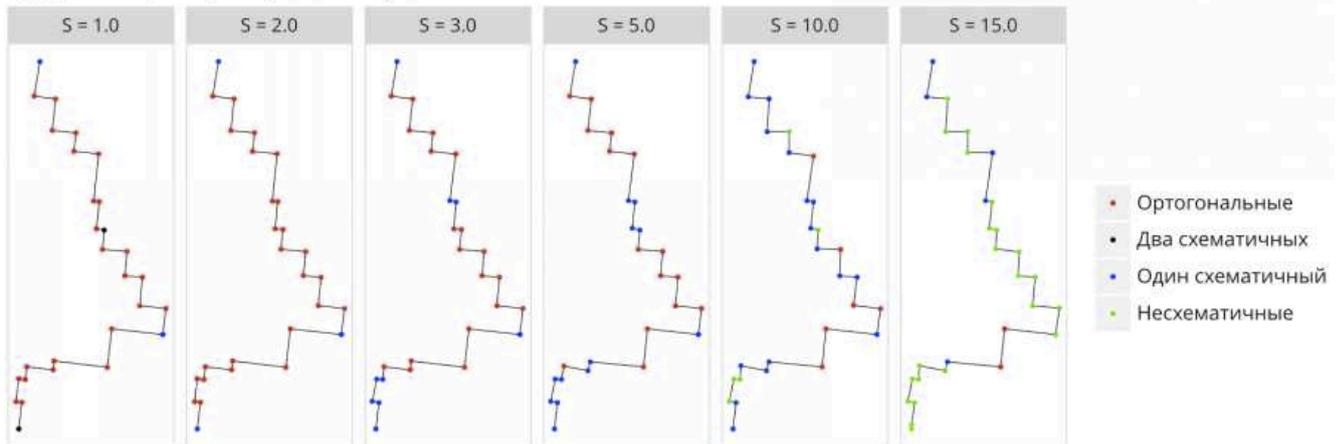
**Медиана** представляет собой вариант упрощения ортогональной конфигурации, при котором удаленное ребро  $e_i$  заменяется новым перпендикулярным ребром, которое соединяет прямые, на которых лежат ребра  $e_{i-2}$  и  $e_{i+2}$ . Медианная замена работает лучше всего, поскольку она имеет тенденцию выравнивать пропорции между соседними изгибами и делает ортогональную структуру более регулярной — преувеличивает малые изгибы за счет уменьшения больших, что ценно для мелкомасштабной генерализации. Такая замена используется по умолчанию для  $U$ -подобных и  $Z$ -подобных конфигураций.

**Срезка** реализуется соединением ребер  $e_{i-1}$  и  $e_{i+2}$  или, альтернативно,  $e_{i-2}$  с  $e_{i+1}$ .

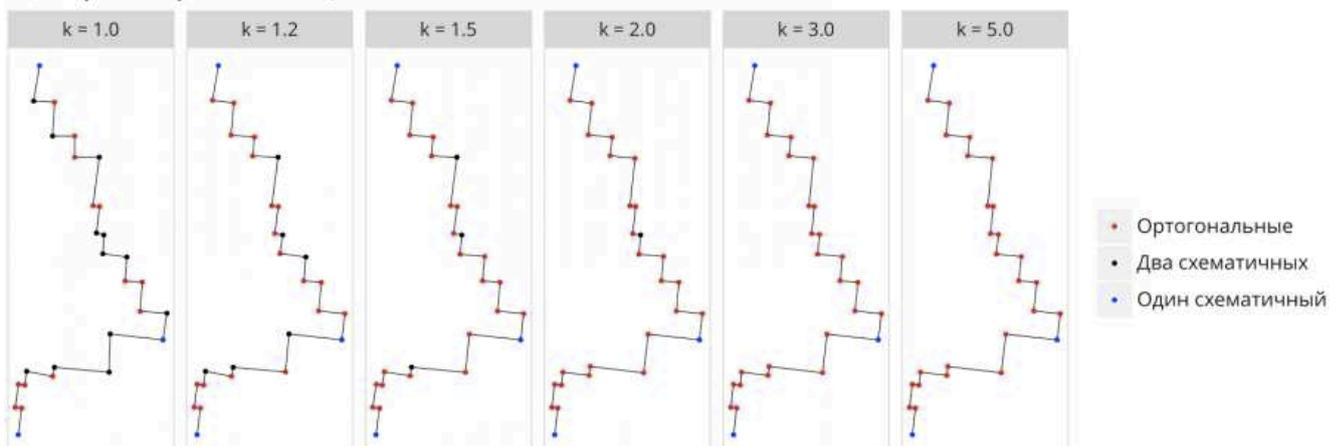
## а) Допуск по углу



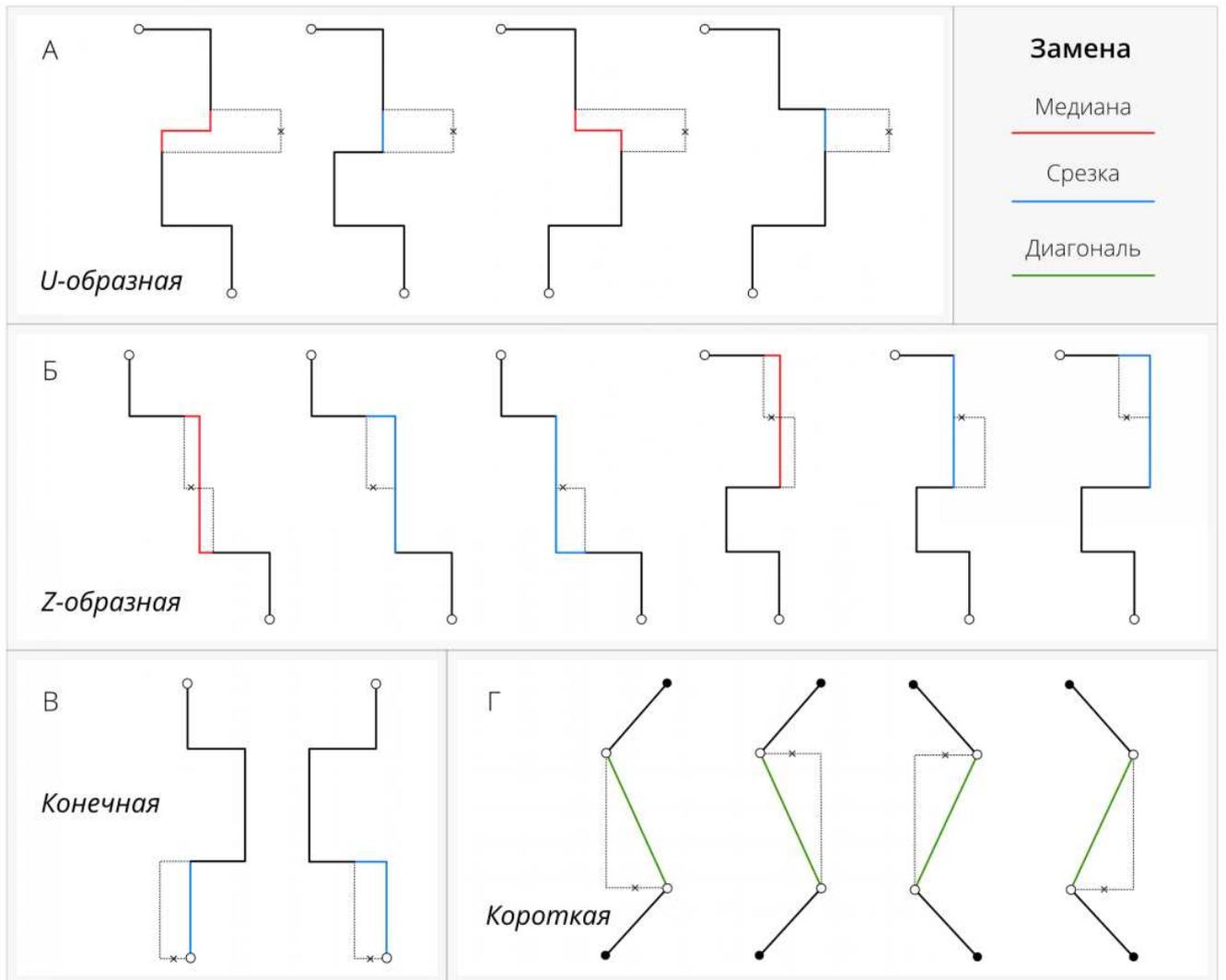
## б) Длина ребра прямого угла



## в) Параметр ингибции



**Рисунок 2.29** – Результат выпрямления углов при изменении: а) допуска по углу  $A$  (при  $k = 3$ ,  $S = 2$ ); б) допуска по длине ребра  $S$  (при  $k = 3$ ,  $A = 15^\circ$ ); в) параметра ингибции  $k$  (при  $A = 15^\circ$ ,  $S = 2$ )

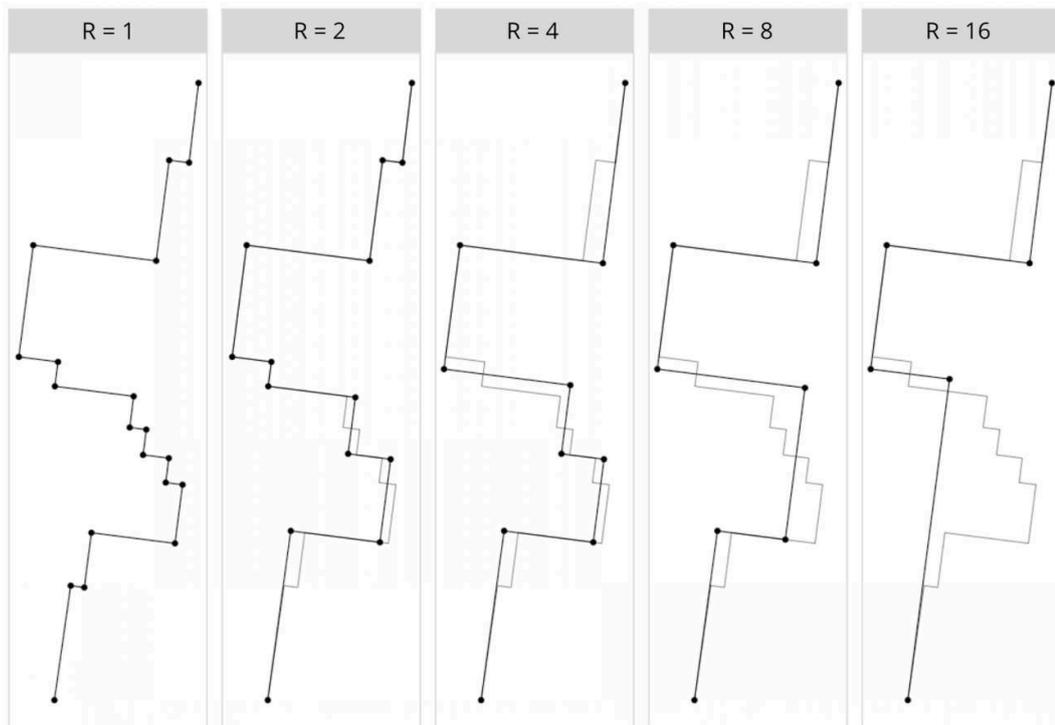


**Рисунок 2.30** – Стратегии замены ребра для различных ортогональных конфигураций. Конечные точки символизируются полыми точками, старая конфигурация – пунктирной линией, сжатый край отмечен крестиком, а новая конфигурация изображена цветом

Срезки применяются в случаях, когда медианная замена приводит к топологической ошибке (например, пересечению). Из двух возможных вариантов срезов выбирается тот, который приводит к наименьшему изменению площади ограничивающего прямоугольника конфигурации и при этом не имеет топологической ошибки. Особый вариант срезки реализуется для первого или последнего ребра в ортогональном сегменте. В этом случае  $e_2$  перемещается вдоль  $e_1$  вплоть до совпадения  $p_2$  с  $p_1$ , при этом  $e_1$  стягивается в точку. Если любой вариант срезки приводит к топологической ошибке, производится попытка выполнить диагональную замену.

**Диагональ** применяется на коротких ортогональных сегментах, состоящих из трех точек, и подразумевает соединение первой и последней точек в сегменте. Если будем бесконечно уменьшать результирующий масштаб, то каждый ортогональный сегмент в конечном итоге будет представлен некоторой трехточечной конфигурацией, которая потребует такой замены. Поскольку данный процесс приводит к неизбежной потере ортогональности, *полученное ребро затем классифицируется как схематичное*, поскольку по своей природе ортогональный рисунок ближе к схематичному. Диагональная замена также выполняется когда и медиана, и срезка приводят к топологическим ошибкам. Если и диагональная замена приводит к топологическим ошибкам, ортогональный сегмент будет оставлен нетронутым.

Пример результата геометрического упрощения ортогонального сегмента с использованием замен представлен на Рисунке 2.31.



**Рисунок 2.31** — Пример геометрического упрощения ортогонального сегмента с использованием техники замены ребра

### Геометрическое упрощение схематичных и несхематичных сегментов

Ранее на Рисунке 2.26 было показано, что алгоритм Ли-Оупеншоу хорошо работает для несхематичных сегментов. Данный алгоритм сохраняет одинаковое количество ребер на единицу площади изображения вследствие того, что основан на регулярном разбиении плоскости. Более того, точки, попадающие в каждую ячейку, могут быть заменены на одну точку, которая усредняет первую и последнюю точку, согласно одному из подходов, предложенных в [Li, Openshaw, 1992]. Полученные точки не являются подмножеством исходных точек, и алгоритм Ли-Оупеншоу в конечном итоге работает как скользящее среднее, что делает его предпочтительным при генерализации гладких несхематичных сегментов.

Одним из преимуществ алгоритма Ли-Оупеншоу является то, что он по определению не может привести к топологическим ошибкам, таким как пересечения с существующими линиями. [Li, 2006, стр. 154] отмечает в своей книге, что *«утвердилось неверное представление о том, что алгоритм Ли-Оупеншоу подходит только для обобщения линий природного характера. В случае если используются точки исходной линии, алгоритм Ли-Оупеншоу можно также использовать для упрощения искусственных линий, например, кадастровых границ»*. Он представил простой пример, используя искусственную линию, в которой в каждой ячейке сохраняется только первая точка. Тем не менее, можно высказать вполне обоснованное предположение, что алгоритм Ли-Оупеншоу неэффективен для обобщения схематичных сегментов, поскольку он не может оценить важность отдельных узлов в контексте всей линии и ее конфигурационных сегментов.

В противоположность этому, алгоритм Дугласа-Пейкера основан на процедуре, которая находит стягивающие хорды между существующими точками. Он стремится максимизировать количество точек, стягиваемых каждой хордой, при этом удовлетворяя критерию допустимого расстояния. Эта тенденция максимизации приводит к форме линии, которая знакома каждому картографу и обычно выглядит чрезмерно обобщенной и угловатой по сравнению с тем, что получается при использовании других алгоритмов. В то же время, это именно тот эффект, который следует использовать при генерализации схематичных линий. Данный метод был использован в качестве алгоритма геометрического упрощения таких линий, с учетом коррекции, предложенной в [Ebisch, 2002].

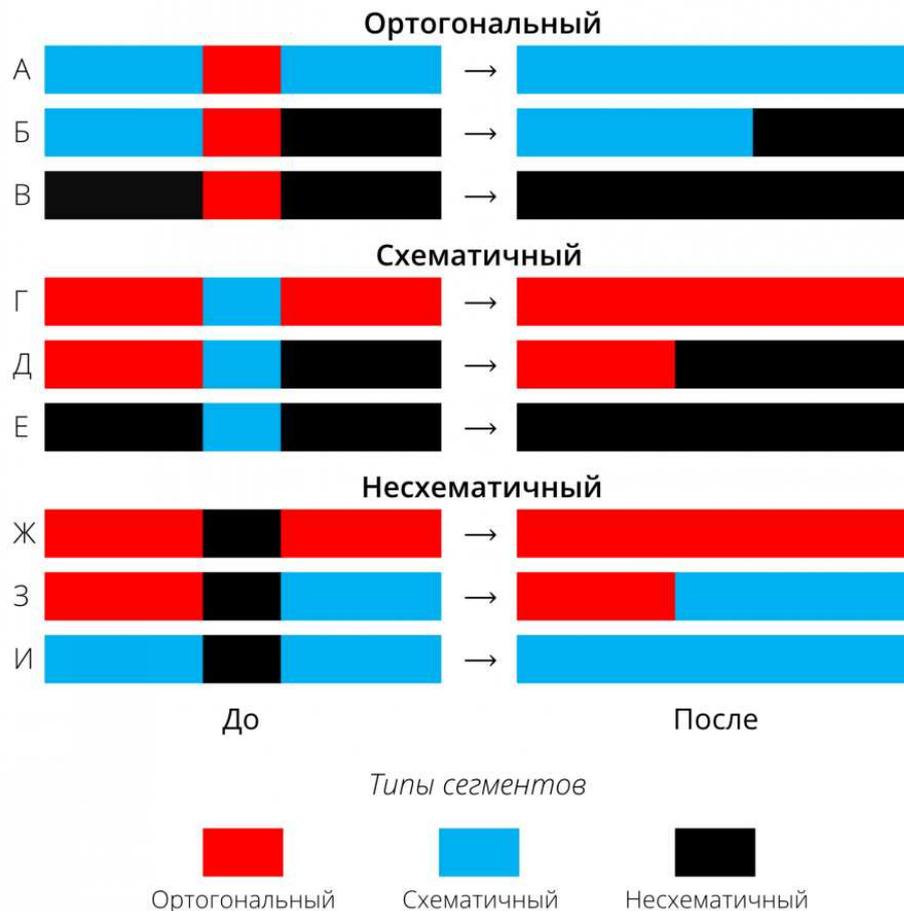
Для того, чтобы упрощение методом Дугласа-Пейкера не приводило к топологическим ошибкам, его рекурсивная процедура выполняется в каждой ветви до тех пор, пока пороговое значение ошибки не будет удовлетворено, либо пока не будет пересечений с существующими сегментами.

Обобщение этих типов сегментов контролируется одним параметром допуска  $T$ , который соответствует разрешению сетки в алгоритме Ли-Оупеншоу и перпендикулярному расстоянию в алгоритме Дугласа-Пейкера. Значение  $T$  определяется в миллиметрах на экране — в этом случае результат зависит от уровня масштабирования, выбранного пользователем.

### Итеративное слияние

Сегментация линии может приводить к визуальным нарушениям, когда некоторые сегменты становятся слишком маленькими в результирующем масштабе. Например, можно представить несхематичный сегмент длиной 1 мм между двумя схематичными сегментами. Поскольку границы сегментов являются фиксированными, данный сегмент останется в виде небольшой детали и будет нарушать общую детализацию изображения.

Данную ситуацию можно обобщить в виде 9 возможных конфигураций, которые представлены на Рисунке 2.32. В левом столбце на этом рисунке показана ситуация после обработки данных. Средний сегмент в каждой строке имеет длину меньше заданного порога  $S'$ . Необходимо разрешить ситуацию, удалив данный сегмент. Предлагается операция *слияния линий* как решение этой проблемы. В зависимости от конкретной конфигурации средний сегмент может быть объединен с одним из соседей или с обоими соседями.



**Рисунок 2.32** – Стратегии слияния коротких сегментов с соседними сегментами. Анализируемый сегмент расположен в центре каждой схемы. Комментарии по стратегиям см. в тексте

Первые три случая (A — C на Рисунке 2.32) представляют ситуацию, в которой должен быть удален короткий ортогональный сегмент. Если он окружен двумя схематичными (A) или несхематичными (C) сегментами, он объединяется с обоими из них, что

приводит к образованию одного сегмента соответствующего типа. Если ортогональный сегмент окружен двумя сегментами разных типов (B), он сливается со схематичным сегментом как с более родственным по характеру.

Вторая группа случаев ( $D - F$  на Рисунке 2.32) относится к удалению короткого схематичного сегмента, окруженного несхематичными и ортогональными сегментами. В случае D короткий схематичный сегмент сжимается двумя перемещениями смежных ортогональных ребер (техника, аналогичная [Buchin *et al.*, 2016]) в середину центрального сегмента. Движения вдоль ребер позволяют сохранять ортогональность результирующего сегмента. В остальных случаях E и F короткий схематичный сегмент интегрируется в структуру соседнего несхематичного сегмента (сегментов).

Третья группа случаев ( $G - I$  на Рисунке 2.32) относится к объединению несхематичных сегментов. Данные конфигурации обрабатываются аналогично предыдущему случаю. Небольшой несхематичный сегмент (конфигурация G) сливается с двумя ребрами смежных ортогональных сегментов. Наконец, небольшой несхематичный сегмент можно свободно интегрировать и в соседние схематичные сегменты (конфигурации H и I).

Операция слияния выполняется сначала для ортогональных сегментов, затем для схематичных сегментов и, наконец, для несхематичных сегментов. В каждой группе сегментов слияние выполняется от более коротких к более длинным сегментам. Эта стратегия позволяет увеличить длину более длинных сегментов, прежде всего, за счет коротких сегментов. Некоторые сегменты становятся длиннее порога после выполнения данной процедуры, что, в конечном счете, приводит к меньшему количеству последующих объединений.

После выполнения слияния геометрическое упрощение сегментов выполняется снова. Затем анализируется возможность нового слияния. Процесс продолжается итеративно до тех пор, пока не останется коротких сегментов, которые требуется подвергнуть слиянию с соседями.

#### 2.4.1.3 Постобработка и обзор модели генерализации

Постобработка аналогична стадии предварительной обработки, но выполняется на уже обработанном (обобщенном) наборе данных и в результирующем масштабе. Этот этап обработки данных был добавлен, чтобы предоставить пользователю возможность сделать изображение более регулярным и выровнять плотность точек вдоль полученных линий. Кроме того, данная процедура позволяет идентифицировать сегменты, которые не были классифицированы как ортогональные в масштабе источника, но в результате геометрического упрощения приобрели форму, близкую к ортогональной. Влияние постобработки описано в разделе, посвященном экспериментальным результатам.

Общая схема разработанной модели генерализации, которая включает в себя несколько этапов и алгоритмов, показана на Рисунке 2.33. Схема иллюстрирует перво-

начальную идею применения различных подходов упрощения к линиям различного характера и сводит ранее представленные этапы обработки линий в единый технологический процесс.

## 2.4.2 Результаты

### 2.4.2.1 Программная реализация

Разработанная технология была реализована в приложении **MapSimplify 2.0**<sup>7</sup> с использованием языка *C#*. Все алгоритмы были реализованы в последовательном режиме, поскольку это упрощает контроль топологических ошибок. Формат *Esri Shapefile* использовался для загрузки исходных данных и сохранения результатов. Пользовательский интерфейс программы представлен на Рисунке 2.34. Результат обобщения контролируется установкой различных параметров, упомянутых выше, а также выбором желаемой комбинации трех алгоритмов обобщения. Постобработка также возможна при установке соответствующего флажка.

Типичный рабочий процесс пользователя следующий. Во-первых, необходимо загрузить данные и установить масштаб предварительной обработки. Установка масштаба влияет на количество обнаруженных ортогональных, схематичных и несхематичных сегментов, поскольку параметры их извлечения определяются в миллиметрах на экране. Затем оптимальная комбинация параметров предварительной обработки может быть найдена путем изменения их значений и экспертной оценки результатов. Аналогичным образом находятся оптимальные параметры геометрического упрощения и постобработки.

### 2.4.2.2 Постановка эксперимента

Границы районов Республики Коми и Архангельской области были выбраны для оценки эффективности предлагаемого метода. Исходными данными были цифровая топографическая карта России масштаба 1 : 1 000 000, предоставленная ком Дата+. Набор данных содержит 116 линий, состоящих в общей сложности из 7090 узлов, в среднем каждая линия содержит 61 узел.

Экспериментальные исследования включили следующие виды работ:

- визуальная и численная оценка результатов сегментации, выполненной при различных величинах параметров предварительной обработки;
- оценка времени выполнения предварительной обработки;
- визуальное сравнение результатов генерализации, полученных различными алгоритмами;

<sup>7</sup>Доступно для загрузки по адресу <http://autolab.geogr.msu.ru/software/MapSimplify>

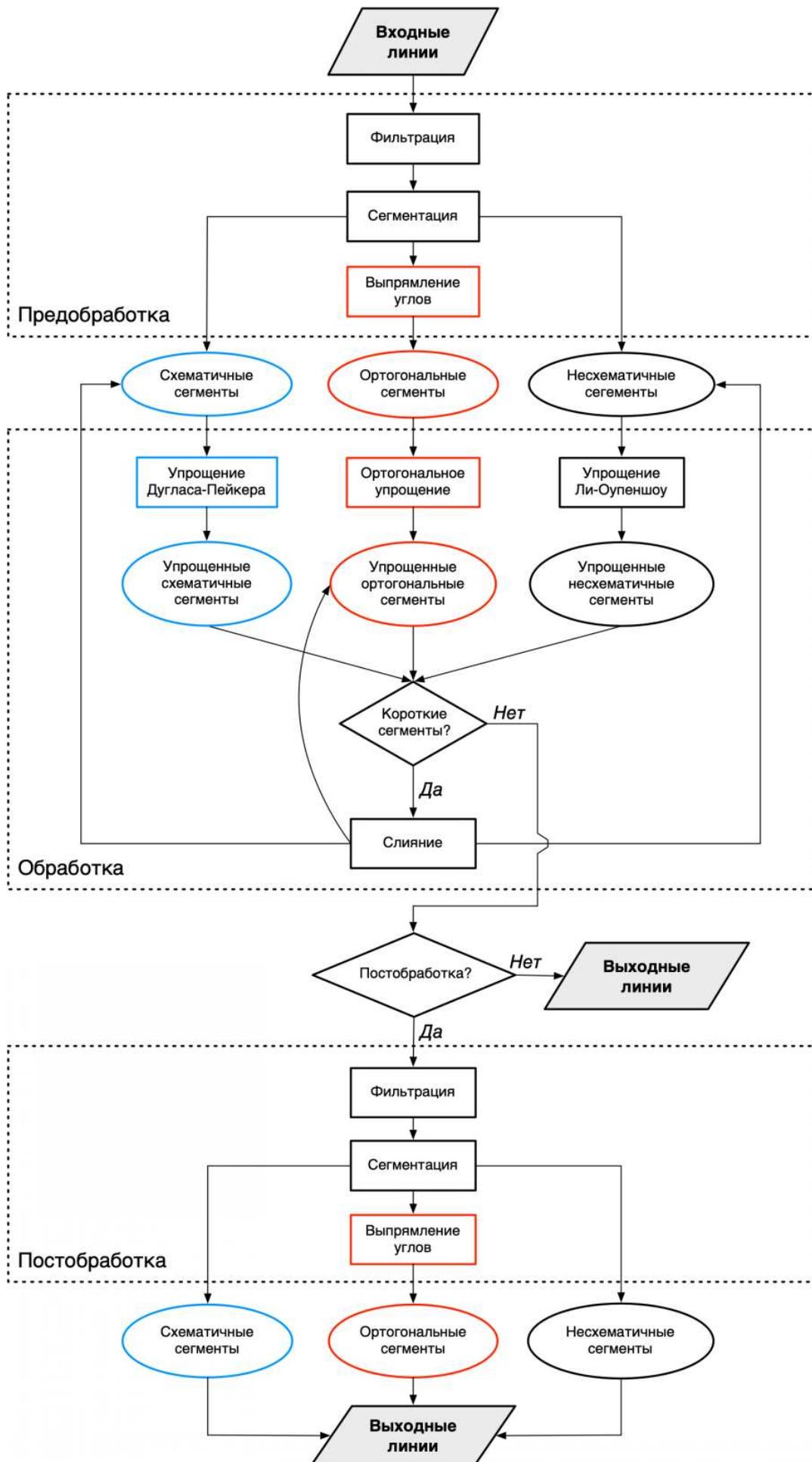


Рисунок 2.33 – Модель геометрического упрощения гетерогенных линий

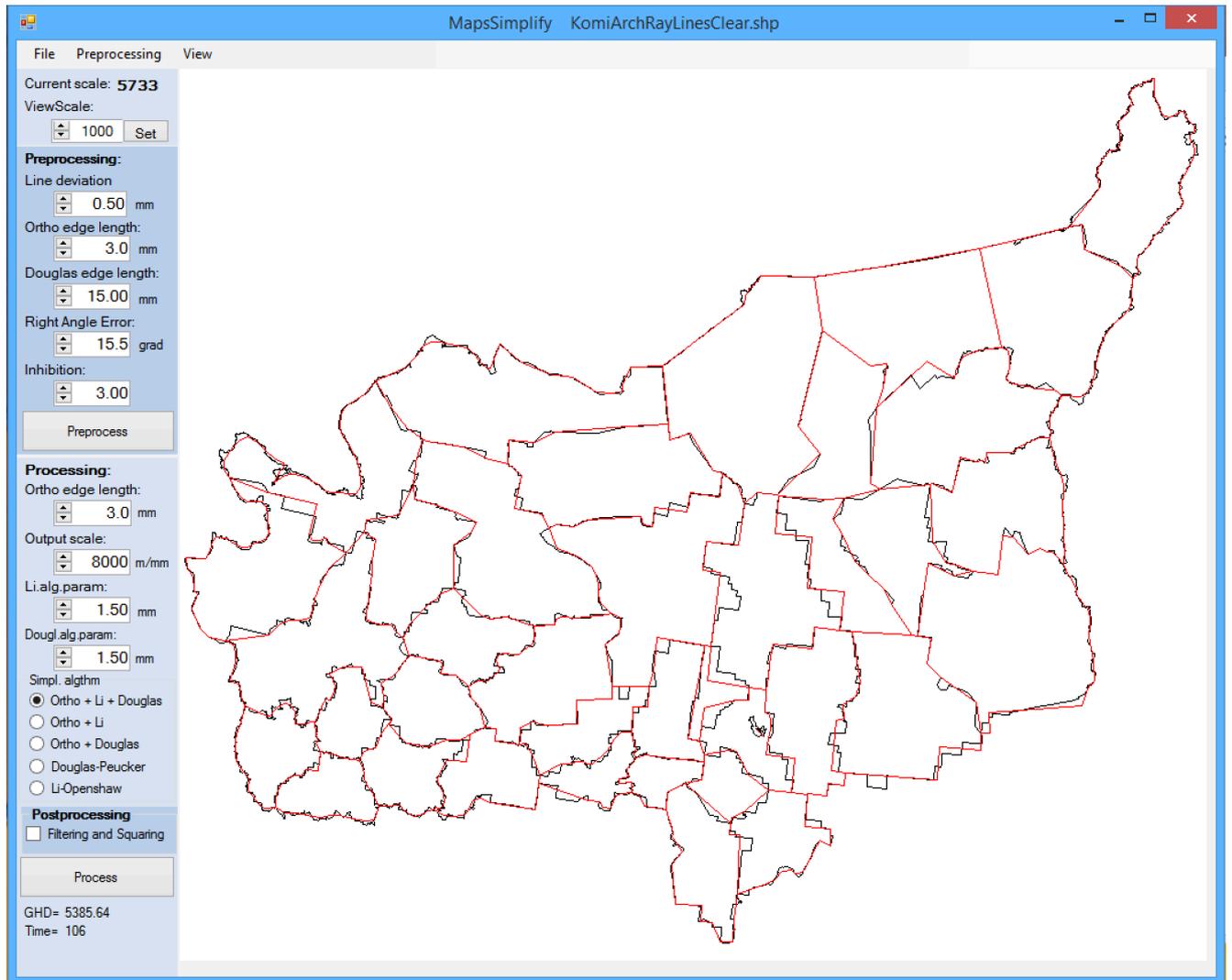


Рисунок 2.34 – Программное приложение MapSimplify 2.0

- численная оценка геометрической точности результатов, полученных различными алгоритмами;
- оценка времени выполнения геометрического упрощения;
- визуальная оценка результатов постобработки.

### 2.4.2.3 Предварительная обработка

Предварительная обработка была выполнена в исходном масштабе данных (1 : 1 000 000) с переменными значениями допусков линейного отклонения  $d$ , углового отклонения  $A$ , минимальной длины ребра ортогонального сегмента  $S$ , минимальной длины ребра схематичного сегмента  $D$  и параметра ингибиции  $k$ .

После предварительной обработки каждая вершина  $p_i$  была атрибутирована значением типа сегмента  $e_i$ , который начинается в  $p_i$ . Пример атрибутирования представлен на Рисунке 2.35. Визуализация типов сегментов по начальным точкам наглядно демонстрирует тот факт, что схематичные сегменты обладают более высокой плотностью точек.

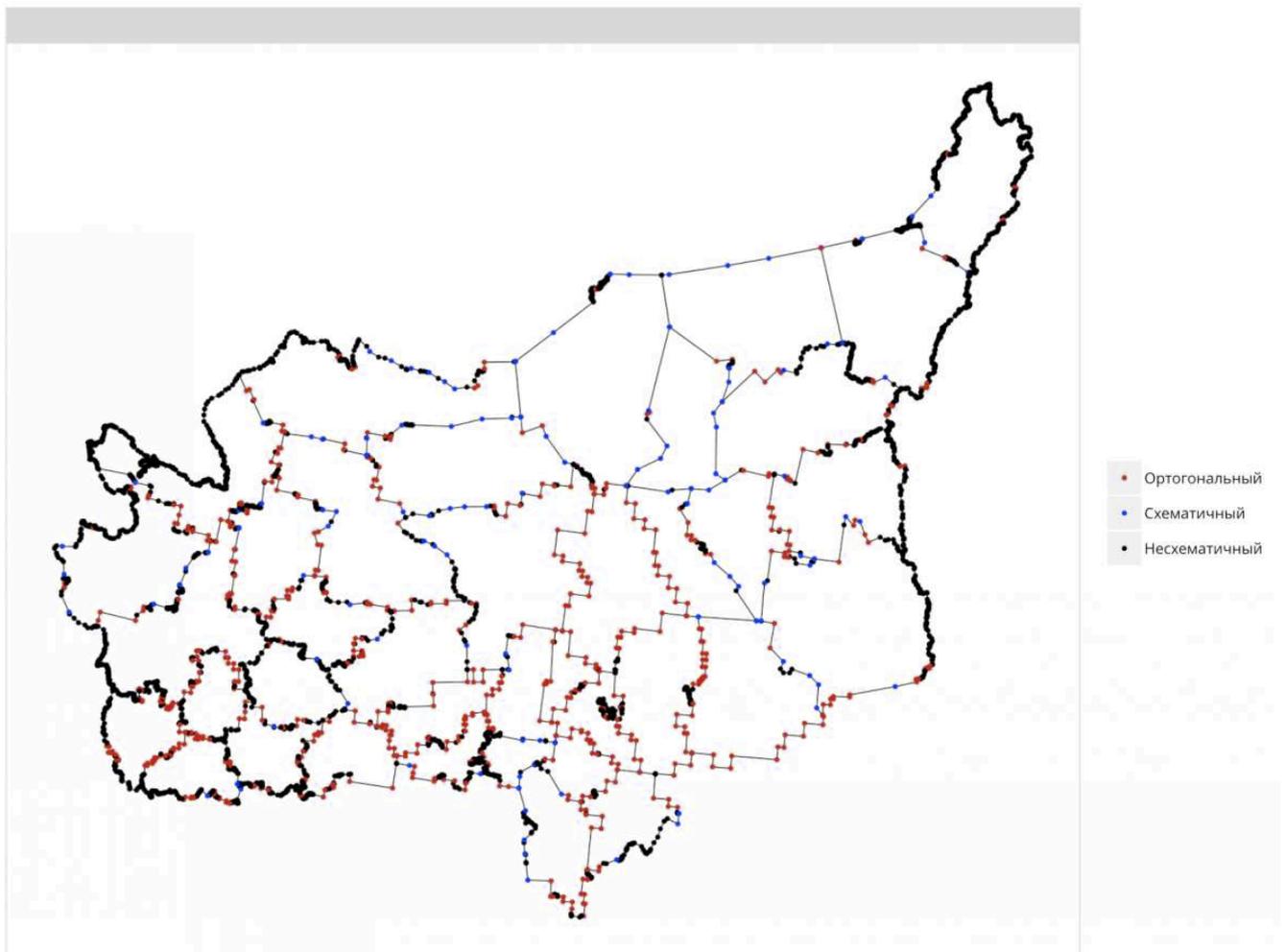


Рисунок 2.35 – Сегментация гетерогенного множества линий. Визуализация по точкам

Итоговая сегментация выполняется по ребрам линий. Результаты, представленные

на Рисунке 2.36 можно сравнить с ручной сегментацией на Рисунке 2.24. Можно заметить, что полученное изображение, особенно на Рисунке 2.36 сверху, близко к ручной интерпретации характера линии. Увеличение допуска на минимальную длину ребра ортогонального сегмента (Рисунок 2.36 снизу) приводит к реклассификации многих мелких ортогональных сегментов в несхематичные, тогда как резкое уменьшение допуска на минимально допустимую длину  $D$  схематичного сегмента приводит к реклассификации многих несхематичных сегментов в схематичные.

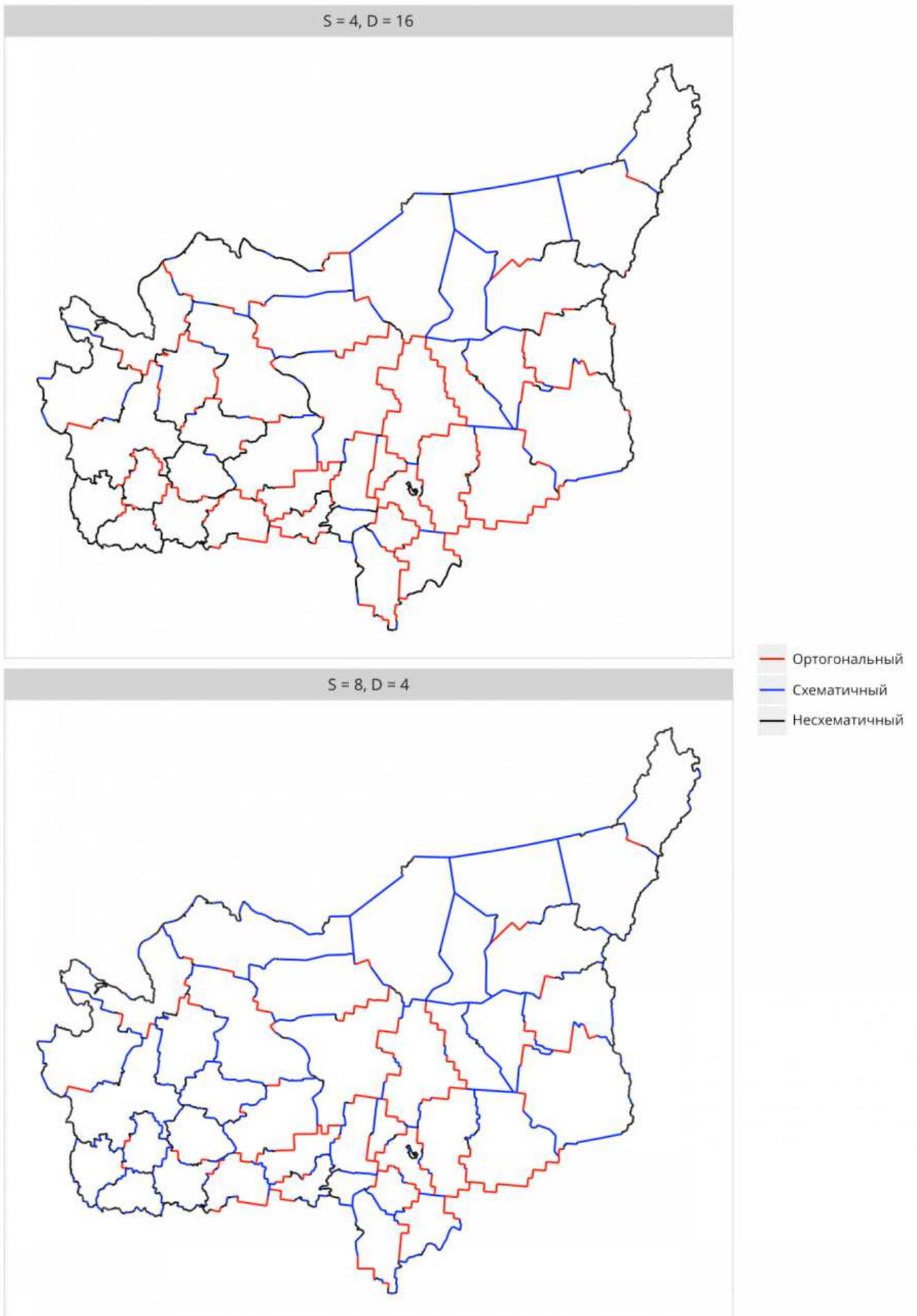
Влияние параметров предварительной обработки показано на Рисунке 2.37. На этом рисунке показаны изменения во времени выполнения и процентном соотношении выделяемых типов линий в результате сегментации при изменении параметров предварительной обработки. Соотношение типов (по суммарной длине) показано вертикальными столбиками под каждым значением параметра, лидирующий тип выделен более темным цветом. *Спарклайны*<sup>8</sup> в правом столбце иллюстрируют тренд в изменении процентной доли и времени выполнения в зависимости от возрастающего значения параметра. Шкала  $y$  аналогична для всех спарклайнов, что позволяет оценить не только тренд, но и его величину.

Увеличение параметра  $d$  приводит к систематическому уменьшению доли несхематичных сегментов. Эта доля в первую очередь поглощается за счет увеличения доли ортогональных сегментов, в меньшей степени — схематичных. Предполагается, что малые значения  $d$  (менее 0,5 мм в масштабе предобработки) эффективно работают в качестве инструмента фильтрации. Более высокие значения параметра приведут к заметному геометрическому упрощению линии, что является нежелательным эффектом на стадии предобработки. Экспертная оценка говорит о том, что значение  $d = 0,25$  мм является оптимальным.

Увеличение пороговой длины ребра ортогональных сегментов  $S$  предсказуемо приводит к уменьшению их доли в общей совокупности линий, тогда как распределение этой доли между схематичными и несхематичными сегментами зависит от значения  $D$ . Ортогональные сегменты переходят в несхематичный класс, если их длина меньше  $D$ . Это означает, что если есть необходимость классифицировать короткие ортогональные сегменты как схематические, необходимо установить  $D$  меньше. Однако малые значения  $D$  могут привести к неправильному обобщению несхематичных сегментов, поэтому выбор параметров должен быть основан на анализе исходных линий (экспертном, либо автоматизированном).

Пороговая длина ребра схематичного сегмента  $D$ , превышающая  $S$ , влияет только на соотношение долей схематичных и несхематичных сегментов, которое уменьшается пропорционально увеличению  $D$ .

<sup>8</sup> *Спарклайны* — термин, введенный Эдвардом Тафти в его книге *Beautiful Evidence* [Tuft, 2006] для обозначения небольших по размеру, но информационно насыщенных графиков. Отличительная особенность спарклайна — его небольшой размер, который должен быть сопоставим с размером печатного слова (*word-sized*).



**Рисунок 2.36** – Сегментация гетерогенного множества линий при разных значениях допуска на минимальную длину ортогонального и схематичного сегмента

Тип сегмента	Параметр предобработки					Тренд
	Отклонение линии ( $d$ )					
	0,125	0,25	0,5	1	2	
Ортогональный	36,5	39,8	45,2	50,3	52,3	
Схематичный	24,7	25,7	26,8	33,4	41,3	
Несхематичный	38,8	34,5	27,9	16,3	6,4	
Соотношение						
Время (мс)	7	8	8	8	8	
Тип сегмента	Длина ортогонального ребра ( $S$ )					Тренд
	0,5	1	2	4	8	
	Ортогональный	46,7	45,6	39,8	32	
Схематичный	23,7	23,9	25,7	28,8	34,8	
Несхематичный	29,6	30,5	34,5	39,2	43,6	
Соотношение						
Время (мс)	10	9	8	6	5	
Тип сегмента	Длина схематичного ребра ( $D$ )					Тренд
	2	4	8	16	32	
	Ортогональный	39,8	39,8	39,8	39,8	
Схематичный	50	36,8	25,7	17,6	10	
Несхематичный	10,2	23,4	34,5	42,6	50,2	
Соотношение						
Время (мс)	8	8	8	8	8	
Тип сегмента	Неопределенность прямого угла ( $A$ )					Тренд
	1	2	4	8	16	
	Ортогональный	5,5	12,9	22,6	32,1	
Схематичный	51,8	45,7	37,9	30,5	25,7	
Несхематичный	42,7	41,4	39,4	37,3	34,5	
Соотношение						
Время (мс)	6	6	7	7	8	
Тип сегмента	Коэффициент ингибиции ( $k$ )					Тренд
	1	1,2	1,5	2	3	
	Ортогональный	27,7	33,8	37,5	39,2	
Схематичный	36	30,9	27,7	26,1	25,7	
Несхематичный	36,3	35,3	34,8	34,7	34,5	
Соотношение						
Время (мс)	7	7	8	8	8	

Рисунок 2.37 – Оценка результатов предобработки

Рост значения углового допуска  $A$  приводит к интенсивному поглощению схематичных сегментов ортогональным классом, тогда как уменьшение доли несхематичных сегментов довольно скромно на этом фоне. Данный результат показывает, что ортогональность в основном типична для простых схематичных сегментов и не является характерной чертой гладких естественных линий (в том виде, в котором они представлены в базах пространственных данных).

Наконец, увеличение параметра ингибиции  $k$  помогает в извлечении большего количества ортогональных сегментов с длинными ребрами. Как и ожидалось, они реклассифицируются в основном из схематичных сегментов, в то время как изменение доли несхематичных сегментов вызывается реклассификацией небольших ребер близких к  $S$ .

Основываясь на данном анализе, пользователь предложенной модели генерализации может осмысленно настраивать параметры для определения наилучших результатов предварительной обработки.

Сложность вычислений различных составляющих предварительной обработки можно оценить, проанализировав итоговое время обработки в нижней строке каждого раздела параметров на Рисунке 2.37. Все параметры можно упорядочить в порядке убывания их влияния на время вычисления следующим образом:  $S \rightarrow A \rightarrow k \rightarrow d \rightarrow D$ . Как правило, обнаружение ортогональных сегментов и выпрямление их углов (которые управляются параметрами  $S$ ,  $A$  и  $k$ ) требуют большого времени обработки. Время обработки увеличивается, если длина ортогонального края становится меньше (это приводит к обнаружению большего количества сегментов). В то же время предварительная обработка также занимает больше времени. Фильтрация и схематический поиск края оказывают незначительное влияние на время обработки. Среднее время предварительной обработки для данного примера составляет 8 мс.

#### 2.4.2.4 Геометрическое упрощение

В процесс анализа результатов геометрического упрощения было выполнено сравнение линий, полученных разработанным методом, с линиями, полученными с применением алгоритмов Дугласа-Пейкера и Ли-Оупеншоу глобально.

Как уже было сказано ранее, этап геометрического упрощения управляется тремя основными параметрами: коэффициентом уменьшения масштаба  $R$ , линейным допуском обобщения  $T$  и результирующей длиной ортогонального сегмента  $S'$ , которая также используется как допуск при выполнении процедуры слияния коротких сегментов. Величина  $T$  действует как мера порогового расстояния для алгоритма Дугласа-Пейкера и как размер ячейки для алгоритма Ли-Оупеншоу.

Мы выполнили обобщение предварительно обработанного набора данных при значениях  $R = 2, 4, 6, 8, 10$ ;  $T = 0, 5, 1, 2$  мм и  $S' = 1, 2, 4$  мм во всех возможных комбинациях. Результаты сначала оценивались визуально, а затем с точки зрения геометриче-

ской точности и времени обработки.

Для каждого подхода были получены четыре цифры:  $R = 2, 4, 8, 16$  для визуального сравнения, иллюстрирующие постепенное упрощение набора данных линии.

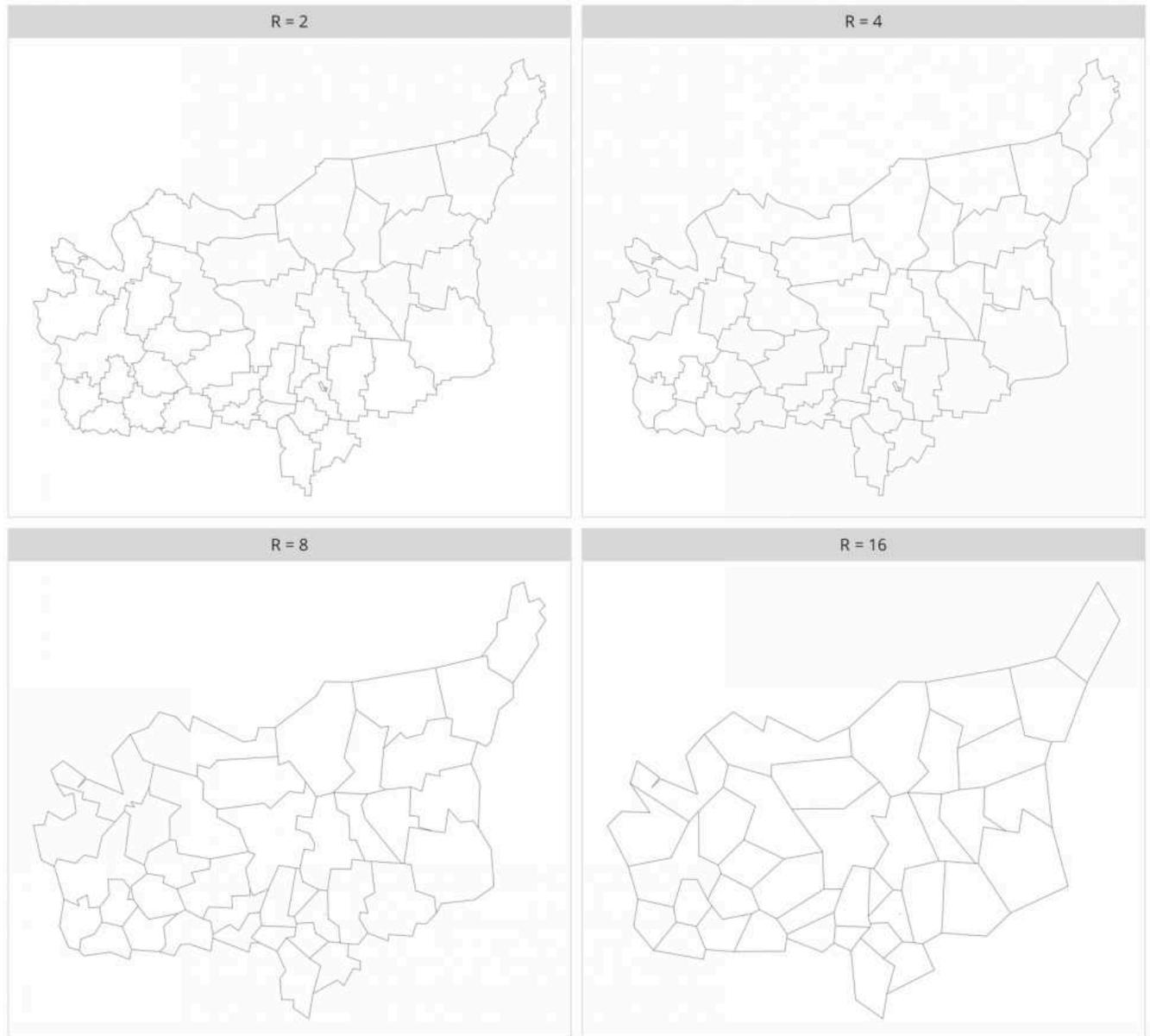
Алгоритм Ли-Оупеншоу (Рисунок 2.38) хорошо работает для несхематичных сегментов, дает гладкий характер линии, который соответствует форме исходной. Последовательное обобщение приводит к более гладким линиям с меньшей детализацией, которые реалистично передают форму природных рубежей, проходящих по рекам и горным хребтам. В то же время, подход Ли-Оупеншоу не сохраняет форму схематичного и ортогонального типа, делая их округлыми. Для сравнения читатель может обратиться к Рисунку 2.24, в котором выделены различные типы сегментов.



**Рисунок 2.38** – Геометрическое упрощение линий алгоритмом Ли-Оупеншоу

Напротив, алгоритм Дугласа-Пейкера (Рисунок 2.39) работает лучше при обработке схематичных сегментов линии, помогает сохранить их резкую и угловатую природу.

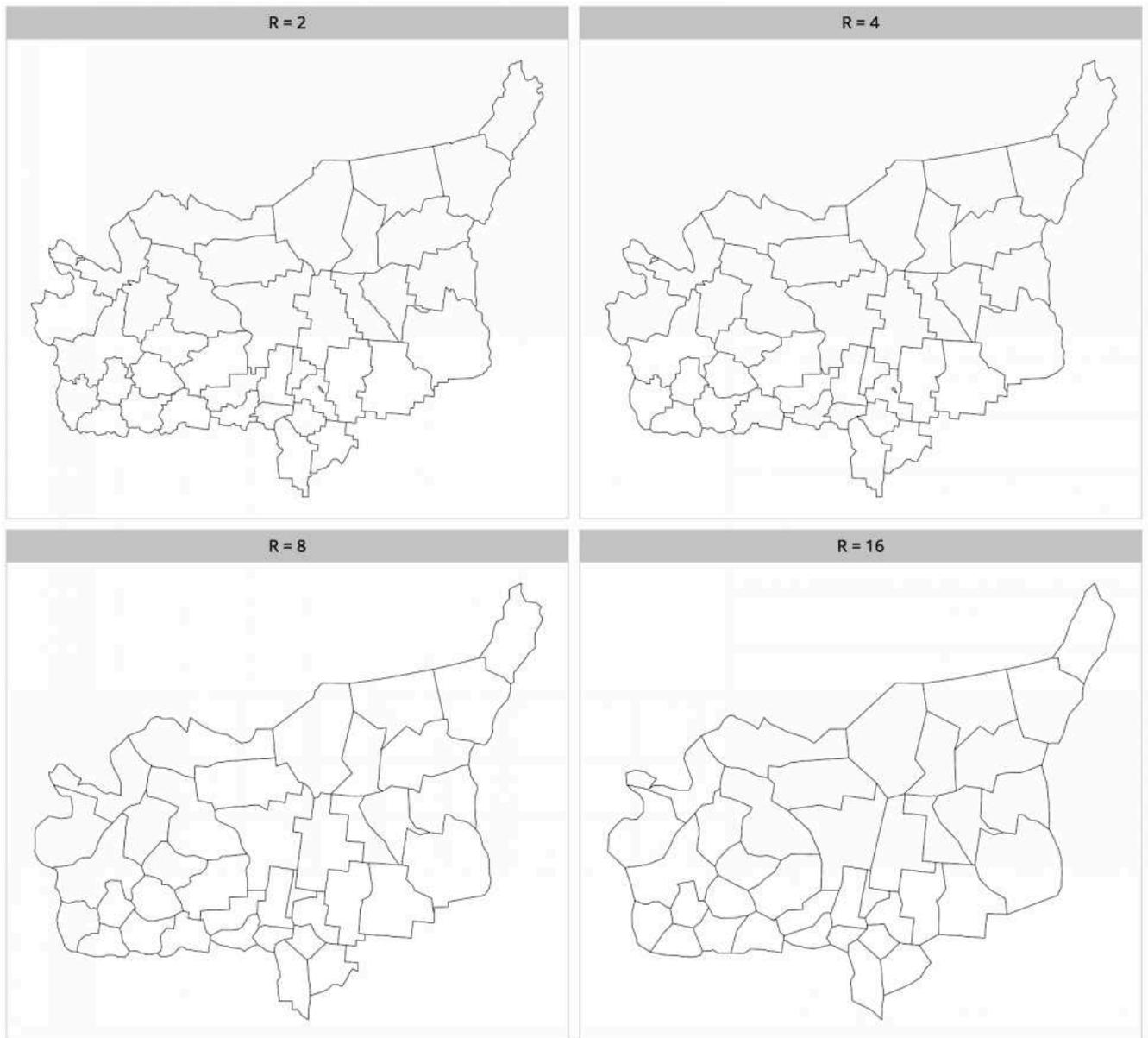
Однако особенности характера линий проявляются, как и в предыдущем случае, повсеместно, что делает несхематичные сегменты линий искусственно угловатыми. Кроме того, алгоритм Дугласа-Пейкера не сохраняет прямых углов, что приводит к искажению формы ортогональных сегментов.



**Рисунок 2.39** – Геометрическое упрощение линий алгоритмом Дугласа-Пейкера

Наконец, Рисунок 2.40 иллюстрирует, что предлагаемая модель обобщения эффективно интегрирует лучшие качества ранее упомянутых алгоритмов, добавляя к ним также специализированный подход для ортогональных сегментов. Подчеркивание ортогонального характера соответствующих частей линии можно рассматривать как процесс типизации (*typification*) ортогонального рисунка, который также является особым алгоритмом генерализации. Несхематичные сегменты обобщаются аналогично тому, как это делается в примере Ли-Опеншоу (Рисунок 2.38), а схематичные сегменты обобщены аналогично тому, как это показано в примере с алгоритмом Дугласа-Пейкера (Рисунок 2.39),

за исключением случаев, когда происходит удаление малых сегментов и слияние линий.



**Рисунок 2.40** – Геометрическое упрощение линий предложенным подходом

Модифицированное расстояние Хаусдорфа ( $d_M$ ), введенное в [Dubuisson, Jain, 1994] и представленное ранее в Разделе 2.3.4, использовано в качестве метрики геометрической точности результатов генерализации. При постановке эксперимента в качестве множеств  $\mathcal{A}$  и  $\mathcal{B}$  в Формуле 2.10 использовались исходные и генерализованные линии соответственно, при этом расстояние вычислялось аналогично модификации для графа от вершин исходной линии к ребрам генерализованной. Среднее значение  $d_M$  для всех линий было использовано как интегральная мера геометрической точности результата генерализации.

Численная оценка пространственной точности представлена на Рисунке 2.41. Вертикальная серия спарклайнов в последнем столбце иллюстрирует, как значение  $d_M$  меняется от алгоритма к алгоритму, тогда как горизонтальная серия иллюстрирует изме-

нения  $d_M$  как функции от степени геометрического упрощения.

Первое наблюдение, которое можно получить из данного рисунка, состоит в том, что  $d_M$  ожидаемо растет с уменьшением масштаба и увеличением параметра  $R$ . При этом наш подход, комбинирующий несколько алгоритмов, характеризуется наименее предсказуемым поведением данного показателя. В некоторых масштабах он имеет аномальные пики, что, скорее всего, является результатом геометрического упрощения ортогональных сегментов, которое управляется не метриками точности, а длиной короткого ребра. Алгоритм Ли-Оупеншоу имеет тенденцию увеличивать  $d_M$ , в то время как Дуглас-Пейкер является в этом плане наиболее предсказуемым алгоритмом, обеспечивая почти линейную зависимость между уменьшением масштаба и геометрической точностью результата.

Второе наблюдение состоит в том, что алгоритмы Ли-Оупеншоу и Дугласа-Пейкера в большинстве случаев приводят к наиболее и наименее точным результатам, соответственно, в то время как предлагаемый метод приводит к промежуточной точности, что можно оценить путем сравнения  $d_M$  для случаев, когда  $S' = T$ .

Время выполнения для разных алгоритмов систематизируется на Рисунке 2.42. В нашей реализации для этого конкретного примера Дуглас-Пейкер оказался самым быстрым алгоритмом с почти постоянным временем выполнения (4 мс). Накладные расходы при поиске самого длинного расстояния на каждой итерации алгоритма Дугласа-Пейкера значительно больше, чем при построении упрощенной линии из выбранных точек. Когда количество начальных точек одинаково для каждого значения параметра, алгоритм выполняет упрощение за практически постоянное время. Фактические же изменения могут быть идентифицированы посредством более тщательного анализа циклов процессора или использования более дробных единиц, чем миллисекунды, однако такой анализ не принесет большой пользы для настоящего исследования.

Алгоритм Ли-Оупеншоу работает быстрее при больших уменьшениях масштаба и крупных разрешениях сетки соответственно. Большая часть времени обработки затрачивается на обрезку линий границами ячеек. В случае значительного уменьшения масштаба количество ячеек сетки стремительно сокращается (если быть точным — пропорционально квадрату степени уменьшения масштаба). Разница во времени обработки по сравнению с алгоритмом Дугласа-Пейкера является поэтому драматичной при небольших уменьшениях масштаба (Ли-Оупеншоу в 9 раз медленнее при  $R = 2$ ,  $T = 0,5$ ); эта разница исчезает в случае значительного уменьшения масштаба, алгоритм Ли-Оупеншоу может работать даже быстрее, чем алгоритм Дугласа-Пейкера (3 мс против 4 мс для  $R = 10$ ,  $T = 2$ ). Таким образом, можно сделать вывод, что *алгоритм Ли-Оупеншоу является вычислительно эффективным при значительной степени обобщения.*

Скорость обработки предложенного подхода зависит также от параметра  $S'$ . Чем больше значение  $S'$ , тем в целом медленнее обработка, поскольку увеличение этого па-

Уменьшение масштаба (R)	Параметр упрощения (l)	Модифицированное расстояние Хаусдорфа, м					Тренд
		Ли-Оупеншоу	Комбинированный метод			Дугласа-Пейкера	
			S' = 1	S' = 2	S' = 4		
2	0,5	266	301	505	905	570	
4		518	740	1811	2277	1222	
6		842	1473	2284	3168	1820	
8		1321	2736	2756	4429	2431	
10		1916	2298	3337	4326	3003	
Тренд							Тренд
2	1	518	553	739	1812	1223	
4		1321	2543	2735	2756	2431	
6		1732	1749	2386	3517	3513	
8		2394	2960	3591	4863	5067	
10		2732	3622	4686	5654	6225	
Тренд							Тренд
2	2	1322	1545	2508	2735	2432	
4		2393	2275	2908	3592	5067	
6		4108	3897	4407	5693	7649	
8		4194	5219	6431	7585	9669	
10		5086	6150	7076	8066	11150	
Тренд							Тренд

Рисунок 2.41 – Оценка геометрической точности генерализации на основе модифицированного расстояния Хаусдорфа ( $d_M$ )

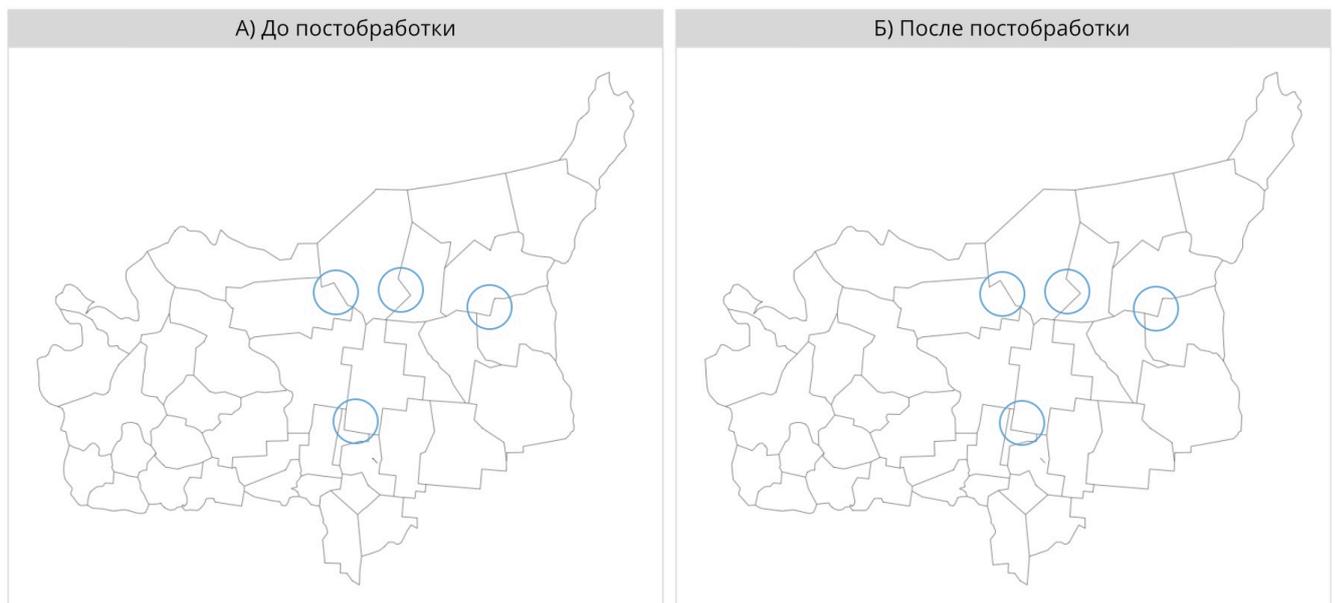
Уменьшение масштаба (R)	Параметр упрощения (L)	Время обработки, мс					Тренд
		Ли-Оупеншоу	Комбинированный метод			Дугласа-Пейкера	
			S' = 1	S' = 2	S' = 4		
2	0,5	326	776	705	638	60	
4		198	406	365	349	54	
6		147	302	275	249	52	
8		121	235	206	191	49	
10		107	191	179	163	45	
Тренд							Тренд
2	1	147	519	411	376	46	
4		100	264	235	214	46	
6		83	193	160	152	45	
8		73	165	135	126	43	
10		70	124	113	109	43	
Тренд							Тренд
2	2	100	423	278	237	56	
4		78	207	175	152	53	
6		68	146	114	108	52	
8		63	121	98	95	51	
10		62	98	89	86	51	
Тренд							Тренд

Рисунок 2.42 – Время обработки различными алгоритмами геометрического упрощения линий

раметра заставляет алгоритм заменять большое число ребер в ортогональных сегментах. Возможны, однако, и некоторые исключения из этого правила ( $S' = 4$  быстрее, чем  $S' = 2$  и даже  $S' = 1$ ). Эти исключения можно отнести на счет ускорения процедуры итеративного слияния, когда ортогональные сегменты значительно уменьшаются в количестве ( $S' = 4$ ). Время обработки разработанного подхода аналогично времени алгоритма Ли-Оупеншоу в случае небольшого обобщения, что означает, что обработка несхематичных сегментов с помощью детальной сетки разбиения требует наибольшего времени. Кроме того, обработка происходит быстрее в случае значительного обобщения и увеличения линейных допусков, что можно объяснить быстрой обработкой несхематичных сегментов алгоритмом Ли-Оупеншоу.

#### 2.4.2.5 Постобработка

Как отмечалось ранее, этап постобработки похож на предварительную обработку и управляется теми же параметрами, но в целевом масштабе, а не в масштабе источника. Пример результата постобработки представлен на Рисунке 2.43. Эффект заключается в том, что линия кажется немного более упрощенной, и почти прямые углы преобразуются в строго прямые (выделены синими кружками на рисунке).



**Рисунок 2.43** – Пример результата после постобработки

Данный этап модели обработки данных может быть полезен для увеличения пространственной упорядоченности и регулярности линий после геометрического упрощения. Недостатком постобработки в ее текущей реализации является то, что она делает линии немного более схематичными и угловатыми. Решение о необходимости постобработки остается за пользователем, поэтому данный этап является в модели необязательным, в отличие от первых двух.

### 2.4.3 Обсуждение

Результаты, представленные в этом разделе, демонстрируют, что с использованием одного интегрированного подхода можно производить геометрическое упрощение линий, в которых смешаны разные пространственные рисунки. Ключом к решению данной проблемы является наблюдение, что каждый алгоритм упрощения наделяет линии определенным характером (рисунком) — сопоставление характера исходной линии со свойствами алгоритма позволяет выбрать наилучший алгоритмический вариант. Полученные результаты позволяют подтвердить данную концепцию. Вместе с тем, представляется важным обсудить несколько технологических вопросов, которые следует учитывать для использования разработанной модели обобщения в ее текущем виде и выполнения будущих исследований.

#### 2.4.3.1 Топологическая корректность

Стадии фильтрации и упрощения разработанной модели обобщения обеспечивают топологическую корректность результатов, их соответствие исходным. Алгоритмы фильтрации, ортогонального упрощения и Дугласа-Пейкера обеспечиваются дополнительными процедурами, которые идентифицируют вновь образованные пересечения линий «на лету» и корректируют выполнение таким образом, который позволяет избежать таких ошибок. Алгоритм Ли-Оупеншоу не может по своему принципу привести к топологическим ошибкам. Следовательно, предлагаемую модель генерализации можно считать обеспечивающей топологическую корректность результатов (по отношению к исходным пространственным данным).

#### 2.4.3.2 Ограничения по масштабам

Предложенная модель генерализации была разработана для мелких масштабов картографирования. Следует отметить, что возможный масштабный диапазон обобщения зависит от пространственной протяженности исходных линий. *Чем больше длина обрабатываемой линии, тем шире диапазон масштабов, на котором предложенная модель может использоваться для геометрического упрощения.* Если линия слишком мала, чтобы быть представленной в целевом масштабе, она должна быть удалена или свернута до точки в зависимости от целей обобщения. Отбор линий также можно использовать для устранения перегруженности в результирующем масштабе.

#### 2.4.3.3 Ограничения по набору данных

Выбранная комбинация характеристик линий (ортогональные, схематичные, несхематичные) типична для административных границ. Хотя другой тип данных может состоять из линий разного характера, общий рабочий процесс, описанный в данном разделе,

останется прежним. Более того, конкретные операции предобработки и обработки могут быть дополнены алгоритмами, необходимыми для правильного упрощения каждого нового типа линий.

#### 2.4.3.4 Выбор параметров

Разработанная модель генерализации инициализируется с использованием множества входных параметров, которые должны быть адаптированы к исходному набору данных и требуемой степени обобщения. Предварительная обработка контролируется пятью геометрическими критериями и масштабом анализа. Для обработки требуются три геометрических параметра и выходной масштаб. Наконец, возможна также постобработка, которая инициализируется параметрами предварительной обработки. Необходимость установки соответствующих значений приводит к 10 параметрам, которые должны быть освоены пользователем. Хотя нам удалось подобрать подходящие значения для многих из них, эти значения специфичны для конкретного набора данных и носят экспертный характер. Поэтому *важным направлением будущих исследований является разработка методики и технологии автоматического выбора большинства параметров генерализации*, по крайней мере, для инициирования процесса настройки модели. Этот выбор может быть реализован путем углубленного анализа характеристик формы линии.

#### 2.4.3.5 Прочие аспекты

Существуют также три проблемы низкого уровня с представленным подходом, связанные с объединением сегментов, точностью ортогонального упрощения и обработкой регулярностей, отличных от ортогональности.

*Проблема объединения* возникает, когда соседние сегменты соединяются под острым углом, что приводит к резкому изменению формы. Это не проблема в случае, когда данные изменения являются свойствами исходного набора данных. Однако, если эти резкости производятся процедурами упрощения (особенно ортогональными) вблизи стыков сегментов, их следует избегать или разрешать. *Проблема точности* требует от нас применять более сложный подход к обобщению ортогональных сегментов, основанный не только на замене малых ребер, но и применении метрических критериев (того же  $d_M$ ), позволяющих подобрать конфигурацию, наиболее точно аппроксимирующую исходную форму. *Проблема регулярности* касается возможности обобщения других типов линий, например, построенных на основе дуг окружностей или углов, кратных величине, отличной от  $90^\circ$ . Этот вопрос актуален преимущественно для крупномасштабного картографирования.

Наконец, необходимо больше экспериментальной работы в широком диапазоне масштабов, чтобы выявить преимущества и недостатки представленного подхода обобщения и дать оценку его применимости в разных географических условиях. В качестве

иллюстрации того, что можно получить удовлетворительные результаты для других территорий, была произведена генерализация границ графств для штата Монтана, США, обладающих схожим сочетанием характеров линий (Рисунок 2.44). Читатель может видеть, что эти характеры сохраняются, за исключением случаев, когда степень уменьшения масштаба (в 8 раз) требует диагональной замены и слияния ортогональных сегментов.

#### 2.4.4 Итоги

В данном параграфе были представлены методика и модель генерализации для геометрического упрощения гетерогенных линейных наборов данных, которая реализует принцип *адаптивности к форме пространственных объектов*. Гетерогенный линейный набор данных состоит из линий, которые имеют разный генезис (как правило, естественные и искусственные линии) и поэтому часто имеют контрастирующий рисунок (характер) линий. Такие наборы данных достаточно трудны для обобщения, поскольку каждый алгоритм генерализации наделяет линии скульптурой в соответствии с заложенными в него процедурами. Подобная гармонизация морфометрических характеристик неизбежно приводит к необходимости применения различных подходов к геометрическому упрощению разнородных линий и их частей.

Поскольку характер линии является качественной, а не количественной характеристикой, целесообразно ввести классификацию, позволяющую систематизировать возможные варианты изменения формы. Предложена классификация, основанная на критериях схематичности, гладкости и регулярности. В настоящей работе были исследованы случаи гладких нерегулярных несхематичных линий, угловатых нерегулярных схематичных линий и ортогональных линий (случай угловатых регулярных схематичных линий). Смешение данных характеров линий типично для административных границ, все они могут сосуществовать в одном наборе данных.

Разработанный подход основан на автоматической сегментации линии с выделением сегментов вышеупомянутых типов и применении соответствующих алгоритмов упрощения к сегментам по отдельности. Алгоритм Ли-Оупеншоу использован для несхематичных сегментов, для схематичных сегментов использован алгоритм Дугласа-Пейкера, а для ортогональных сегментов был разработан специальный алгоритм упрощения. Все части метода реализованы в модели генерализации, которая состоит из этапов предварительной обработки, обработки и постобработки. Результаты демонстрируют эффективность подхода в сохранении характера линии при обобщении.

Многое в данной теме, однако, остается неисследованным. Наиболее важным направлением исследований видится внедрение распознавания других характеров линий и видов регулярности. Также необходимы экспериментальные исследования в более широком диапазоне масштабов, особенно в крупных картографических масштабах. Кроме



**Рисунок 2.44** – Геометрическое упрощение границ графств штата Монтана, США

того, параллельная обработка должна быть реализована для всех возможных элементов разработанной модели обобщения, что потребует специальной переработки конкретных этапов. Подобное развитие может заложить основу для интеллектуальных экспертных систем, которые выбирают наиболее подходящий алгоритм для обобщения входного набора данных или его морфологически однородных сегментов.

## 2.5 Адаптивная генерализация цифровых моделей рельефа с сохранением структурных линий<sup>9</sup>

**Пространственная структура** относится к числу важнейших характеристик географических явлений сплошного распространения. К таким явлениям относится рельеф земной поверхности, который принято подразделять на формы различного иерархического порядка [Симонов, 1998]. Анализ алгоритмов генерализации ЦМР в Главе 1 показал, что ни один из них не удовлетворяет полностью географическим принципам обобщения рельефа. Многие алгоритмы реализованы в виде авторских программ, и, как правило, доступны только непосредственным разработчикам, что препятствует их использованию в научной среде. Также алгоритмы нуждаются в их разностороннем тестировании на предмет выполнения географических принципов генерализации. В целях сохранения основных структурных линий и форм рельефа в процессе обобщения был разработан специализированный метод генерализации цифровых моделей рельефа [Лурье, Самсонов, 2010; Samsonov, 2011].

### 2.5.1 Требования

Предварительно, с опорой на принципы гипсометрического картографирования [Заруцкая, 1958] и общие требования к генерализации ЦМР [Weibel, 1987] были выдвинуты следующие требования к алгоритму:

- эффективность в средних и мелких масштабах картографирования;
- возможность удалять мелкие формы рельефа и оставлять крупные, объединяя с ними удаленные формы;
- возможность использовать объекты гидрографии (линейные и площадные) в качестве опорного каркаса;
- сохранение положения и характера (резкости) структурных и водораздельных линий, актуальных в целевом масштабе картографирования;

<sup>9</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Лурье, Самсонов, 2010; Samsonov, 2011]

- сохранение общего профиля поверхности, характерных перегибов, поперечных и продольных профилей долин;
- возможность расширения долин и междуречных пространств для улучшения читаемости горизонталей;
- возможность реализации на базе стандартных средств ГИС.

Многие предложенные ранее методы генерализации ЦМР содержат полезные идеи и алгоритмы, эффективно, на наш взгляд, решающие одну из подзадач генерализации. Авторская методика опирается на следующие из них:

- выделение тальвегов (потенциальных водотоков) алгоритмом *D8* и построение водоразделов для них [O'Callaghan, Mark, 1984];
- генерализация сети тальвегов методом подъема в направлении минимальной положительной разности аккумуляции стока [Leonowicz, Jenny, Hurni, 2009];
- заполнение отрицательных форм рельефа путем построения триангуляции по водораздельным линиям [Jordan, 2007].

## 2.5.2 Морфологическое упрощение поверхности

Общая схема алгоритма представлена на Рисунке 2.45. Первоначально на основе ЦМР выделяются тальвеги, значимые в исходном масштабе, которые затем генерализуются в соответствии с целевым масштабом. Для основных тальвегов строятся их водоразделы. Далее отбираются те тальвеги, которые непосредственно впадают<sup>10</sup> в основные, и для них также строятся водоразделы (второго порядка). На заключительном этапе поверхность восстанавливается путем триангуляции основных водоразделов и тальвегов, а также водоразделов второго порядка, которые «заполняют» малые отрицательные формы.

Похожие алгоритмы предлагались и ранее [Jordan, 2007; Ai, Li, 2010]. Основное отличие разработанного алгоритма заключается в том, что он не использует классификацию тальвегов с присвоением им порядков, а вместо этого генерализует их на основе критерия суммарной длины, как это сделано в работе [Leonowicz, Jenny, Hurni, 2009]. Из каждой ячейки растра вверх по склону прослеживается полная линия тока вплоть до истока тальвега 1-го порядка (в точках слияния тальвегов алгоритм идет в сторону тальвега с большей аккумуляцией стока<sup>11</sup>). Это дает более достоверный результат генерализации,

<sup>10</sup>С физической точки зрения это не совсем точно, поскольку впадать может только водоток, а тальveg лишь задает потенциальное местоположение для водотока (который может как быть, так и не быть в этом месте). Однако любой тальveg характеризуется уклоном и направлением, а сеть тальвегов образует структуру из соподчиненных по направлению элементов, что естественным образом отражает структуру гидрографической сети. Поэтому в рамках данной работы считается допустимым применение понятия *впадения* к тальвегам.

<sup>11</sup>В геоинформатике термин «аккумуляция стока» (*flow accumulation*) устойчиво закрепился за показателем, который на самом деле обозначает площадь поверхностного (орографического) водосбора, выраженную в количестве дренируемых ячеек ЦМР или их суммарной площади. Термин не вполне удачный, однако его ис-

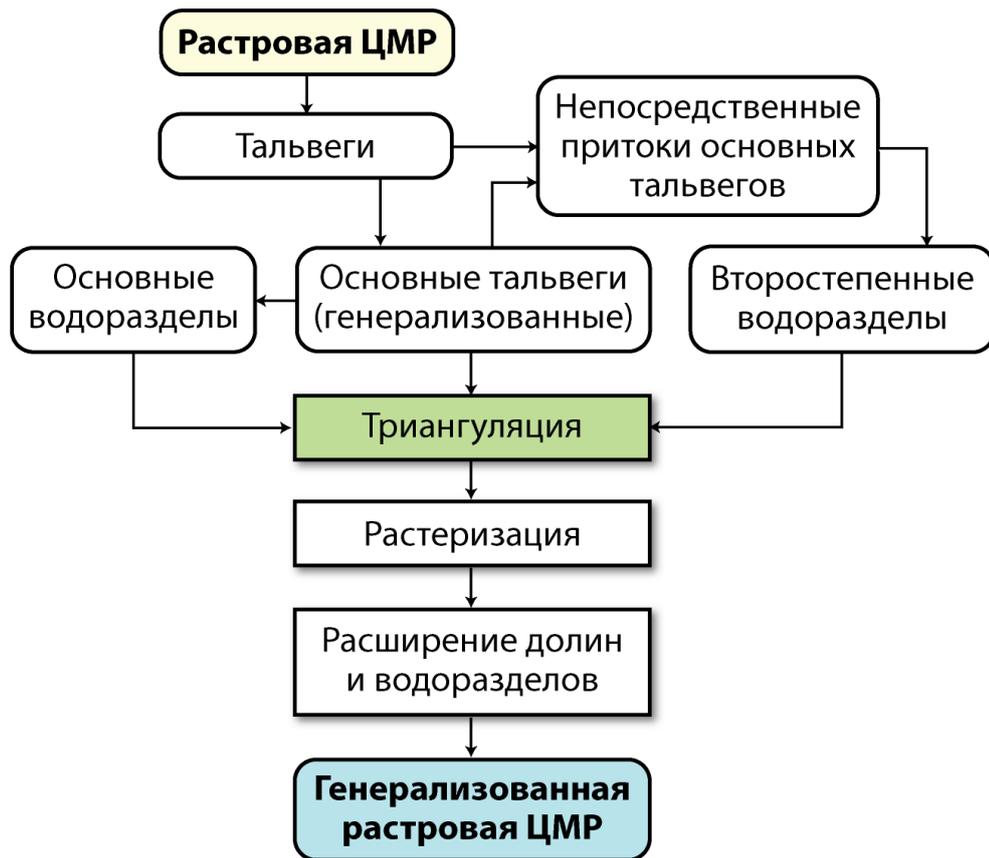


Рисунок 2.45 – Общая схема авторского алгоритма генерализации ЦМР

т.к. удаление тальвегов на основе их порядка может привести к исключению значимых отрицательных форм с низким порядком тальвега, а также исключению коротких истоков рек.

Основным исходным параметром алгоритма является минимальный значимый размер отрицательной формы расчленения, задаваемый длиной ее тальвега и величиной накопления тока в начале тальвега (точке истока). Длина тальвега определяется исходя из целевого масштаба картографирования. В частности, И. П. Заруцкая и Т. Г. Сваткова рекомендуют для форм эрозионного расчленения использовать ценз 3 мм [Заруцкая, Сваткова, 1982]. Для электронных карт ввиду их худшей читаемости имеет смысл увеличить ценз в 1,5 – 2 раза, установив его равным, например, 5 мм.

Если длина тальвега зависит от целевого масштаба, то пороговая величина аккумуляции стока зависит от степени обобщения. Чем меньше минимальный порог аккумуляции, тем выше по склону будет расположено начало тальвега. Величина подходящей аккумуляции стока зависит от характера исходной ЦМР, степени уменьшения, морфологического типа рельефа и устанавливается экспериментальным путем.

пользование в целом не противоречит гидрологической терминологии, поскольку там по отношению к стоку используются такие понятия как *объем, слой, модуль и коэффициент* [Михайлов, Добролюбов, 2017]. Таким образом, при упоминании *аккумуляции* стока читателю должно быть понятно, что речь идет о геоинформационном показателе, отражающем потенциал накопления стока ячейкой ЦМР и выраженном в виде площади ее поверхностного водосбора.

Каждому выделенному тальвегу можно поставить в соответствие бассейн его устьевой точки. Границей бассейна является линия водораздела. Тальвег и водораздел представляются в виде трехмерных линий, имеющих в каждой точке не только плоские координаты, но и высоту. Это позволяет восстанавливать на основе линий тальвегов и водоразделов трехмерный каркас поверхности.

Очевидно, что если выделить по исходной ЦМР тальвеги, построить водоразделы между ними, а затем использовать все эти линии в качестве ребер триангуляции для восстановления поверхности, исказится профиль склонов: все долины станут V-образными, а положительные формы приобретут острые вершины. Следовательно, требуется добавить точки и линии, которые будут характеризовать профиль поверхности на склонах выделенных бассейнов.

В процессе поиска решения этой проблемы была выдвинута гипотеза о том, что общий профиль склона может быть восстановлен по водораздельным линиям врезанных в него отрицательных форм рельефа. Эта идея органически переплетается с задачей удаления мелких форм при обобщении рельефа.

Решение состоит из двух этапов: сначала выделяется сеть тальвегов исходя из ценза аккумуляции тока, но без учета их длины. При этом можно установить ценз немного меньше того, что использовался для выделения основных тальвегов. Это позволит получить водоразделы в верховьях основных тальвегов и тем самым сохранить характер профиля ложбины в районе водосборной воронки. Полученная сеть содержит как основные, так и удаляемые тальвеги. После чего путем оверлея из полученного слоя удаляются основные тальвеги. В результате получается слой тальвегов, которые необходимо вместе с их бассейнами удалить из ЦМР.

Выделенная сеть второстепенных тальвегов (назовем их тальвегами-притоками) может иметь несколько порядков вложенности. В нашей задаче интерес представляют только те тальвеги, которые непосредственно впадают в основные (сохраняемые), т.е. являются притоками первого порядка. Это связано с тем, что водораздельная линия по своей природе всегда начинается у пересечения двух тальвегов. Если построить водоразделы для притоков второго порядка, они будут включать точки рядом с линией притока первого порядка, в то время как эта линия должна быть исключена из ЦМР. Это приведет к тому, что на склоне проявятся ненужные ложбины. То же самое можно сказать о притоках третьего и далее порядков.

Чтобы отобрать притоки первого порядка, необходимо найти те из них, устьевые точки которых расположены на основных тальвегах. Для этого можно построить вокруг основных тальвегов буферную зону, ширина которой равна разрешению исходной ЦМР, и далее отобрать с помощью нее устьевые точки. Для устьевых точек притоков первого порядка строятся бассейны, которые и дают искомые водораздельные линии для восстановления профиля склона.

Еще одна гипотеза заключается в том, что данный алгоритм должен обеспечивать

косвенную генерализацию положительных форм. Крупные положительные формы рельефа обеспечивают значительную аккумуляцию стока на своих склонах. Из этого следует, что при выделении полной сети тальвегов на основе аккумуляции стока без учета их длины склоны значимых положительных форм будут покрыты водораздельными линиями первого и второго порядка, которые и обеспечат каркас для их генерализации. Если склоны формы рельефа оказались не охвачены сетью водораздельных линий, это означает, что данная форма является незначительной при заданном допуске аккумуляции стока и при генерализации ее можно полностью удалить.

В результате работы алгоритма по исходной ЦМР должны быть выделены (Рисунок 2.46):

1. Основные тальвеги, сохраняемые в масштабе картографирования.
2. Водораздельные линии между основными тальвегами.
3. Водораздельные линии между удаляемыми тальвегами.

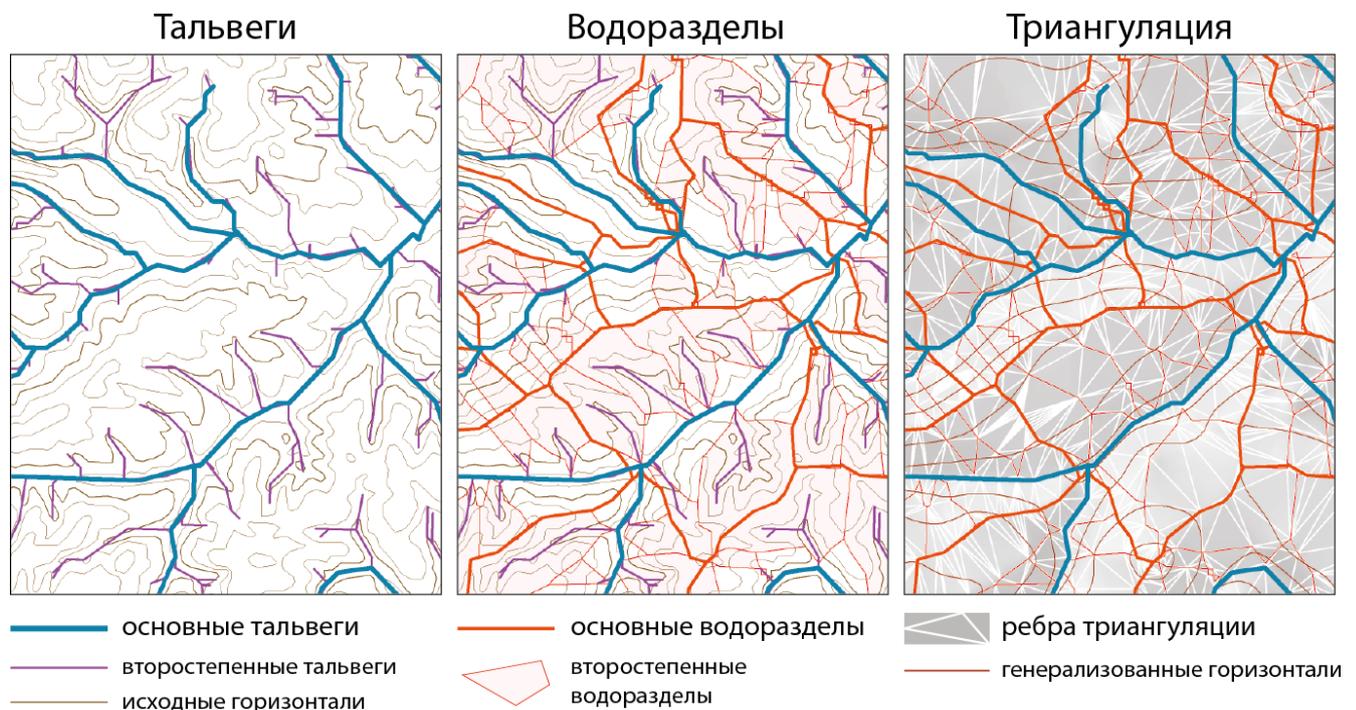


Рисунок 2.46 – Построение триангуляции на основе структурных линий при генерализации ЦМР

### 2.5.3 Постобработка ЦМР

Поскольку местоположение каркасных линий (тальвегов, водоразделов) при генерализации не меняется, днища долин перестают выражаться в масштабе, что приводит к появлению вдоль тальвегов разорванных горизонталей. Для устранения этого эффекта полученную модель необходимо подвергать постобработке в целях расширения долин и междуречий, что по своему действию аналогично сдвигам горизонталей, используемым

при рукописном обобщении рельефа [Заруцкая, 1958]. Сдвиги горизонталей призваны улучшить читаемость форм рельефа за счет их преувеличения.

Разработанный метод расширения долин и междуречий опирается на усовершенствованный вариант алгоритма генерализации ЦМР Леонович-Йенни [Leonowicz, Jenny, Hurni, 2009]. В данном алгоритме генерализация поверхности достигается путем преувеличения крупных долин и междуречий, в ходе которого крупными формами рельефа поглощаются осложняющие их формы рельефа более низкого порядка. Это достигается путем применения следующего алгоритма:

1. Выделение сети тальвегов по цифровой модели.
2. Генерализация сети тальвегов с сохранением наиболее значимых. Построение серии буферных зон вокруг тальвегов (в примере использовалось 5 буферных зон с шагом расстояния равным разрешению ЦМР).
3. Сглаживание исходной ЦМР с помощью фильтра нижней квантили. Результатом является значение, расположенное на уровне 25% (меньше медианы) среди значений в пределах плавающего окна. На полученной ЦМР-1 преувеличены (углублены и расширены) отрицательные формы рельефа, срезаны положительные формы.
4. Сглаживание исходной ЦМР с помощью фильтра верхней квантили. Результатом является значение, расположенное на уровне 75% (больше медианы) среди значений в пределах плавающего окна. На полученной ЦМР-2 преувеличены (приподняты и расширены) положительные формы рельефа, заполнены отрицательные формы.
5. Взвешенный оверлей сглаженных моделей на основе буферных зон вокруг тальвегов. В пределах первой буферной зоны используются значения ЦМР-1, за пределами последней буферной зоны используются значения ЦМР-2. В промежуточных буферных зонах происходит взвешенное осреднение высот (переходная зона).

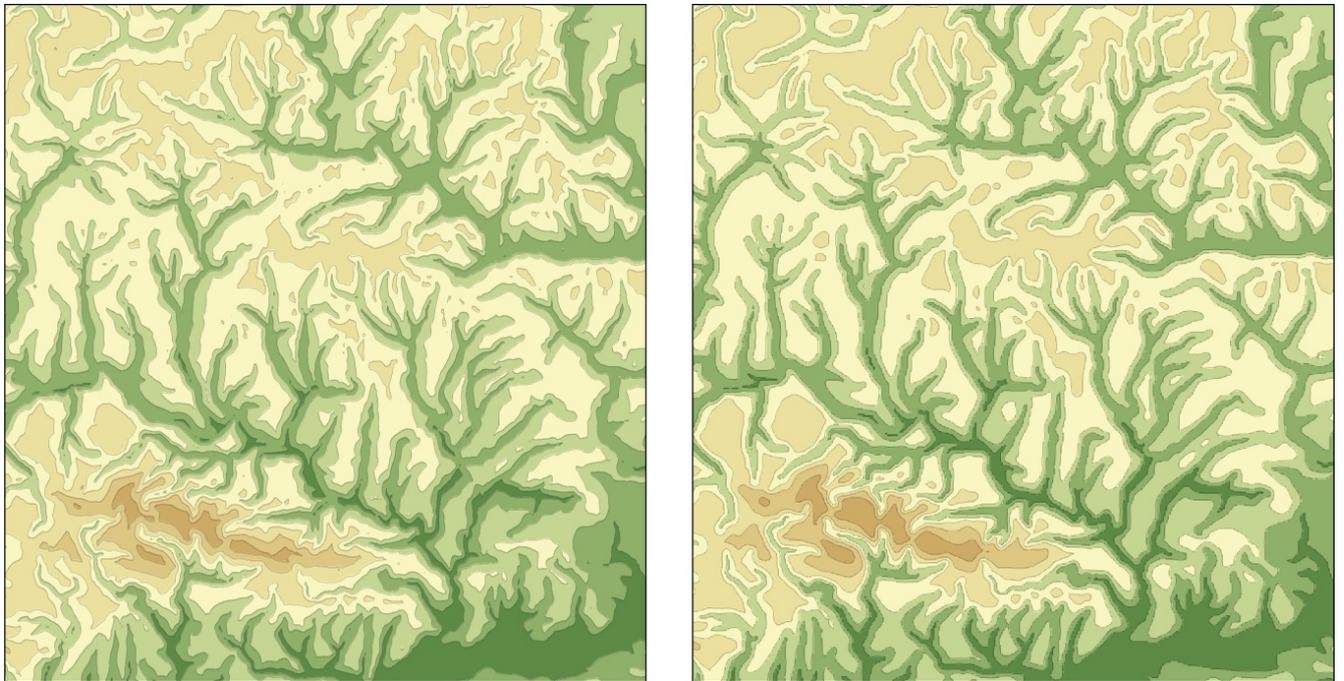
Алгоритм обладает двойным действием: он приводит к удалению малых форм рельефа и преувеличению крупных форм. Однако, поскольку данный метод генерализации основан на итеративной фильтрации, он не гарантирует выполнение основного условия мелкомасштабного картографирования — удаления или сохранения (преувеличения) формы рельефа целиком, а не частично. Удаляются только формы рельефа, которые достаточно мелки, чтобы быть полностью срезаны/заполнены за 7 — 10 итераций сглаживающих фильтров. При этом отрицательные формы рельефа, достаточно крупные в размере, но не учтенные при выбранном допуске выделения тальвегов, оказываются заполненными только отчасти, что приводит к рассогласованию рисунка соседних горизонталей.

В то же время, остается возможность использовать второе свойство алгоритма —

преувеличивать крупные формы, что и используется в нашей методике. Для этих целей алгоритм был усовершенствован, путем использования:

- исходной модели для участков, соответствующих склонам;
- растра евклидовых расстояний вместо векторных буферных зон для вычисления весов;
- фильтров минимума и максимума вместо квантильных для усиления эффекта расширения.

Если при большом количестве итераций (7-10) обработки ЦМР с помощью фильтров нижней и верхней квантили происходит ее генерализация, то при нескольких итерациях (1-3) усовершенствованный алгоритм дает расширение долин и междуречий (Рисунок 2.47).



**Рисунок 2.47** – Расширение долин и междуречий. Слева – исходная ЦМР, справа – ЦМР после обработки специальным алгоритмом

Поясним принцип расширения на примере, представленном на Рисунке 2.48. Исходный поперечный профиль речной долины показан на Рисунке 2.48б прямой пунктирной линией. В процессе расширения профиль приобретает выпукло-вогнутую форму, в которой вогнутый участок соответствует расширенному днищу долины, а выпуклый – расширенной поверхности междуречья. При этом поверхность склона сокращается и становится более крутой. Точка перегиба определяется расстоянием  $D$ , которое отсчитывается от водотока (талвега) данной долины. Это расстояние на местности зависит от масштаба картографирования, а в единицах измерения карты определяется требованиями и допусками на минимальный поперечный размер формы рельефа. Например, если составитель карты желает, чтобы формы рельефа были в ширину не менее 3 мм, следует

установить  $D = 3$  мм. Поскольку расстояние  $D$  откладывается в обе стороны от водотока, это обеспечит гарантированную ширину днища долины, а также плавный переход к поверхности склона. Во всех точках, располагающихся на расстоянии  $D$  высоты поверхности сохраняются, а на более дальних расстояниях высоты преувеличиваются в сторону междуречий.

Эффект трансформации профилей форм рельефа достигается путем вычисления в каждой ячейке ЦМР средневзвешенной суммы трех поверхностей:

1.  $Z_0$  — генерализованная поверхность.
2.  $Z_{val}$  — с преувеличенными отрицательными и срезанными положительными формами рельефа, полученная обработкой  $Z_0$  фильтром минимума.
3.  $Z_{ridge}$  — с преувеличенными положительными и заполненными отрицательными формами рельефа, полученная обработкой  $Z_0$  фильтром максимума.

В каждой ячейке ЦМР этим поверхностям соответствуют веса  $W_{val}$ ,  $W_{ridge}$  и  $W_0$  соответственно, подчиняющиеся закону  $W_{val} + W_{ridge} + W_0 = 1$ . Результирующее значение в ячейке получается путем выпуклой линейной комбинации значений  $Z$  с соответствующими весами:

$$\tilde{Z} = Z_{val}W_{val} + Z_{ridge}W_{ridge} + Z_0W_0 \quad (2.16)$$

Поскольку все варианты  $Z$  — это растровые модели, это означает, что и веса  $W$  должны быть упорядочены в виде растра той же размерности. Их вычисление осуществляется следующим образом:

1. Для тальвегов, сохраненных при генерализации ЦМР, строится растр евклидовых расстояний  $E$  с размером ячейки, таким же как и у  $Z_0$ .
2. Поверхность  $W_{val}$  получается из последовательности выражений алгебры карт:

$$W^- = (D - E)/D, \quad (2.17)$$

$$W_{val} = \max [0, W^-]. \quad (2.18)$$

3. Поверхность  $W_{ridge}$  получается из последовательности выражений алгебры карт:

$$W^+ = \max [0, -W^-], \quad (2.19)$$

$$W_{ridge} = \min [1, W^+]. \quad (2.20)$$

4. Поверхность  $W_0$  получается выражением алгебры карт:

$$W_0 = 1 - W_{val} - W_{ridge}. \quad (2.21)$$

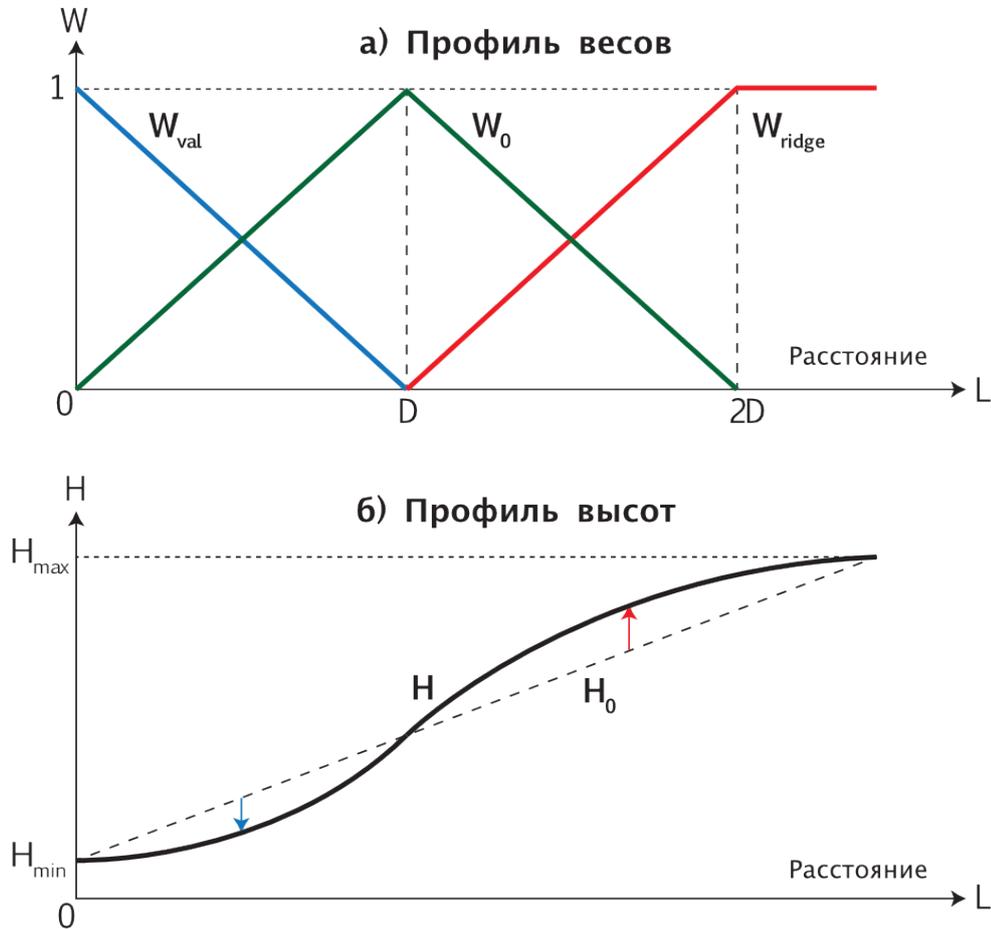
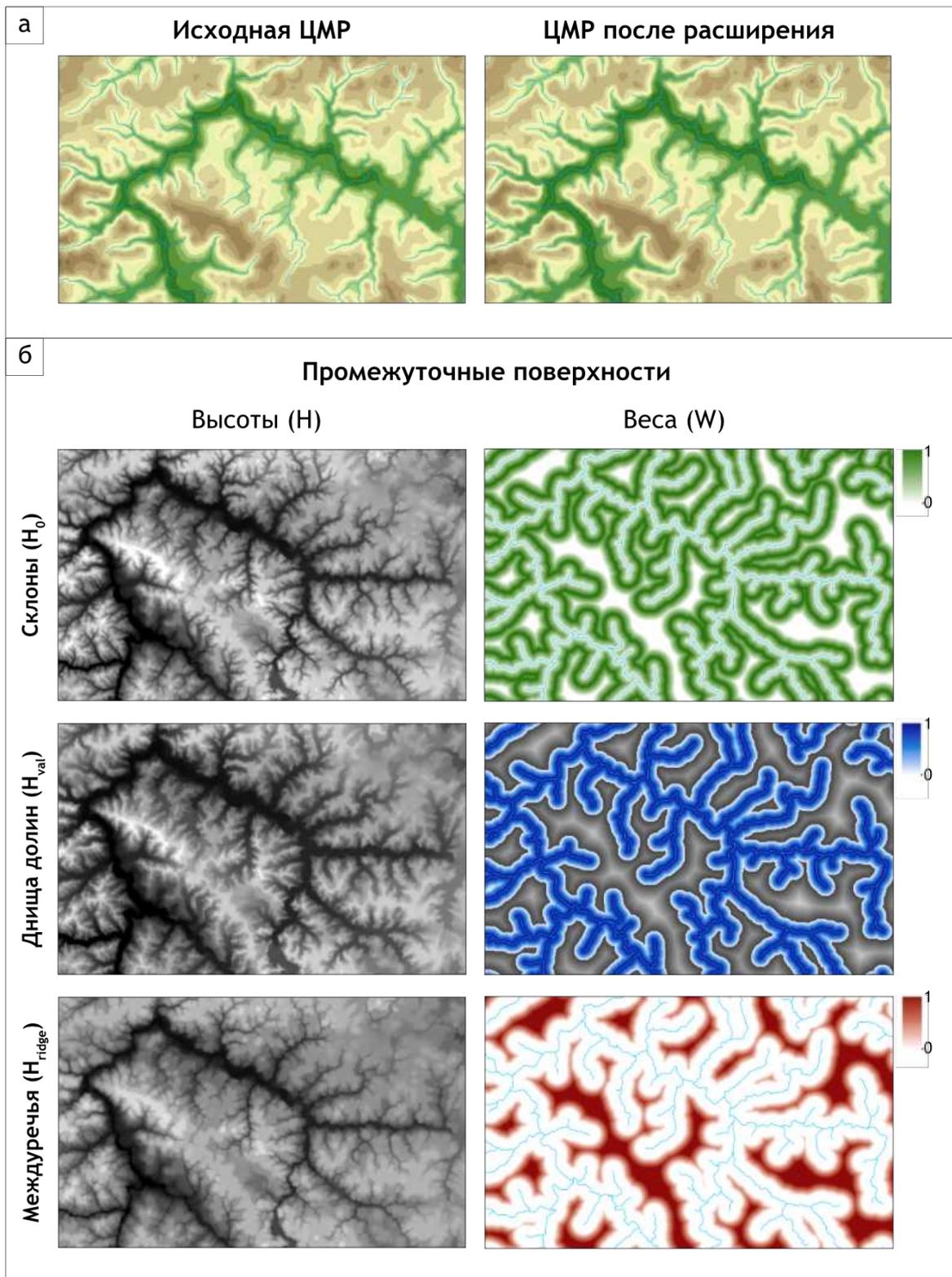


Рисунок 2.48 – Принцип расширения долин и междуречий

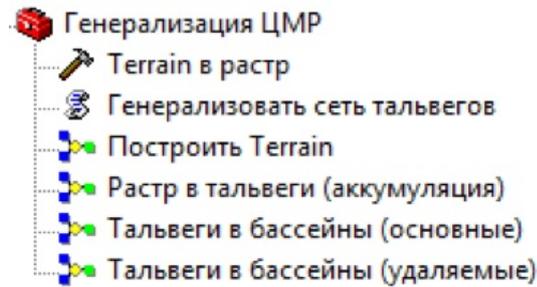
## 2.5.4 Реализация алгоритма

Предложенный алгоритм генерализации ЦМР не является стандартным функциональным средством ГИС. Для его апробации и применения при обобщении рельефа потребовалась проектирование многостадийных моделей обработки данных, а также программирование на языке Python. ArcGIS Desktop предоставляет набор готовых инструментов обработки данных, которые можно выстраивать в алгоритмические цепочки с помощью специального модуля Model Builder (среда визуального программирования). Построенную цепочку можно сохранить как модель обработки данных со списком параметров и далее использовать как любой инструмент, изначально имеющийся в ArcGIS Desktop. Построение моделей значительно облегчает многостадийную обработку данных.



**Рисунок 2.49** – Поверхности, используемые при расширении долин и междуречий: (а) пример расширения, (б) разложение результирующей поверхности на составляющие

В целях снижения большого количества входных и выходных параметров алгоритм был разбит на несколько подзадач. Далее путем экспериментов на небольших тестовых ЦМР с параллельной оценкой достоверности генерализации была выработана наиболее оптимальная последовательность операций для решения каждой подзадачи. После этого были построены и собраны в готовый набор инструментов модели обработки данных (Рисунок 2.50, Таблица 2.5, Таблица 2.6).



**Рисунок 2.50** – Набор инструментов генерализации ЦМР для ArcGIS Desktop

Задачу генерализации сети тальвегов оказалось невозможно решить стандартными средствами ArcGIS Desktop, в связи с чем был написан небольшой скрипт на языке программирования *Python*, который реализует алгоритм трассировки тальвегов методом наименьшей отрицательной разности аккумуляции стока, использованный в работе [Leonowicz, Jenny, Hurni, 2009]. Алгоритм последовательно перебирает все ячейки модели аккумуляции стока и из каждой «пускает» тальвег вверх по склону. Среди 8-ми соседних ячеек ищется та, в которой значение аккумуляции стока меньше, чем в текущей, но при этом наиболее близко к нему. Операция повторяется для выбранной ячейки и т.д., пока не будет достигнуто пороговое минимальное значение аккумуляции стока, при котором тальвег не считается выраженным, т.е. алгоритм добрался до истока. Если количество накопленных ячеек превышает порог длины, они помечаются в новом растре как принадлежащие к тальвегам.

Разработанные модели представлены на Рисунках 2.51 – 2.53. Их преимущество заключается в простоте и удобстве использования. Все параметры моделей документированы. Любой пользователь, владеющий ArcGIS и кратко ознакомившийся с сутью предложенного алгоритма, может сразу же перейти к генерализации имеющихся у него ЦМР. При этом не потребуются освоение нового программного обеспечения.

**Таблица 2.5 – Инструменты генерализации ЦМР для ArcGIS Desktop**

Название инструмента	Решаемая задача	Тип	Входные параметры	Выходные параметры
Растр в тальвеги (аккумуляция)	Выделение сети тальвегов по ЦМР на основе порогового значения аккумуляции тока	Модель	ЦМР, пороговое значение аккумуляции тока	Растровый слой тальвегов
Генерализовать сеть тальвегов	Генерализация сети тальвегов по модели аккумуляции тока на основе критериев длины и минимальной аккумуляции тока	Скрипт на языке Python	ЦМ аккумуляции тока, пороговое значение аккумуляции тока, минимальная длина тальвега	Растровый слой тальвегов
Тальвеги в бассейны (основные)	Построение бассейнов для основных тальвегов	Модель	Векторный слой тальвегов, ЦМ направления тока, ЦМ аккумуляции тока, ширина буфера для устьевых точек, основной допуск привязки устьевых точек	Векторный слой бассейнов
Тальвеги в бассейны (удаляемые)	Построение бассейнов для удаляемых тальвегов	Модель	Векторный и растровый слой основных тальвегов, растровый слой всех тальвегов, ЦМ направления тока, ЦМ аккумуляции тока, основной и расширенный допуск привязки устьевых точек, ширина буферной зоны для отбора устьевых точек	Векторный слой бассейнов
Построить TIN	Построение триангуляционной модели на основе полученных линий	Модель	Исходная ЦМР, основные тальвеги, основные водоразделы, удаляемые водоразделы, средний интервал между точками, определение пирамидных слоев	Триангуляционный слой
TIN в растр	Преобразование триангуляции в растровую модель	Инструмент модуля 3D Analyst	Триангуляционный слой, тип данных (целочисленный, с плавающей точкой), метод интерполяции (линейный, естественного соседа), разрешение выходного растра	Растровая ЦМР

### 2.5.5 Тестирование алгоритма

Тестирование алгоритма производилось в средних и мелких масштабах картографирования, на разных морфологических типах рельефа, при этом уменьшение масштаба производилось в 4-10 раз. Параметры тестирования представлены в Таблице 2.7.

Оценка достоверности генерализации производилась путем сравнительного анализа изображений и профилей, некоторые из которых представлены на Рисунках 2.54, 2.55 и 2.56. В качестве образцов для каждой территории представлено 3 изображения способом горизонталей с послойной окраской и отмывкой, в основе которых лежат:

- *Модель 1*, полученная передискретизацией исходной модели.
- *Модель 2*, полученная сглаживанием Модели 1 для удаления деталей.
- *Модель 3*, полученная авторским алгоритмом генерализации.

*Модель 2* сглаживалась таким образом, чтобы ее детализация приблизительно равнялась детализации *Модели 3*. Для каждого примера также построено несколько профилей.

Анализ изображений позволяет сделать следующие выводы:

1. Передискретизация в качестве средства обобщения неприменима. Изображение загромождено деталями, пластика отсутствует, рельеф читается очень плохо.
2. При сглаживании моделей пропадает согласованность горизонталей, поскольку формы рельефа не удаляются целиком. Уменьшаются площади положительных и отрицательных форм, увеличиваются площади склонов, теряется острота гребней, скругленное замыкание горизонталей на равнинном рельефе не соответствует интенсивному характеру овражного расчленения. Обобщенный профиль осредняет исходную поверхность, нивелируя морфологические особенности рельефа.
3. Модели, полученные авторским алгоритмом генерализации, характеризуются хорошей согласованностью горизонталей, сохранением профилей склонов и характера междуречий (острые гребни, плоскогорья). При этом, если форма расчленения удаляется целиком, то все горизонтали, соответствующие ей, также выравниваются. Мелкие положительные формы на водоразделах объединяются. Профиль поверхности, как правило, близок к требуемому: мелкие неровности удаляются полностью, при этом профиль идет по водораздельным точкам. Наиболее удачным является пример обобщения цифровой модели GDEM до масштаба 1 : 1 000 000: полученные горизонтали очень близки к рукописному аналогу с топографической карты масштаба 1 : 1 000 000.

**Таблица 2.6** – Комментарии к некоторым параметрам инструментов генерализации

Параметр	Комментарий
ЦМ направления тока	Растровая цифровая модель, которая в каждой ячейке содержит информацию о направлении тока жидкости из нее в соседние ячейки. Строится на основе ЦМР.
ЦМ аккумуляции стока	Растровая цифровая модель, которая в каждой ячейке содержит информацию о количестве дренируемых ею ячеек выше по склону. Строится на основе ЦМ направления стока.
Пороговое значение аккумуляции стока	Минимальное значение аккумуляции стока, при котором ячейка считается принадлежащей к тальвегу
Ширина буфера для устьевых точек	Поскольку каждая точка слияния двух тальвегов содержит 2 устьевых точки, их необходимо разделить для того, чтобы построить отдельные бассейны для обоих тальвегов. Для этого вокруг устьевых точек строятся небольшие буферные зоны (экспериментально установлено, что достаточно радиуса в 2 ячейки ЦМР). Далее тальвеги обрезаются полученными кругами, в результате чего устьевые точки слегка смещаются выше по течению (расходятся в стороны друг относительно друга), и появляется возможность построить отдельные бассейны для двух тальвегов.
Допуск привязки устьевых точек	В результате растрово-векторных преобразований положение устьевых точек смещается относительно растровой линии в пределах величины размера ячейки. Для того чтобы корректно построить бассейны, необходимо “вернуть” устьевые точки на исходное положение. Для этого устанавливается допустимая величина смещения и на основе ЦМ аккумуляции стока точки возвращаются на свои места.

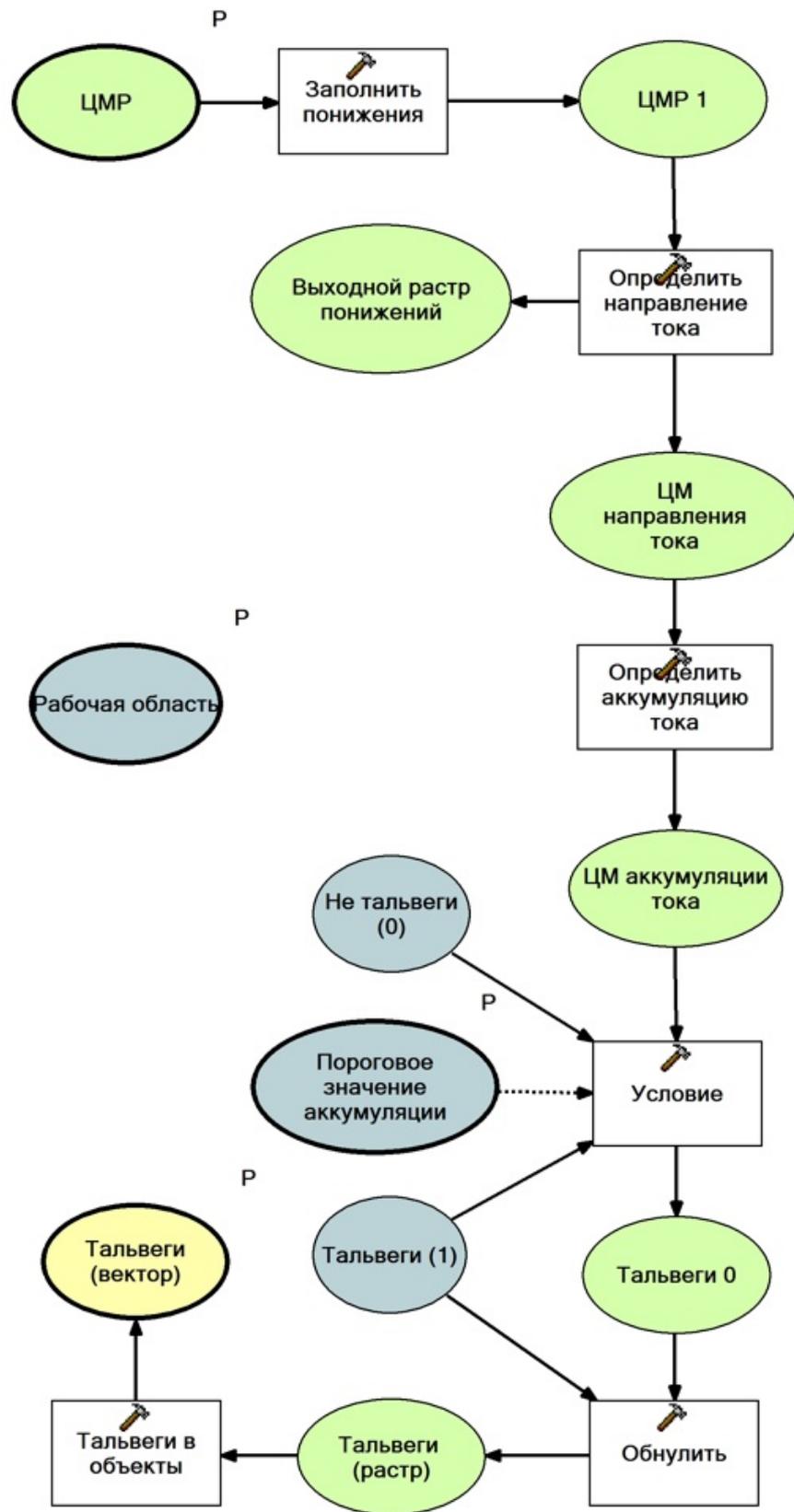


Рисунок 2.51 – Модель геообработки для выделения тальвегов по аккумуляции тока

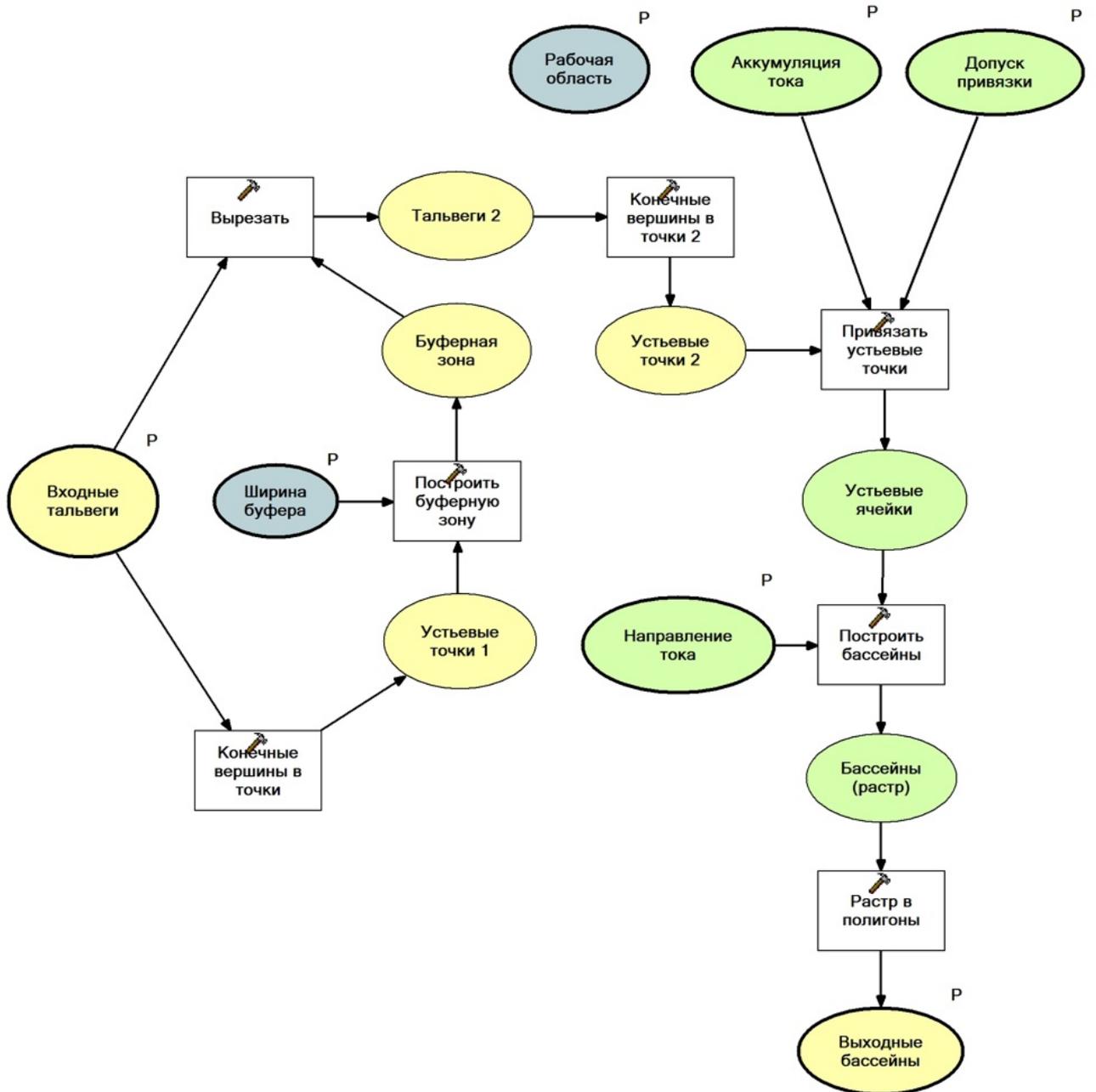


Рисунок 2.52 – Модель геообработки для построения бассейнов для основных тальвегов

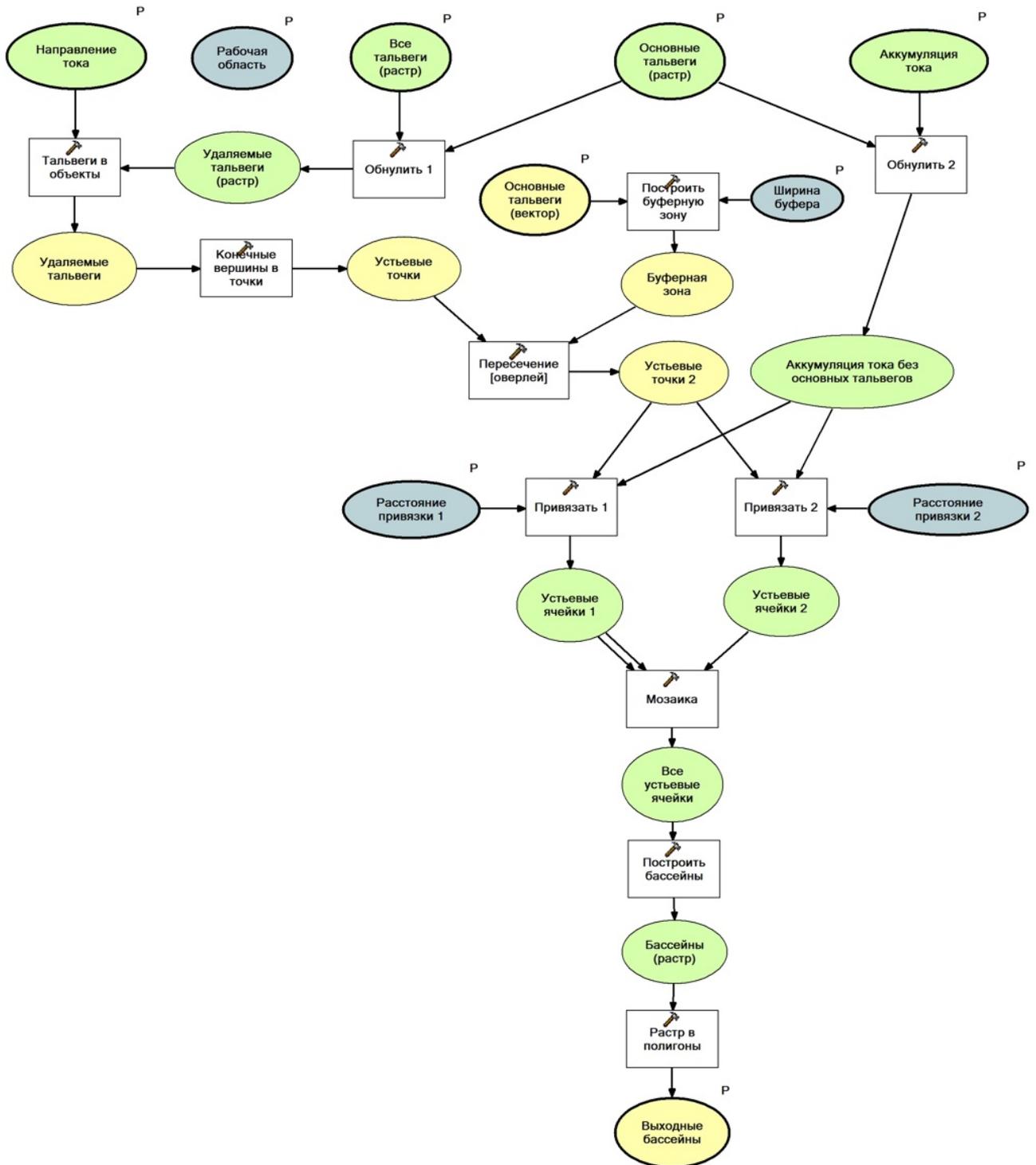
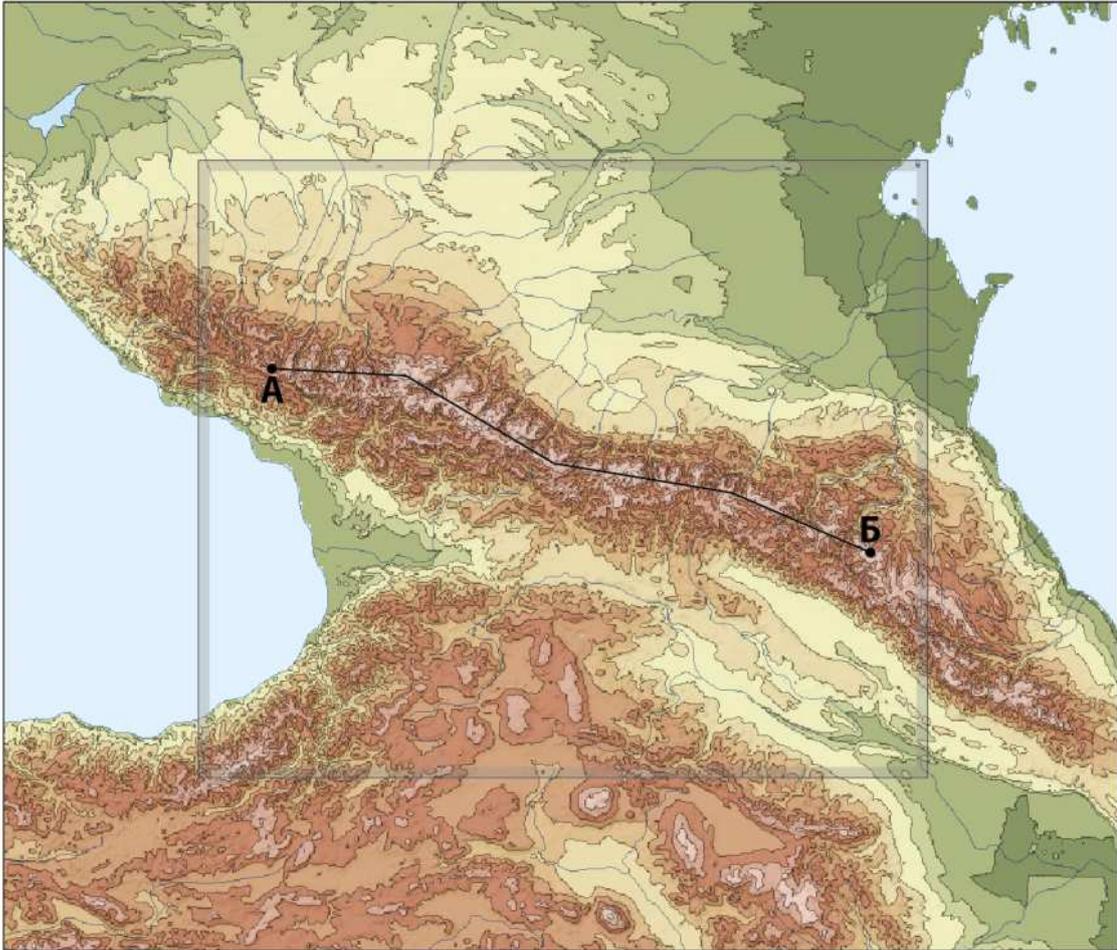
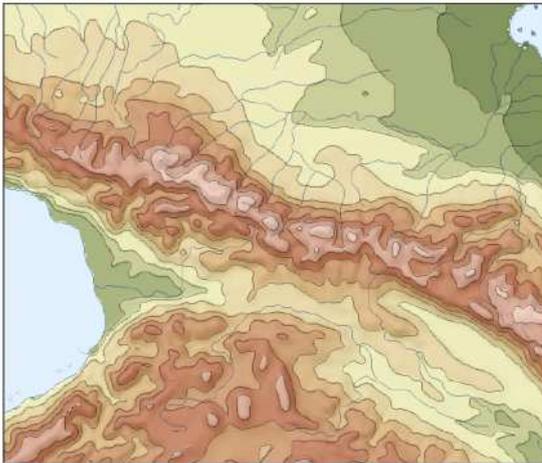


Рисунок 2.53 – Модель геообработки для построения бассейнов второстепенных тальвегов

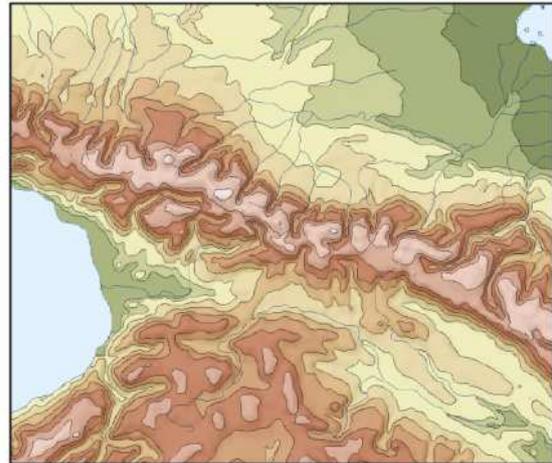
## Исходная ЦМР (передискретизация)



## Сглаживание



## Новый алгоритм



## Профиль по линии А—Б

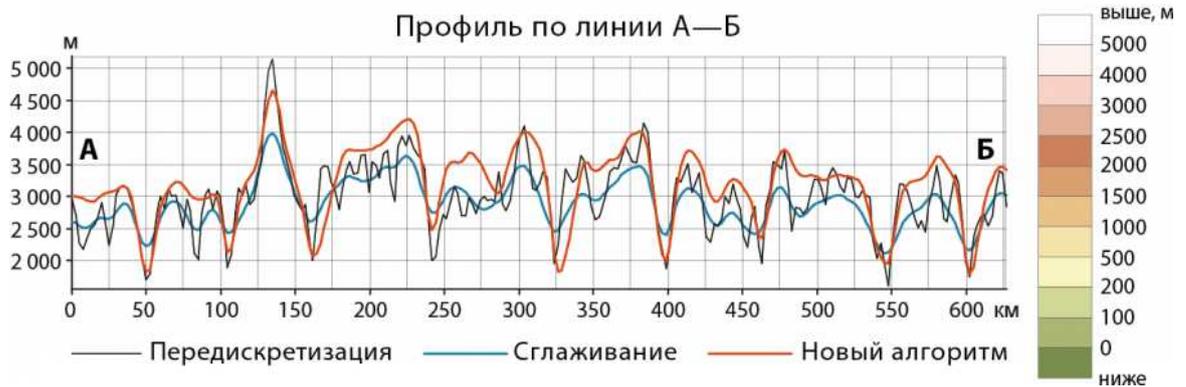
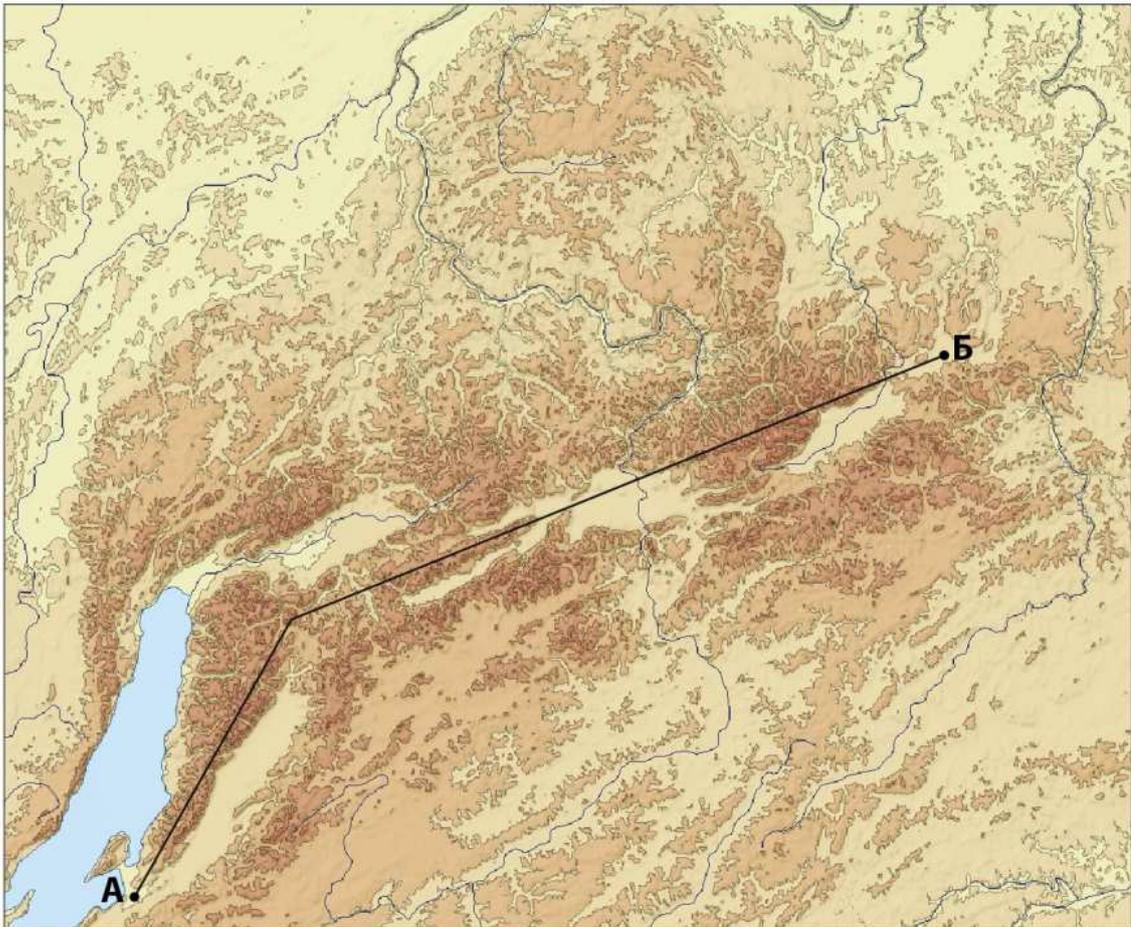
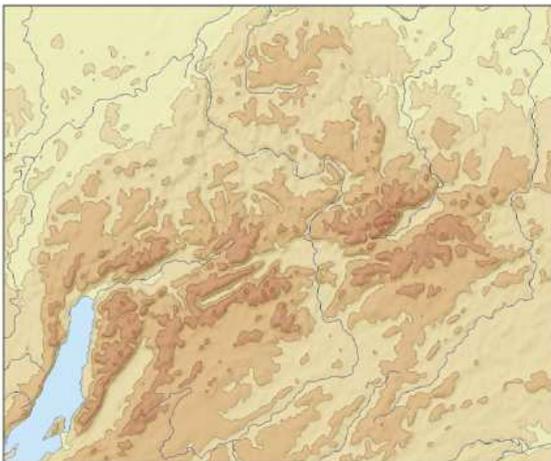


Рисунок 2.54 – Генерализация высокогорья (1 : 1 000 000 → 1 : 10 000 000)

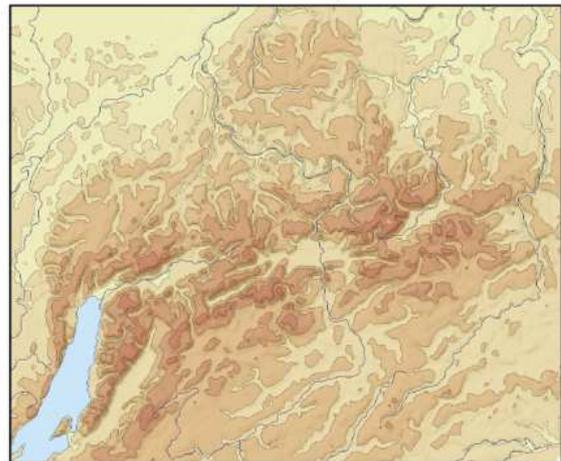
## Исходная ЦМР (передискретизация)



## Сглаживание



## Новый алгоритм



## Профиль по линии А—Б

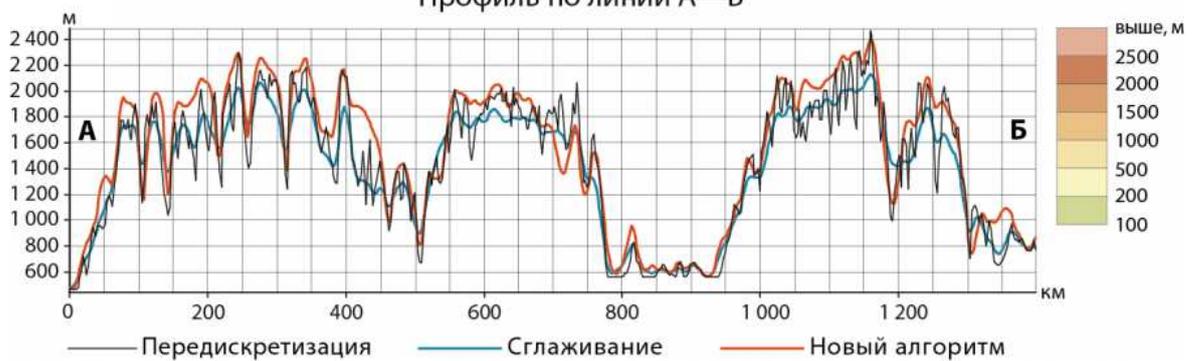


Рисунок 2.55 – Генерализация нагорья (1 : 1 000 000 → 1 : 10 000 000)

Таблица 2.7 – Участки тестирования алгоритма генерализации

Масштаб	Источник (исходная ЦМР)	Разрешение, м (исходная ЦМР)	Масштаб (исходная ЦМР)	Разрешение, м (целевая ЦМР)	Масштаб (целевая ЦМР)	Территория
Мелкий	ГТОРО	1000	1:1 000 000	2500	1:10 000 000	Кавказ, Среднерусская возвышенность, Плато Путорана, Становое нагорье
Средний	ASTER GDEM	30	1:100 000	250	1:1 000 000	Кавказ (Эльбрус)
Средний	Топокарта 1:200 000	200	1:200 000	250	1:1 000 000	Среднерусская возвышенность, плато Путорана, Становое нагорье

### Результаты генерализации



### Попарное наложение горизонталей



Рисунок 2.56 – Генерализация равнинно-эрозионного рельефа (1 : 200 000 → 1 : 1 000 000)

### 2.5.6 Итоги

В настоящем разделе диссертации представлен алгоритм структурной генерализации цифровых моделей рельефа. Алгоритм удовлетворяет изначально выдвинутым требованиям и в целом обеспечивает сохранение морфологического правдоподобия ЦМР. При этом сохраняются не только водораздельные, но и все значимые структурные линии.

Как известно, структурные линии рельефа не всегда совпадают с водораздельными [Заруцкая, 1958]. Такими линиями могут быть, например, резкие бровки склонов. В то время как междуречье с водораздельной линией может быть пологим и морфологически невыразительным. Использование водораздельных линий поперечных бассейнов, которые отражают профиль поверхности, дает основание полагать, что все характерные перегибы поверхности и структурные линии с высокой степенью вероятности будут включены в результирующую ЦМР. Таким образом, достигается эффект, идентичный проведению ведущих горизонталей.

Несмотря на положительные результаты тестирования, необходимо отметить недостатки полученных моделей, которые требуют совершенствования алгоритма:

1. Наличие разорванных горизонталей по днищам долин, несмотря на их расширение при постобработке. Эта проблема связана не с алгоритмом генерализации, а с качеством исходных ЦМР, в которые уже заложены ошибочные замкнутые локальные понижения (*spurious pits*). Для решения этой проблемы можно, как перед генерализацией, так и на стадии постобработки, использовать оптимальный алгоритм удаления депрессий, предложенный [Soille, 2004]. Данный алгоритм удаляет понижения либо путем их заполнения до наиболее низкой отметки борта (*filling*), либо прорезанием тальвега шириной в одну ячейку от днища понижения до ближайшей отметки с той же высотой (*carving*). Критерием, который влияет на выбор одного из двух методов, является количество ячеек ЦМР, которые необходимо подвергнуть изменению высоты, чтобы достичь требуемого результата. В качестве альтернативного критерия Сои предлагает использовать объем поверхности.
2. Угловатые горизонталы в областях с неявными превышениями. Наиболее явно это видно на междуречьях равнинно-эрозионного рельефа. Горизонталы в этих областях должны быть менее детализированы.
3. Присутствуют некоторые детали, неадекватные масштабу, выражающиеся в мелких замкнутых горизонталях.
4. При генерализации тальвегов не учитывается расстояние между ними. В горных районах это часто приводит к тому, что сеть расчленения по горизонтали излишне подробна и требуется объединение соседних форм рельефа.

5. В результате постобработки профиль поверхности получается излишне сглаженным, форма вершин в некоторых случаях искажается. Для решения этой проблемы, по всей видимости, необходимо локально адаптировать параметры фильтрации: вес, размер окна, число итераций для фильтров верхней и нижней квантили. Дальнейшая работа над алгоритмом должна дать пути решения этих проблем.

## 2.6 Агрегирование разнородных пространственных данных для построения баз данных характеристик городской среды<sup>12</sup>

**Пространственные соотношения** характеризуют пропорции, в которых объекты разных типов присутствуют на местности. Генерализация пространственных данных методом агрегирования позволяет вычислять подобные соотношения для моделей городской погоды и климата [Baklanov *et al.*, 2018; Garuma, 2018; Lipson *et al.*, 2022].

**Целью** настоящего раздела является разработка методики, обеспечивающей возможность расчета параметров городской среды путем генерализации (агрегирования) детальных пространственных данных с учетом их соотношения в модельных ячейках, а также ее апробация на примере расчета этих параметров по модельной сетке, покрывающей территорию Московского мегаполиса с размером ячейки 500 м [Samsonov, Konstantinov, Varentsov, 2015].

Недавние исследования для Москвы показали перспективность применения модели COSMO с параметризацией TERRA\_URB [Wouters *et al.*, 2016] и шагом сетки 1 км для климатических исследований [Varentsov *et al.*, 2018], а также оперативного прогноза погоды в составе системы COSMO-Ru Гидрометцентра России [Rivin *et al.*, 2019]. В рамках развития данной системы ведется работа по использованию меньших шагов сетки, например, 500 м, что требует расчета параметров городской среды с соответствующей детальностью.

В данном разделе речь будет идти о получении топографических характеристик, которые можно условно разделить на две категории: характеристики геометрии городской застройки, в том числе параметры городских каньонов, а также распределение типов подстилающей поверхности, описывающих характер поверхностей, слагающих городскую среду: здания, дороги, лес, водные объекты и т.д. И те, и другие характеристики могут быть выражены в виде физических констант и переменных в специальных блоках параметризации урбанистических эффектов, используемых в метеорологических моделях.

<sup>12</sup> При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov, Konstantinov, Varentsov, 2015; Samsonov, Varentsov, 2020]

Изложение построено следующим образом: сначала описывается подготовка данных, обеспечивающая топологически корректное покрытие с правильным соотношением типов подстилающей поверхности и городских каньонов. Затем описывается процедура генерализации этих данных методом агрегирования, при которой вводятся необходимые коррекции на вычисление доли урбанизированной территории.

## 2.6.1 Методика

### 2.6.1.1 Геометрия городской застройки

#### Объектно-ориентированный подход

В рамках работы предложен *объектно-ориентированный* подход к анализу геометрии городской застройки [Samsonov, Konstantinov, Varentsov, 2015]. Ключевыми характеристиками застройки для метеорологического моделирования являются характеристики городских каньонов. При этом каньоны отсутствуют в топографических базах данных, что требует разработки специализированных алгоритмов для извлечения их границ.

В классической постановке каньон определяется парой зданий [Oke, 1987]. В то же время, когда соседние каньоны имеют общую ориентацию, они формируют один каньон более высокого порядка, который создает особые условия для физических процессов в городской среде. Такие каньоны формируются вдоль городских улиц и называются *направленными* [Samsonov, Konstantinov, Varentsov, 2015]. При этом остальные, обычно внутриквартальные, пространства между зданиями состоят из ненаправленных каньонов. С некоторой долей условности можно говорить о том, что *длина направленного каньона значительно больше его ширины*. Помимо этого, направленный каньон имеет постоянное или постепенно, но не резко меняющееся направление. В работе [Samsonov, Konstantinov, Varentsov, 2015] предложена иерархическая классификация каньонов, которая может быть использована для моделирования взаимодействия атмосферы и городской подстилающей поверхности в разных масштабах:

1. *Микро-каньоны*. Это базовый уровень, на котором каньон определяется каждой парой зданий (городской каньон в его классическом понимании). Этот уровень является атомарным и должен быть использован для микроклиматических исследований, таких как оценка энергетического баланса каньона, эффективности расходования энергии в зданиях, исследование биоклиматической комфортности. Типичный линейный масштаб микро-каньонов составляет порядка десятков метров.
2. *Мезо-каньоны* представляют собой промежуточный уровень обобщения понятия городского каньона и определяются как цепочка каньонов между перекрестками, вытянутых вдоль осевой линии одной улицы. Длинные мезо-каньоны могут быть разбиты на секции, имеющие однородный по высоте, материалам и прочим

свойствам тип застройки. Данный уровень является атомарным для исследования сети направленных каньонов города и может быть использован для климатического и метеорологического моделирования на локальном и мезомасштабном уровнях. Возможные направления приложений включают прогноз погоды и климата в городах, исследований энергетического баланса в масштабах города. Типичный линейный масштаб мезо-каньонов составляет порядка сотен метров.

3. *Макро-каньоны* соответствуют длинным цепочкам мезо-каньонов вдоль улиц. Примером макро-каньона может быть любая улица — в Москве, например, Ленинский проспект. Макро-каньоны используются для решения специальных задач таких как исследование основных коридоров вентиляции (продувания) города. Типичный линейный размер макро-каньонов имеет масштаб порядка километров.

Под выделением каньонов будем понимать выделение плоских участков между зданиями, которые соответствуют днищам искомых каньонов (при этом трехмерный объем каньона может быть получен далее на основе информации о высотах зданий). Для этого требуется следующая информация:

1. Геометрия плановых очертаний отдельных зданий.
2. Линии улично-дорожной сети (для определения направленных каньонов).
3. Типы наземного покрова (нужны для правильного учета высокой растительности между зданиями).

Поскольку каньоны представляют собой пространства между зданиями, необходима некая структура данных, которая позволила бы связать между собой каньонобразующие объекты (здания и растительность) и выделять между ними участки, принадлежащие каньонам. В качестве такой структуры удобно использовать *триангуляцию Делоне*. При поиске треугольников, принадлежащих каньонам необходимо наложить условие, при котором такой треугольник должен иметь хотя бы одну вершину, принадлежащую зданию, в противном случае в выборку попадут еще и чисто «растительные» каньоны (просеки, аллеи), которые в данном случае не представляют интереса.

Следующей стадией является извлечение иерархии каньонов в соответствии с ранее определенными уровнями (микро-, мезо- и макро-) на основе их отношений с отдельными зданиями, сегментами улиц и улицами.

1. *Микро-каньоны* могут быть извлечены путем нахождения треугольников, соединяющих общую для них пару зданий.
2. *Мезо-каньоны* могут быть извлечены путем нахождения треугольников, пересекающих один и тот же сегмент улично-дорожной сети между двумя перекрестками. Треугольники, пересекаемые несколькими сегментами, помечаются как принадлежащие перекресткам.
3. *Макро-каньоны* могут быть извлечены путем нахождения треугольников, имею-

щих одинаковый идентификатор городской магистрали, которая пересекает их.

Пример результата выделения каньонов различного уровня приведен на Рисунке 2.57. Днище микроканьона состоит из треугольников, соединяющих одну пару зданий, расположенных друг напротив друга. Днище мезо-каньона формируется между зданиями, расположенными на участке улицы от одного перекрестка до другого. Макро-каньоны соответствуют треугольникам, выделенным вдоль всей улицы.



Рисунок 2.57 – Городские каньоны разного иерархического ранга

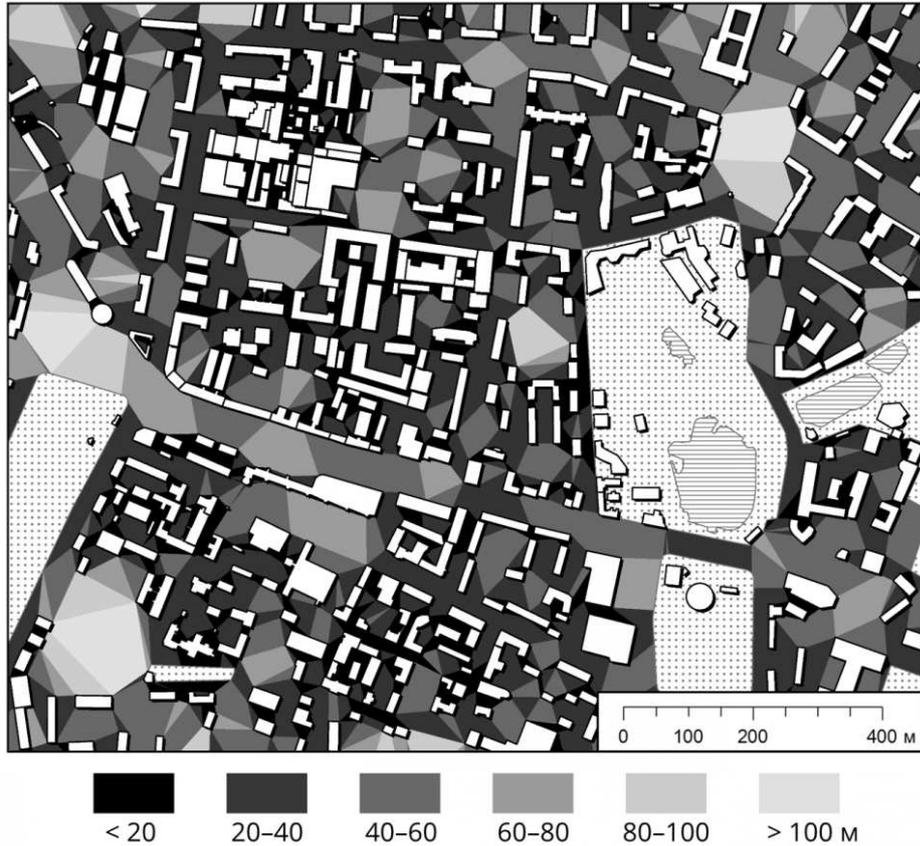
За локальную ширину каньона удобно принять высоту треугольника, соответствующую одной из его сторон:

$$w_i = \frac{2S_i}{f(e_{1i}, e_{2i}, e_{3i})}, \quad (2.22)$$

где  $S_i$  — площадь  $i$ -го треугольника,  $e_{1i}, e_{2i}, e_{3i}$  — длины ребер  $i$ -го треугольника,  $f$  — функция выбора длины одной из трех сторон, которая имеет различный вид в зависимости от типа каньона. В случае ненаправленного каньона берется кратчайшая сторона:  $f = \min$ . Для треугольников, принадлежащих направленным каньонам, не являющихся перекрестками, берется длина стороны, не пересекающейся с линиями дорожной сети. Наконец, для перекрестков (треугольников, пересекаемых несколькими улицами) берется средняя длина стороны:  $f = \text{avg}$ . Пример результата вычисления локальной ширины городского каньона данным методом представлен на Рисунке 2.58.

При агрегировании данных о каньонах средняя ширина каньона может быть найдена путем взвешенного осреднения по всем треугольникам, занятым каньонами в пределах модельной ячейки:

$$\bar{w} = \frac{\sum_i^n S_i w_i}{\sum_i^n S_i}, \quad (2.23)$$



**Рисунок 2.58** — Локальная ширина городского каньона (Пресненский район, Москва). Белым цветом отмечены здания

где  $n$  — число треугольников в модельной ячейке.

В качестве характеристики, описывающей среднюю высоту следует использовать средневзвешенную по периметру высоту зданий:

$$\bar{h} = \frac{\sum_j^k l_j h_j}{\sum_j^k l_j}, \quad (2.24)$$

где  $n$  — число зданий внутри зоны,  $l_j$  — периметр  $j$ -го здания, касающийся каньона  $h_j$  — высота  $i$ -го здания.

Средние пропорции каньона внутри расчетной зоны определяются отношением средней ширины и высоты:

$$HW = \bar{w} / \bar{h}. \quad (2.25)$$

Для агрегирования основных (модальных) направлений каньонов в пределах ячейки при объектно-ориентированном подходе используется эмпирическая функция плотности распределения направлений улиц в области определения  $[0, 180)^\circ$ . В дискретном варианте рассчитывается гистограмма направлений с интервалом  $h$ , находится интервал, образующий максимальный пик гистограммы, и определяется модальное направ-

ление  $\theta_1$  по формуле:

$$\theta_1 = \alpha_k + h \frac{L_k - L_{k-1}}{2L_k - L_{k-1} - L_{k+1}}, \quad (2.26)$$

где  $\alpha_k$  — нижняя граница  $k$ -го интервала гистограммы,  $L_k$  — суммарная длина улиц, направление которых попадает в  $k$ -й интервал гистограммы. Для нахождения второй по значимости моды  $\theta_2$  ищется следующий по высоте пик гистограммы, выполняются аналогичные расчеты.

Альтернативно для поиска модальных направлений может использоваться метод ядерного сглаживания [Rosenblatt, 1956; Parzen, 1962]. Математически операция ядерной оценки плотности в точке  $x$  определяется как:

$$\hat{f}_h(x) = \frac{1}{nh} \sum_{i=1}^n K\left(\frac{x - x_i}{h}\right), \quad (2.27)$$

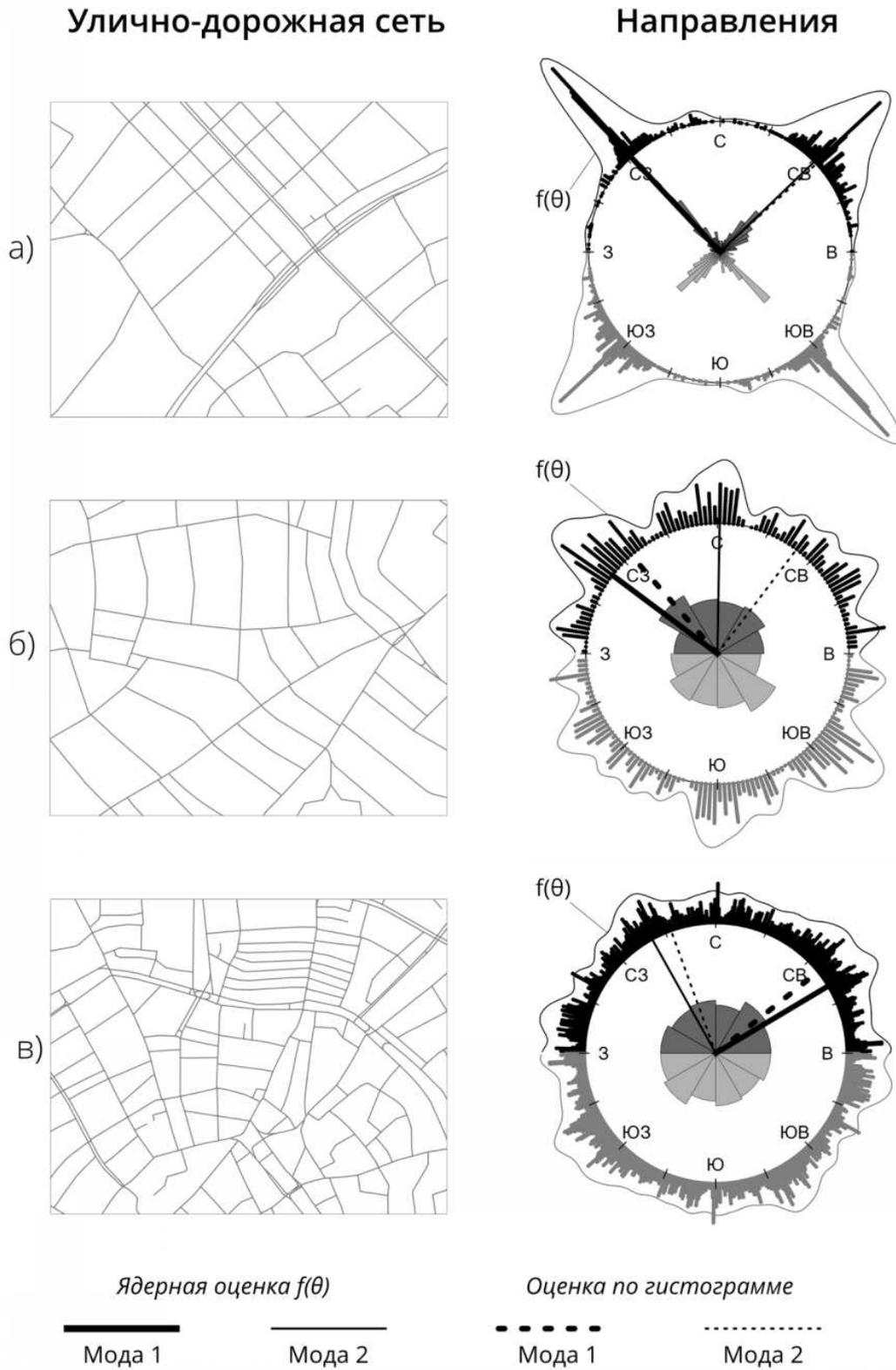
где  $K$  — ядерная функция,  $h > 0$  — сглаживающий параметр,  $x_i$  — элементы выборки,  $n$  — размер выборки. Ядерная функция должна удовлетворять двум критериям:

$$\begin{cases} K(x) \geq 0; \\ \int_{-\infty}^{+\infty} K(x) dx = 1. \end{cases}$$

Когда подобрана функция ядра и ее параметры, оценка плотности распределения (вычисление функции  $\hat{f}_h(x)$ ) для круговых данных делается либо для исходных направлений  $\theta_j$ , либо с равным (достаточно малым) интервалом — например, через 1 градус [Pewsey, Neuhauser, Ruxton, 2013]. После того как произведена оценка, для решения нашей задачи могут быть выбраны направления, в которых функция плотности распределения достигает локального максимума — первого и второго по величине. Эти направления и будут соответствовать первой и второй моде распределения направлений городских каньонов.

На Рисунке 2.59 показаны примеры оценки плотности распределения направлений методами построения гистограмм и ядерной оценки. Исходные измерения сгруппированы по интервалам шириной 1 градус и показаны точками вдоль радиус-векторов, отложенных из центра круга. Гистограмма направлений в данном случае представлена в виде розы-диаграммы в центре каждого графика, что позволяет избежать эффекта некорректной интерпретации, связанного с ее разрывом. Результат ядерной оценки плотности распределения показан сплошной линией, окаймляющей исходные измерения. Для поиска оптимального количества интервалов гистограммы, а также параметра концентрации в ядерной оценке использованы методы, предложенные в [Oliveira, Crujeiras, Rodríguez-Casal, 2012; Knuth, 2013].

Видно, что в случае ярко выраженной моды (Рисунок 2.59а) оптимальное количе-



**Рисунок 2.59** – Преобладающие направления каньонов по линиям улично-дорожной сети. Примеры: (а) ул.Тверская–Садовое кольцо, (б) ул.Пречистенка–Старокопюшенный пер., (в) ул. Большая Лубянка–Бульварное кольцо

ство интервалов гистограммы достаточно велико и позволяет достаточно точно определить направление, которое практически совпадает с направлением, выделяемым методом ядерной оценки. В то же время, если моды распределения выражены менее явно (Рисунок 2.59в), оптимальное количество интервалов гистограммы радикально снижается и тем сильнее могут отличаться оценки, полученные методом гистограмм и ядерного сглаживания. Можно говорить о том, что метод ядерного сглаживания дает более локализованную оценку и является предпочтительным для выявления основных направлений городских каньонов.

### Растровый подход

Недостатком объектно-ориентированного подхода является зависимость вычисления направлений каньонов от направления улиц, что делает невозможным оценку при неполных данных об улично-дорожной сети. Помимо этого, оценка ширины каньона получается довольно дискретизированной (см. пример на Рисунке 2.58). Для избежания этих недостатков была также разработана растровая технология вычисления геометрических параметров городских каньонов, кратко представленная в [Kosheleva et al., 2023].

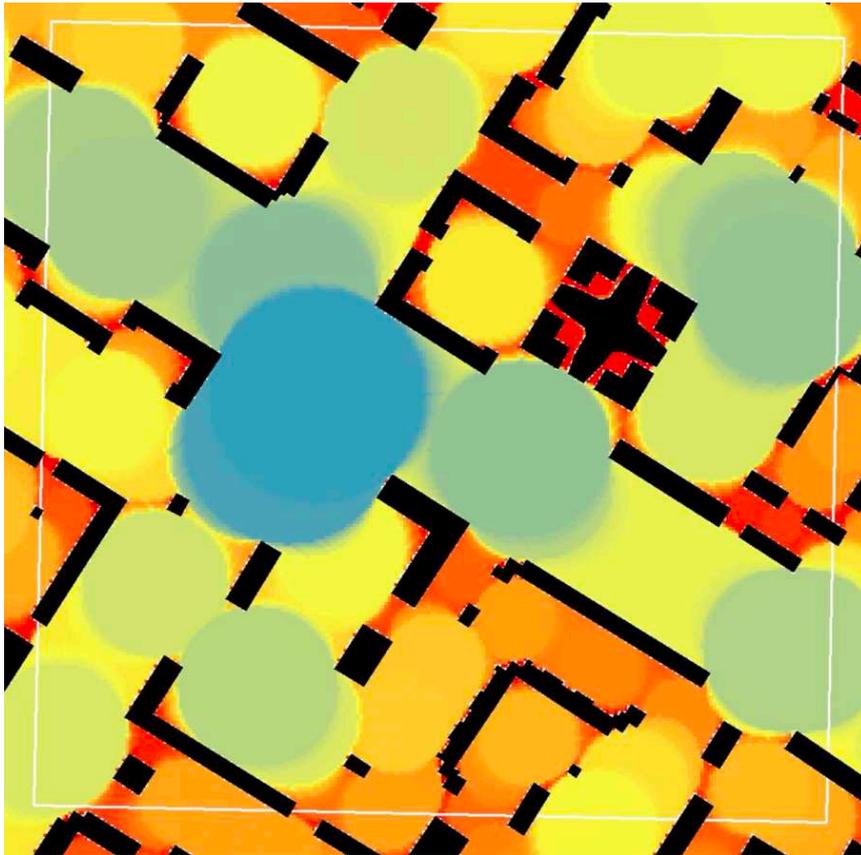
Ширина локального каньона для данной точки определяется как диаметр окружности наибольшего радиуса (доминантной окружности), вписанной между зданиями вокруг и содержащей данную точку внутри. Далее расчеты производятся по всей территории исследования на растровой сетке заданного разрешения (в нашем случае 5 м). Таким образом, для каждой точки пространства, занятого каньонами, становится известна ширина локального каньона ( $w$ ) с дискретностью, определяемой разрешением раstra. Более подробное описание вычисления ширины в растровом режиме см. в Разделе 3.2.1.2. Пример детальной оценки ширины городского каньона растровым методом представлен на Рисунке 2.60.

Высота локального каньона ( $h$ ) определяется как среднее арифметическое высот зданий, затрагиваемых доминантной окружностью. Далее на основе этих характеристик рассчитываются пропорции каньона ( $h/w$ ). Для определения длины ( $L$ ) и ориентации ( $\theta$ ) макроканьона для каждого пикселя доминантной окружности находят отрезок максимальной длины, проходящий через центр пикселя и примыкающий к стенам зданий. Среди всех таких отрезков берется максимальный.

#### 2.6.1.2 Соотношение типов подстилающей поверхности

Соотношение типов подстилающей поверхности определялось на основе данных OpenStreetMap (сокращенно OSM) — открытого и свободно распространяемого набора пространственных данных. Исходная структура данных OSM организована в формате XML — Extensible Markup Language<sup>13</sup>. Все пространственные объекты и связи между ни-

<sup>13</sup> Подробное описание структуры данных OSM доступно на сайте [http://wiki.openstreetmap.org/wiki/OSM\\_XML](http://wiki.openstreetmap.org/wiki/OSM_XML)



**Рисунок 2.60** – Пример детальной оценки ширины городского каньона растровым методом

ми описаны в терминах узлов (точечных объектов), путей (линейных объектов) и отношений (связей между объектами). Площадные объекты создаются путем векторизации их границ в виде *замкнутых* путей, поэтому отдельный тип площадного объекта в OSM отсутствует. Узлы, пути и отношения могут иметь произвольное количество свойств, которые именуется *тегами*. Каждый тег состоит из ключа (названия тега) и значения. Например, жилое здание в структуре OSM представляет собой замкнутый путь, который имеет тег `building = 'residential'`, где `building` — это ключ, а `residential` — его значение. Наличие тега `building` говорит о том, что данный замкнутый путь следует интерпретировать как площадной объект.

Данные OSM требуют предварительной обработки для решения задачи оценки распределения типов подстилающей поверхности. Во-первых, XML-структура должна быть приведена к *реляционной* (табличной) модели данных, которая используется в геоинформационных системах. В реляционной модели данных каждому объекту соответствует строка в таблице, а по столбцам идут их атрибуты, причем количество атрибутов строго фиксировано. Данные группируются в таблицы по принципу однородности (отдельно гидрография, здания, дороги и т.д.), что позволяет определить ограниченный, но полный набор атрибутов, достаточный для кодирования основных свойств объектов каждой группы. В число атрибутов входит также геометрия объекта (его координаты), поэтому в геоинформатике такую модель называют *геореляционной*, тем самым уточняя, что

Таблица 2.8 – Структура данных OSM (выделены слои, используемые в работе)

Слой	Описание	Тип
poi-point	Точки интереса	point
railway-line	Железнодорожные, трамвайные пути и метро	line
settlement-point	Населенные пункты	point
railway-station-point	Железнодорожные станции	point
landuse-polygon	Землепользование	polygon
railway-platform-polygon	Железнодорожные платформы	polygon
highway-line	Дорожная сеть (осевые линии)	line
nature_reserve-polygon	Охраняемые природные территории	polygon
vegetation-polygon	Растительность	polygon
building-point	Здания	point
building-polygon	Здания	polygon
settlement-polygon	Населенные пункты	polygon
boundary-polygon	Административные границы	polygon
surface-polygon	Поверхность	polygon
poi-polygon	Области интереса	polygon
water-line	Водотоки	line
water-polygon	Водоёмы	polygon

она поддерживает хранение географических данных, расширяет традиционную реляционную модель. Во-вторых, должны быть устранены топологические ошибки (взаимные перекрытия, пересечения, дыры) между объектами, которые вносят искажения в статистику соотношения типов подстилающей поверхности. В-третьих, все разнообразие типов объектов (а их в базе данных OSM тысячи, как и на топографических картах) должно быть сведено к ограниченному набору классов подстилающей поверхности, которые могут быть учтены в метеорологических расчетах<sup>14</sup>.

Одним из источников данных OSM в геореляционной форме является интернет-портал NextGIS (<https://data.nextgis.com/>). Данные OSM на этом портале преобразованы в формат *Esri Shapefile*, который подходит для решения поставленных задач. Преобразованные данные имеют следующий состав слоев (Таблица 2.8). В таблице выделены строки слоев, которые используются в дальнейшей работе.

Следующим этапом после получения данных OSM в геореляционной форме является интеграция всех данных в единое покрытие (один слой), в котором произведено устранение топологических ошибок. Чтобы учесть площадь, занимаемую дорогами, необходимо на основе линейного представления дорог в виде осевых линий получить их площадное представление. Данные о ширине дорожного полотна в атрибутах OSM, как правило, отсутствуют. В то же время, имеются данные о классе (значимости) каждой линии и количестве полос, хранящихся в тегах *highway* и *lanes* соответственно. Следовательно, оценить ширину дорожного полотна можно путем табуляции (табличного сопоставления)

<sup>14</sup>С классификатором OpenStreetMap можно подробно ознакомиться по ссылке [http://wiki.openstreetmap.org/wiki/Map\\_Features](http://wiki.openstreetmap.org/wiki/Map_Features)

Таблица 2.9 – Ширина полосы движения для данных OpenStreetMap

Ширина полосы, м	Классы OpenStreetMap	Категории дорожной сети (РФ)
1	'footway', 'steps', 'path'	В городах: пешеходные улицы
1.5	'cycleway', 'road', 'track', 'bridleway', 'unclassified'	В городах: велосипедные дорожки За городом: дорога обычного типа (V)
3	'living_street', 'residential', 'service'	В городах: улицы и дороги местного значения, проезды
3.5	'primary', 'primary_link', 'secondary', 'secondary_link', 'tertiary', 'tertiary_link'	В городах: магистральные улицы За городом: дорога обычного типа (IВ, II, III, IV)
4	'trunk', 'trunk_link', 'motorway', 'motorway_link', 'raceway'	В городах: магистральные дороги За городом: автомагистраль (IA), скоростная дорога (IБ)
5	'pedestrian'	В городах: пешеходные улицы (ширина по проекту)

класса и ширины, которую необходимо установить для каждого класса. Для этого можно использовать нормативные документы ГОСТ Р 52399–2005 «Геометрические элементы дорог» и СП 42.13330.2011 «Градостроительство. Планировка и застройка городских и сельских поселений». Сопоставление классификатора OSM с классами дорог в данных документах позволяет оценить примерную ширину одной полосы движения для дороги каждого класса (Таблица 2.9) с учетом обочин, разделительной полосы и прочих элементов. Далее полученное значение умножается на количество полос, привязанное к каждой осевой линии, чтобы получить оценку фактической ширины дорожного полотна.

Зная ширину каждого сегмента улично-дорожной сети, можно трансформировать линейное представление дорог в площадное. Это преобразование осуществляется путем построения вокруг дорог буферных зон соответствующего радиуса. После того как получено площадное представление дорог, можно производить интеграцию всех слоев в единое покрытие. Для этого предложена технология *Слияние – Вырезание – Объединение* (Dissolve-Erase-Merge), позволяющая устранить топологические ошибки внутри каждого слоя и избежать их возникновения при слиянии слоев (Рисунок 2.61). Смысл процедуры заключается в следующем:

1. Объекты текущего слоя, относящиеся к одному типу (например, полигоны леса в слое vegetation-polygon), подвергаются процедуре *слияния* (**Dissolve**), что позволяет устранить перекрытия между ними.
2. Объектами текущего слоя *вырезаются* (**Erase**) области в следующем по порядку слое. При выполнении данной операции из вырезаемого слоя удаляются те части объектов, которые пересекаются с объектами вырезающего слоя. Это позволяет устранить перекрытие между слоями.
3. Текущий слой *объединяется* с последующим (**Merge**). При этом создается новый слой, в который включаются объекты всех объединяемых слоев. Объединенный

слой становится текущим. Множество атрибутов объединенного слоя включает в себя все атрибуты объединяемых слоев.

Шаги 2-3 повторяются до тех пор, пока есть последующие слои.

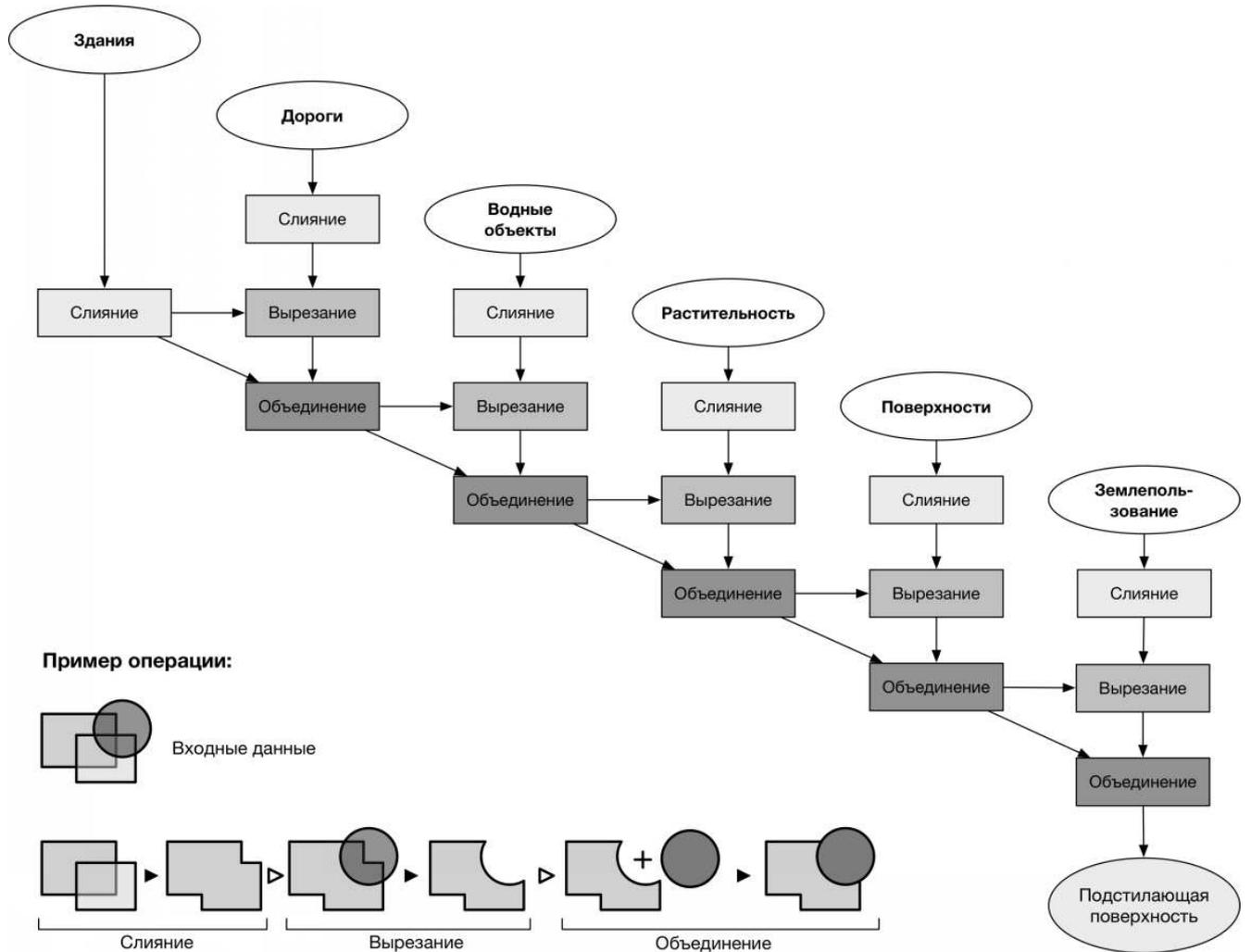
Процедура выполнялась в следующей последовательности, отражающей приоритетность слоев: **Здания** → **Дороги** → **Водные объекты** → **Растительность** → **Поверхности** → **Землепользование**.

В данной последовательности не участвует слой населенных пунктов, он используется при построении баз данных отдельно и служит для определения принадлежности расчетной зоны (ячейки) к населенному пункту того или иного типа (крупный город, город, сельский населенный пункт). Общая последовательность действий представлена на Рисунке 2.61. В левом нижнем углу рисунка проиллюстрированы описанные выше операции слияния, вырезания и объединения. Для каждой операции слева от знака ► показаны входные объекты, а справа — результат выполнения операции. Значок ▷ обозначает переход от одной операции к другой. На выходе получается топологически корректный слой, в котором множество атрибутов состоит из атрибутов всех входных слоев, то есть является их объединением. При этом сохраняется все многообразие классов объектов, которые хранились в исходных слоях, а сам слой геометрически уже пригоден для агрегирования соотношений.

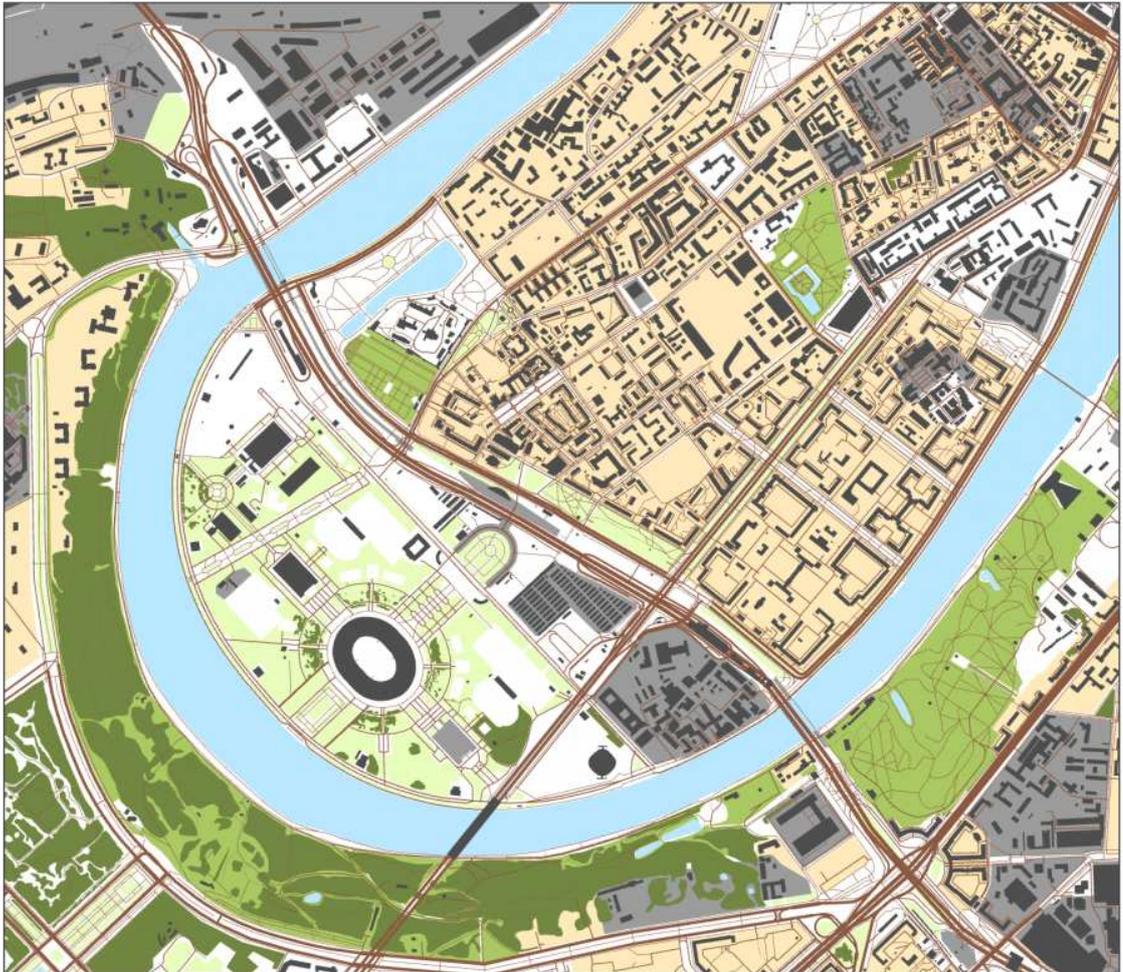
Заключительным этапом подготовки данных является приведение всей совокупности объектов к ограниченному количеству классов, ориентированных на решение задач климатического моделирования. В работе [Samsonov, Konstantinov, 2014] был предложен следующий классификатор типов подстилающей поверхности, а также запрос на языке SQL, позволяющий данный тип поверхности извлечь Таблица 2.10.

Запрос применяется к атрибутивным полям слоя подстилающей поверхности, полученного на предыдущем этапе. После выполнения каждого запроса производится заполнение поля, в котором хранится название типа подстилающей поверхности в соответствии с классификатором. В заключение следует также отметить, что поскольку объекты OSM не покрывают область данных целиком, полученный слой также содержит пустые области, которые при дальнейшей классификации должны быть отнесены к категории «Прочее». Таким образом, каждая точка анализируемой территории оказывается отнесенной к одному из возможных типов подстилающей поверхности, что позволяет перейти к вычислению их распределения по расчетным зонам (ячейкам). На Рисунке 2.62 представлено картографическое представление результата проведенной обработки данных OSM — топологически корректного покрытия с типами подстилающей поверхности.

Имея в распоряжении данные о геометрии городских каньонов и распределении типов подстилающей поверхности, можно получить комплексную характеристику городской территории для целей климатического моделирования. При этом возможно пойти как по пути анализа, рассчитав описательные статистики застройки и подстилающей по-



**Рисунок 2.61** – Стратегия «слияние-вырезание-объединение» при создании слоя подстилающей поверхности



	Здания		Высокая растительность
	Дороги		Смешанная растительность
	Селитебные территории		Кустарники
	Промышленные территории		Трава
	Прочие территории		Водные объекты

————— 500 м

**Рисунок 2.62** – Типы подстилающей поверхности по данным OSM

**Таблица 2.10** – Запросы для извлечения разных типов землепользования из OpenStreetMap

Тип поверхности	SQL-запрос
Здания (buildings)	"BUILDING" is not NULL
Дороги (roads)	"HIGHWAY" is not NULL
Трава (grass)	"LANDUSE" in ( 'grass', 'grassland', 'meadow' ) or "SURFACE" = 'grass' OR "NATURAL" = 'grassland'
Высокая растительность (high vegetation)	"LANDUSE" in ( 'forest', 'logging' ) OR "NATURAL" in ( 'tree', 'tree_row', 'wood' ) OR "WOOD" IS NOT NULL
Низкая растительность (low vegetation)	"LANDUSE" in ( 'orchard', 'vineyard', 'scrub', 'farm', 'farmland', 'greenfield' ) or "NATURAL" = 'scrub'
Смешанная растительность (mixed vegetation)	"LANDUSE" in ( 'cemetery', 'garden', 'park', 'natural_reserve' )
Твердые поверхности и грунты (surface)	"LANDUSE" = 'parking' or ("SURFACE" in ( 'asphalt', 'asphalt;compacted', 'asphalt;concrete', 'asphalt;gravel', 'brick', 'cobblestone', 'concrete', 'concrete:lanes', 'concrete:plates', 'granite', 'paved', 'paving_stones', 'pebblestone', 'stone' ) AND "HIGHWAY" IS NULL)
Водные объекты (water)	"NATURAL" in ( 'water', 'waterway', 'wetland' ) OR "WATERWAY" IS NOT NULL OR "LANDUSE" = 'reservoir'
Промышленные территории (industrial)	"LANDUSE" in ( 'construction', 'garage', 'garages', 'industrial', 'military', 'railway', 'retail', 'depot', 'commercial' )
Жилые территории (residential)	"LANDUSE" in ( 'residential', 'allotments' )

верхности в пределах заданных зон, так и по пути синтеза, сформировав зоны по принципу устойчивости сочетания застройки и подстилающей поверхности. В первом случае результатом будет являться база данных характеристик городской среды, во втором случае — локальные, так называемые климатические зоны (ЛКЗ) города. Поскольку преобладающий тип ЛКЗ также может быть включен в состав базы данных, рассмотрим решение этой задачи в первую очередь, и завершим наше изложение описанием структуры базы данных для решения задач городской климатологии.

### 2.6.1.3 Локальные климатические зоны

Локальные климатические зоны (ЛКЗ) представляют комбинации типов застройки и подстилающей поверхности, которые отличаются особым режимом процессов в приземном слое атмосферы [Stewart, Oke, 2012]. Классификация ЛКЗ показана на Рисунке 2.63. Поскольку информация об ЛКЗ также используется в метеорологическом моделировании, возникает задача выделения ЛКЗ на территорию исследования и агрегирования результатов на модельную сетку.

В настоящем исследовании было проведено построение локальных климатических зон на территорию Москвы методом пиксельной классификации [Самсонов, Тригуб, 2018]. Данная работа выполнялась в рамках реализации проекта WUDAPT [Ching et al., 2014], который призван объединить усилия специалистов со всего мира в решении задачи построения карт ЛКЗ на крупнейшие города. Охват территории, а также даты используемых снимков Landsat 8 (29 марта 2014 года и 21 сентября 2014 года) были

### Типы локальных климатических зон застройки (built-up)

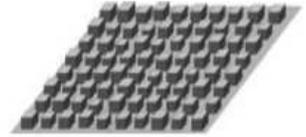
1. Компактная высокоэтажная застройка (Compact high-rise)



2. Компактная среднеэтажная застройка (Compact midrise)



3. Компактная малоэтажная застройка (Compact low-rise)



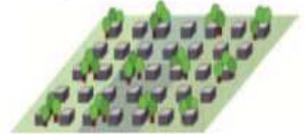
4. Открытая высокоэтажная застройка (Open high-rise)



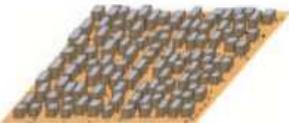
5. Открытая среднеэтажная застройка (Open midrise)



6. Открытая малоэтажная застройка (Open low-rise)



7. Облегченная малоэтажная застройка (Lightweight low-rise)



8. Крупногабаритная малоэтажная застройка (Large low-rise)



9. Рассеянная застройка (Sparsely built)

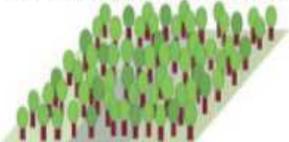


10. Тяжелая промышленность (Heavy industry)



### Типы локальных климатических зон земельного покрова (land cover)

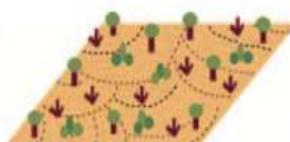
A. Частые деревья (лесные и парковые массивы) (Dense trees)



B. Редкие деревья (Scattered trees)



C. Кустарники (Bush, scrub)



D. Низкорослая растительность (Low plants)



E. Обнаженные горные породы/искусственное покрытие (Bare rock or paved)



F. Обнаженные почвы / песок (Bare soil or sand)



G. Водные объекты (Water)

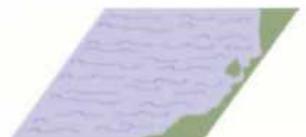


Рисунок 2.63 – Типы локальных климатических зон согласно [Stewart, Oke, 2012]

установлены координаторами проекта с целью сопоставимости карт ЛКЗ, получаемых на разные города. Разрешение исходных данных Landsat 8 составляет 30 метров. Использование разносезонных снимков при классификации методом деревьев решений позволяет использовать не только статичное состояние разных классов подстилающей поверхности, но и межсезонную изменчивость их отражательной способности при дифференциации получаемых ЛКЗ. Процесс получения ЛКЗ можно представить в виде трех этапов: набор эталонов, автоматизированная классификация, генерализация полученных зон.

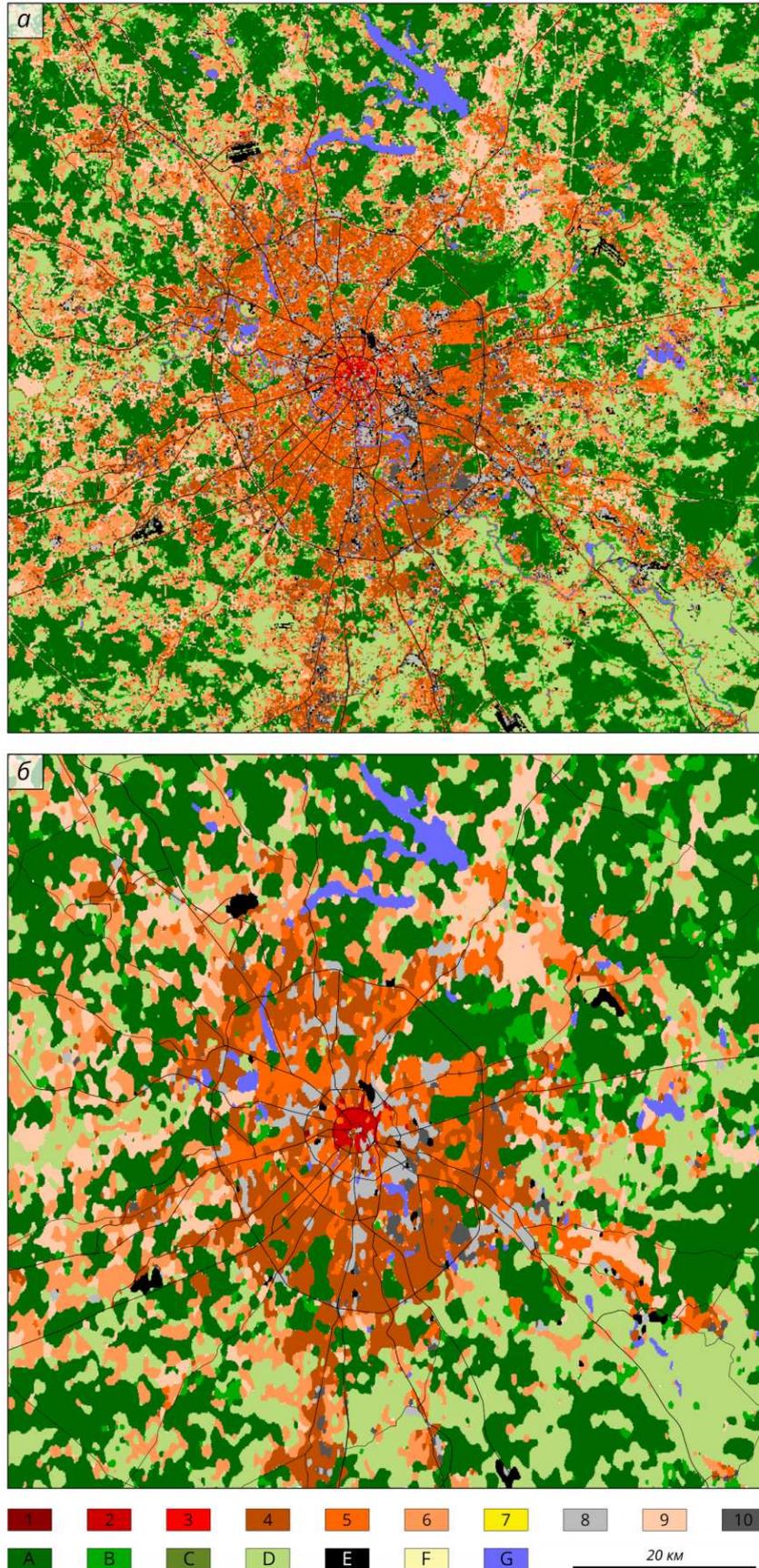
Набор эталонов для дешифрирования каждого из типов ЛКЗ производился на основе снимков сверхвысокого разрешения, доступных в Google Earth с привлечением данных об этажности строений, доступных в OpenStreetMap. Совмещение спутниковых и картографических данных о застройке позволило при наборе эталонов четко разделить классы ЛКЗ застройки. В качестве эталонных участков для каждого класса ЛКЗ было набрано по  $20 \pm 5$  полигонов шириной не менее 200 метров. Поскольку каждый класс ЛКЗ обладает некоторой вариабельностью даже в пределах одного города, большое количество эталонов позволило учесть разнообразие возможных вариантов каждого класса.

На втором этапе в программном пакете SAGA GIS использованием набранных эталонов проводилась классификация с обучением для получения локальных климатических зон на всю картографируемую территорию. Для проведения классификации использовались следующие каналы снимков Landsat 8: 1 (Coastal / Aerosol, Синий), 2 (Blue), 3 (Зеленый), 4 (Красный), 5 (Ближний инфракрасный, NIR), 6 (Коротковолновый инфракрасный, SWIR 2), 7 (Коротковолновый инфракрасный, SWIR 3); 10 (Тепловой инфракрасный, TIR1), 11 (Тепловой инфракрасный, TIR2). Перед выполнением классификации, производится передискретизация исходных данных Landsat 8 на сетку с разрешением 100 метров. Данное разрешение рекомендовано авторами методики [Bechtel *et al.*, 2015] как дающее оптимальное соотношение устойчивости и пространственной детализации выделяемых ЛКЗ. Результаты классификации имеют аналогичное разрешение.

В результате дешифрирования снимков при создании эталонов было выделено 14 ЛКЗ в пределах Московской агломерации из 17 возможных: 9 типов застройки и 5 типов наземного покрова. Из типов застройки не удалось достоверно дешифрировать облегченную низкоэтажную застройку (7), которая получается в виде спорадических пикселей, соответствующих гаражным кооперативам. Ввиду их малой площади и рассеянности, данный класс было решено не дешифрировать. Из типов наземного покрова не выделились кустарники (С) и открытые почвы/песок (F), которые не занимают сплошных однородных площадей необходимого для ЛКЗ размера. Полученная карта ЛКЗ Москвы и сопредельных территорий представлена на Рисунке 2.64.

Устойчивость полученных зон по отношению к набранным эталонам была оценена методом кросс-валидации (бутстреппинга). Классификация проведена 25 раз с использованием 50 % случайно выбранных эталонов из каждого класса. Оставшиеся 50 % этало-

## ЛОКАЛЬНЫЕ КЛИМАТИЧЕСКИЕ ЗОНЫ МОСКВЫ



**Рисунок 2.64** – Локальные климатические зоны Москвы: а – схема дешифрирования с разрешением 100 м; б – карта локальных климатических зон после генерализации методом большинства (majority). Обозначения классов даны в соответствии с Рисунком 2.63

нов сравнивались с результатом дешифрирования, при этом в пределах каждого эталонного участка оценивался преобладающий тип ЛКЗ на схеме дешифрирования. Точность классификации составила 0,776 (это означает, что в среднем 78 эталонов из 100 получили правильный тип ЛКЗ). Точность отделения ЛКЗ застройки (1–10) от ЛКЗ природного типа (А–F, за исключением E) составила 0,942. Точность взаимного различения классов застройки – 0,608. При этом классы «перепутываются» преимущественно с родственными им классами (ЛКЗ 2 и 3, ЛКЗ 4 и 5), что вполне естественно с учетом того, что в реальности нет чётких пространственных границ между зонами застройки разной высоты и этажности, а существует некий континуум, в пределах которого один тип застройки сменяет другой в бесконечном числе возможных сочетаний. Исходя из этого, полученные результаты можно признать удовлетворительными.

Детальный анализ полученных ЛКЗ, проведенный также в [Самсонов, Тригуб, 2018] показал, преимущественное соответствие характера распределения типов подстилающей поверхности в полученных зонах их определению; преобладание в Москве открытой средне- и высокоэтажной застройки; недостаточной и неравномерной обеспеченности разных типов ЛКЗ станциями метеонаблюдений, которые необходимы для детального исследования связей характеристик подстилающей поверхности и формируемых ей особенностей микроклимата каждой зоны.

#### 2.6.1.4 Агрегирование данных

Поскольку создаваемая база данных ориентирована прежде всего на использование в задачах метеорологического и климатического моделирования, она получается путем вычисления описательных статистик распределения типов подстилающей поверхности, городских каньонов и локальных климатических зон по ячейкам, структура которых определяется принимающей метеорологической моделью. Геометрически это простая операция, в рамках которой данные обрезаются контуром ячейки и производится подсчет показателей по оставшимся объектам: доли зданий в площади ячейки, средней высоты направленных каньонов, преобладающей локальной климатической зоны и т.д.

Использование сеток различного пространственного разрешения обеспечивает возможность даунскейлинга в процессе метеорологического моделирования. Изначально в исследовании [Samsonov, Konstantinov, Varentsov, 2015] были использованы сетки с разрешением 1000, 500, 200 метров, построенные в проекции UTM (Universal Transverse Mercator), зона 37, ориентированные на использование в модели URB\_MOS. В дальнейших исследованиях был также произведен расчет по ячейкам модели COSMO\_CLM [Варенцов *et al.*, 2017], также имеющим разрешение порядка 1 км, но заданным на градусной сетке и имеющим небольшой разворот относительно плоской системы координат проекции UTM. *Изменение параметров сетки никак не влияет на порядок расчетов с точки зрения геоинформационных алгоритмов, данные могут быть извлечены*

при любом заданном разрешении.

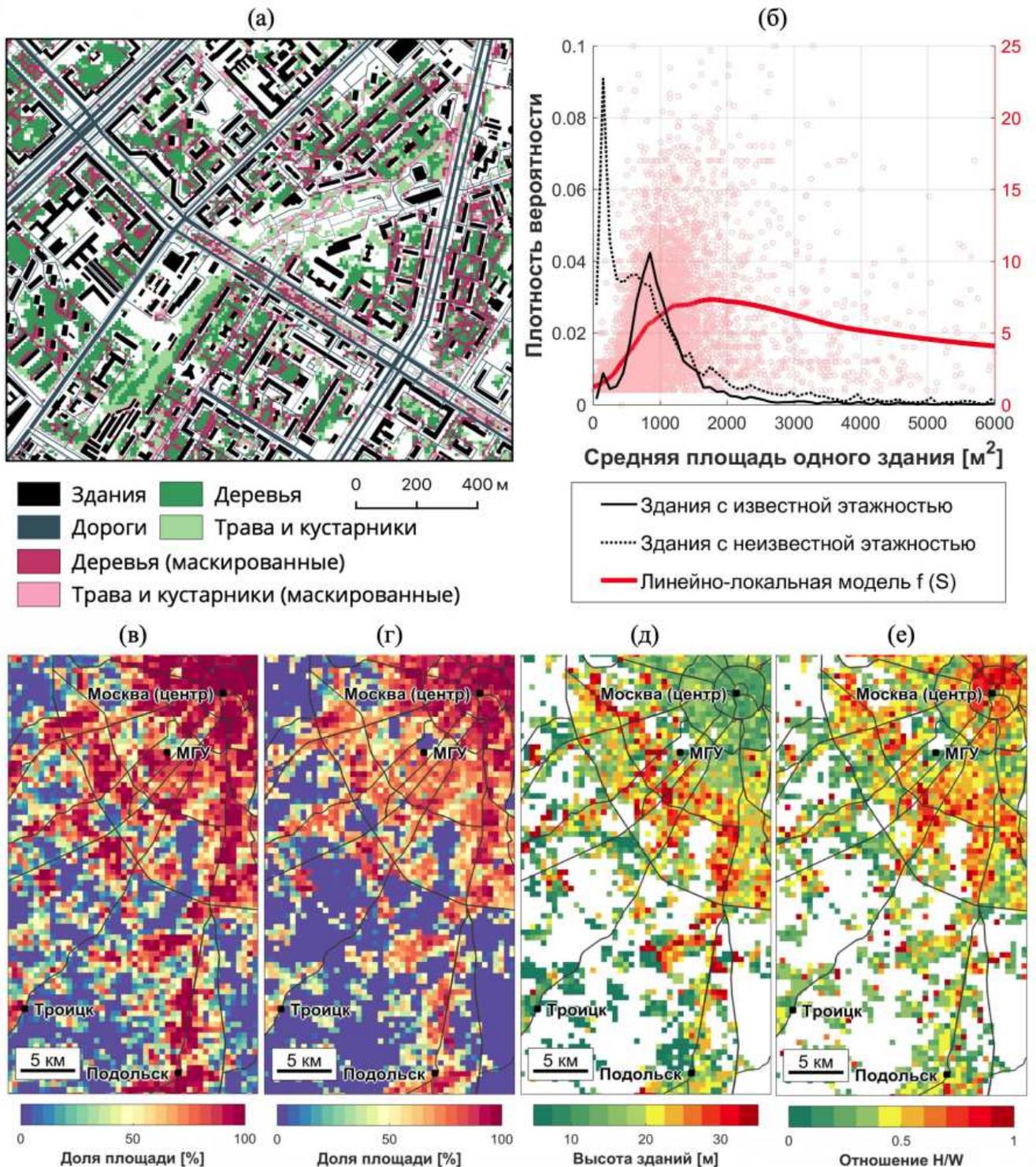
При использовании данных о типах подстилающей поверхности, полученных вышеописанным методом, возникает проблема с оценкой доли площади урбанизированной территории  $\lambda_{\text{urb}}$ , которая не может быть явно определена по данным OSM, что вынуждает прибегать к ряду эмпирических допущений [Варенцов *et al.*, 2017]. Чтобы этого избежать, для оценки  $\lambda_{\text{urb}}$  использованы данные Copernicus Global Land Cover (CGLC) [Buchhorn *et al.*, 2020].

Уникальной особенностью CGLC является кодирование доли площади, занятой различными типами наземного покрова, в том числе застроенной территорией. Однако, такая территория по данным CGLC включает растительность внутри застройки (газоны, дворовое озеленение и т.д.). Поскольку городская поверхность в TERRA\_URB считается полностью запечатанной, необходима коррекция данных CGLC путем исключения растительности из площади урбанизированной территории. Для идентификации растительности в городской среде использованы спутниковые данные Sentinel-2 (снимок за 28 августа 2019 г.). По пороговому значению индекса NDVI (Normalized Difference Vegetation Index)  $> 0.6$  определены пиксели растительности, а затем произведено их разделение на деревья и траву/кустарники по пороговому значению яркости зеленого канала ( $\leq 500$  для деревьев и  $> 500$  для травы/кустарников).

Предложенный алгоритм определения необходимых параметров включает три этапа [Samsonov, Varentsov, 2020]:

1. Подготовка данных OSM инструментами: конвертация линейных объектов дорожно-уличной сети в полигональные с учетом класса и количества полос, генерализация типов поверхности, устранение их взаимных пересечений согласно методике, описанной в предыдущем параграфе.
2. Маскирование (выделение) пикселей растительности, пересекающихся с дорогами и зданиями из-за их перекрытия кронами и в силу дискретности данных (Рисунок 2.65а). Такие пиксели отмечаются как не относящиеся к растительному покрову.
3. Определение набора промежуточных характеристик (Таблица 2.11) по ячейкам сетки (нами использована сетка модели COSMO с шагом 500 м размером  $200 \times 200$  км, включающая Москву и большую часть Московской области) и расчет на их основе искомых параметров.

Значения  $\lambda_{\text{urb}}$  определены по данным CGLC, скорректированным с учетом информации о растительности и застройке. Учитывалось, что  $\lambda_{\text{urb}}$  не может превышать долю площади, свободной от растительности, и не может быть меньше доли площади дорог и зданий:



**Рисунок 2.65** — Примеры: (а) маскирования растительности, пересекающейся со зданиями и дорогами; (б) плотность распределения вероятности зданий с различной площадью с известной и неизвестной этажностью, и зависимость этажности от площади одного здания; (в) доля площади урбанизированной территории по исходным и (г) скорректированным данным CGLC; (д) средняя высота зданий; (е) среднее отношение Н/В на сетке с шагом 500 м для сегмента области исследования (юго-запад Москвы). Темно-серыми линиями на (в-е) показаны основные автомобильные магистрали

**Таблица 2.11** – Параметры для ячеек модели, полученные на основе геоинформационного анализа

Обозначение в тексте	Исходные данные	Описание параметра
$\lambda_{\text{CGLC URB}}$	CGLC	доля площади урбанизированной (застроенной) территории
$\lambda_{\text{OSM BLD}}$	OSM	доля площади зданий
$\lambda_{\text{OSM ROADS}}$	OSM	доля площади дорог
$\lambda_{\text{OSM GREEN}}$	OSM	доля площади зеленых насаждений по данным
$L_{\text{OSM}}$	OSM	средняя этажность зданий, для которых она известна
$\lambda_{\text{OSM LV}}$	OSM	доля площади зданий с известной этажностью в их общей площади
$N_{\text{OSM BLD}}$	OSM	суммарное число зданий в ячейке сетки
$N_{\text{OSM LV}}$	OSM	число зданий с известной этажностью в ячейке сетки
$\lambda_{\text{SENT GRASS}}$	Sentinel-2 и OSM	доля площади, занятая травяной растительностью
$\lambda_{\text{SENT TREES}}$	Sentinel-2 и OSM	доля площади, занятая древесной растительностью

$$\lambda_{\text{urb}} = \max \left[ \lambda_{\text{OSM BLD}} + \lambda_{\text{OSM ROADS}}; \min \left[ \lambda_{\text{CGLC URB}}; 1 - \max \left[ \lambda_{\text{OSM GREEN}}; \lambda_{\text{SENT GRASS}} + \lambda_{\text{SENT TREES}} \right] \right] \right]. \quad (2.28)$$

Высота зданий определена по их этажности  $L_{\text{OSM}}$ , доступной как один из атрибутов в OSM. Однако она известна лишь для 20% зданий в пределах области исследования. Как правило, этажность не известна для частных домов, гаражей и технических сооружений небольшой высоты и площади (Рисунок 2.65б). Для восстановления информации об этажности сначала в пределах каждой ячейки сетки определена средняя площадь здания с известной этажностью:

$$S_{\text{LV}} = \frac{S_0 \cdot \lambda_{\text{OSM BLD}} \cdot \lambda_{\text{OSM LV}}}{N_{\text{OSM LV}}}, \quad (2.29)$$

где  $S_0$  – площадь ячейки сетки. Аналогично определена средняя площадь здания с неизвестной этажностью  $S_{\text{noLV}}$ . По данным  $L_{\text{OSM}}$  и  $S_{\text{LV}}$  получена модель  $L = f(S)$ , представляющая собой линейную локально-регрессионную зависимость этажности здания от его площади (Рисунок 2.65б). На ее основе по данным  $S_{\text{noLV}}$  восстановлена этажность зданий, для которых она была не известна.

Несмотря на большой разброс точек на Рисунок 2.65б, она позволяет отделить малоэтажную застройку с площадью здания 100 – 300 м<sup>2</sup> от более высотной. Средняя высота зданий  $H$  определена при допущении, что высота одного этажа  $H_L = 3$  м, высота фундамента и чердака  $H_0 = 2$  м:

$$H = H_L \cdot [L_{\text{OSM}} \cdot \lambda_{\text{OSM LV}} + f(S_{\text{noLV}}) \cdot (1 - \lambda_{\text{OSM LV}})] + H_0. \quad (2.30)$$

Для оценки отношения аспекта городских каньонов ( $H/W$ ) также необходимо оце-

нить их среднюю ширину  $W$ . Оригинальная методика идентификации каньонов в реальной застройке [Samsonov, Konstantinov, Varentsov, 2015] использовалась для оценки их ширины в предшествующих работах авторов. Альтернативные методы, часто используемые в зарубежной практике [Masson et al., 2020], основаны на более простых аналитических соображениях. Например, при допущении, что здания в ячейке сетки расположены регулярным образом в пределах ее урбанизированной части, имеют одинаковый размер и квадратную форму, ширина каньона определяется как:

$$W = \frac{\sqrt{S_0} \cdot (\sqrt{\lambda_{urb0}} - \sqrt{\lambda_{OSM BLD}})}{\sqrt{N_{OSM BLD}}}, \quad (2.31)$$

где  $\lambda_{urb0} = \max[\lambda_{CGLC URB}; \lambda_{OSM BLD} + \lambda_{OSM ROADS}]$ . Последняя использована вместо  $\lambda_{urb}$ , т.к. исключение из площади урбанизированной территории растительности ведет к уменьшению ширины заполняемых ею каньонов и искажению геометрических характеристик застройки. Сравнение оценки  $W$  по Формуле 2.31 с шириной каньонов, идентифицированных в реальной застройке по методике из [Samsonov, Konstantinov, Varentsov [2015]], показало хорошее согласие с коэффициентом корреляции 0.81 (на примере наиболее урбанизированных ячеек сетки с шагом 1 км). Поэтому, а также по причине невысокой чувствительности модели к параметру  $H/W$  [Wouters et al., 2016], для его определения на сетке с шагом 500 м использована оценка  $W$  по формуле 2.31, как более простая для массового расчета. Доля площади зданий в доле площади урбанизированной территории,  $\lambda_{bid}$ , определена как  $\lambda_{OSM BLD}/\lambda_{urb0}$  (использование  $\lambda_{urb0}$  объяснено выше).

Пространственное распределение полученных значений  $\lambda_{urb}$ ,  $H$  и  $H/W$  показано на Рисунке 2.65в-е. для фрагмента территории. Неопределенность полученных оценок зависит от ряда факторов, включающих как точность исходных данных, так и условность применяемых в методике допущений и пороговых значений. Исходные данные OpenStreetMap были верифицированы ранее путем сравнения с эталонной информацией из официальных источников [Samsonov, Konstantinov, 2014], которое показало их согласие в части представления застройки. Поскольку эталонные величины для производных параметров, как и пороговых значений, используемых при их вычислении, отсутствуют, возможна лишь их качественная оценка, в том числе в форме сравнения с альтернативными источниками данных. Полученные значения  $\lambda_{urb}$  и  $Q_H$  демонстрируют более детальную и реалистичную картину по сравнению с используемыми в COSMO по умолчанию значениями из глобальных баз данных [Варенцов et al., 2017]. Значения  $\lambda_{urb}$  лучше отражают запечатанность по сравнению с данными CGLC (Рисунок 2.65в, г). Так,  $\lambda_{CGLC URB}$  достигает 100% для спальных районов на юго-западе города и более 50% для ряда загородных территорий с дачной застройкой. Значения  $\lambda_{urb}$  для таких территорий меньше и составляют 70 – 80% и 10 – 30% соответственно, что представляется более реалистичным. Для параметров  $H$  и  $H/W$  альтернативные оценки отсутствуют, можно

лишь отметить соответствие пространственной картины ожиданиям: значения  $H$  минимальны в историческом центре и в пригородах, отношение  $H/W$  максимально в центре и в районах с новой высотной застройкой (Рисунок 2.65д,е). О качестве полученных данных, влиянии их неопределенности, можно будет судить и по итогам анализа результатов моделирования с их использованием.

## 2.6.2 Результаты

Итогом геоинформационных исследований структуры городской среды является база пространственных данных, включающая необходимые характеристики застройки, подстилающей поверхности, локальных климатических зон, методика расчета которых была рассмотрена в настоящей главе. Состав полей базы данных представлен в Таблице 2.12 и Таблице 2.13. Все показатели (42 шт.) условно разделены на 5 разделов: общая информация, административная структура, структура подстилающей поверхности, параметры застройки, локальные климатические зоны.

*Общая информация* (6 параметров) содержит данные о геометрии самой ячейки, ее координатах и средней высоте, полученной по цифровой модели рельефа SRTM. Раздел *административной структуры* (3 параметра) включает в себя информацию о принадлежности определенных долей ячейки к населенным пунктам трех возможных классов: крупному городу, городу и сельскому населенному пункту. Данная информация нужна для того чтобы в случае нехватки данных о городской застройке и подстилающей поверхности можно было как-то аппроксимировать эти данные в процессе моделирования. Например, если ячейка принадлежит сельскому населенному пункту, но ни одно из зданий не имеет известной этажности, логично сделать предположение, что все здания имеют высоту  $2 \pm 1$  этажа, что уже существенно лучше полного отсутствия данных об этажности. Раздел *структуры подстилающей поверхности* (9 параметров) содержит информацию о соотношении типов подстилающей поверхности в пределах ячейки, полученных по данным OpenStreetMap. Раздел *параметров застройки* (22 параметра) наиболее массивен и содержит детальную информацию о геометрии городской среды — параметрах зданий и городских каньонов. Высоты зданий получены из этажности OpenStreetMap путем их умножения на 3. Расстояние от пола до потолка жилого здания согласно СП 54.13330.2011 «Здания жилые многоквартирные» (СНиП 31-01-2003) в климатических условиях Москвы должна составлять не менее 2,7 м, а для нежилых помещений согласно СНиП 31-06-2008 «Общественные здания и сооружения» (СНиП 2.08.02-89) — не менее 3 м. Таким образом, используемая формула дает минимально возможную высоту здания при заданной этажности. Наконец, раздел *локальных климатических зон* (2 параметра) включает в себя тип наиболее представительной ЛКЗ в пределах данной ячейки, а также долю, которую занимает данная ЛКЗ.

**Таблица 2.12** – Структура полей базы данных характеристик подстилающей поверхности и геометрии городской застройки (характеристики 1-18)

№	Название поля	Содержание
Общая информация о ячейке		
1	ID	Идентификатор
2	LAT	Широта центра (WGS84)
3	LONG	Долгота центра (WGS84)
4	X	Прямоугольная координата X центра (проекция UTM37)
5	Y	Прямоугольная координата Y центра (проекция UTM37)
6	ZMEAN	Средняя высота рельефа
Административная структура (% от площади ячейки)		
7	CITY	Большой город (более 100 тыс.чел.)
8	TOWN	Средний или малый город, ПГТ (менее 100 тыс.чел.)
9	VILLAGE	Сельский населенный пункт
Структура подстилающей поверхности (% от площади ячейки)		
10	BUILDINGS	Здания
11	ROADS	Дороги
12	GREENTALL	Высокая растительность (лес)
13	GREENLOW	Низкая растительность (трава, кустарники)
14	GREENMIXED	Смешанная растительность (состав и структура неизвестны)
15	INDUSTRIAL	Промышленная территория
16	RESIDENTIAL	Жилая территория
17	WATER	Водные поверхности и болота
18	OTHER	Прочее (тип поверхности неизвестен)

**Таблица 2.13** – Структура полей базы данных характеристик подстилающей поверхности и геометрии городской застройки (характеристики 19-45)

№	Название поля	Содержание
Параметры застройки (здания и городские каньоны)		
19	BLDCOUNT	Общее количество зданий
20	LEVELS	Средняя этажность зданий
21	LEVELSRATIO	Доля зданий с известной этажностью (%)
22	HEIGHTS	Средняя высота зданий
23	SVF	Средний индекс открытости небосвода (Sky View Factor)
24	MDCHEIGHT	Средняя высота направленного каньона
25	MUCHEIGHT	Средняя высота ненаправленного каньона
26	MDCWIDTH	Средняя ширина направленного каньона
27	MUCWIDTH	Средняя ширина ненаправленного каньона
28	MDCRATIO	Доля направленных каньонов в площади ячейки
29	MUCRATIO	Доля ненаправленных каньонов в площади ячейки
30	MDCHWRATIO	Среднее соотношение высоты и ширины для направленных каньонов
31	MUCHWRATIO	Среднее соотношение высоты и ширины для ненаправленных каньонов
32	HWRATIO	Среднее соотношение высоты и ширины для всех каньонов
33	DIR1	Основное направление направленных каньонов
34	DIR2	Второстепенное направление направленных каньонов
35	DIRRATIO	Соотношение частот основного и второстепенного направлений
36	FINDEX1	Фронтальный индекс для направлений С и Ю
37	FINDEX2	Фронтальный индекс для направлений ССВ и ЮЮЗ
38	FINDEX3	Фронтальный индекс для направлений СВ и ЮЗ
39	FINDEX4	Фронтальный индекс для направлений ВСВ и ЗЮЗ
40	FINDEX5	Фронтальный индекс для направлений В и З
41	FINDEX6	Фронтальный индекс для направлений ВЮВ и ЗСЗ
42	FINDEX7	Фронтальный индекс для направлений ЮВ и СЗ
43	FINDEX8	Фронтальный индекс для направлений ЮЮВ и ССЗ
Локальные климатические зоны (ЛКЗ)		
44	LCZTYPE	Преобладающая ЛКЗ в ячейке
45	LCZRATIO	Доля ячейки, занятая преобладающей ЛКЗ

### 2.6.3 Итоги

В настоящем разделе диссертации предложена методика расчета параметров городской среды основанная на геоинформационном анализе открытой картографической и спутниковой информации. С ее использованием получен набор параметров для Московского региона на сетках с шагом 200, 500 и 1000 м. Использование исходных данных с глобальным покрытием позволяет легко применить методику и для других регионов и расчетных сеток.

Полученные с использованием предложенной методики сеточные базы данных характеристик подстилающей поверхности и геометрии городской застройки получили применения в целом ряде исследований. Так, в работе [Varentsov *et al.*, 2020] использование полученной базы данных позволило сократить среднеквадратическую ошибку прогноза температуры на 12 – 18% в среднем и более чем на 30% для урбанизированных типов подстилающей поверхности по сравнению со стандартными параметризациями. В статье [Varentsov *et al.*, 2021] показано, что пространственная изменчивость городского острова тепла определяется как локальными (сотни метров), так и мезомасштабными (2 – 20 км) факторами. А исследование [Kosheleva *et al.*, 2023] впервые выявило, что накопление дорожной пыли имеет статистически значимую связь с характеристиками городских каньонов разных пространственных масштабов, достигая максимума в коротких каньонах ( $L < 500$ ). Полученные результаты нашли также свое внедрение в оперативной системе численного прогноза погоды Гидрометцентра [Rivin *et al.*, 2019; Rivin *et al.*, 2020].

## 2.7 Результаты и выводы

Во второй главе диссертации были представлены оригинальные методики адаптивной генерализации пространственных данных для целей картографирования, позволяющие учитывать пространственное распределение, вертикальную и горизонтальную связность, форму, пространственную структуру и пространственные соотношения, проявляющиеся в географической среде. Разнообразие представленных примеров (населенные пункты, дороги, административные границы, рельеф, земельный покров, городская застройка) наглядно демонстрируют широту тематического охвата разработанных методов и технологий.

Основные выводы по Главе 2 заключаются в следующем:

1. Районирование территории по густоте размещения населенных пунктов и дорожной сети для целей генерализации может быть осуществлено на основе диаграммы Вороного (вспомогательного разбиения пространства) и поверхности плотности распределения, вычисленных по цифровым моделям соответствующих объектов. Автоматизированный отбор (прореживание) с

индивидуальными критериями для каждого района, а также с привлечением методов сетевого анализа позволяет сохранить и подчеркнуть территориальные контрасты.

2. Использование графового подхода, а также сегментации линейных объектов на морфологически однородные участки в сочетании со специализированными алгоритмами геометрического упрощения позволяет обеспечить воспроизведение пространственных отношений линейных объектов, которые могут быть распознаны по их цифровым моделям.
3. Структурный подход является ключом к морфологически корректной генерализации цифровых моделей рельефа, при которой сохраняются крупнейшие положительные и отрицательные формы рельефа, а также появляется возможность их преувеличения на основе адаптивной статистической фильтрации для обеспечения читаемости в мелких масштабах картографирования.
4. Агрегирование соотношения площадей разных типов пространственных объектов по ячейкам регулярной сети при использовании разнородных пространственных данных в качестве источников требует их взаимного согласования и устранения перекрытий методами оверлея для обеспечения корректных значений.

Обобщение этих выводов позволяет сформулировать второе защищаемое положение:

#### **Защищаемое положение №2**

*Адаптивность* генерализации пространственных данных обеспечивает сохранение и воспроизведение пространственного распределения, пространственных отношений, характера формы и основных структурных элементов картографируемых объектов, а также соотношения их площадей, представленных в исходных цифровых моделях. Реализация принципа адаптивности основывается на включении в геоинформационные алгоритмы генерализации операций морфометрического, сетевого, оверлейного и статистического анализа, а также вспомогательных разбиений пространства, определяемых множеством генерализуемых объектов.

МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ  
имени М. В. ЛОМОНОСОВА

ГЕОГРАФИЧЕСКИЙ ФАКУЛЬТЕТ

*На правах рукописи*

**Самсонов Тимофей Евгеньевич**

**Генерализация пространственных данных  
и ее картографические приложения**

1.6.20 — Геоинформатика, картография

ДИССЕРТАЦИЯ  
на соискание учёной степени  
доктора географических наук

**ТОМ 2**

Москва — 2025

# Содержание

<b>3 Детальность пространственных данных . . . . .</b>	<b>4</b>
3.1 Плотность пространственных объектов на цифровых топографических картах и ее связь с географическим положением . . . . .	4
3.2 Гранулярность цифровых моделей рельефа и ее связь с масштабом картографирования . . . . .	28
3.3 Контроль детальности линейных пространственных объектов при обработке разными алгоритмами геометрического упрощения . . . . .	56
3.4 Повышение детальности изображения при визуализации цифровых моделей рельефа с помощью дополнительных горизонталей . . . . .	75
3.5 Картографическая визуализация цифровых моделей рельефа с помощью штрихов и управление его детальностью . . . . .	99
3.6 Конфляция гидрографических данных и цифровых моделей рельефа разной детальности . . . . .	116
3.7 Взаимная координатная привязка цифровых геоизображений разной детальности . . . . .	154
3.8 Результаты и выводы . . . . .	170
<b>4 Мультимасштабное картографирование . . . . .</b>	<b>173</b>
4.1 Основные понятия и принципы мультимасштабного картографирования .	173
4.2 Вычисление параметров визуализации цифровых моделей дорожной сети и структуры населенных пунктов на мультимасштабных картах . . . . .	178
4.3 Мультимасштабное гипсометрическое картографирование на основе цифровых моделей рельефа . . . . .	188
4.4 Мультимасштабная визуализация цифровых моделей характеристик ветрового волнения . . . . .	206
4.5 Построение цифровых моделей полимагистралей при мультимасштабном картографировании транспортных сетей . . . . .	223

4.6	Визуализация пространственных данных способами картограмм и карто- диаграмм на мультимасштабных картах . . . . .	233
4.7	Результаты и выводы . . . . .	249
<b>Заключение . . . . .</b>		<b>251</b>
	Публикации по теме диссертации . . . . .	253
<b>Список литературы . . . . .</b>		<b>262</b>
<b>Приложение . . . . .</b>		<b>315</b>

## 3 Детальность пространственных данных

В главе раскрыт принцип *детальности*, который определяет необходимость оценки подробности цифровых моделей географических объектов не масштабом, а количественными характеристиками их геометрии и атрибутов. Исследованы три основных аспекта детальности пространственных данных и карт: количественная оценка детальности, ее управляемое изменение, а также интеграция данных и карт разной детальности. Разработанные подходы повышают объективность процесса картографической генерализации, позволяют алгоритмизировать получение пространственных данных и карт необходимой подробности, а также совместное использование гетерогенных данных, различающихся своей детальностью.

**Оценка** детальности пространственных данных исследована для таких ключевых ее параметров как плотность и гранулярность.

**Управление** детальностью исследовано на геометрическом уровне (при генерализации пространственных данных) и на уровне изображения (при визуализации цифровых моделей рельефа).

**Интеграция** разнородных разнодетальных пространственных данных рассмотрена на примере конфляции (пообъектного совмещения) и координатной привязки.

### 3.1 Плотность пространственных объектов на цифровых топографических картах и ее связь с географическим положением<sup>1</sup>

**Плотность**, или количество информации на единицу площади, является одной из общепринятых характеристик детальности [Ruas, Bianchin, 2002; Lemmens, 2011]. Она может относиться как непосредственно к самим объектам, так и к их элементам (например, числу вершин в составе линий и полигонов). При этом разнообразие географических ландшафтов приводит к существенным различиям в рисунке местности и, как следствие, необходимости в пространственной дифференциации используемых методов картографической генерализации [Touya, Duchêne, Ruas, 2010; Mitropoulos, Nakos,

---

<sup>1</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov et al., 2023]

2011]. Плотность данных также является отражением сложности ландшафта, а это означает, что ее изменения должны наблюдаться не только по вертикали (при смене уровней детализации), но также по горизонтали (в пространстве). До настоящего момента исследования горизонтальной изменчивости плотности пространственных данных и ее факторов не получили достаточного внимания. С целью решения этой проблемы был разработан метод, который позволяет оценивать и предсказывать изменения в плотности пространственных данных в зависимости от характеристик местоположения [Якимова *et al.*, 2021; Samsonov *et al.*, 2023].

### 3.1.1 Методика

#### 3.1.1.1 Общие принципы

В основе метода лежит понятие относительной плотности — отношения плотности выбранного участка к средней плотности для всего уровня детализации. Регрессионная модель между относительной плотностью точек, линий, пересечений в базе пространственных данных и предикторами (характеристиками местности, полученными по альтернативным источникам) позволяет получить нормированное (приведенное к среднему) значение и тем самым получить оценку наиболее вероятного подходящего уровня детализации для заданного набора данных (Рисунок 3.1).

Пусть задана некая метрика плотности данных и вычислено ее значение  $d$  для каждого из фрагментов обучающей выборки. Поскольку ожидаемое значение  $d$  варьируется по территории, нормируем его на среднее по всем фрагментам  $\bar{d}$ , чтобы получить величину относительной плотности  $\hat{d}$  (Рисунке 3.1б):

$$\hat{d} = d/\bar{d}. \quad (3.1)$$

Относительная плотность показывает, насколько текущий фрагмент более плотно насыщен пространственными данными на фоне всего уровня детализации.

Представим относительную плотность как функцию от параметров географического положения:

$$g_i(\hat{d}_i) = \sum_j \beta_{ij} f_{ij}(l_j) + \epsilon, \quad (3.2)$$

где  $\hat{d}_i$  — относительная величина  $i$ -й метрики плотности,  $l_j$  — величина  $j$ -й характеристики местоположения,  $\beta_{ij}$  — коэффициенты,  $\epsilon$  — ошибка,  $f$  и  $g$  — функции преобразования, которые приводят зависимость между  $\hat{d}_i$  и  $l_j$  к линейному виду (например, логарифмирование).

Далее введем следующее предположение:

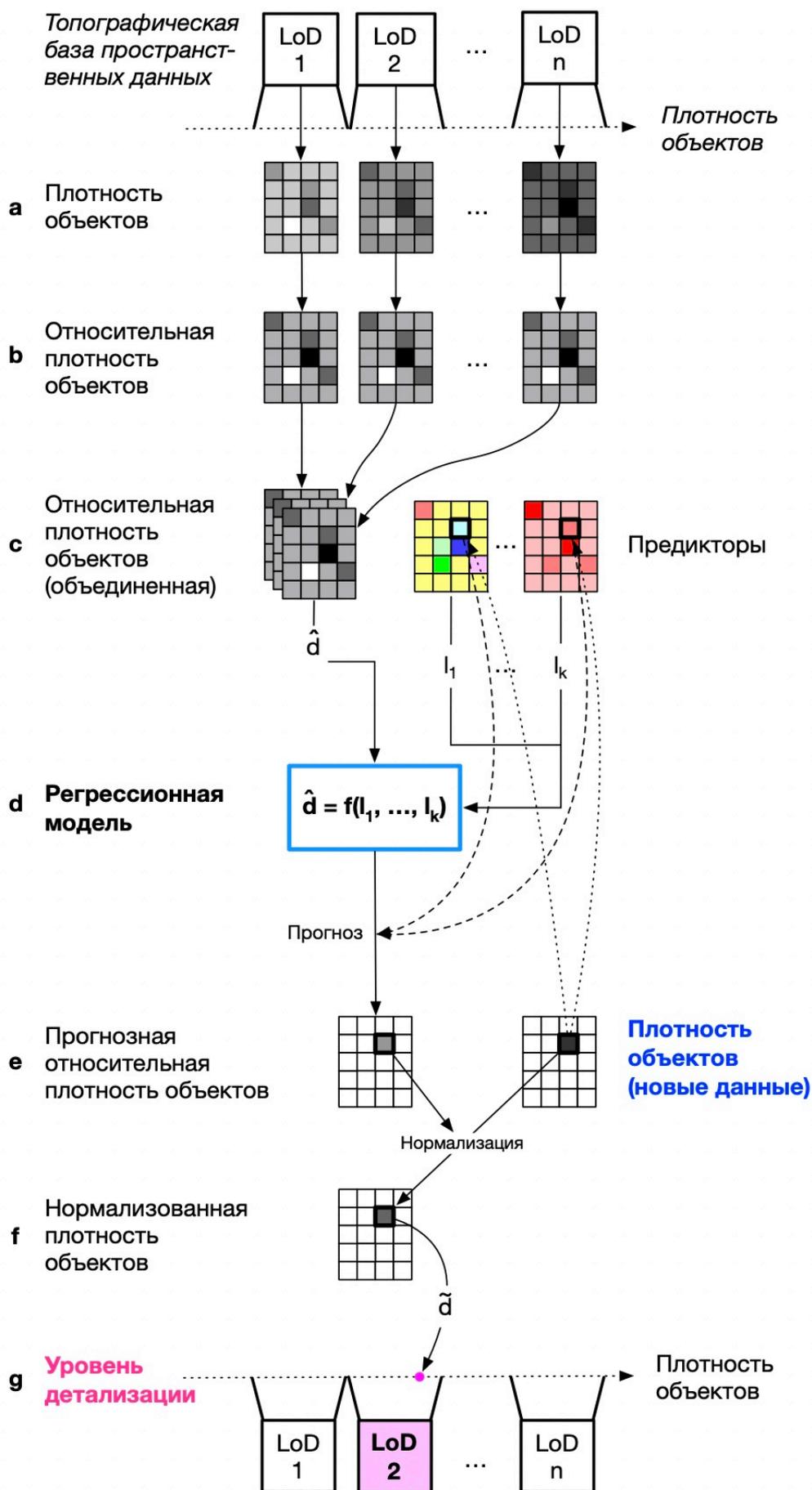


Рисунок 3.1 – Технология определения уровня детализации для данных топографических карт [Samsonov et al., 2023]

*Пространственное распределение относительной плотности одинаково для разных уровней детализации*

Это предположение позволяет объединить относительные плотности для разных уровней детализации в одну выборку (Рисунок 3.1в) и затем использовать для определения коэффициентов модели.

Имея значения  $l_j$  для заданного фрагмента пространственных данных, можно вычислить значение  $\widehat{d}_i$  путем взятия обратной функции:

$$\widehat{d}_i = g^{-1} \left( \sum_j \beta_{ij} f(l_j) + \epsilon \right). \quad (3.3)$$

Например, если модель имеет вид  $\ln(\widehat{d}_1 + 1) = 0.74 \ln(l_1 + 1) + 0.2$  и для нового фрагмента данных  $l_1 = 2$ . В этом случае предсказанное моделью значение будет в  $\widehat{d}_1 = e^{0.74 \ln 3 + 0.2} - 1 \approx 1.753$  раз выше чем среднее  $\overline{d}_1$  по уровню детализации.

Предсказанное значение относительной плотности  $\widehat{d}_i$  позволяет, имея в распоряжении ее абсолютную величину  $d_i$ , получить нормированное значение  $\widetilde{d}_i$ :

$$\widetilde{d}_i = d_i / \widehat{d}_i. \quad (3.4)$$

Если Модель 3.3 эффективна, то ожидается, что  $\widetilde{d}_i \approx \overline{d}_i$  для данных с соответствующего уровня детализации. Поскольку реальные значения будут отклоняться от истинных, для оценки вероятности принадлежности данных определенному уровню детализации можно использовать  $Z$ -значение (при условии, что данные распределены нормально):

$$z_i = \frac{\widetilde{d}_i - \overline{d}_i}{s_i}, \quad (3.5)$$

где  $s_i$  — стандартное отклонение величин  $i$ -й метрики плотности.

Если в базе пространственных данных есть несколько уровней детализации, то данные будут рекомендованы для включения в уровень  $k$ , если для любого  $m \neq k$  выполняется условие:

$$|z_{ik}| < |z_{im}|, \quad (3.6)$$

где  $z_{in}$  —  $Z$ -значение набора данных относительно распределения метрики плотности для  $n$ -го уровня детализации.

Последняя гипотеза является основной идеей, проверяемой в настоящем исследовании.

### 3.1.1.2 Частная реализация

Для данного исследования в качестве объектов, плотность которых моделируется, были выбраны точки, линии и перекрестки. Соответствующие показатели плотности рассчитываются следующим образом:

- *Плотность точек ( $d_P$ )*. Каждый элемент пространственных данных преобразуется в точечные элементы. Для линейных и многоугольных объектов выделяются их вершины. Общее количество полученных точек делится на площадь, покрытую данными.
- *Плотность линий ( $d_L$ )*. Общая длина всех линейных объектов и общий периметр полигональных объектов суммируются, а затем делятся на площадь, покрытую данными.
- *Плотность пересечений ( $d_I$ )*. Рассчитывается наложение всех линейных и полигональных объектов. Результирующая геометрия задается как точечная — это означает, что выводятся все пересечения между линейными и полигональными слоями. Количество пересечений делится на площадь, покрытую данными.

Если плотность точек и линий характеризует общее количество пространственных данных, покрывающих территорию, то плотность пересечений отражает сложность топологических связей между объектами в базе данных: большее количество пересечений указывает на более сложную структуру объектов.

Свойства местоположения выступают в качестве предикторов плотности. Для этой цели рассматривались две группы мер:

- *Непространственные меры* описывают общие свойства местоположения и не учитывают форму и пространственную структуру географических объектов, покрывающих территорию. В частности, были использованы соотношения площадей, занимаемых различными классами почвенно-растительного покрова, такими как водный, лесной, городской и другие.
- *Пространственные показатели* характеризуют местоположение через геометрию и форму географических объектов, покрывающих территорию. Для этого использовались меры сложности ландшафта (совместная энтропия, индекс заражения, фрактальная размерность) и индексы расчлененности рельефа, описанные ниже.

Совместная энтропия описывает общую сложность ландшафтного узора [Nowosad, Stepinski, 2019]:

$$\text{jointent} = - \sum_{i=1}^K \sum_{j=1}^K p_{ij} \log_2 p_{ij}, \quad (3.7)$$

где  $p_{ij}$  — вероятность того, что  $i$ -й и  $j$ -й классы наблюдаются в соседних растровых

ячейках, а  $K$  — общее количество классов наземного покрова.

Индекс распространенности рассчитывается аналогичным образом, но учитывает количество классов [Riitters *et al.*, 1996]:

$$\text{contag} = 1 + \frac{\sum_{i=1}^K \sum_{j=1}^K p_{ij} \ln p_{ij}}{2 \ln K}. \quad (3.8)$$

Данный индекс указывает на вероятность принадлежности двух случайно выбранных ячеек одному классу.

Индекс *пересеченности рельефа* — это среднее квадратическое разности значений высот между ячейками окна  $3 \times 3$  и центральной ячейкой  $z_{00}$ , рассчитанное по растровой цифровой модели рельефа [Riley, De Gloria, Elliot, 1999]:

$$\text{tri} = \sqrt{\sum_{i=-1}^1 \sum_{j=-1}^1 (z_{ij} - z_{00})^2}, \quad (3.9)$$

где  $z_{00}$  — центральная ячейка плавающего окна.

## 3.1.2 Апробация

### 3.1.2.1 Подготовка данных

Экспериментальная оценка метода проводилась на мультимасштабной топографической базе данных с 3 уровнями детализации, соответствующими масштабам 1 : 200 000, 1 : 500 000 и 1 : 1 000 000 (далее — УД 200, 500 и 1000). База данных представляет собой слои цифровых топографических карт России соответствующих масштабов, составленных Федеральной службой государственной регистрации, кадастра и картографии (Росреестр) путем генерализации более крупномасштабных карт. Каждый УД представлен в формате хранения базы геоданных *Esri* и содержит 47/47/40 слоев для УД 200/500/1000 соответственно. Слои в каждой БД сгруппированы в восемь наборов данных: административные границы (3/2/2 слоя), экономические (7/9/6 слоев), геодезия (3/3/2 слоя), гидрография (11/11/10 слоев), рельеф (5/5/5 слоев), населенные пункты (6/5/4 слоя), транспорт (6/6/6 слоев) и растительность/земля (6/6/5 слоев). Количество слоев незначительно отличается между УД, поскольку некоторые типы объектов удаляются или добавляются между масштабами.

Из каждого уровня детализации было извлечено 33 фрагмента выборки, центрированных на населенных пунктах, расположенных в различных географических условиях, в результате чего всего было получено 99 фрагментов данных. Каждый фрагмент был обрезан по прямоугольнику  $100 \times 100$  км, а затем спроецирован в *азимутальную равновеликую проекцию Ламберта* с соответствующим центром. Данная проекция была выбрана потому, что она позволяет каждому фрагменту покрыть одинаковую площадь и сохранить

соотношение площадей различных типов наземного покрова. Другие искажения (углов, длин) незначительны в пределах каждого фрагмента. Карта расположения фрагментов выборки представлена на Рисунке 3.2. Фрагменты были разделены на обучающую и тестовую выборки, что описано далее в разделе «Экспериментальная работа». О разнице в плотности данных между фрагментами можно судить по Рисунку 3.3. Хорошо видно, что высокоурбанизированный фрагмент Москвы характеризуется значительно большей плотностью данных на каждом УД.

Для описания характеристик местоположения были использованы внешние источники данных, полученные независимо от топографических данных, с более высокой детальностью и не подвергнутые картографической генерализации. Основным источником данных является Copernicus Global Land Cover (CGLC) [Buchhorn *et al.*, 2020], который представляет собой глобальное растровое покрытие с разрешением 100 м, полученное путем классификации спутниковых снимков. Из этого набора данных были извлечены 33 фрагмента, которые покрывают ту же площадь и имеют ту же проекцию, что и фрагменты из базы данных. Классифицированные изображения наземного покрова представлены в Приложении на Рисунках A.1 — A.3. Цвета соответствуют официальной легенде CGLC в Приложении в Таблице A.1 которая включает 23 класса. Эти рисунки иллюстрируют разнообразие ландшафтов, выбранных для исследования. В частности, есть фрагменты с малым (например, Алдан, Токур) и большим (например, Иркутск, Тюмень, Санкт-Петербург) количеством классов наземного покрова. Их также можно охарактеризовать малым (например, Комсомольск-на-Амуре, Нарьян-Мар) или большим (например, Элиста, Рубцовск) размером пятен наземного покрова. Также наблюдается четкое различие между преимущественно природными (например, Лесосибирск, Ухта) и городскими/сельскохозяйственными (например, Уфа, Ульяновск) ландшафтами.

Перед расчетом характеристик местоположения данные CGLC были переклассифицированы в меньшее число классов. Этот шаг предварительной обработки был выполнен для того, чтобы избежать избыточности данных в тех случаях, когда разнообразие классов наземного покрова значительно превышает аналогичное в топографических данных. В частности, все классы густых лесов (111-116) и редких лесов (121-126) были объединены всего в два класса: *cforest* (густой лес) и *oforest* (редкий лес), что соответствует классификации, применяемой на топографических картах. Аналогично, открытое море (20) и постоянные водоемы (80) были объединены в один класс *воды*. Кроме того, из рассмотрения были исключены классы «нет данных» (0), «открытый грунт» (60), «снег» (70) и «мох» (100), поскольку на большинстве фрагментов они занимают ничтожную площадь. Полученная 8-элементная классификация наземного покрова, помимо редкого леса, густого леса и воды, включала исходные классы *кустарников* (20), *травы* (30), *сельскохозяйственных земель* (40), *урбанизированных территорий* (50) и *болот* (90).

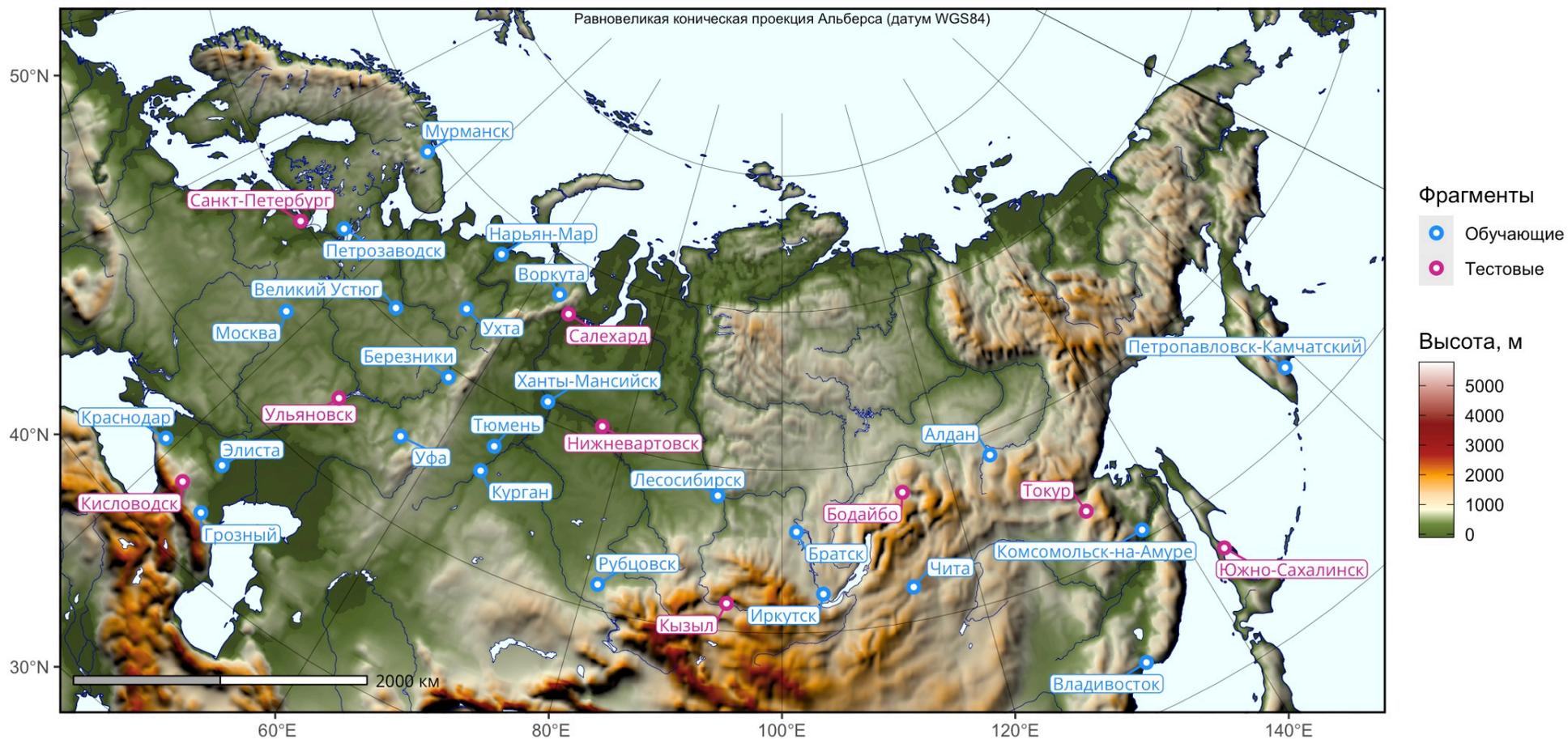
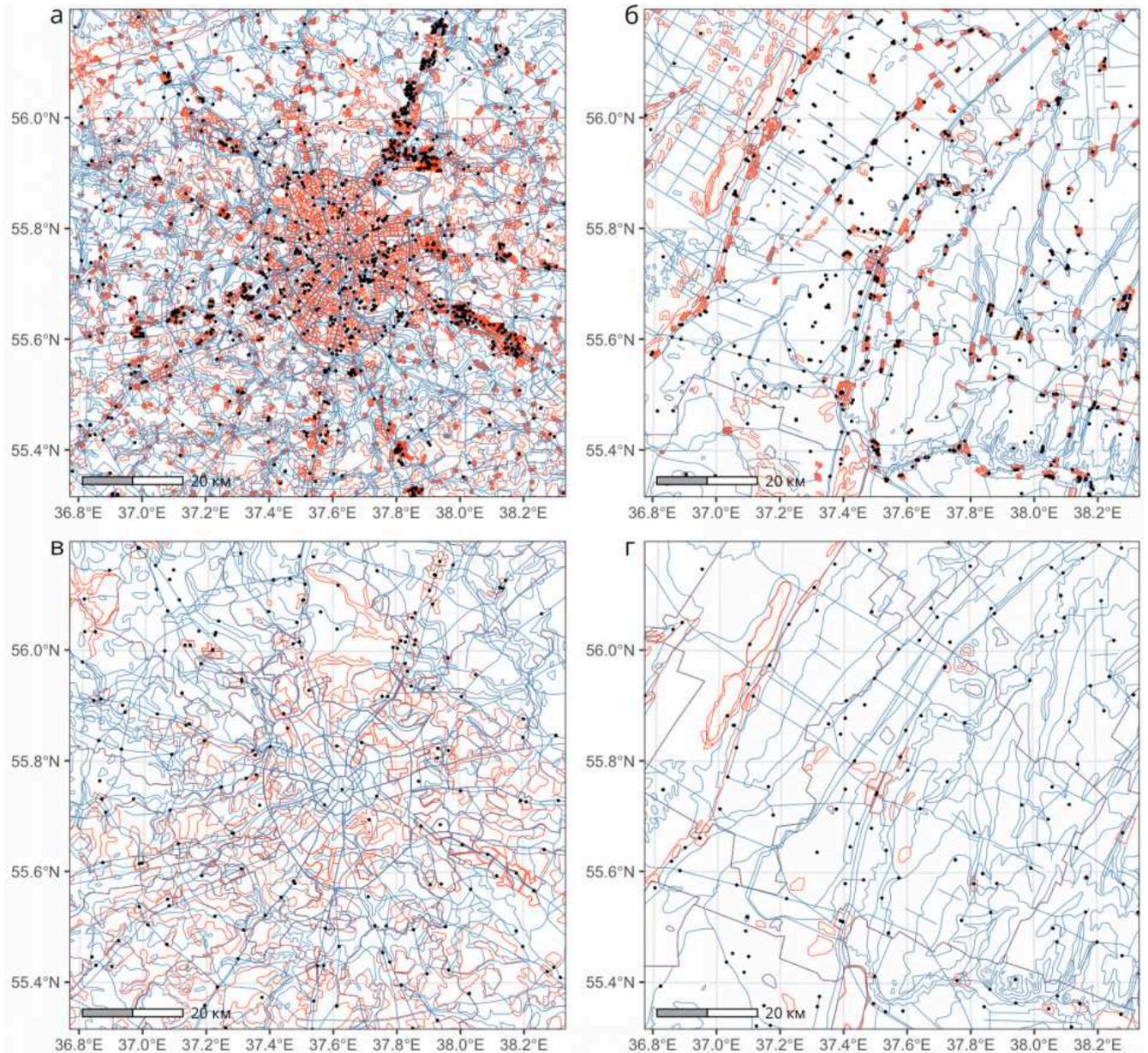


Рисунок 3.2 – Расположение фрагментов данных



**Рисунок 3.3** – Примеры уровней детализации 500 и 1000 для двух фрагментов: (а) Москва, УД 500; (б) Рубцовск, УД 500 ; (с) Москва, УД 1000; (д) Рубцовск, УД 1000. Точечные, линейные и полигональные объекты показаны черным, синим и красным цветом соответственно. Каждый фрагмент имеет размер 100 на 100 км. Равновеликая азимутальная проекция Ламберта

Для оценки расчлененности рельефа использовалась глобальная растровая цифровая модель рельефа *GMTED\_2010* с разрешением  $7.5''$  (около 140 м на  $50^\circ$  с. ш.). Поскольку 100-метровое покрытие наземного покрова не позволяет адекватно воспроизвести линейные объекты, дополнительно были использованы независимые данные о густоте гидрографической и дорожной сети. Данные о гидрографической сети были получены из глобальной БД *MERIT Hydro-Vector* [Lin et al., 2021] с разрешением  $3''$  (примерно 60 м на  $50^\circ$  с.ш.). Таким образом, источники данных о земельном покрове, рельефе и густоте гидрографической сети имеют пространственное разрешение порядка 100 м. Кроме того, плотность дорожной сети была получена из базы данных *GRIP* [Meijer et al., 2018] с разрешением  $5'$  (примерно 6 км на  $50^\circ$  с.ш.), которая была уточнена с использованием

официальной российской статистики протяженности дорог по муниципалитетам [Росстат, 2024].

### 3.1.2.2 Экспериментальная работа

Для целей моделирования 24 фрагмента (73% от общего количества) были использованы в качестве обучающей выборки, а 9 фрагментов (27% от общего количества) — в качестве тестовой, что близко к общепринятому соотношению 70/30. Эти выборки были сформированы таким образом, чтобы обеспечить сопоставимое разнообразие типов ландшафтов в обеих из них.

Для каждого фрагмента были рассчитаны  $d_P$ ,  $d_L$  и  $d_I$  как переменные для моделирования. Для обучающей выборки было рассчитано среднее значение каждого показателя, а относительные плотности  $\widehat{d}_P$ ,  $\widehat{d}_L$  и  $\widehat{d}_I$  были получены путем деления исходных значений на соответствующие средние.

На основе реклассифицированных данных *CGLC* были рассчитаны доли площади каждого фрагмента, занятые каждым классом наземного покрова, а также показатели *contag*, *pafrac* и *jointent*. Показатель *tri* был рассчитан на основе данных *GMTED2010*. Переменные *roads* и *rivers* — это густота дорожной и гидрографической сети соответственно, полученные путем деления длин из *GRIP* и *MERIT Hydro-Vector* на площадь фрагмента (10 000 км<sup>2</sup>).

Поскольку набор рассчитанных предикторов состоит из 14 переменных, было принято решение провести анализ данных методом главных компонент (*PCA* — *principal component analysis*) с целью сокращения размерности пространства признаков. Перед применением *PCA* переменные были логарифмированы, центрированы относительно нуля и масштабированы до единичной дисперсии. Лог-преобразование выполнялось в виде  $\ln(l_j + 1)$ , где  $l_j$  — значение свойства местоположения. Слагаемое +1 используется для того, чтобы исключить возможность преобразования  $l_j = 0$  в  $-\infty$ .

После проведения *PCA* была подобрана следующая модель для каждого показателя относительной плотности:

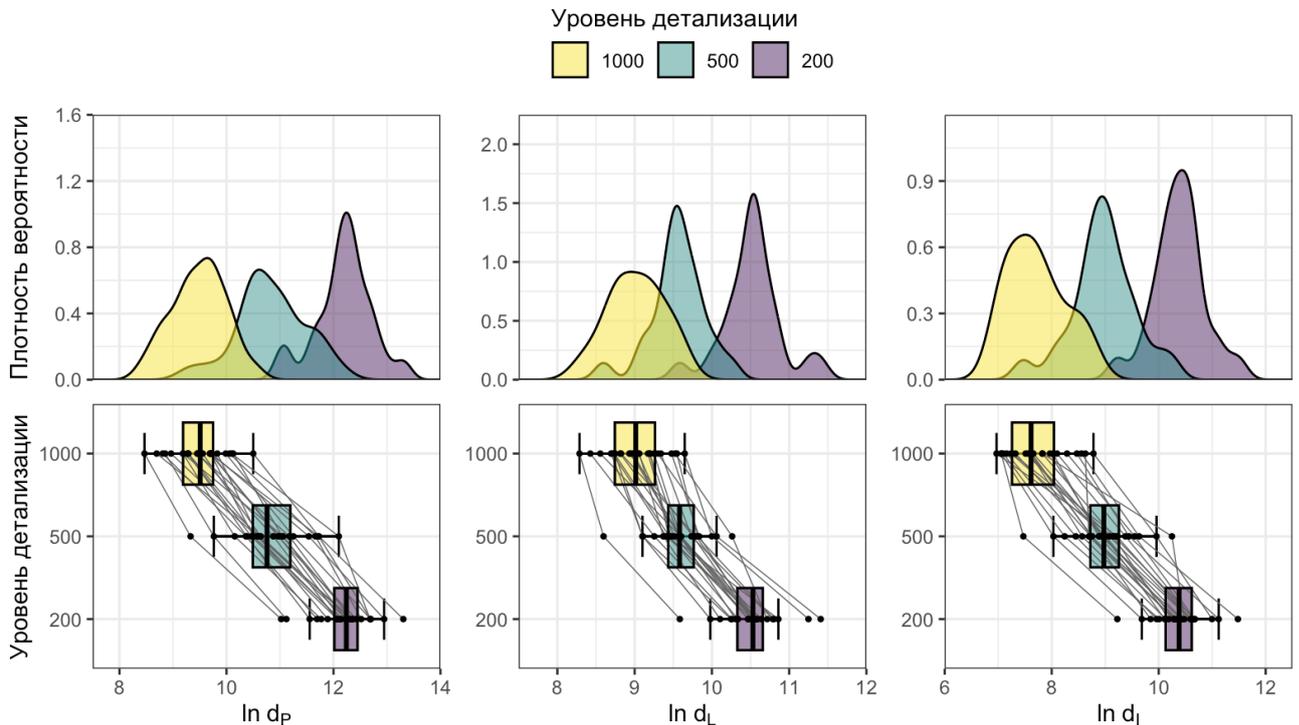
$$\ln(\widehat{d}_i + 1) = \sum_j \beta_{ij} PC_j + \epsilon_i, \quad (3.10)$$

где  $PC_j$  —  $j$ -я главная компонента. Для всех полученных результатов были рассчитаны  $Z$ -значения, которые оценивались для отнесения набора данных к тому или иному уровню детализации. Расчет характеристик плотности производился с помощью программного модуля *MapAnalyzer QGIS*, написанного на языке программирования Python ([Якимова *et al.*, 2021]). Все остальные аналитические расчеты, включая *PCA*, построение регрессионной модели и проведение статистических тестов, были запрограммированы на языке *R*, включая пакет *landscapemetrics* ([Hesselbarth *et al.*, 2019]).

### 3.1.3 Результаты и их анализ

#### 3.1.3.1 Абсолютная плотность

Анализ результатов по обучающей выборке начнем с визуализации оценок плотности вероятности для абсолютных (ненормированных) значений плотности (Рисунок 3.4). Каждая переменная плотности на каждом УД имеет логнормальное распределение, что подтверждают значения  $p > 0.05$  для теста Шапиро-Уилка применительно к логарифмированным значениям переменных. Это позволяет применять  $Z$ -значения в дальнейшем анализе.



**Рисунок 3.4** – Плотность распределения показателей плотности пространственных данных по уровням детализации: (а) плотность точек  $d_P$ , (б) плотность линий  $d_L$ , (в) плотность пересечений  $d_I$ . Все величины логарифмированы

Как видно из линий, соединяющих УД каждого фрагмента на Рисунке 3.4 абсолютная величина плотности закономерно уменьшается при соответствующем снижении детальности. Это наблюдение справедливо и для медианных значений по всем фрагментам, однако распределения значительно пересекаются.

Чтобы проверить, можно ли безошибочно отнести каждый из фрагментов обучающей выборки к его уровню детализации с помощью показателей плотности, вычислим  $Z$ -значения, а затем сравним их по всем УД. Процедура сравнения выглядит следующим образом:

- По всем фрагментам  $j$ -го уровня детализации вычисляются среднее значение и стандартное отклонение  $i$ -й меры плотности  $d_{ij} = d_i(L_j)$ . Для обучающей вы-

борки в нашем случае это  $3_i \times 3_j = 9$  эталонных комбинаций.

- Для каждого  $m$ -го фрагмента  $k$ -го уровня детализации вычисляются  $Z$ -значения  $i$ -й меры плотности  $d_{ikm} = d_i(L_k, F_m)$  относительно эталонного среднего значения и стандартного отклонения каждого  $d_{ij}$ . Для обучающей выборки в нашем случае это  $3_i \times 3_k \times 24_m = 216$  комбинаций, каждая из которых соотносится с 9-ю эталонными, итого  $216 \times 9 = 1944$   $Z$ -значений.
- Ожидается, что абсолютное значение  $Z$ -значения должно быть минимальным, когда  $k = j$ , т.е. когда  $Z$ -значение метрики плотности вычисляется по отношению к среднему значению и стандартному отклонению ее собственного уровня детализации. Если это так, то абсолютные значения плотности можно использовать для определения уровня детализации, которому соответствует исследуемый фрагмент.

Результаты вычисления  $Z$ -значений представлены на Рисунке 3.5а и Рисунке 3.6а, где строки соответствуют  $d_{km}$ , а столбцы соответствуют  $d_{ij}$ . Все значения сгруппированы в матрицы размером  $3 \times 3$ , которые соответствуют одной группе вычислений.  $Z$ -оценки плотности относительно собственных распределений находятся на антидиагонали каждой матрицы. Таким образом, если результаты соответствуют ожиданиям, эти оценки должны иметь наименьшие абсолютные значения в каждой строке матрицы (окрашены синим цветом). Если же наименьшее  $Z$ -значение получилось относительно другого уровня детализации, оно окрашено в таблице оранжевым цветом.

Матрицы на Рисунке 3.5а и Рисунке 3.6а показывают, что для 14 фрагментов из 24 есть случаи, когда ожидания не оправдываются. В частности, для 17% всех  $d_{ikm}$  комбинаций (37 из 216) наименьшее абсолютное значение  $Z$  получено для другого УД. Это означает, что каждый 6-й фрагмент статистически ближе к другому УД по плотности пространственных данных, так как она значительно меньше или больше среднего значения для «родного» УД. Например, Рубцовск расположен в равнинном и засушливом, однородном малонаселенном ландшафте, и данные УД 200 и 500 на окружающую его территорию более близки по плотности к УД 500 и 1000 соответственно.

Такая значительная изменчивость плотности пространственных данных ставит вопрос о том, определяется ли она свойствами географического положения и может ли быть предсказана. Чтобы ответить на этот вопрос, применим метод, предложенный ранее.

### 3.1.3.2 Главные компоненты

Результаты PCA полученных моделей обобщены на Рисунке 3.7, Рисунке 3.8, и Рисунке 3.9. В соответствии с Рисунком 3.7 первые две компоненты объясняют 53.3% общей дисперсии, в то время как большая часть дисперсии (96%) объясняется первыми восемью компонентами, которые и были выбраны для дальнейшего анализа.

	а Точки									б Точки									Уровень детализации
	Точки			Линии			Пересечения			Точки			Линии			Пересечения			
Алдан	6.87	4.075	0.477	6.606	2.881	0.773	5.341	2.526	-0.376	9.055	4.88	-0.496	8.955	3.02	-0.338	7.841	3.488	-1	1000
	3.263	1.085	-1.721	5.357	1.446	-0.767	2.722	0.185	-2.431	4.992	0.861	-4.459	8.272	0.788	-3.448	3.481	-0.083	-3.758	500
	1.437	-1.256	-4.724	2.362	-1.241	-3.28	0.465	-2.746	-6.056	1.343	-4.157	-11.241	2.415	-3.761	-7.256	0.203	-3.969	-8.27	200
Березники	5.106	2.761	-0.044	4.183	1.52	0.484	5.039	2.291	-0.297	7.954	4.45	0.26	6.972	2.728	1.077	8.281	4.031	0.029	1000
	1.888	0.06	-2.128	2.814	0.018	-1.07	2.45	-0.028	-2.36	3.903	0.435	-3.712	5.771	0.42	-1.662	3.842	0.362	-2.916	500
	-0.283	-2.523	-5.227	0.019	-2.557	-3.56	0.121	-3.014	-5.966	-0.107	-4.724	-10.245	0.351	-4.065	-5.783	0.625	-3.448	-7.284	200
Братск	5.208	2.155	-0.383	4.197	1.613	-0.564	4.158	1.356	-1.032	7.22	2.659	-1.132	6.197	2.08	-1.389	6.835	2.502	-1.192	1000
	1.968	-0.413	-2.392	2.828	0.115	-2.171	1.656	-0.87	-3.023	3.177	-1.338	-5.089	4.795	-0.398	-4.773	2.657	-0.891	-3.916	500
	-0.164	-3.107	-5.553	0.032	-2.468	-4.574	-0.884	-4.08	-6.805	-1.074	-7.084	-12.079	-0.455	-4.74	-8.35	-0.761	-4.914	-8.454	200
Чита	5.609	2.296	0.144	4.483	1.545	-0.102	4.765	1.902	-0.163	8.057	3.108	-0.107	6.272	1.589	-1.034	8.052	3.624	0.431	1000
	2.28	-0.303	-1.981	3.129	0.043	-1.685	2.203	-0.378	-2.239	4.005	-0.893	-4.074	4.888	-1.017	-4.325	3.654	0.028	-2.587	500
	0.222	-2.971	-5.045	0.309	-2.534	-4.126	-0.191	-3.457	-5.813	0.029	-6.492	-10.728	-0.377	-5.25	-7.98	0.406	-3.838	-6.899	200
Элиста	3.247	0.558	-1.246	2.655	0.241	-0.892	3.646	0.586	-1.215	9.146	5.128	2.433	7.413	3.567	1.76	8.988	4.256	1.47	1000
	0.439	-1.658	-3.065	1.21	-1.325	-2.516	1.194	-1.564	-3.188	5.082	1.106	-1.56	6.328	1.477	-0.801	4.421	0.546	-1.735	500
	-2.055	-4.646	-6.385	-1.459	-3.794	-4.891	-1.468	-4.959	-7.014	1.463	-3.831	-7.381	0.81	-3.193	-5.073	1.303	-3.233	-5.902	200
Грозный	5.705	2.201	0.682	4.894	1.981	0.89	6.306	3.202	1.398	8.129	2.894	0.624	6.873	2.23	0.492	9.029	4.228	1.44	1000
	2.355	-0.377	-1.561	3.56	0.501	-0.644	3.592	0.794	-0.832	4.075	-1.105	-3.351	5.646	-0.208	-2.4	4.454	0.523	-1.761	500
	0.315	-3.063	-4.527	0.707	-2.112	-3.167	1.566	-1.975	-4.032	0.123	-6.775	-9.765	0.248	-4.583	-6.392	1.342	-3.259	-5.932	200
Иркутск	5.252	2.021	0.433	4.136	1.172	0.508	4.479	1.579	-0.216	7.278	2.45	0.079	5.487	0.765	-0.293	6.504	2.019	-0.757	1000
	2.002	-0.518	-1.755	2.764	-0.348	-1.045	1.946	-0.669	-2.287	3.233	-1.544	-3.89	3.898	-2.056	-3.391	2.387	-1.286	-3.559	500
	-0.122	-3.237	-4.766	-0.027	-2.894	-3.536	-0.518	-3.826	-5.874	-0.998	-7.359	-10.483	-1.194	-6.108	-7.209	-1.078	-5.376	-8.037	200
Ханты-Мансийск	5.453	3.129	-0.584	4.279	2.145	-0.536	4.579	2.656	-1.383	7.454	3.983	-1.564	6.695	3.294	-0.978	7.14	4.165	-2.082	1000
	2.158	0.347	-2.549	2.915	0.674	-2.142	2.036	0.302	-3.34	3.408	-0.027	-5.516	5.422	1.134	-4.255	2.907	0.471	-4.644	500
	0.071	-2.168	-5.747	0.112	-1.952	-4.547	-0.404	-2.598	-7.205	-0.766	-5.34	-12.648	0.063	-3.476	-7.922	-0.469	-3.32	-9.307	200
Комсомольск-на-Амуре	6.02	4.324	1.141	4.933	2.196	1.461	4.851	2.157	0.57	7.803	5.269	0.515	6.611	2.25	1.08	7.074	2.908	0.454	1000
	2.601	1.278	-1.203	3.601	0.727	-0.044	2.28	-0.148	-1.578	3.753	1.246	-3.459	5.317	-0.184	-1.66	2.853	-0.558	-2.568	500
	0.618	-1.017	-4.084	0.744	-1.904	-2.614	-0.094	-3.167	-4.977	-0.306	-3.644	-9.909	-0.024	-4.563	-5.781	-0.532	-4.524	-6.876	200
Краснодар	4.536	1.331	-0.585	3.395	0.232	-0.75	5.341	2.232	0.465	6.915	2.127	-0.736	6.02	0.98	-0.585	7.486	2.677	-0.055	1000
	1.444	-1.055	-2.549	1.986	-1.335	-2.366	2.722	-0.081	-1.674	2.875	-1.864	-4.696	4.571	-1.785	-3.758	3.191	-0.747	-2.985	500
	-0.812	-3.901	-5.748	-0.744	-3.804	-4.754	0.465	-3.082	-5.097	-1.476	-7.785	-11.556	-0.64	-5.884	-7.512	-0.137	-4.746	-7.364	200
Курган	4.11	1.777	-0.41	3.007	0.978	0.116	3.948	1.927	-0.216	7.824	4.338	1.07	6.255	3.021	1.648	6.701	3.576	0.261	1000
	1.112	-0.708	-2.413	1.579	-0.552	-1.457	1.466	-0.355	-2.287	3.774	0.324	-2.909	4.867	0.789	-0.943	2.547	-0.012	-2.725	500
	-1.222	-3.471	-5.579	-1.119	-3.082	-3.916	-1.124	-3.429	-5.874	-0.279	-4.872	-9.177	-0.395	-3.76	-5.189	-0.89	-3.885	-7.061	200
Лесосибирск	5.133	1.946	-0.412	4.021	1.366	0.088	4.246	0.974	-1.148	7.841	3.078	-0.444	6.286	2.056	0.02	7.177	2.117	-1.165	1000
	1.91	-0.576	-2.414	2.643	-0.144	-1.487	1.735	-1.214	-3.127	3.79	-0.922	-4.408	4.907	-0.428	-2.996	2.938	-1.206	-3.894	500
	-0.236	-3.309	-5.581	-0.139	-2.707	-3.943	-0.784	-4.516	-6.937	-0.257	-6.532	-11.172	-0.362	-4.764	-6.884	-0.433	-5.283	-8.428	200

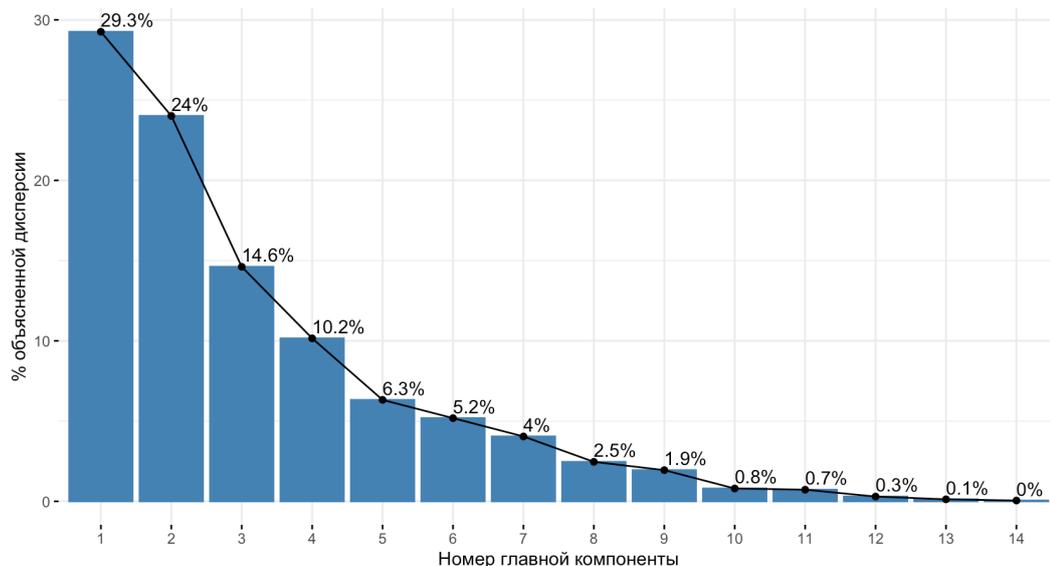
Рисунок 3.5 – Z-значения характеристик плотности (обучающая выборка): (а) плотность распределения; (б) относительная плотность распределения. Каждая строка представляет собой уникальную комбинацию фрагмента и уровня детализации. Каждый столбец – уникальная комбинация меры плотности и уровня детализации

	а									б									Уровень детализации
	Точки			Линии			Пересечения			Точки			Линии			Пересечения			
Москва	5.705	3.101	1.276	4.173	2.258	1.271	6.075	4.154	1.971	7.666	3.777	1.05	5.609	2.557	0.984	6.958	3.986	0.611	1000
	2.355	0.325	-1.098	2.803	0.792	-0.244	3.384	1.652	-0.315	3.618	-0.231	-2.93	4.053	0.203	-1.78	2.758	0.325	-2.439	500
	0.314	-2.195	-3.955	0.009	-1.844	-2.799	1.303	-0.889	-3.379	-0.486	-5.611	-9.204	-1.067	-4.243	-5.88	-0.644	-3.491	-6.726	200
Мурманск	7.581	4.45	2.018	6.168	2.711	1.742	6.972	3.552	1.556	9.437	4.759	1.126	7.957	2.449	0.905	9.779	4.49	1.404	1000
	3.818	1.377	-0.52	4.898	1.268	0.251	4.192	1.109	-0.69	5.37	0.741	-2.854	7.013	0.067	-1.879	5.069	0.738	-1.79	500
	2.123	-0.895	-3.239	1.939	-1.405	-2.342	2.326	-1.576	-3.852	1.847	-4.316	-9.104	1.376	-4.356	-5.962	2.061	-3.008	-5.966	200
Нарьян-Мар	6.369	4.443	0.41	4.136	2.033	-0.236	5.12	4.688	0.21	8.573	5.696	-0.329	6.801	3.45	-0.166	7.537	6.869	-0.057	1000
	2.873	1.371	-1.773	2.765	0.556	-1.826	2.523	2.134	-1.903	4.515	1.668	-4.294	5.556	1.329	-3.23	3.232	2.685	-2.986	500
	0.954	-0.902	-4.789	-0.026	-2.061	-4.256	0.213	-0.279	-5.388	0.709	-3.082	-11.021	0.173	-3.314	-7.076	-0.088	-0.729	-7.366	200
Петропавловск-Камчатский	6.348	5.19	1.218	4.693	3.438	1.4	5.476	3.394	1.057	6.826	5.096	-0.836	5.175	3.176	-0.071	6.844	3.625	0.011	1000
	2.856	1.953	-1.143	3.349	2.031	-0.109	2.844	0.967	-1.139	2.787	1.075	-4.796	3.506	0.984	-3.111	2.665	0.029	-2.93	500
	0.934	-0.182	-4.01	0.512	-0.702	-2.674	0.619	-1.755	-4.421	-1.593	-3.872	-11.689	-1.518	-3.599	-6.978	-0.752	-3.837	-7.301	200
Петрозаводск	5.839	2.935	0.425	4.354	1.088	-0.84	5.479	2.683	0.199	8.727	4.389	0.639	7.594	2.389	-0.683	8.932	4.609	0.766	1000
	2.46	0.195	-1.762	2.994	-0.436	-2.461	2.847	0.326	-1.913	4.668	0.375	-3.336	6.556	-0.009	-3.882	4.375	0.835	-2.312	500
	0.443	-2.355	-4.775	0.184	-2.976	-4.841	0.623	-2.566	-5.401	0.911	-4.804	-9.745	0.998	-4.418	-7.615	1.249	-2.894	-6.577	200
Рубцовск	3.054	-0.307	-2.014	1.567	-1.161	-2.02	2.798	-0.455	-1.223	7.579	2.558	0.008	5.215	0.869	-0.5	6.047	1.017	-0.172	1000
	0.288	-2.332	-3.664	0.067	-2.797	-3.699	0.43	-2.503	-3.195	3.532	-1.437	-3.961	3.556	-1.925	-3.651	2.012	-2.107	-3.08	500
	-2.24	-5.48	-7.125	-2.512	-5.151	-5.982	-2.436	-6.146	-7.023	-0.601	-7.217	-10.577	-1.477	-5.999	-7.424	-1.516	-6.337	-7.476	200
Тюмень	4.385	1.749	-1.338	3.645	1.193	-0.775	4.782	2.391	-0.88	7.487	3.55	-1.062	6.485	2.578	-0.558	7.576	3.878	-1.18	1000
	1.326	-0.729	-3.136	2.249	-0.326	-2.392	2.218	0.063	-2.886	3.441	-0.456	-5.019	5.157	0.23	-3.724	3.265	0.236	-3.906	500
	-0.958	-3.498	-6.473	-0.502	-2.874	-4.777	-0.172	-2.9	-6.631	-0.722	-5.91	-11.986	-0.156	-4.221	-7.484	-0.05	-3.595	-8.443	200
Уфа	5.276	2.976	0.446	4.006	1.579	0.646	5.431	3.076	0.709	8.565	5.129	1.35	6.849	2.982	1.494	7.844	4.202	0.542	1000
	2.021	0.227	-1.745	2.628	0.079	-0.901	2.803	0.68	-1.453	4.507	1.107	-2.633	5.616	0.739	-1.137	3.484	0.501	-2.496	500
	-0.099	-2.316	-4.754	-0.152	-2.5	-3.403	0.568	-2.119	-4.818	0.698	-3.83	-8.809	0.223	-3.801	-5.349	0.206	-3.284	-6.792	200
Ухта	4.779	2.208	-1.562	3.431	0.654	-1.634	5.016	1.883	-0.729	7.967	4.126	-1.505	6.389	1.964	-1.68	8.414	3.567	-0.471	1000
	1.633	-0.372	-3.311	2.024	-0.892	-3.294	2.43	-0.395	-2.749	3.915	0.115	-5.458	5.037	-0.544	-5.14	3.95	-0.018	-3.325	500
	-0.578	-3.056	-6.689	-0.709	-3.396	-5.608	0.095	-3.48	-6.458	-0.091	-5.151	-12.57	-0.255	-4.86	-8.652	0.752	-3.893	-7.763	200
Великий Устюг	5.485	2.247	-1.029	4.312	1.587	-1.277	5.353	2.777	-0.837	9.509	4.673	-0.222	7.702	3.359	-1.204	9.361	5.377	-0.212	1000
	2.183	-0.341	-2.896	2.949	0.088	-2.92	2.733	0.411	-2.847	5.441	0.656	-4.188	6.692	1.215	-4.54	4.726	1.463	-3.113	500
	0.102	-3.018	-6.176	0.144	-2.493	-5.263	0.479	-2.46	-6.582	1.942	-4.43	-10.88	1.111	-3.408	-8.157	1.66	-2.158	-7.515	200
Владивосток	6.371	3.505	0.998	5.09	1.256	0.431	4.757	1.813	1.689	8.497	4.216	0.469	8.054	1.944	0.629	6.667	2.115	1.923	1000
	2.875	0.64	-1.315	3.766	-0.26	-1.126	2.195	-0.458	-0.57	4.439	0.203	-3.504	7.136	-0.57	-2.228	2.52	-1.208	-1.365	500
	0.956	-1.805	-4.222	0.896	-2.813	-3.611	-0.201	-3.559	-3.7	0.608	-5.033	-9.969	1.477	-4.881	-6.249	-0.922	-5.285	-5.468	200
Воркута	5.65	3.356	-0.058	3.564	1.576	-0.184	4.436	2.365	-0.11	7.339	3.913	-1.188	5.362	2.195	-0.61	6.029	2.826	-1.001	1000
	2.312	0.524	-2.138	2.164	0.077	-1.772	1.906	0.04	-2.191	3.294	-0.097	-5.144	3.741	-0.253	-3.79	1.997	-0.625	-3.759	500
	0.261	-1.949	-5.24	-0.58	-2.503	-4.206	-0.567	-2.929	-5.752	-0.918	-5.432	-12.153	-1.324	-4.62	-7.539	-1.534	-4.603	-8.271	200

Рисунок 3.6 – Z-значения характеристик плотности (обучающая выборка, продолжение): (а) плотность распределения; (б) относительная плотность распределения. Каждая строка представляет собой уникальную комбинацию фрагмента и уровня детализации. Каждый столбец – уникальная комбинация меры плотности и уровня детализации

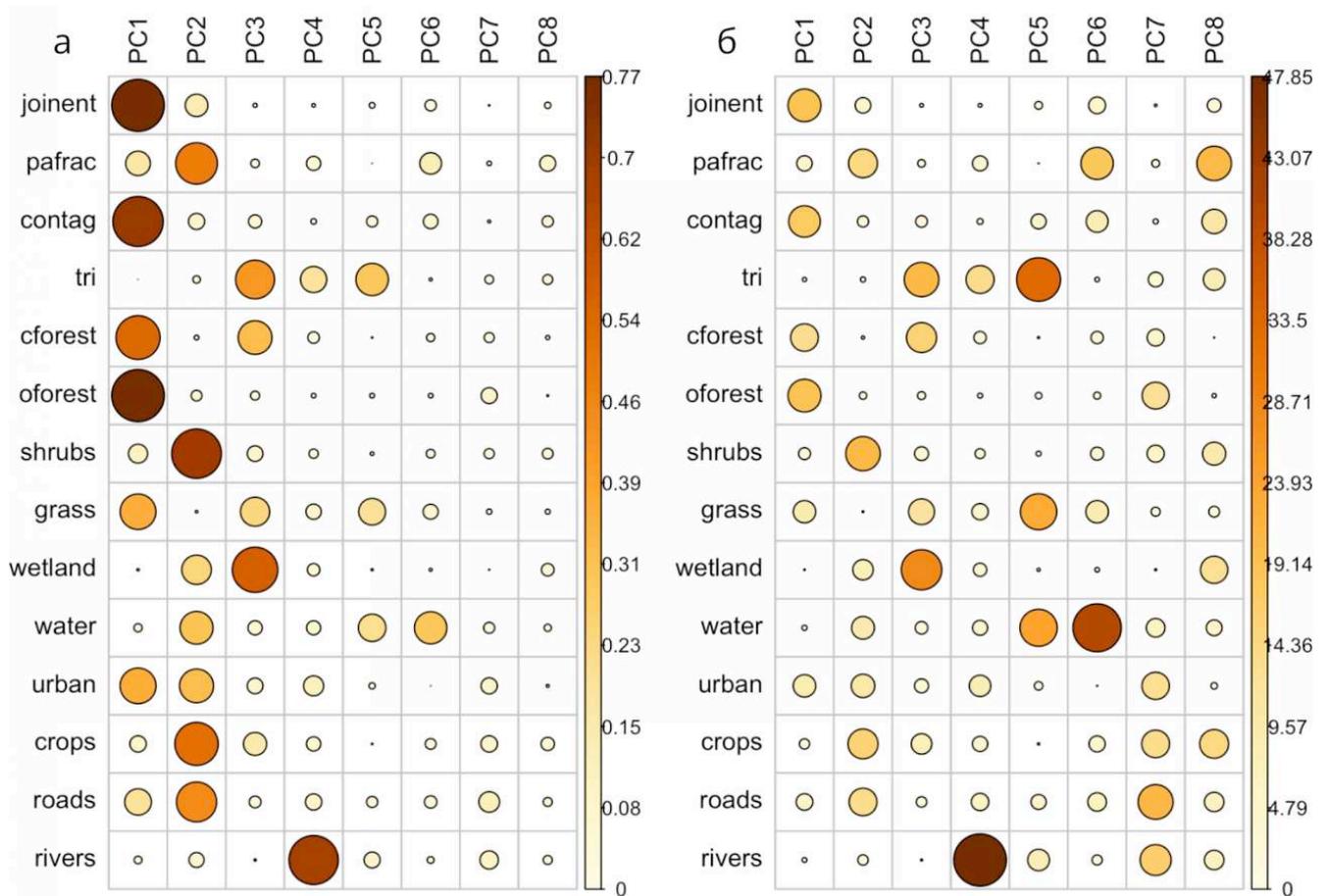
Качество каждой компоненты и вклад переменных можно оценить по Рисунку 3.8. В частности, компонента 1 формируется в основном за счет ландшафтных метрик *jointent* и *contag*, а также доли обоих типов леса (*cforest* и *oforest*). В компоненте 2 доминируют *pafrac*, *shrubs*, *crops* и *roads*. Компоненту 3 формируют предикторы *tri*, *cforest*, *grass* и *wetlands*. В компоненту 4 значительный вклад вносят *rivers* и *tri*. В компоненте 5 доминируют *tri*, *grass* и *water*, а компонента 6 сформирована в основном переменными *pafrac*, *contag* и *water*. Компонента 7 состоит в основном из предикторов *forest*, *urban*, *crops*, *roads* и *rivers*. Наконец, компонента 8 имеет значительный вклад от *pafrac*, *contag*, *tri*, *shrubs*, *wetlands*, *crops*, *roads* и *rivers*. Каждая переменная внесла значительный вклад хотя бы в одну компоненту, в то время как большинство из них вносят значительный вклад в 2–3 компоненты.

Примечательно, что, вопреки ожиданиям, доля городских территорий не вносит столь значительного вклада, как другие свойства местоположения, ни в одну конкретную компоненту. Однако это единственное свойство местоположения, которое вносит одинаковый вклад в наиболее важные компоненты 1-2, что приводит к значительному вкладу в целом при дальнейшем анализе.

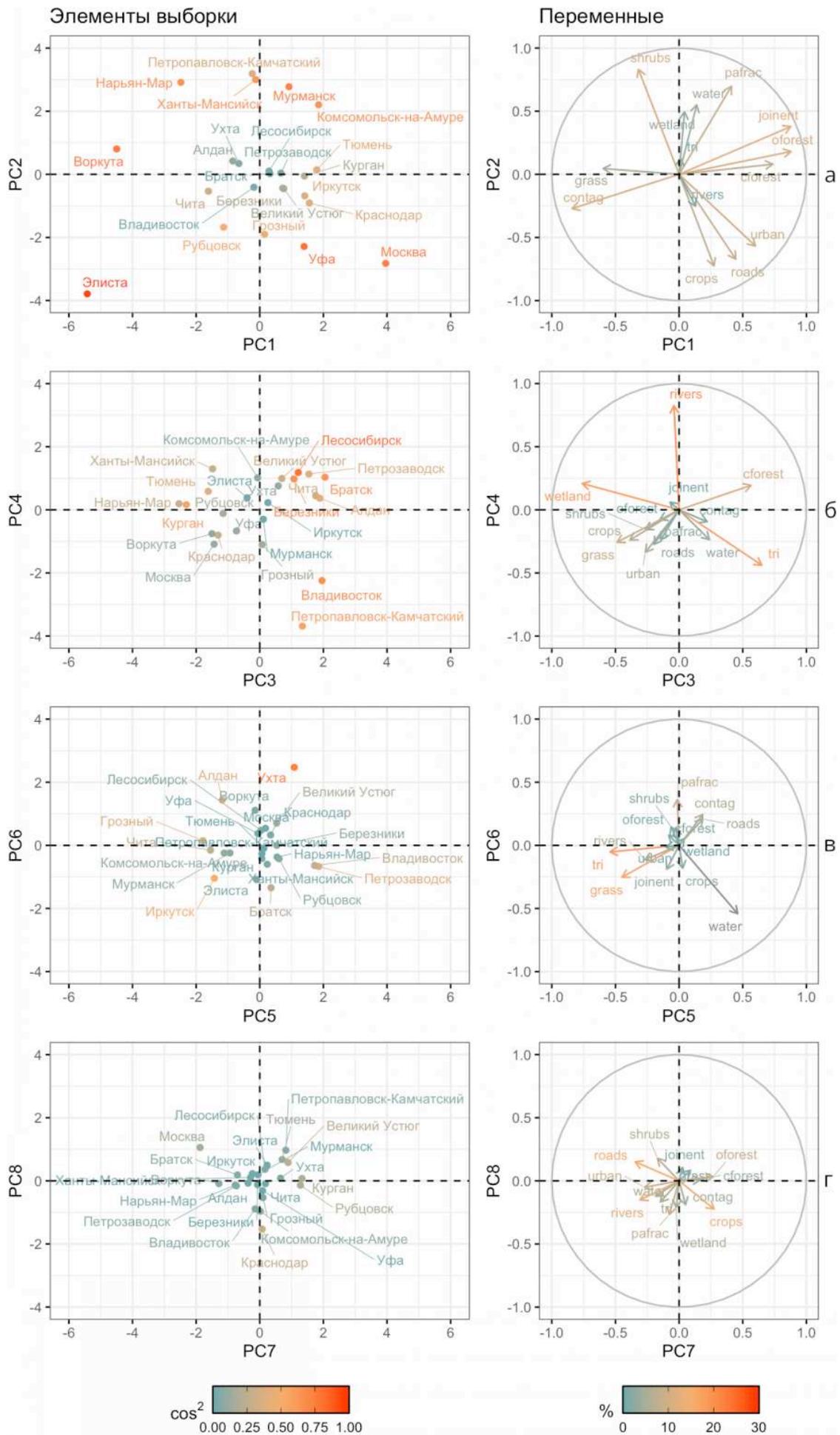


**Рисунок 3.7** — Доля общей дисперсии, объясненная каждой главной компонентой (график собственных значений)

Рисунок 3.9 показывает детали PCA с точки зрения распределения фрагментов и вкладов переменных по первым восьми компонентам. На Рисунке 3.9а видно, что по компонентам 1-2 различаются городские/сельскохозяйственные районы (нижняя правая четверть) и преимущественно природные (верхняя левая четверть). На Рисунке 3.9б согласно компонентам 3–4 фрагменты упорядочиваются от равнинных территорий (левая часть) к возвышенным (правая часть) и от менее (нижняя часть) к более (верхняя часть) развитой речной сети. На Рисунке 3.9в видим дифференциацию преимуществен-



**Рисунок 3.8** – Сводный анализ главных компонент: (а) качество представления,  $\cos^2$  ; (б) вклад переменных, %



**Рисунок 3.9** — Анализ главных компонент: (а) PC1-PC2; (б) PC3-PC4; (в) PC5-PC6; (г) PC7-PC8.  $\cos^2$  — качество представления, % — доля объясненной дисперсии

но по доле крупных водных объектов согласно компонентам 5–6, при этом прибрежные фрагменты расположены в нижней правой четверти графика. Наконец, вклад компонент 7–8 в дифференциацию фрагментов, представленный на Рисунке 3.9г, существенно менее отчетлив, чем для предыдущих компонент. Графики, представленные на Рисунке 3.9 в правой части подкрепляют сделанные ранее выводы векторами вклада каждой компоненты.

### 3.1.3.3 Относительная плотность

Таблица 3.1 обобщает характеристики регрессионных моделей для каждого показателя относительной плотности. Как видно, компоненты 1–5 и 7 вносят значительный вклад в  $\widehat{d}_P$  и  $\widehat{d}_L$ , в то время как только компоненты 1, 4 и 7 оказывают значительное влияние на  $\widehat{d}_I$ . Компоненты 6 и 8 оказывают незначительное влияние на результат во всех моделях. Эти отношения между главными компонентами и смоделированными относительными мерами плотности можно оценить визуально по Рисунку 3.10 и Рисунку 3.11. В частности, линии тренда для компонент 6 и 8 имеют менее выраженный наклон, что свидетельствует о слабой корреляции с относительной плотностью пространственных данных.

Полученная регрессионная модель позволяет рассчитать нормированные плотности  $\widetilde{d}_i = d_i / \widehat{d}_i$ , которые учитывают неоднородность ландшафта и приводят плотности к их ожидаемым (средним) значениям для соответствующих уровней детализации. Распределения нормированных плотностей представлены на Рисунке 3.12. По сравнению с их абсолютными значениями на Рисунке 3.4 видно, что нормированные распределения имеют больший эксцесс и существенно меньшую долю пересечений, чем абсолютные значения, что означает, что они более точно определяют уровень детализации пространственных данных.

Чтобы проверить эту гипотезу, были выполнены  $Z$ -тесты для нормированных значений плотности по отношению к среднему и стандартному отклонению абсолютных значений. Результаты представлены на Рисунке 3.5б и Рисунке 3.6б. Теперь только для 4.6% из всех  $d_{ikm}$  (10 из 216)  $Z$ -оценка ложно указывает на другой уровень детализации. Полученный результат говорит о том, что ошибка определения УД значительно снижается после обработки предложенным методом, и примерно для 95% элементов обучающей выборки УД определяется корректно.

Проверка точности была осуществлена для 9 фрагментов из тестовой выборки (всего проведено  $3_i \times 3_k \times 9_m = 81$  тестов). Как видно из Рисунка 3.13, разработанный подход уменьшает ошибку распознавания УД с 12.3% (10 из 81) до 2.5% (2 из 81), что дает значительный прирост точности. Как и в случае с обучающей выборкой, для более чем 95% из общего числа фрагментов уровень детализации был распознан корректно.

Таблица 3.1 – Характеристики регрессионных моделей

Компонента	Точки ( $\ln \widehat{d}_P$ )			Линии ( $\ln \widehat{d}_L$ )			Пересечения ( $\ln \widehat{d}_I$ )		
	$\beta$	95% ДИ**	*р- значение	$\beta$	95% ДИ**	*р- значение	$\beta$	95% ДИ**	*р- значение
Свободный член	0.66	0.63, 0.70	● ● ●	0.67	0.64, 0.69	● ● ●	0.66	0.62, 0.71	● ● ●
PC1	0.02	0.00, 0.04	●	0.02	0.01, 0.03	● ●	0.04	0.02, 0.06	● ● ●
PC2	-0.08	-0.10, -0.06	● ● ●	-0.04	-0.05, -0.02	● ● ●	-0.02	-0.05, 0.00	○
PC3	-0.04	-0.06, -0.01	● ●	-0.03	-0.05, -0.02	● ● ●	0.01	-0.02, 0.05	○
PC4	0.08	0.05, 0.11	● ● ●	0.04	0.02, 0.06	● ● ●	0.09	0.06, 0.13	● ● ●
PC5	0.05	0.02, 0.09	● ●	0.07	0.05, 0.10	● ● ●	0.03	-0.02, 0.08	○
PC6	0.01	-0.04, 0.05	○	0.01	-0.02, 0.04	○	0.04	-0.02, 0.09	○
PC7	-0.13	-0.18, -0.08	● ● ●	-0.07	-0.10, -0.03	● ● ●	-0.13	-0.19, -0.06	● ● ●
PC8	0.00	-0.06, 0.06	○	-0.02	-0.07, 0.02	○	-0.01	-0.09, 0.07	○

\* р-значения: ● ● ● < 0.001, ● ● < 0.01, ● < 0.05, ○ ≥ 0.05

\*\* ДИ - доверительный интервал

### 3.1.4 Обсуждение

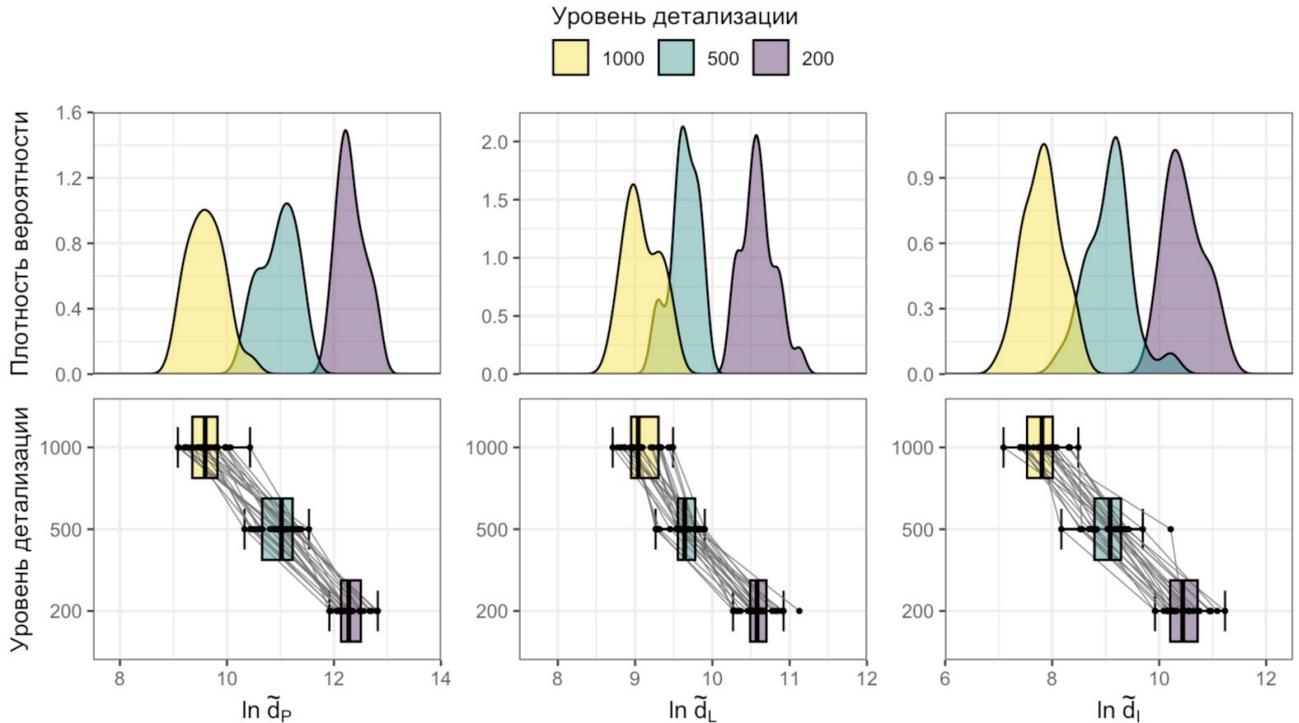
Разработанная методика имеет несколько потенциальных применений:

1. Распознавание уровня детализации пространственных данных при отсутствии соответствующих метаданных.
2. Сопоставление двух и более наборов топографических данных на предмет соответствия их детальности. Например, когда обновленный фрагмент данных необходимо интегрировать в существующий уровень детализации.
3. Контроль соотношения детальности для разных территорий при использовании коллаборативной генерализации [Touya, Duchêne, Ruas, 2010].

Следует, однако, подчеркнуть, что эти потенциальные применения нуждаются в дальнейшей практической оценке их эффективности.







**Рисунок 3.12** — Распределения нормированной плотности признаков по масштабам:  $\ln \tilde{d}_P$  — плотность точек;  $\ln \tilde{d}_L$  — плотность линий;  $\ln \tilde{d}_I$  — плотность пересечений

Полученные результаты имеют ряд ограничений, которые необходимо преодолеть в ходе дальнейших работ. Текущее исследование основано на экспериментальной выборке, которая в настоящее время состоит из 99 фрагментов пространственных данных. Скромный размер выборки и тот факт, что фрагменты не охватывают все возможные типы ландшафтов, могут потенциально ограничить эффективность полученных главных компонент и регрессионной модели. В частности, ожидаемое (среднее) значение каждого показателя плотности изменится, если выборка будет расширена за счет включения дополнительных фрагментов.

Наиболее надежный способ улучшить текущий результат — разделить общую площадь покрытия данными регулярной сеткой, и проводить построение и тестирование модели на основе полного набора фрагментов. Поскольку площадь России составляет примерно  $1.7 \times 10^7$  км<sup>2</sup>, а площадь каждого фрагмента составляет  $10^4$  км<sup>2</sup> то в результате получится примерно 1700 фрагментов. Это количество можно увеличить в несколько раз, извлекая более мелкие фрагменты данных (например,  $50 \times 50$  км) и применяя аугментацию (преобразование) данных. Обучающая и тестовая выборка при этом могут формироваться случайным путем в серии экспериментов. Следует, однако, иметь в виду, что такой подход будет значительно более требователен к вычислительным ресурсам.

	а									б									Уровень детализации
	Точки			Линии			Пересечения			Точки			Линии			Пересечения			
Бодайбо	5.86	3.637	1.49	5.502	3.059	2.026	5.129	2.536	0.567	6.758	3.438	0.23	6.117	2.225	0.578	7.142	3.133	0.088	1000
	2.476	0.743	-0.931	4.198	1.634	0.549	2.531	0.194	-1.581	2.719	-0.566	-3.741	4.693	-0.215	-2.292	2.909	-0.374	-2.868	500
	0.463	-1.679	-3.748	1.295	-1.068	-2.068	0.223	-2.734	-4.98	-1.683	-6.057	-10.284	-0.539	-4.589	-6.302	-0.467	-4.309	-7.227	200
Кисловодск	5.872	2.238	0.492	5.38	2.044	0.725	6.564	3.048	1.07	7.973	2.544	-0.064	7.223	1.906	-0.195	9.224	3.788	0.728	1000
	2.485	-0.348	-1.709	4.071	0.567	-0.818	3.825	0.656	-1.128	3.921	-1.451	-4.032	6.087	-0.617	-3.267	4.614	0.162	-2.344	500
	0.475	-3.027	-4.71	1.177	-2.051	-3.327	1.86	-2.15	-4.407	-0.082	-7.236	-10.672	0.612	-4.92	-7.107	1.529	-3.681	-6.614	200
Кызыл	5.764	2.433	0.375	4.651	1.874	0.754	4.896	2.093	-0.631	8.021	3.045	-0.03	6.416	1.992	0.207	7.65	3.315	-0.897	1000
	2.401	-0.196	-1.801	3.305	0.389	-0.787	2.321	-0.206	-2.661	3.969	-0.955	-3.998	5.071	-0.509	-2.759	3.325	-0.225	-3.674	500
	0.371	-2.839	-4.822	0.471	-2.215	-3.298	-0.043	-3.24	-6.347	-0.019	-6.576	-10.626	-0.227	-4.831	-6.688	0.02	-4.134	-8.171	200
Нижневартовск	5.549	2.982	-0.326	4.494	2.009	-0.54	4.948	1.438	-0.305	7.554	3.72	-1.223	6.867	2.908	-1.154	7.241	1.812	-0.883	1000
	2.234	0.232	-2.347	3.14	0.531	-2.145	2.367	-0.797	-2.367	3.507	-0.287	-5.178	5.639	0.646	-4.476	2.99	-1.455	-3.662	500
	0.164	-2.309	-5.498	0.319	-2.084	-4.55	0.016	-3.987	-5.975	-0.634	-5.686	-12.198	0.242	-3.878	-8.105	-0.372	-5.575	-8.157	200
Салехард	5.474	3.546	0.356	4.097	2.236	0.219	4.675	2.818	-0.539	6.487	3.606	-1.159	5.341	2.377	-0.837	6.256	3.384	-1.807	1000
	2.175	0.671	-1.816	2.723	0.77	-1.348	2.121	0.448	-2.579	2.451	-0.4	-5.116	3.715	-0.024	-4.077	2.183	-0.168	-4.42	500
	0.092	-1.767	-4.841	-0.065	-1.865	-3.816	-0.295	-2.413	-6.242	-2.04	-5.836	-12.115	-1.346	-4.431	-7.775	-1.316	-4.068	-9.044	200
Санкт-Петербург	5.095	2.65	-0.545	3.855	1.553	-0.221	5.617	3.343	1.183	6.881	3.229	-1.545	6.035	2.368	-0.459	6.752	3.234	-0.106	1000
	1.88	-0.026	-2.518	2.469	0.053	-1.81	2.971	0.921	-1.026	2.841	-0.774	-5.497	4.59	-0.034	-3.599	2.589	-0.291	-3.027	500
	-0.273	-2.629	-5.709	-0.299	-2.525	-4.241	0.78	-1.814	-4.278	-1.521	-6.333	-12.623	-0.624	-4.44	-7.381	-0.841	-4.212	-7.413	200
Токур	7.41	5.176	1.806	5.994	2.955	1.98	5.082	2.629	-0.146	10.231	6.894	1.86	8.056	3.213	1.661	8.559	4.765	0.474	1000
	3.684	1.943	-0.685	4.715	1.524	0.501	2.489	0.277	-2.224	6.155	2.854	-2.128	7.139	1.031	-0.927	4.069	0.962	-2.552	500
	1.957	-0.195	-3.443	1.771	-1.169	-2.112	0.17	-2.629	-5.794	2.892	-1.504	-8.137	1.479	-3.56	-5.176	0.891	-2.745	-6.857	200
Ульяновск	4.06	1.615	-1.045	3.199	0.428	-0.611	4.554	2.091	-0.068	6.59	2.938	-1.036	6.245	1.831	0.175	7.131	3.322	-0.017	1000
	1.072	-0.834	-2.908	1.78	-1.129	-2.22	2.013	-0.208	-2.154	2.552	-1.061	-4.993	4.855	-0.712	-2.8	2.9	-0.219	-2.953	500
	-1.271	-3.627	-6.191	-0.934	-3.614	-4.619	-0.433	-3.242	-5.705	-1.905	-6.716	-11.952	-0.405	-4.999	-6.722	-0.477	-4.128	-7.327	200
Южно-Сахалинск	5.975	3.29	0.647	4.335	1.265	0.482	4.824	2.512	0.77	7.552	3.541	-0.407	6.324	1.433	0.185	6.723	3.147	0.454	1000
	2.565	0.472	-1.589	2.973	-0.25	-1.072	2.256	0.172	-1.398	3.505	-0.464	-4.371	4.955	-1.214	-2.788	2.566	-0.362	-2.568	500
	0.574	-2.013	-4.56	0.166	-2.804	-3.562	-0.124	-2.762	-4.749	-0.636	-5.921	-11.123	-0.323	-5.413	-6.712	-0.868	-4.295	-6.877	200
	200	500	1000	200	500	1000	200	500	1000	200	500	1000	200	500	1000	200	500	1000	

Рисунок 3.13 – Z-значения для показателей плотности пространственных данных (тестовая выборка): (а) абсолютная плотность; (б) нормированная плотность. Каждая строка представляет собой уникальную комбинацию фрагмента и уровня детализации. Каждый столбец – уникальная комбинация показателя плотности и уровня детализации.

Еще одним ограничением предложенной методики является то, что текущее исследование опирается на конкретный геоинформационный *продукт* — базу данных цифровых топографических карт с тремя уровнями детализации. При этом для каждого продукта и уровня детализации в его составе используются определенные правила обобщения данных. То же касается и конкретных источников данных, используемых для прогнозирования. В частности, использованные в настоящем исследовании источники не подходят для оценки плотности детальных (крупномасштабных) пространственных данных из-за их разрешения. Кроме того, каждый из источников данных имеет свои погрешности и ограниченную точность, что неизбежно сказывается на качестве результатов. Однако оценка распространения подобных ошибок выходит за рамки настоящей работы.

Более завуалированная проблема кроется в предположении, что относительная плотность пространственных данных одинакова распределена для каждого фрагмента независимо от уровня детализации. Если различия существуют, то такая информация не может быть напрямую интегрирована в разработанную модель. Кроме того, следует подчеркнуть, что в данном исследовании рассматриваются только характеристики УД, основанные на плотности. Безусловно, другие дескрипторы УД, такие как гранулярность и разрешение (как пространственное, так и атрибутивное), должны быть изучены аналогичным образом.

Наконец, вместо простой линейной регрессии, опирающейся на метод главных компонент, можно использовать более сложные методы построения моделей распознавания УД, такие как глубокое обучение (*deep learning*). Это может потенциально повысить надежность и устойчивость полученной модели. Применение глубокого обучения, однако, будет требовать значительно более крупной обучающей выборки.

### 3.1.5 Итоги

Цифровые топографические карты принято представлять в виде мультимасштабных пространственных баз данных с несколькими уровнями детализации (УД), соответствующими масштабам карты. Последовательность УД, покрывающих одну и ту же территорию, характеризуется монотонным изменением характеристик детальности, таких как гранулярность и плотность пространственных данных: чем меньше целевой масштаб карты, тем меньше детальность УД. Однако свойства одного УД могут иметь существенную пространственную дифференциацию. Например, горизонтالي высот для сильно расчлененного рельефа будут естественным образом сложнее, чем для слабо расчлененного. Изменчивость формы отдельных объектов а также рельефа земной поверхности приводят к неоднородности каждого УД. Таким образом, показатели, описывающие уровень детализации, могут иметь пространственную вариацию. Это затрудняет интеграцию фрагментов топографических данных, охватывающих разные территории, в единую базу данных. Необходим автоматизированный метод, который позволяет оценить

плотность пространственных данных на соответствие целевому уровню детализации.

В данном разделе диссертации представлен подход к решению этой задачи с помощью статистического обучения на основе мультимасштабной топографической базы данных, уровни детализации которой получены ручным обобщением. Для каждого из трех УД проанализировано, насколько плотность точек, линий и пересечений в каждом фрагменте выше или ниже средней плотности УД. Это соотношение, называемое относительной плотностью, затем моделируется как функция свойств местоположения, зависящая от ряда пространственных и непространственных характеристик. Полученные на их основе главные компоненты показали значительный вклад в варьирование плотности данных как показателей сложности рисунка ландшафта, так и соотношения различных классов наземного покрова. Оказались различимы городские и природные, горные и равнинные территории, а также характеризующиеся менее и более развитой речной сетью. Результаты показали, что нормирование плотности пространственных данных на предсказанную по модели относительную плотность для той же территории приводит к получению величины, близкой к среднему значению уровня детализации. Использование нормированных значений позволяет снизить ошибку распознавания УД в несколько раз. Это подтверждает гипотезу о том, что свойства местоположения обуславливают изменчивость плотности пространственных данных, используемых при построении топографических карт.

Результаты, полученные в данном разделе диссертации, могут быть использованы в процессах интеграции и обобщения пространственных данных — как при коллаборативном обобщении, либо при объединении данных, полученных независимо друг от друга для разных районов. В будущих исследованиях можно использовать более репрезентативную и большую выборку данных и различные методы машинного обучения для улучшения предсказательной силы предложенного подхода.

## 3.2 Гранулярность цифровых моделей рельефа и ее связь с масштабом картографирования<sup>2</sup>

*Гранулярность* относится к одной из важнейших характеристик детальности данных и характеризует типичный размер объектов или их элементов [Ruas, Bianchin, 2002]. В данном разделе представлен вычислительный подход, формализующий определение гранулярности ЦМР и ее пригодности для картографирования в определенном масштабе [Samsonov, 2022]. На первом этапе расчет гранулярности производится с помощью новой метрики, называемой *шириной формы рельефа*. На следующем этапе оптимальная

<sup>2</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov, 2022]

гранулярность определяется статистически на основе ЦМР, восстановленных по топографическим картам на территории с различными типами рельефа. Далее представлены результаты экспериментальной работы по достижению оптимальной детальности картографического изображения рельефа в заданном масштабе путем генерализации ЦМР с ограничением по гранулярности. В заключительном эксперименте оценивается устойчивость разработанной метрики ширины формы рельефа к увеличению пространственного разрешения ЦМР. Таким образом, в настоящем разделе диссертации затрагиваются два аспекта детальности: оценка и управление.

### 3.2.1 Методика

#### 3.2.1.1 Общая идея

Разработанный подход направлен на создание надежной методики оценки гранулярности ЦМР, удовлетворяющего следующим требованиям:

1. **Отсутствие параметров:** метод должен быть пригоден для оценки гранулярности ЦМР без каких-либо параметров.
2. **Учет высоты:** расчет гранулярности ЦМР должен учитывать не только горизонтальные, но и вертикальные размеры форм рельефа. Более высокие формы при прочих равных условиях должны оказывать большее влияние на величину гранулярности.
3. **Устойчивость к повышающей передискретизации:** изменения гранулярности должны быть незначительными при увеличении пространственного разрешения ЦМР (уменьшении размера ячейки).

Первое требование направлено на то, чтобы сделать методику полностью автоматической. Второе требование учитывает тот факт, что более высокие формы рельефа сильнее выражены в его картографическом изображении. Третье требование гарантирует, что переход к меньшему размеру ячейки путем передискретизации не изменит гранулярность, поскольку никакие формы рельефа в данном случае не добавляются и не удаляются. Аналогичное требование для понижения разрешения не имеет смысла, поскольку такой переход неизбежно приведет к уничтожению мелких форм рельефа и, следовательно, уменьшению гранулярности.

#### 3.2.1.2 Ширина формы рельефа

Как было определено ранее, гранулярность определяется типичным размером географического объекта, представленного в пространственных данных. Для ЦМР таким объектом является форма рельефа. В Разделе [3.4.2.3](#) представлен алгоритм оценки ширины области между горизонталями основного сечения для построения дополнительных

горизонталей при изображении рельефа. Этот же подход используется для оценки размера формы рельефа через ее ширину, которую можно формально определить следующим образом:

**Ширина формы рельефа** в каждой точке — это диаметр наибольшего круга, который охватывает точку и вписан в границу формы рельефа.

Главный вопрос заключается в том, можно ли построить такой круг, не имея точной границы формы рельефа. Поскольку при расчете евклидова расстояния упомянутым методом будет искаться ближайшая точка вдоль границы формы рельефа, можно не искать линейные границы, а аппроксимировать их ограниченным числом **характерных точек**, которые располагаются вдоль границ. Это наблюдение значительно облегчает оценку ширины. Для этой цели можно использовать ячейки ЦМР, располагающиеся вдоль тальвегов — *дренажные ячейки* — поскольку они разделяют положительные формы рельефа и таким образом располагаются на их границах. При этом противоположные по топографическому положению ячейки водоразделов, а также другие типы особых точек (такие, как, например, седловины), как правило, не столь выражены в поверхности, если она не сильно расчленена, и поэтому ненадежны для аппроксимации границ форм рельефа.

Дренажные ячейки ЦМР традиционно выделяются путем пороговой фильтрации раstra аккумуляции стока (площади водосбора), полученного на основе модели направлений стока D8 [Mark, 1984]. Гидрологический подход, однако, не подходит для наших целей, так как нарушается первое требование к методу расчета гранулярности: в качестве параметра необходимо значение минимальной величины аккумуляции стока. В то же время, для наших целей не требуется связанная сеть дренажных линий, достаточно иметь некоторый набор ячеек, расположенных вдоль них. Это послабление открывает возможность использовать чисто морфометрический подход к выделению дренажных ячеек по ЦМР. В частности, подходящей оказалась вычислительная схема на основе скользящего окна размером  $2 \times 2$ , разработанная в [Band, 1986]. В этом методе окно центрируется по северо-западной ячейке, и отмечается ячейка с наибольшей высотой. После обработки всего раstra с помощью такого окна все немаркированные ячейки представляют собой дренажные.

Пример оценки ширины *положительных* форм рельефа по дренажным ячейкам представлен на Рисунке 3.14а. Дренажные ячейки окрашены в синий цвет. Красные точки — местоположения, выбранные вручную для иллюстрации подхода на основе кругов. Черным цветом отмечены доминантные круги, рассчитанные для красных точек. Видно, что каждый доминантный круг не обязательно центрирован относительно местоположения, которое он покрывает. Также, несмотря на то, что методика выделения дренажных ячеек пропускает тальвеги некоторых незначительных отрицательных форм, а полученный набор ячеек не является связным, можно четко видеть, что он обеспечивает достаточно плотный набор характерных точек для аппроксимации границ

положительных форм рельефа. Результирующая поверхность ширины формы рельефа показана на Рисунке 3.14б. Дополнительные поверхности, представленные на Рисунке 3.14в-г, объясняются в следующем разделе, посвященном расчету гранулярности.

### 3.2.1.3 Расчет гранулярности

После того как ширина формы рельефа оценена для каждой ячейки ЦМР, ее необходимо агрегировать, чтобы получить одно числовое значение, характеризующее гранулярность ЦМР. Наивный подход мог бы заключаться в вычислении среднего или медианного значения ширины по всем ячейкам ЦМР. Однако полученное значение будет нерепрезентативным, если ЦМР покрывает морфологически разные области или если она зашумлена.

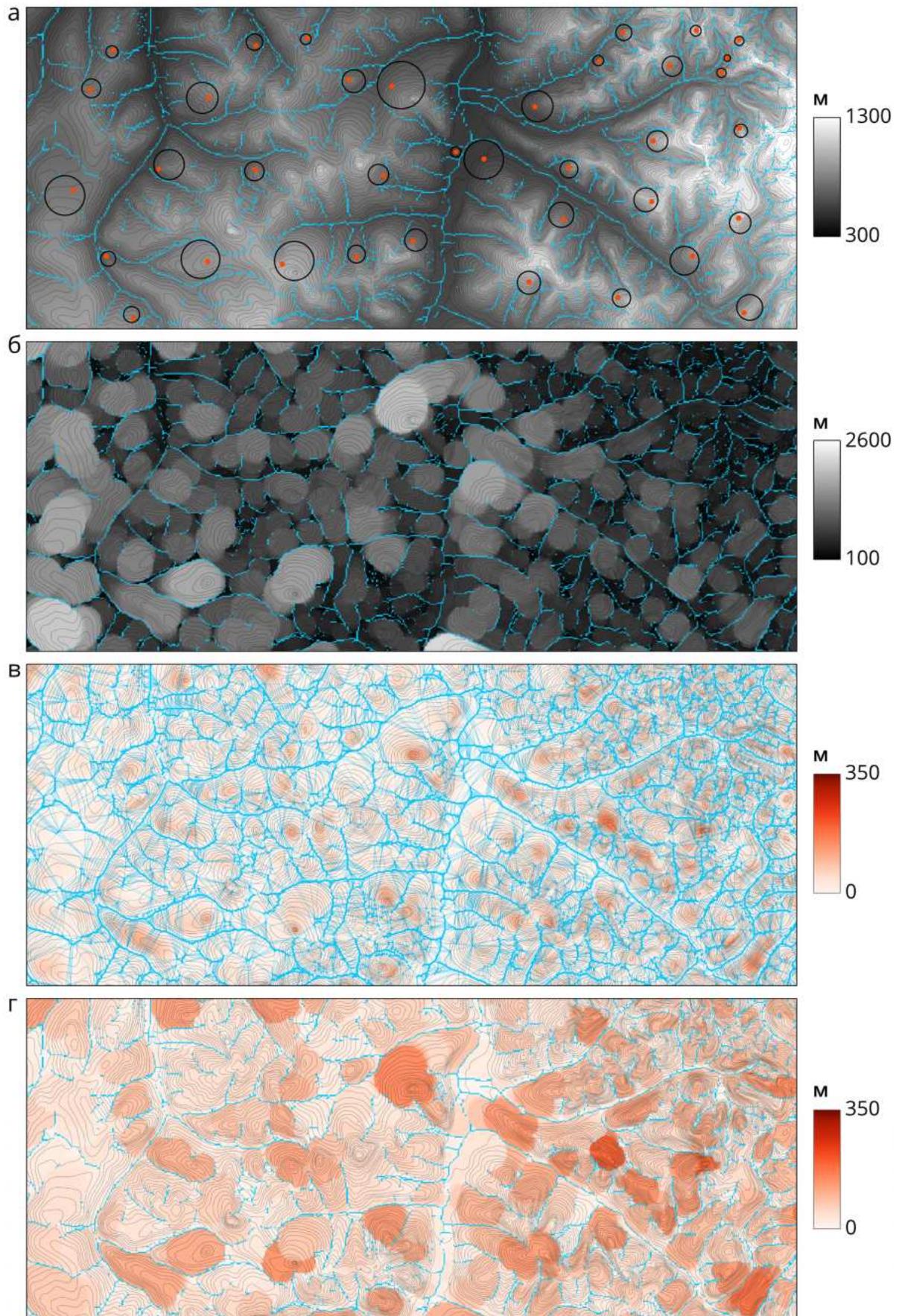
Типичный пример первой ситуации — картографирование предгорий: в этом случае сильно расчлененная местность с плотным размещением четко распознаваемых форм рельефа в одной части карты соседствует с равнинной областью, где число и выразительность форм рельефа существенно меньше. Простое усреднение по всей территории будет смещать оценку гранулярности в сторону более крупных в плане равнинных форм рельефа, несмотря на их более слабую вертикальную выраженность. И напротив, зашумленные ЦМР обычно содержат множество мелких, но незначительных искусственных неровностей, которые будут смещать гранулярность в сторону меньшей величины.

Можно отметить, что в обоих случаях формы рельефа, которые смещают результирующую гранулярность в сторону нерепрезентативных значений, имеют малые вертикальные размеры. Следовательно, для устранения такой смещенности гранулярность  $G$  должна быть рассчитана как *средневзвешенное* значение ширины формы рельефа, где *высота формы*  $H$  в точке  $p$  является весовым коэффициентом:

$$G = \frac{\sum_p W(p)H(p)}{\sum_p H(p)}. \quad (3.11)$$

Для расчета  $H$  необходима базисная поверхность «нулевого» уровня  $Z_0$ , относительно которого можно оценить высоту формы рельефа в вертикальном измерении. Поскольку в нашем случае рассматриваются положительные формы рельефа, такая поверхность может быть реконструирована по высотам дренажных ячеек. Метод интерполяции для этой цели должен быть локальным, т.е. использовать высоты только тех дренажных ячеек, которые ограничивают текущую форму рельефа. *Интерполяция на основе триангуляции* удовлетворяет этому свойству и поэтому применяется в данной работе. Проинтерполировав  $Z_0$ , получаем в каждой точке относительную высоту над базисной поверхностью  $Z(p) - Z_0(p)$ . Эта переменная показана на Рисунке 3.14в.

Пусть точка  $p$  находится внутри доминантного круга  $\widehat{C}$ , который определяет ло-



**Рисунок 3.14** – Предварительные этапы расчета гранулярности: (а) места расположения образцов и их доминантные окружности, вписанные в набор дренажных ячеек; (б) ширина рельефа; (в) относительная высота - разница между высотой и базовой поверхностью, восстановленной по триангуляции дренажных ячеек; (г) высота рельефа

кальную ширину формы рельефа. Будем считать высотой этой формы рельефа максимальную относительную высоту  $Z(q) - Z_0(q)$  по всем точкам  $q$  в пределах того же доминантного круга:

$$H(p) = \max_q [Z(q) - Z_0(q)], \quad q \in \widehat{C}, \quad p \in \widehat{C} \quad (3.12)$$

Полученное значение распространяется на все ячейки, которые охватываются доминантным кругом. Результирующая поверхность высоты формы рельефа представлена на Рисунке 3.14г. Ее использование в качестве веса при осреднении ширины формы рельефа позволяет удовлетворить второе требование к вычислению гранулярности: учет высоты форм рельефа. Переменные высоты и ширины должны вычисляться одновременно, чтобы уменьшить необходимость дублирования операции распространения ширины и высоты в пределах доминантных кругов.

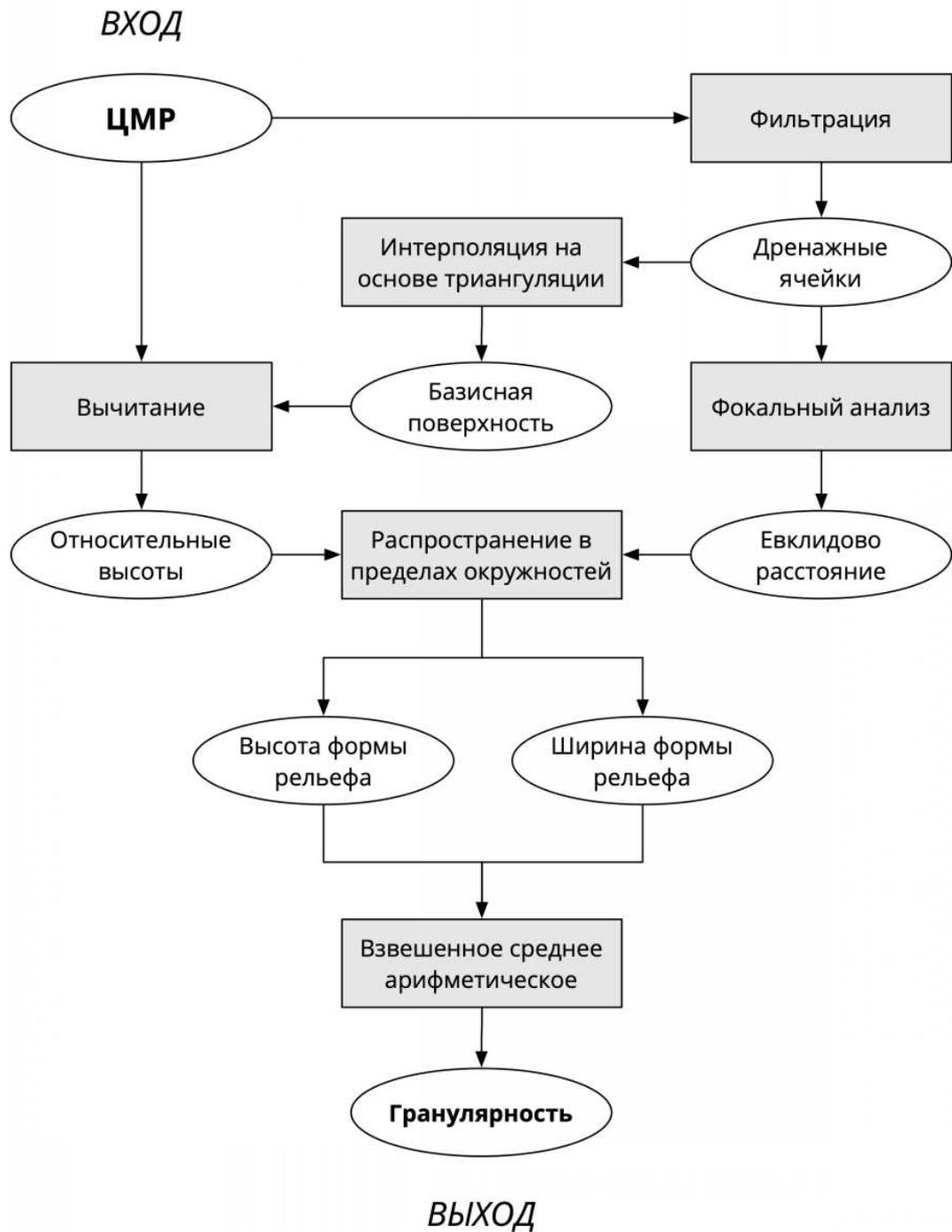
Итоговый алгоритм расчета гранулярности, который обобщает изложенные принципы, представлен на Рисунке 3.15.

Поскольку качество графического представления обычно оценивается в единицах картографического изображения (мм или см), целесообразно различать *абсолютную* гранулярность в единицах системы координат ЦМР, и *относительную* гранулярность в единицах картографического изображения. Например, если абсолютная гранулярность ЦМР составляет 1500 м, то ее относительная величина в масштабах 1 : 200 000 и 1 : 500 000 будет равна 7.5 и 3 мм соответственно. При дальнейшем изложении для краткости будем использовать термин «гранулярность» для обозначения абсолютной гранулярности, в то время как относительная гранулярность будет обозначаться в явном виде.

#### 3.2.1.4 Устойчивость к повышению разрешения

Пространственное горизонтальное разрешение (далее — просто разрешение) и гранулярность являются неотъемлемыми свойствами детальности. В частности, разрешение определяет *нижнюю границу* детальности: одна ячейка ЦМР — это наименьшее пространство, доступное для представления информации о рельефа. Понижение разрешения (увеличение размера ячейки ЦМР) неизбежно влияет на детальность, поскольку это деструктивная операция, загроубляющая представление рельефа. Противоположный процесс — повышение разрешения (уменьшение размера ячейки ЦМР) не является деструктивным, но может привнести новые тонкие детали в поверхность, вызванные интерполяцией. С точки зрения пользователя ЦМР нежелательно, чтобы метод оценки гранулярности был чувствителен к этим несущественным изменениям. Для подобной чувствительности используется коэффициент изменения  $C_v$ :

$$C_v = \frac{|v_0 - v_1|}{\min(v_0, v_1)} \times 100\%, \quad (3.13)$$



**Рисунок 3.15** – Алгоритм вычисления гранулярности ЦМР. Овалами обозначены растровые данные, прямоугольниками – операции

Таблица 3.2 – Сечение рельефа, м

Уровень детализации	Исходный масштаб	Равнины	Горы
200	1:200 000	20	40
500	1:500 000	50	100
1000	1:1 000 000	50, 100	100, 200

где  $v_0$  и  $v_1$  – значения переменной  $v > 0$  до и после преобразования. Влияние повышения разрешения на гранулярность будем считать незначительным, если значение  $C_v$  для гранулярности хотя бы на порядок (в 10 раз) меньше значения  $C_v$  для разрешения растра. Например, если размер ячейки ЦМР уменьшен с  $R_0 = 100$  до  $R_1 = 50$  м ( $C_R = 100\%$ ), а гранулярность исходной ЦМР составляет  $G_0 = 500$  м, то результирующее значение  $G_1$  должно быть в диапазоне 455 – 550 м чтобы обеспечить  $C_G \leq 10\%$ . Выполнение этого условия проверяется путем передискретизации ЦМР в сторону меньшего размера ячейки и выполнения соответствующих Уравнению 3.13 расчетов.

## 3.2.2 Реализация методики и обработка данных

### 3.2.2.1 Генерация ЦМР

Для настоящего исследования использовалась топографическая база данных с тремя уровнями детализации (УД), созданная на основе цифровых топографических карт масштабов 1 : 200 000, 1 : 500 000 и 1 : 1 000 000 (введем для них обозначения: УД 200, 500 и 1000), составленных в 2008–2010 гг. подведомственными учреждениями Федеральной службы государственной регистрации, кадастра и картографии (Росреестр) на основе генерализации крупномасштабных карт. Каждая карта содержит набор слоев, отражающих рельеф, гидрографию, растительность, транспорт, населенные пункты и другие стандартные элементы содержания топографических карт. В частности, рельеф представлен горизонталями с сечением рельефа согласно Таблице 3.2.

Поскольку горизонталы на этих картах вручную генерализованы картографами для достижения оптимальной детальности изображения, они были использованы для восстановления эталонных цифровых моделей рельефа, пригодных для исследования гранулярности. Было выделено 36 фрагментов данных с центрами относительно населенных пунктов РФ, расположенных в пределах территорий с разными типами рельефа (с учетом трех уровней детализации всего получилось 108 фрагментов данных). Каждый фрагмент был обрезан по прямоугольнику  $100 \times 100$  км, а затем спроецирован в азимутальную равновеликую проекцию Ламберта с соответствующим центром. Карта расположения фрагментов данных представлена на Рисунке 3.16.

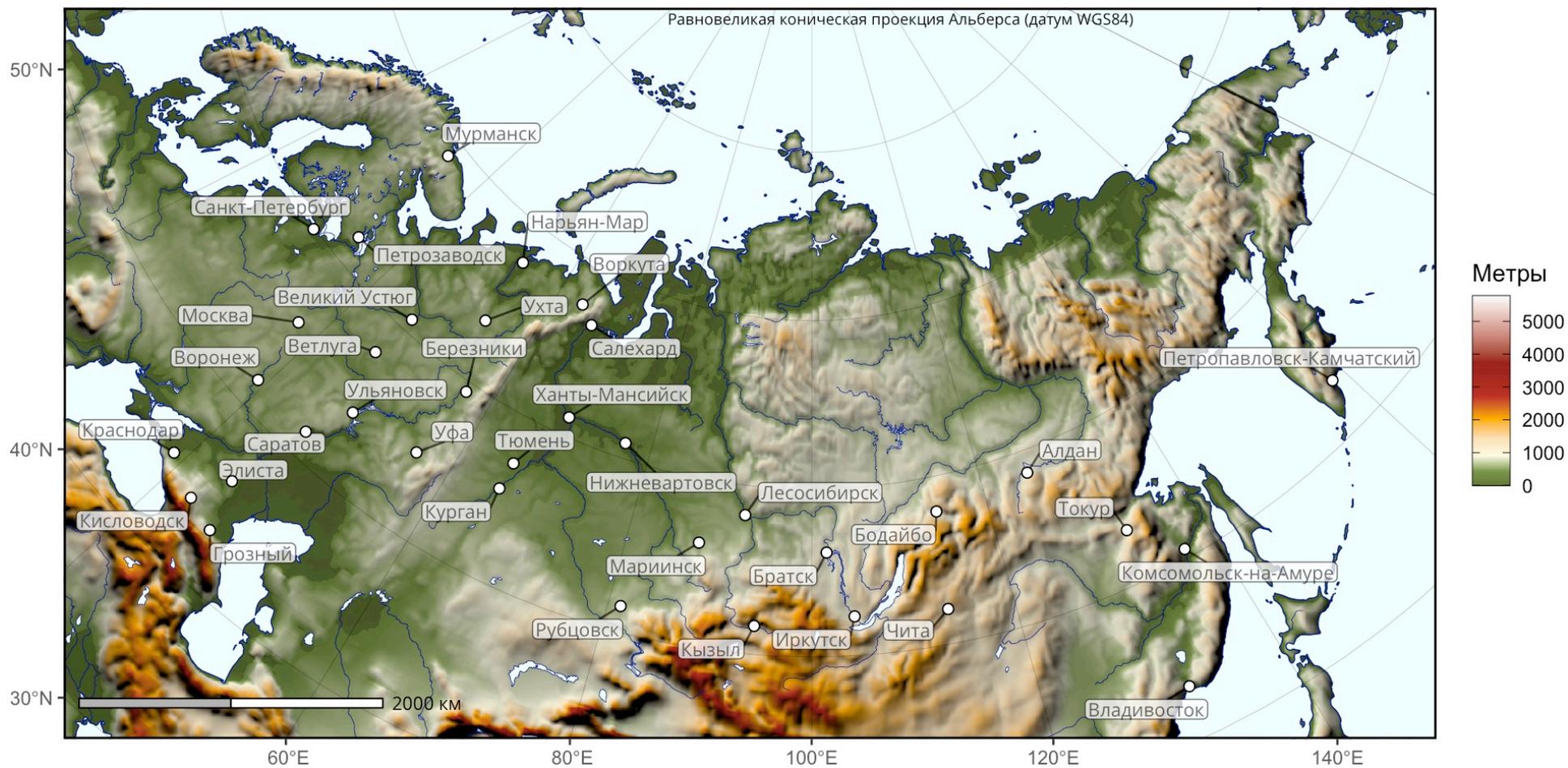


Рисунок 3.16 – Местоположение фрагментов ЦМР

Растровые ЦМР были реконструированы на основе горизонталей и точечных отметок высот с использованием подхода, разработанного С. М. Кошелем [Koshel, 2012], в основе которого лежит «волновой» алгоритм вычисления расстояний от интерполируемой точки до двух ближайших горизонталей. Поскольку назначение данных ЦМР — картографическое изображение рельефа, они должны обладать достаточно высоким разрешением, чтобы отражать резкие переломы поверхности рельефа, такие как острые гребни. В частности, для печатных карт, с которыми работают с типичного расстояния 30 см, необходимо обеспечить графическую точность 0,1 мм [Kennelly et al., 2021] (что, например, означает, что для карты в масштабе 1 : 500 000 необходимо использовать ЦМР с разрешением 50 м). Для карт на экранах компьютеров, которые просматриваются с расстояния в 2 — 3 раза больше, можно использовать пропорционально более грубое разрешение. Исходя из этих соображений разрешение ЦМР было установлено равным 0,25 мм в масштабе картографирования, а разрешение (размеры) ЦМР для УД 200, 500 и 1000 — 50 м (2000 × 2000 ячеек), 125 м (800 × 800 ячеек) и 250 м (400 × 400 ячеек) соответственно.

### 3.2.2.2 Эксперименты

Для каждой ЦМР был выполнен расчет растров ширины и высоты формы рельефа и оценка гранулярности на их основе согласно разработанной методике. Обработка заняла в среднем 47, 7 и 3 секунд для ЦМР уровней детализации 200, 500 и 1000. Картографические представления полученных ЦМР для УД 1000, а также рассчитанные растры ширины формы рельефа приведены в Приложении на Рисунках A.4 — A.14. При визуализации использована классическая комбинация горизонталей, градиентной гипсометрической окраски и аналитической отмывки. Сечение рельефа горизонталями выбрано в соответствии с топографическими картами (Таблица 3.2), цветовая шкала одинакова для всех фрагментов вне зависимости от размаха высот.

Гранулярность полученных ЦМР была подвергнута статистическому анализу с целью выявления ее оптимальных величин для представления рельефа на мелкомасштабных картах. В частности, для каждого уровня детализации были рассчитаны медиана и интерквартильный размах гранулярности. Полученные значения также были интерпретированы в контексте радикального закона Топфера [Töpfer, Pillewizer, 1966] с целью количественной оценки интенсивности отбора форм рельефа при уменьшении масштаба.

Для оценки эффективности гранулярности как одного из критериев детальности ЦМР был проведен следующий эксперимент. Были извлечены два фрагмента ЦМР NASADEM [NASADEM, 2020], имеющие тот же пространственный охват, что и фрагменты *Бодайбо* и *Грозный*. Пространственное разрешение фрагментов составляет 1", что составляет  $\approx 16.5$  м для фрагмента *Бодайбо* и  $\approx 23$  м для фрагмента *Грозный*. Каждый фрагмент был проецирован в ту же проекцию, что и фрагмент данных топографических карт (*Азимутальная равновеликая проекция Ламберта*) с аналогичным центром. Пере-

проецированные фрагменты ЦМР NASADEM были генерализованы с помощью метода структурного упрощения [Samsonov, 2011] и сглаживания с сохранением особенностей [Sun et al., 2007] таким образом, что их гранулярность стала приблизительно равной ( $\pm 5\%$ ) гранулярности ЦМР, созданной на основе топографической карты.

Упрощение использовалось для уменьшения детальности представления рельефа, а сглаживание — для устранения шума в данных и сглаживания разрывов поверхности, возникающих при интерполяции структурных линий методом триангуляции. Разрешение генерализованных ЦМР было установлено равным 50, 125 и 250 м соответственно. Для сравнения исходные фрагменты NASADEM были просто передискретизированы (без обобщения) с помощью кубической интерполяции до *такого же разрешения*  $R = 50, 125, 250$  м и отдельно до *такой же гранулярности*  $G$ , как и эталонные ЦМР. Передискретизированные модели сравнивались с генерализованными, чтобы проверить, можно ли использовать только лишь разрешение или гранулярность для достижения необходимой детальности изображения, или же генерализация является обязательным шагом.

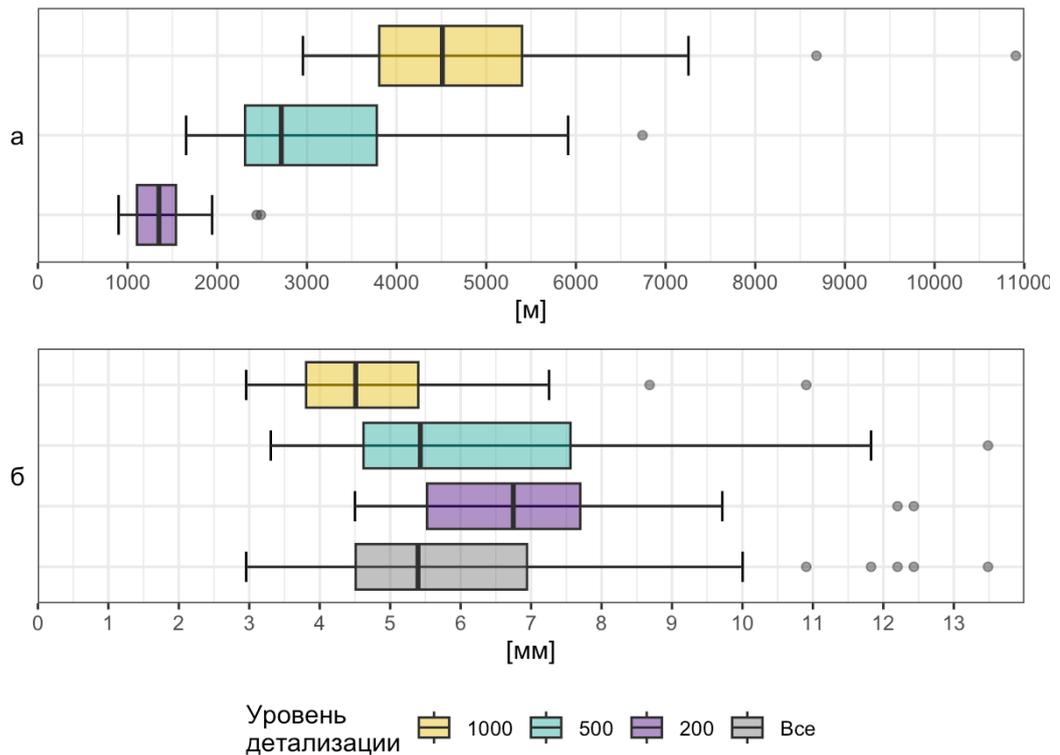
Наконец, генерализованные фрагменты ЦМР NASADEM *Бодайбо* и *Грозный* были использованы для оценки чувствительности предложенного метода расчета гранулярности к повышению разрешения ЦМР. Для этого генерализованные ЦМР были передискретизированы в 2 и 4 раза с помощью бикубической интерполяции и их гранулярность сравнивалась с гранулярностью исходной модели. Чтобы убедиться, что изменения не зависят от типа рельефа, аналогичный эксперимент был проведен для всех фрагментов эталонной ЦМР УД 1000.

### 3.2.3 Результаты

#### 3.2.3.1 Оптимальная гранулярность

Распределение гранулярности для разных уровней детализации представлено на Рисунке 3.17. В частности, глобальное медианное значение гранулярности (обозначено толстым вертикальным отрезком на каждой диаграмме) составляет 1265, 2714 и 4510 м для УД 200, 500 и 1000 соответственно. Относительная гранулярность в масштабе картографирования равна 6.75, 5.43 и 4.51 мм соответственно.

Поскольку целью эксперимента является поиск оптимальной гранулярности ЦМР, необходимо перейти к поиску значения, которое является наиболее репрезентативным для всех уровней детализации. Для этого можно использовать медианное значение относительной гранулярности 5.38 мм и ее интерквартильный интервал  $Q1 - Q3$ , который составляет 4.5 — 6.9 мм для всех 108 фрагментов ЦМР (серый прямоугольник на Рисунке 3.17б). Обратим также внимание на то, что интервал значений 5 — 6 мм не только включает в себя медиану по всем УД, но также пересекается с интерквартильными размахами всех УД по отдельности. Отсюда можно вывести следующее эмпирическое



**Рисунок 3.17** – Гранулярность ЦМР по данным топографических карт разной детальности: (а) абсолютная гранулярность; (б) относительная гранулярность (в масштабе карты) [Samsonov, 2022]

правило:

**Оптимальная относительная гранулярность** цифровой модели рельефа для средне- и мелкомасштабного картографирования составляет 5 — 6 мм.

Наблюдаемое снижение величины относительной гранулярности при уменьшении масштаба свидетельствует о том, что при генерализации изображение рельефа становится более густо заполненным различными формами рельефа (а их размер в масштабе карты уменьшается). Эта тенденция характеризует интенсивность отбора форм рельефа и может быть связана с радикальным законом Топфера, который ранее уже использовался для оценки отбора населенных пунктов в Разделе 2.1.5 и автодорог в Разделе 2.2.4. Напомним, что данный закон гласит, что в среднем при уменьшении масштаба карты в 2 раза количество изображенных на ней объектов уменьшается в  $\sqrt{2}$  раз [Töpfer, Pillewizer, 1966]. Взяв в расчет, что количество объектов обратно пропорционально площади, а площадь круга равна  $\pi r^2$ , где  $r = W/2$  для ширины формы рельефа, можем рассчитать для перехода между уровнями детализации 200 → 500 (2,5 раза):  $\frac{2^2}{\pi 5.43^2} / \frac{2^2}{\pi 6.75^2} = \left(\frac{6.75}{5.43}\right)^2 \approx 1.55 \approx \sqrt{2.39}$ . Аналогично для перехода 500 → 1000 (в 2 раза):  $\left(\frac{5.43}{4.51}\right)^2 \approx 1.45 \approx \sqrt{2.10}$ .

Таблица 3.3 – Гранулярность ЦМР

*УД	Разрешение, м	Эталон, м	Генерализация, м	Передискретизация по G, м	Передискретизация по R, м
Бодайбо					
200	50	1595	1611	1606	1073
500	125	2282	2275	2278	1407
1000	250	3678	3780	3864	1750
Грозный					
200	50	1392	1436	1339	501
500	125	2992	2966	3109	927
1000	250	4510	4448	4689	1662

\* УД – Уровень детализации

Данные расчеты показывают, что снижение детальности происходит нелинейно: переход 200 → 500 более консервативен (сохраняет больше форм рельефа, чем ожидается), а 500 → 1000 более избирателен (сохраняет меньше форм рельефа, чем ожидается). Подобная нелинейность отбора вполне естественна для топографических карт с учетом назначения карт разных масштабов. Полный переход между уровнями детализации 200 → 1000 (в 5 раз) характеризуется  $\left(\frac{6.75}{4.51}\right)^2 \approx 2.24 \approx \sqrt{5.01}$ -кратным уменьшением числа форм рельефа, что практически полностью соответствует ожидаемому значению  $\sqrt{5}$ .

Это наблюдение можно обобщить в виде следующего важного результата:

**Интенсивность отбора форм** при генерализации изображения рельефа на топографических картах подчиняется *радикальному закону Топфера*.

### 3.2.3.2 Генерализация ЦМР и оптимальная детальность

Значения гранулярности для эталонных, генерализованных и передискретизированных ЦМР, полученных в ходе второго эксперимента, представлены в Таблице 3.3.

Фрагмент *Бодайбо* охватывает горную местность с генетически однородным рельефом. Генерализация такой ЦМР — в целом несложный процесс, поскольку формы рельефа в пределах территории характеризуется схожими морфометрическими характеристиками. Сравнение генерализованных и эталонных ЦМР представлено на Рисунке 3.18. Дополнительные растры ширины формы рельефа для тех же ЦМР показаны на Рисунке 3.19. Видно, что автоматическая генерализация дает немного более гладкий результат, нежели эталон, но различия в детальности весьма незначительны, что говорит об эффективности меры гранулярности в качестве ограничивающего критерия при генерализации ЦМР. Этот вывод можно дополнить, проанализировав растры ширины форм рельефа, представленные на Рисунке 3.19, которые также характеризуются схожим ри-

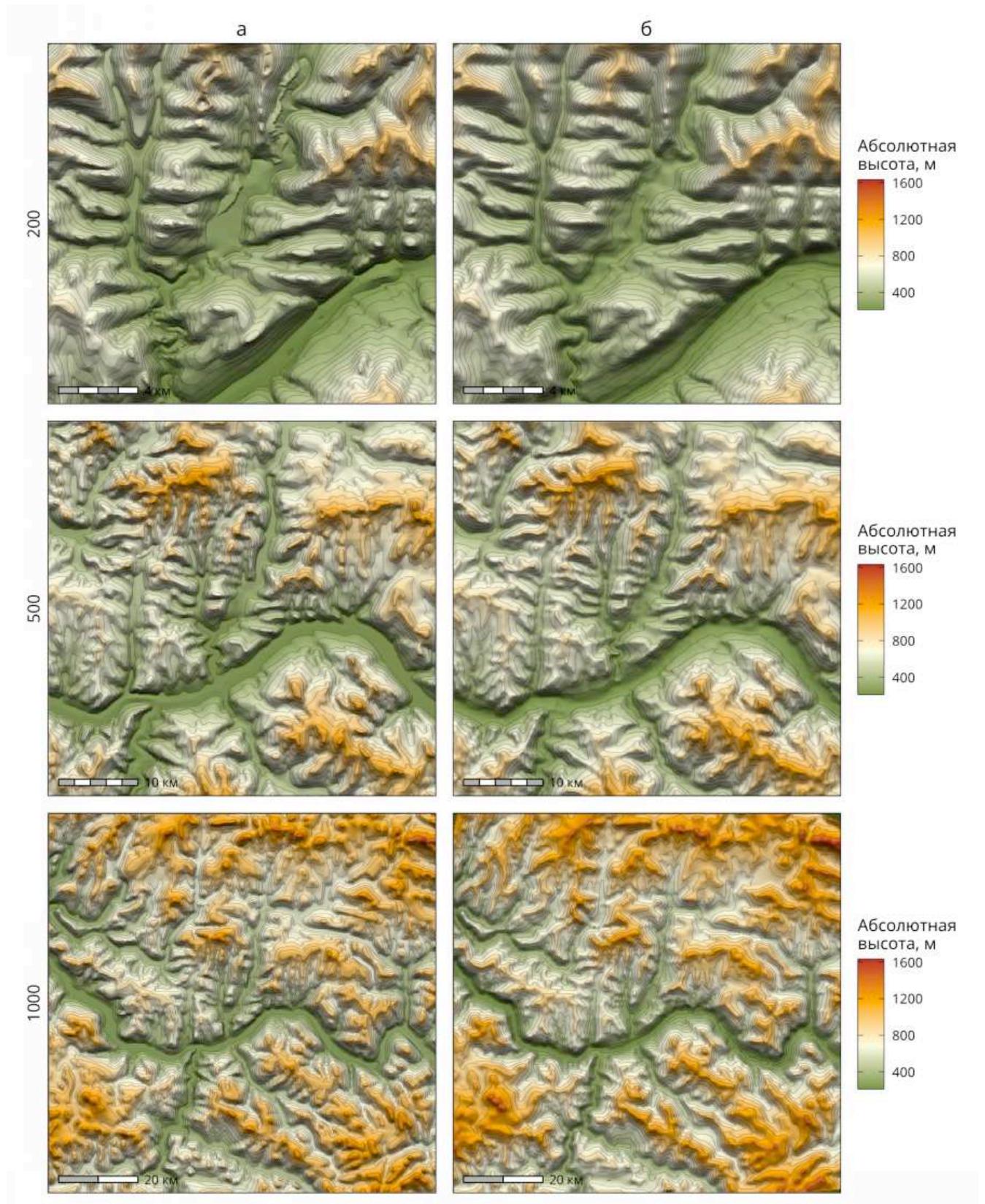
сунком доминантных кругов.

На Рисунке 3.20 представлена передискретизация фрагмента NASADEM на территорию Бодайбо до разрешения эталонных ЦМР. Очевидно, что во всех случаях детализация такой модели существенно выше, чем у эталона, причем разница становится более существенной при более грубых разрешениях, о чем свидетельствуют меньшие значения гранулярности в Таблице 3.3 (столбец *Передискретизация по R*). На Рисунке 3.21 представлена передискретизация NASADEM до гранулярности эталонных ЦМР. Полученные разрешения растров составляют 200, 425 и 850 м для УД 200, 500 и 1000 соответственно. Размер форм рельефа в полученных таким путем ЦМР довольно близок к эталонным ЦМР, что означает, что гранулярность эффективно выполняет свою функцию. Однако графическое качество представления низкое из-за слишком большого размера ячеек.

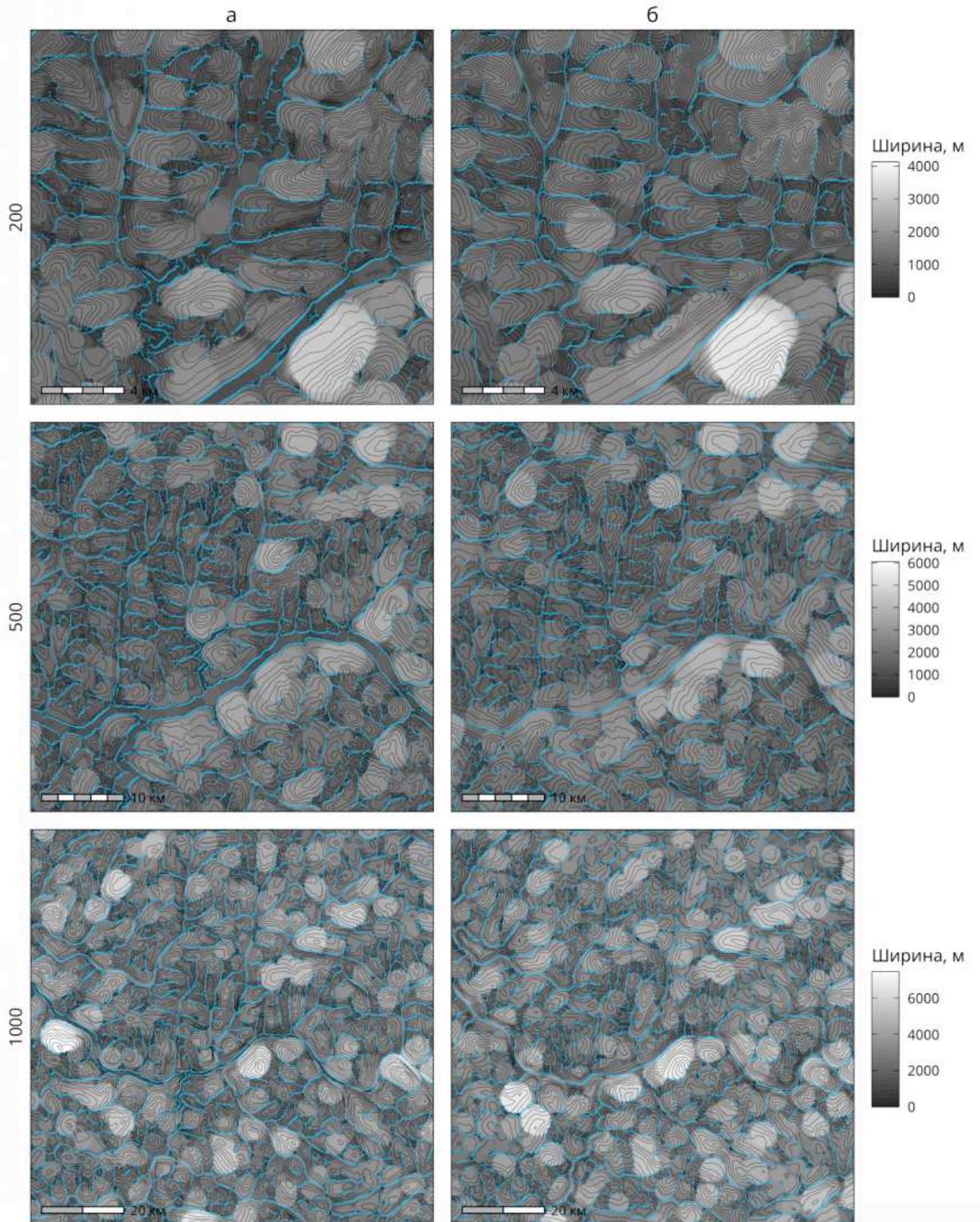
Фрагмент *Грозный* охватывает генетически неоднородное сочетание равнинной, предгорной и горной местности, что делает его довольно сложным для генерализации. Кроме того, исходный фрагмент NASADEM, покрывающий эту территорию, зашумлен. Результат генерализации NASADEM для этого фрагмента для трех уровней детализации можно сравнить с эталонными ЦМР на Рисунке 3.22. Дополнительные растры ширины формы рельефа представлены на Рисунке 3.23. В соответствии с ожиданиями генерализованные ЦМР содержат большое количество мелких форм рельефа на зашумленных участках. Однако, поскольку эти формы рельефа имеют небольшую высоту, их вес в оценке гранулярности незначителен. Таким образом, автоматически генерализованные ЦМР лишь немногим более гладкие, чем эталонные, имеют визуально схожий уровень детализации.

На Рисунке 3.24 представлен результат передискретизации фрагмента NASADEM до разрешений эталонных ЦМР. На Рисунке 3.25 показан результат передискретизации NASADEM до гранулярности эталонных ЦМР. Полученные разрешения растров составляют 225, 425, и 575 м для УД 200, 500 и 1000 соответственно. Можно отметить, что для УД 1000 разрешение (575 м) существенно выше, чем для Бодайбо (850 м). Это означает, что из-за морфометрических особенностей рельефа понижение разрешения сильнее изменяет его гранулярность. Визуальный же анализ результатов приводит к тем же выводам, что и в случае с Бодайбо: передискретизация до одинакового разрешения приводит к чрезмерной детальности, в то время как передискретизация до схожей гранулярности дает схожий размер форм рельефа, но излишне грубое разрешение ЦМР оказывается недостаточным для обеспечения высокого картографического качества изображения рельефа.

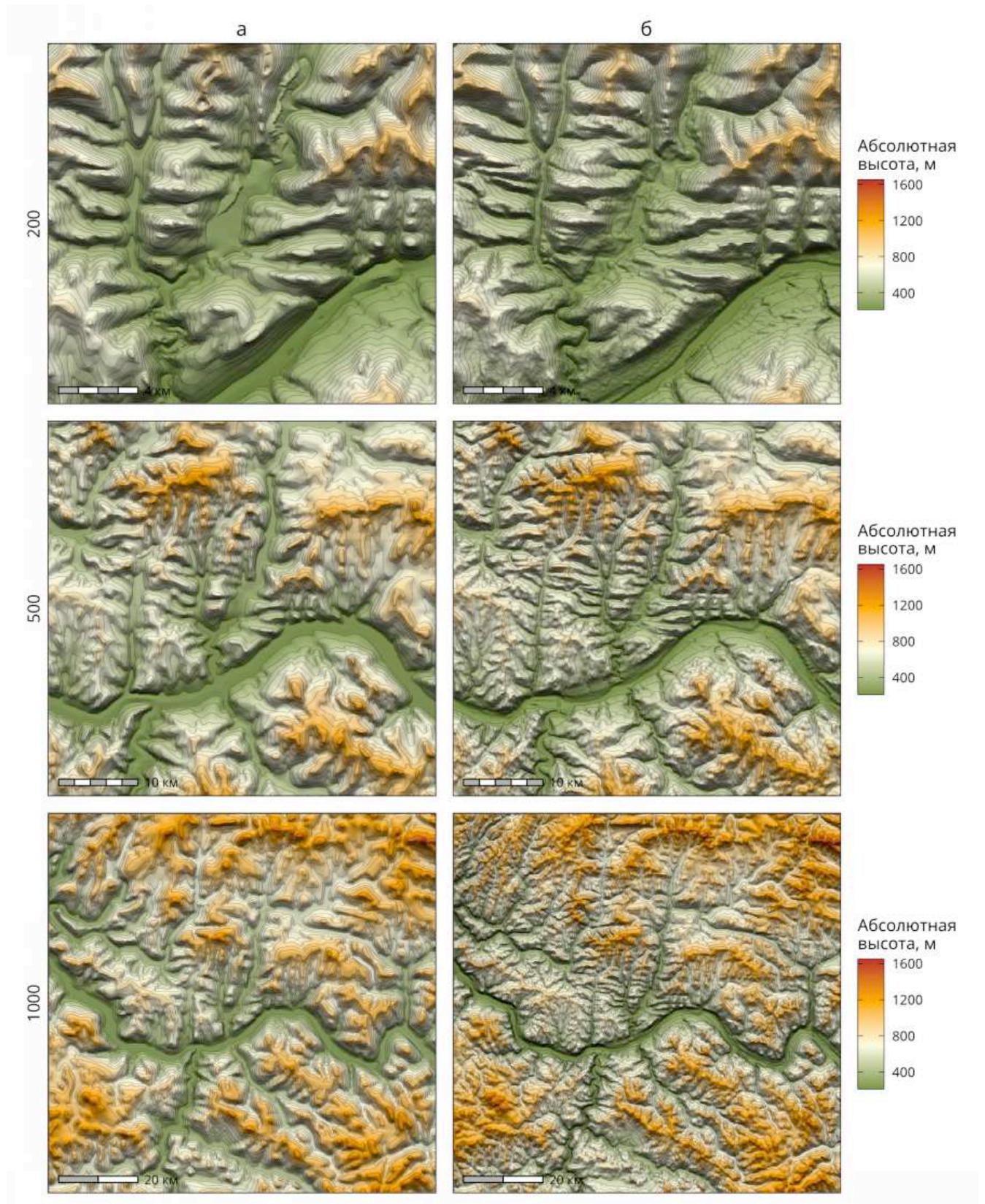
В дополнение к визуальным сравнениям различий в детализации экспериментальных ЦМР была проведена статистическая оценка ширины растров рельефа. На Рисунке 3.26 показаны распределения ширины формы рельефа для фрагментов Бодайбо (а) и Грозный (б). Анализ этих рисунков показывает, что результаты, ограниченные гранулярностью (обобщение и передискретизация до  $G$ ), значительно более близки к



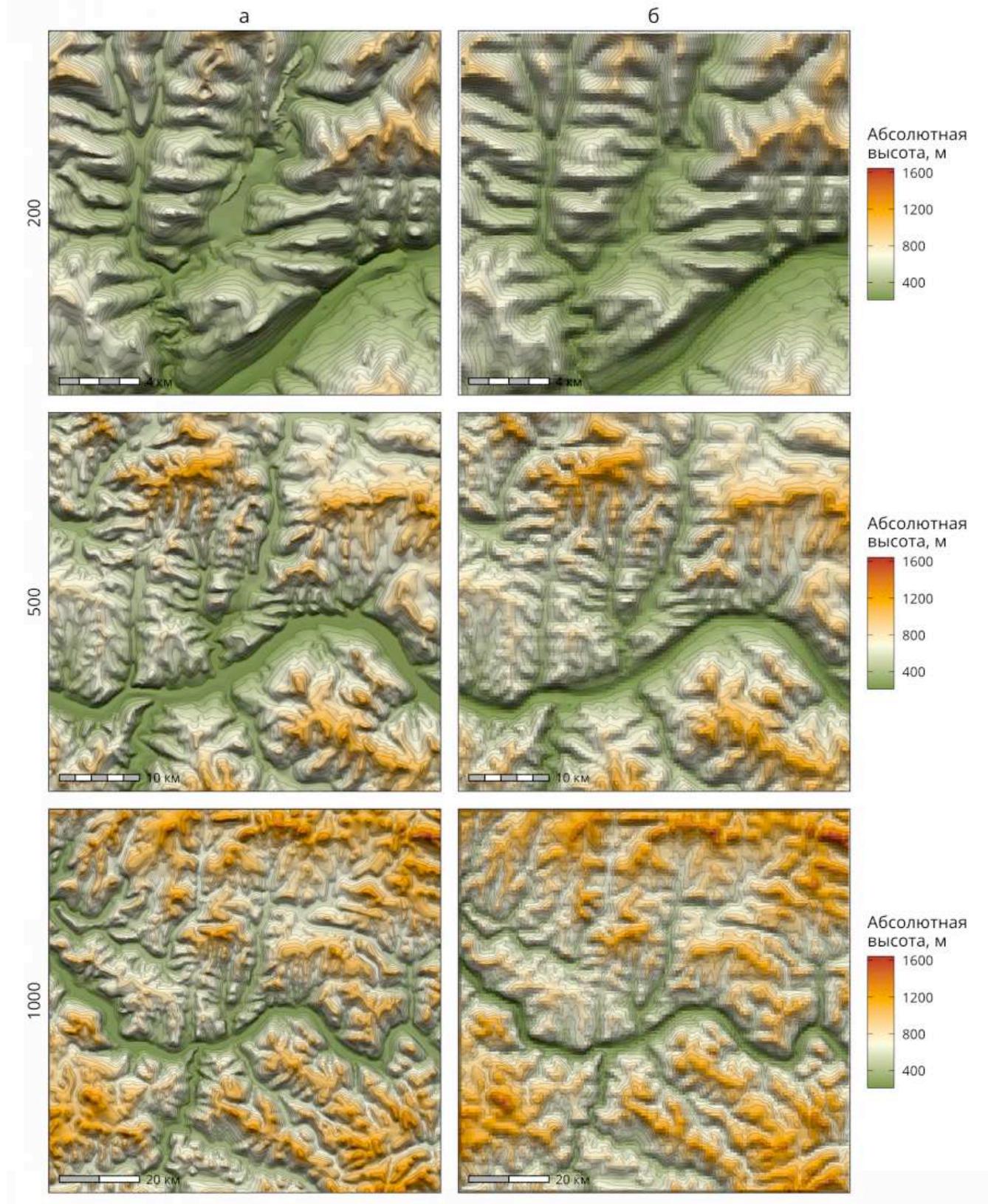
**Рисунок 3.18** – Сравнение абсолютных высот для (а) эталонной ЦМР, построенной по цифровой топографической карте (ЦТК) и (б) фрагмента NASADEM, генерализованного до той же гранулярности. Фрагмент Бодайбо. По строкам указаны уровни детализации ЦТК.



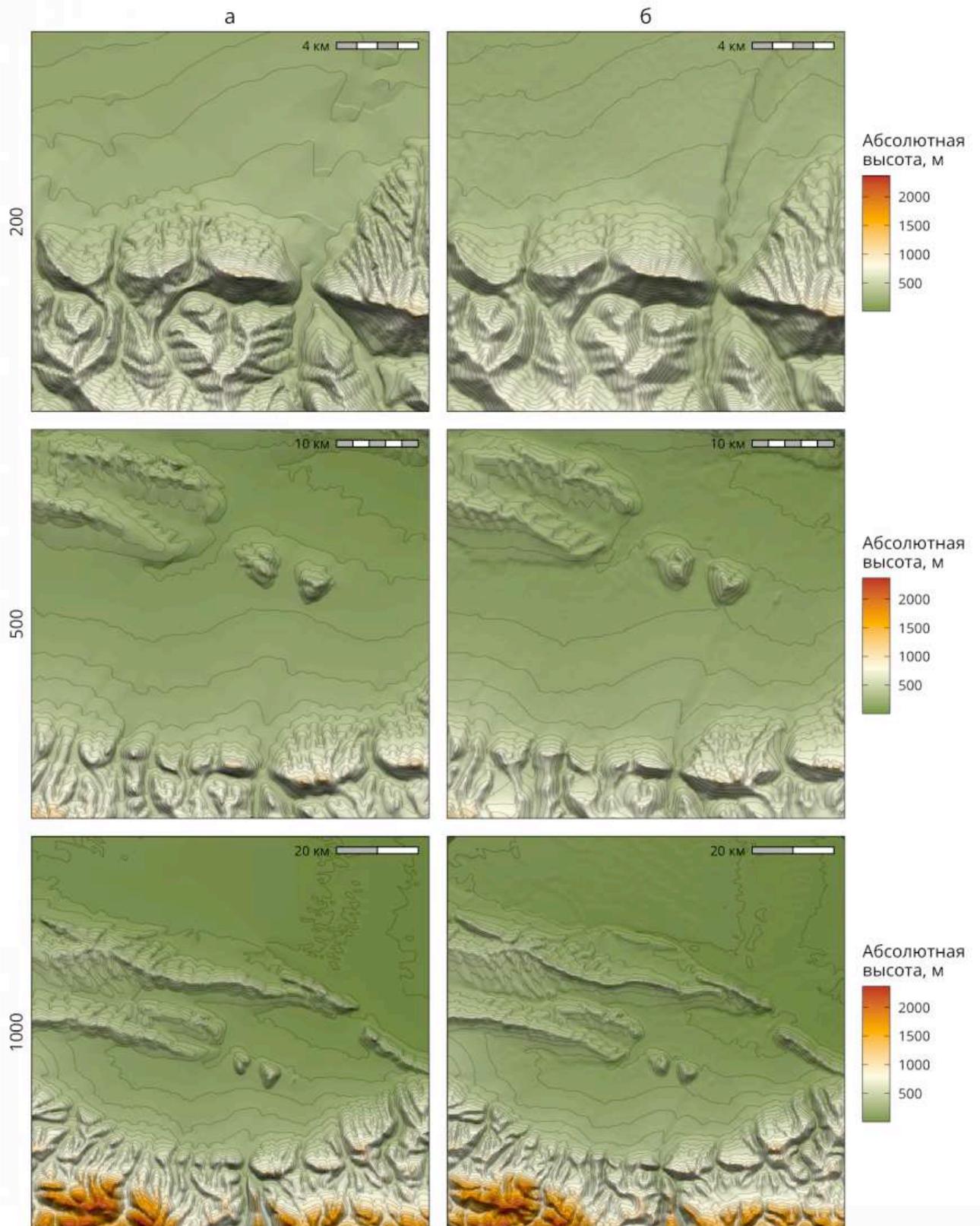
**Рисунок 3.19** – Сравнение ширины формы рельефа для (а) эталонной ЦМР, построенной по цифровой топографической карте (ЦТК) и (б) фрагмента NASADEM, генерализованного до той же гранулярности. Фрагмент Бодайбо. По строкам указаны уровни детализации ЦТК.



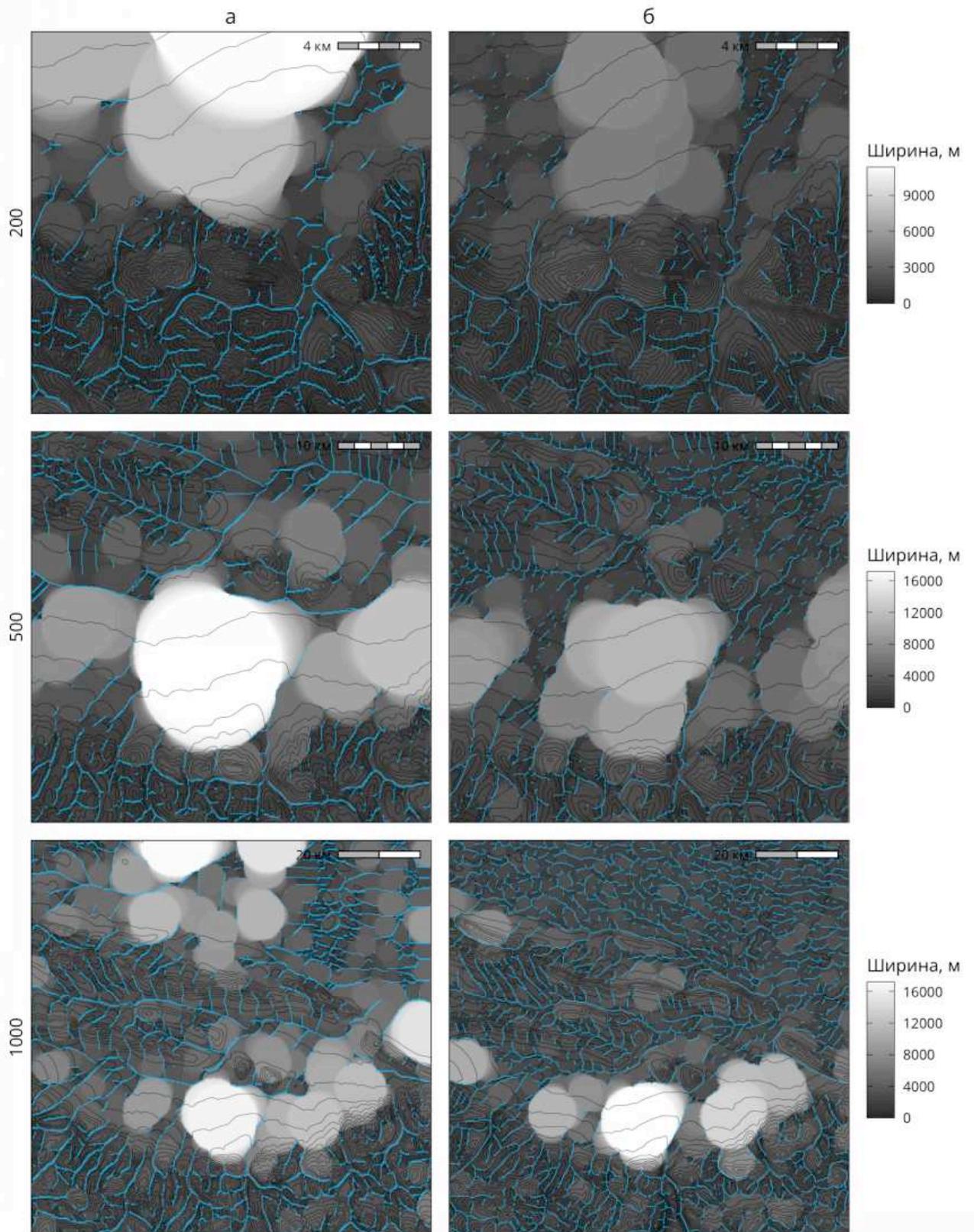
**Рисунок 3.20** – Сравнение абсолютных высот для (а) эталонной ЦМР, построенной по цифровой топографической карте (ЦТК) и (б) фрагмента NASADEM, передискретизированного до такого же разрешения. Фрагмент Бодайбо. По строкам указаны уровни детализации ЦТК



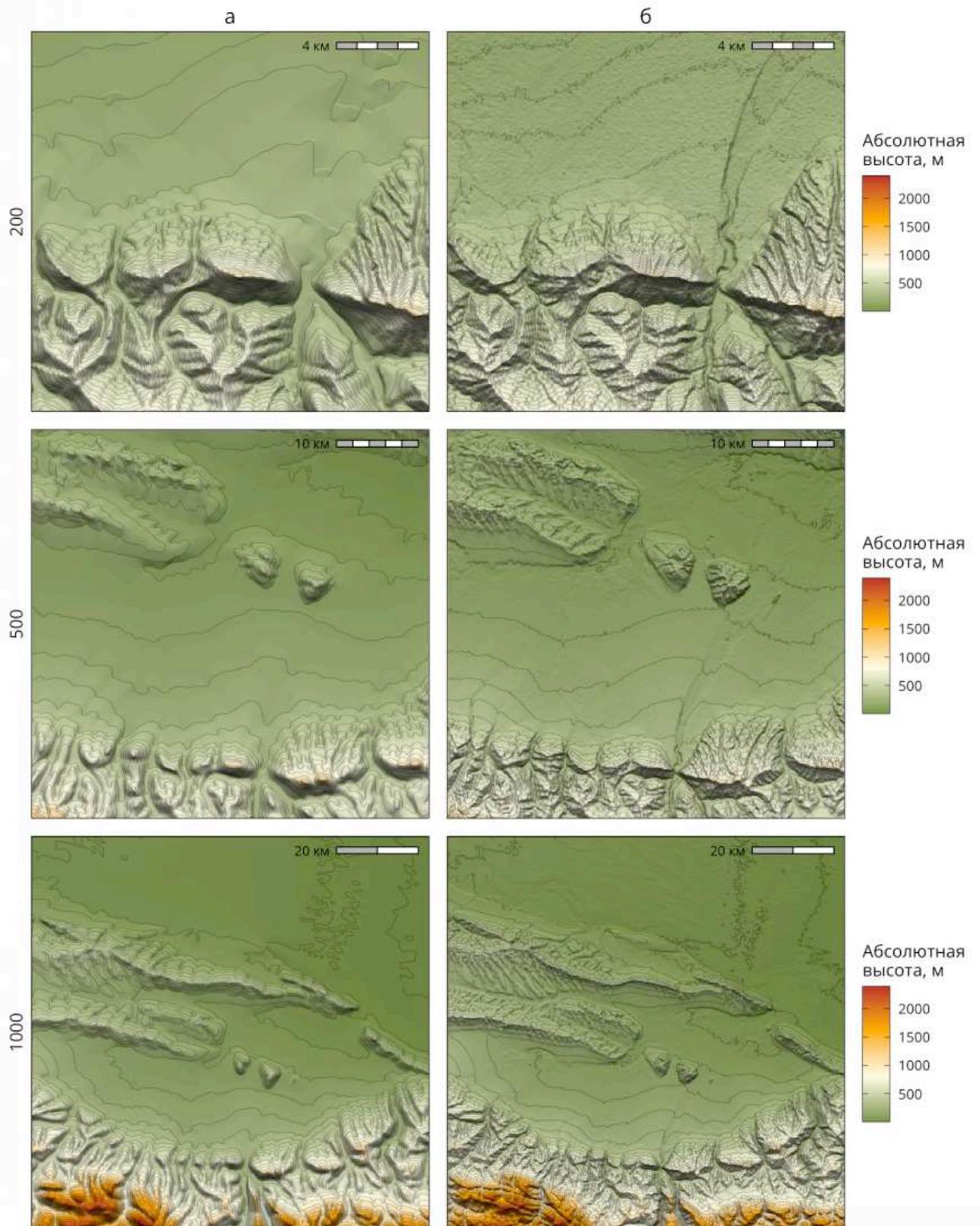
**Рисунок 3.21** – Сравнение абсолютных высот для (а) эталонной ЦМР, построенной по цифровой топографической карте (ЦТК) и (б) фрагмента NASADEM, передискретизированного до аналогичной гранулярности. Фрагмент Бодайбо. По строкам указаны уровни детализации ЦТК



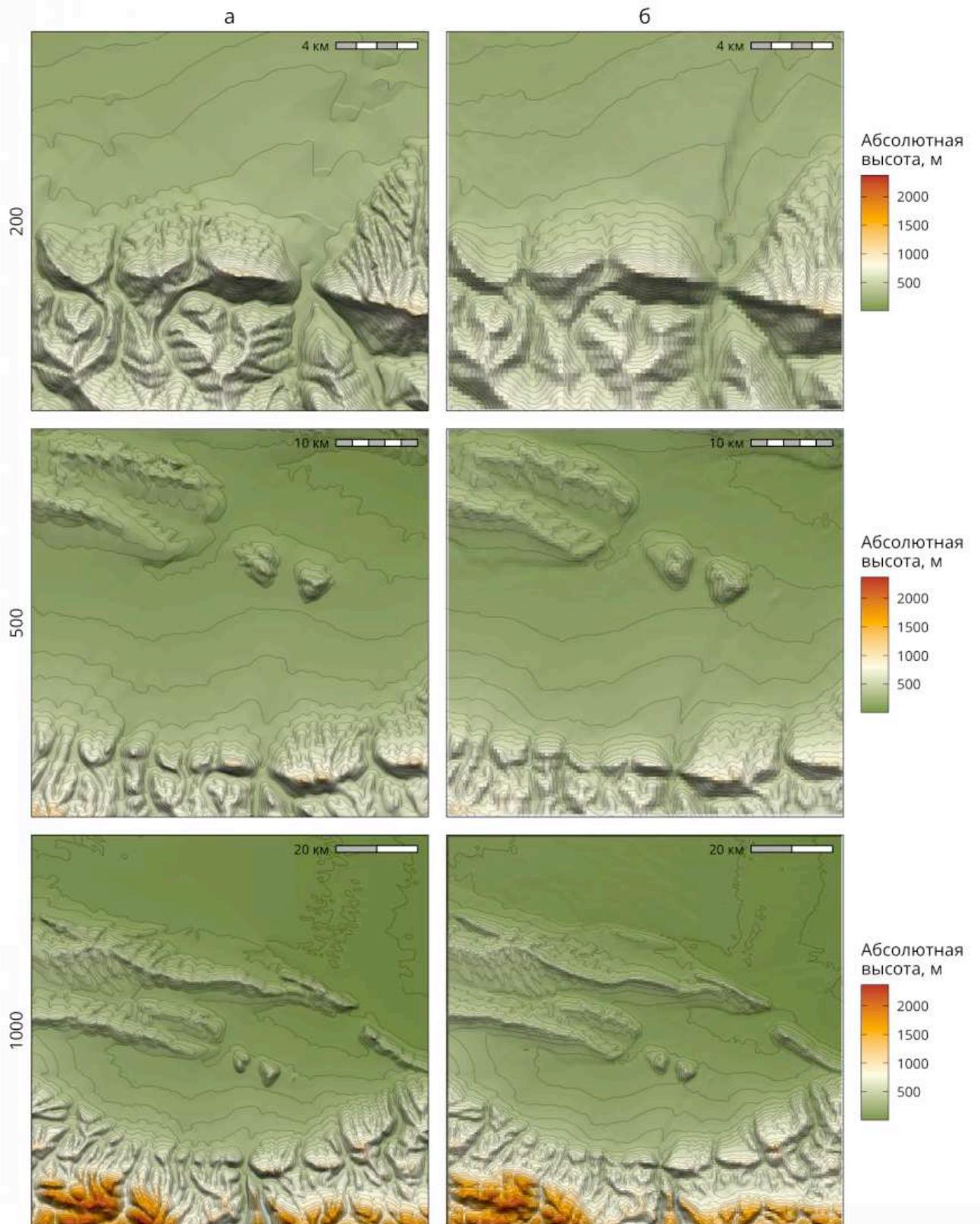
**Рисунок 3.22** – Сравнение абсолютных высот для (а) эталонной ЦМР, построенной по цифровой топографической карте (ЦТК) и (б) фрагмента NASADEM, генерализованного до той же гранулярности. Фрагмент Грозный. По строкам указаны уровни детализации ЦТК.



**Рисунок 3.23** – Сравнение ширины формы рельефа для (а) эталонной ЦМР, построенной по цифровой топографической карте (ЦТК) и (б) фрагмента NASADEM, генерализованного до той же гранулярности. Фрагмент Грозный. По строкам указаны уровни детализации ЦТК.



**Рисунок 3.24** – Сравнение абсолютных высот для (а) эталонной ЦМР, построенной по цифровой топографической карте (ЦТК) и (б) фрагмента NASADEM, передискретизированного до такого же разрешения. Фрагмент Грозный. По строкам указаны уровни детализации ЦТК.



**Рисунок 3.25** – Сравнение абсолютных высот для (а) эталонной ЦМР, построенной по цифровой топографической карте (ЦТК) и (б) фрагмента NASADEM, передискретизированного до аналогичной гранулярности. Фрагмент Грозный. По строкам указаны уровни детализации ЦТК

обобщенной вручную эталонной ЦМР, чем передискретизация до того же разрешения  $R$ . Это отражается в форме кривой плотности распределения. При этом на диаграммах размаха («ящиках с усами») для обоих фрагментов видно значительное количество выбросов, которые свидетельствуют о неоднородном характере рельефа: преобладание мелких форм при заметном количестве крупных форм. На генетически разнородном Грозненском фрагменте это свойство проявляется ярче.

Статистическая проверка различий средних и дисперсий между распределениями с помощью тестов Стьюдента (Уэлча) и Фишера показывает высокосignификантные ( $p < 0,001$ ) различия при любом сочетании распределений, что связано, прежде всего, с высокой мощностью статистики: количество ячеек ЦМР для УД 200, 500 и 1000 составляет около 4 000 000, 600 000 и 150 000 соответственно. Интегральная оценка различий распределений была выполнена с помощью статистики Колмогорова-Смирнова  $D$ , значения которой представлены в Таблице 3.4. Результаты также высоко значимы ( $p < 0.001$ ), что означает, что распределения отличаются друг от друга. Это дает нам статистическое обоснование того, что генерализация ЦМР превосходит передискретизацию до заданной гранулярности  $G$ , с точки зрения близости к ручному эталону. В худшем случае УД 200 для фрагмента Грозный генерализованная ЦМР в  $0.3015/0.1820 \approx 1.66$  раз ближе к эталонной, в то время как в лучшем случае УД 200 и 500 фрагмента Бодайбо  $D < 0.0001$  для генерализованной ЦМР, а передискретизация по  $G$  дает значения больше 0.01 в обоих случаях. Как и ожидалось, передискретизация по  $R$  дает наихудшую статистику близости к исходному распределению высот.

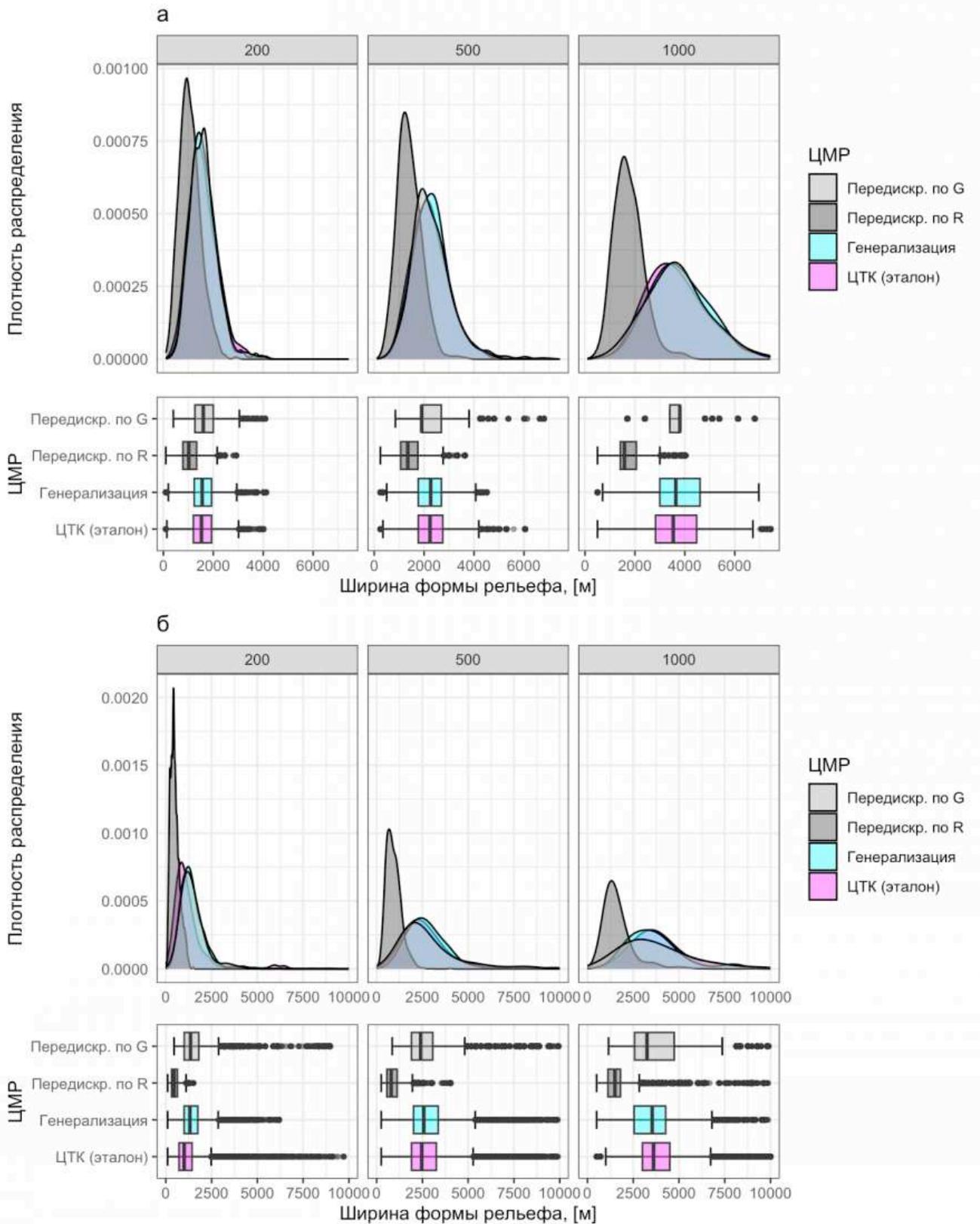
Результаты, полученные в этой части эксперимента, позволяют сформулировать следующее утверждение:

**Оптимальная детальность** ЦМР для мелкомасштабного картографирования рельефа требует сочетания генерализации ЦМР до необходимой гранулярности и высокого разрешения результирующей ЦМР, которое обеспечивает высокое качество визуализации. Передискретизация ЦМР без выполнения генерализации для этой цели не достаточна.

### 3.2.3.3 Устойчивость к повышению разрешения

Коэффициенты изменения разрешения ЦМР для передискретизации  $\times 2$  и  $\times 4$  (на примере УД 200) составляют  $\frac{|50-25|}{\min(50, 25)} \times 100\% = 100\%$  и  $\frac{|50-12.5|}{\min(50, 12.5)} \times 100\% = 300\%$  соответственно. Следуя правилу «на порядок меньше», установленному в Разделе 3.2.1.4, коэффициенты изменения гранулярности будут считаться несущественными, если они меньше 10% и 30% соответственно.

Гранулярность генерализованных фрагментов «Бодайбо» и «Грозный», подвергнутых повышающей передискретизации, представлена в Таблице 3.5. Наихудшие коэффициенты изменения гранулярности при передискретизации в 2 и 4 раза составляют 7.85



**Рисунок 3.26** – Абсолютная ширина рельефа (взвешенная по высоте) для фрагментов Бодайбо (а) и Грозного (б). Распределения для эталонной, обобщенной, повторно дискретизированной с той же гранулярностью ( $G$ ) и разрешением ( $R$ ) ЦМР показаны разными цветами заливки

**Таблица 3.4** – Статистика Колмогорова-Смирнова для сравнения исходного и генерализованного распределения высот

*УД	Генерализация	Передискретизация по G	Передискретизация по R
Бодайбо			
200	0.0000	0.0430	0.3493
500	0.0000	0.1377	0.4650
1000	0.0141	0.1654	0.7090
Грозный			
200	0.1820	0.3015	0.5263
500	0.0278	0.1638	0.7730
1000	0.0565	0.1976	0.7455

\*УД – Уровень детализации

и 14.43 соответственно, что, согласно приведенным выше критериям, является незначительным.

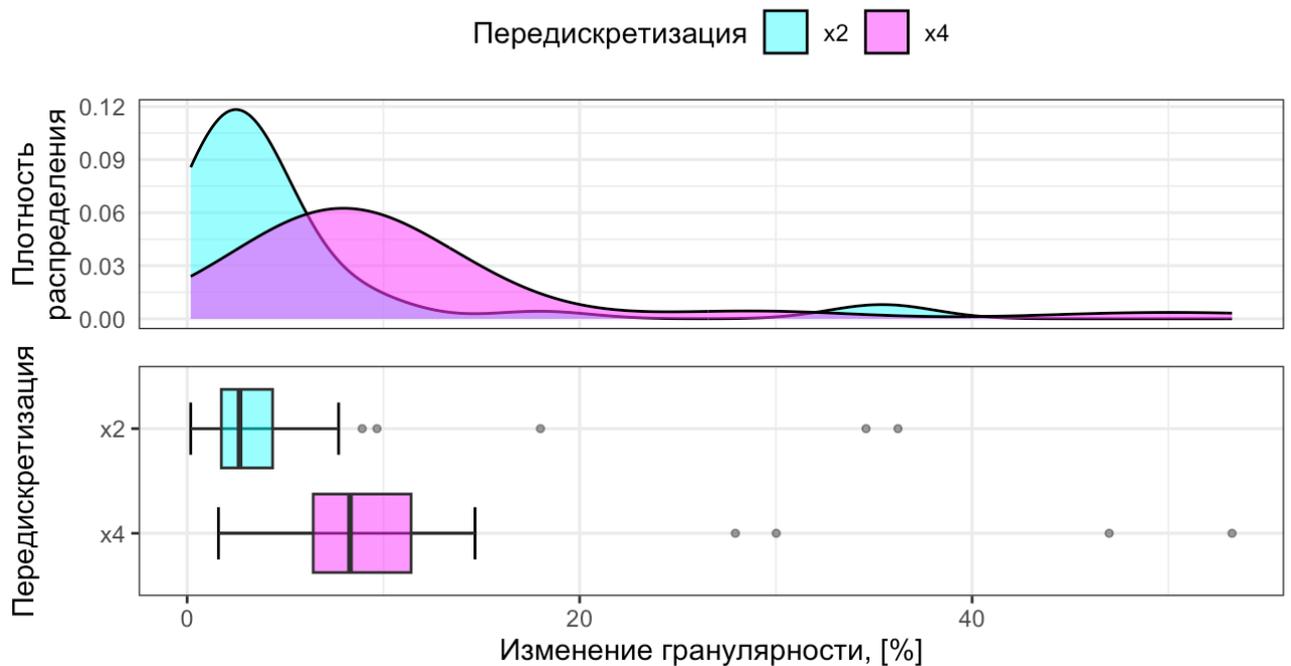
Распределения коэффициентов изменения гранулярности всех фрагментов ЦМР УД 1000 при передискретизации представлены на Рисунке 3.27. Медианные значения и интерквартильный размах для  $\times 2$  и  $\times 4$  составляют 2.7 (1.7–4.4) % и 8.3 (6.4–11.4) % соответственно. Это означает, что в целом коэффициент изменения гранулярности незначителен, как и в предыдущем эксперименте. Тем не менее, есть несколько выбросов. Так, для передискретизации  $\times 2$  такими фрагментами являются Братск (18.0%), Петропавловск-Камчатский (36.2%) и Владивосток (34.6%). Для передискретизации  $\times 4$  эти Братск (27.9%), Петропавловск-Камчатский (47.0%), Ухта (30.0%) и Владивосток (53.2%). Подобная сильная изменчивость характерна для ЦМР, содержащих существенную долю плоских территорий (равнины, плато, водоемы и т.д.). На этих территориях небольшие вариации высот, вызванные передискретизацией, могут иметь такой же порядок величины, что и естественные вариации высот, представленные в исходной ЦМР. Передискретизация может привести к появлению новых характерных точек, которые уменьшают среднюю ширину формы рельефа до 2 – 3 раз. Подобная чувствительность должна быть преодолена в будущих исследованиях с помощью более надежных методов выделения характерных точек поверхности.

### 3.2.4 Обсуждение

Предложенный подход к оценке гранулярности ЦМР обладает рядом важных свойств, которые необходимо учитывать для его практического применения. Некоторые из этих свойств вытекают из методологического замысла и реализации метода, другие вытекают из экспериментальных результатов, полученных в ходе настоящего исследования. Наиболее важные из них можно обобщить следующим образом.

Таблица 3.5 – Устойчивость показателя гранулярности к увеличению разрешения ЦМР

Передискр.	Разрешение, м	Бодайбо		Грозный	
		Гранулярность, м	Изменение	Гранулярность, м	Изменение
УД 200					
×4	12.5	1588	1.448	1390	3.309
×2	25	1597	0.877	1408	1.989
×1	50	1611	0.000	1436	0.000
УД 500					
×4	31.25	2147	5.962	2693	10.137
×2	62.5	2208	3.034	2750	7.855
×1	125	2275	0.000	2966	0.000
УД 1000					
×4	62.5	3611	4.680	3887	14.433
×2	125	3681	2.689	4144	7.336
×1	250	3780	0.000	4448	0.000



**Рисунок 3.27** – Изменение гранулярности при повышающей передискретизации в 2 и 4 раза для уровня детализации 1000. На верхнем графике представлена плотность распределения, на нижнем – диаграммы размаха.

- Гранулярность разработана и исследована как свойство ЦМР, важное для задач картографирования. Она помогает обеспечить требуемую детальность картографического изображения рельефа. Хотя гранулярность может быть потенциально ценной характеристикой в задачах анализа ЦМР, подобное применение выходит за рамки настоящей работы и требует отдельного исследования.
- Гранулярность не эквивалентна уровню детализации и должна использоваться в сочетании с другими свойствами для всесторонней характеристики ЦМР. Горизонтальное и вертикальное разрешение ЦМР, точность высот также являются существенными компонентами детальности ЦМР. Горизонтальное разрешение ЦМР играет особую роль при оценке пригодности ЦМР для мелкомасштабного картографирования, поскольку оно влияет на графическое качество изображения, возможность четко воспроизводить резкие перепады высот, характерные структурные линии. Влияние других перечисленных свойств на общую детальность ЦМР более важно для аналитических задач, а также картографирования в крупных и средних масштабах.
- Предлагаемый подход к расчету гранулярности ЦМР основан на оценке ширины формы рельефа, которая определяется расстоянием между характерными точками поверхности. В настоящее время для этой цели используются только дренажные ячейки. Однако методика не привязана к способу определения этих точек. В частности, если будет разработан объективный алгоритм определения границ форм рельефа, то полученные линии или полигоны могут быть использованы в качестве ограничивающих объектов при оценке их ширины.
- Дренажные ячейки ЦМР могут быть довольно спорадическими. В этом случае дополнительная фильтрация (например, удаление единичных ячеек-выбросов) может улучшить качество оценки гранулярности и снизить восприимчивость расчетов к шуму. В настоящем исследовании такая обработка не применялась, чтобы продемонстрировать эффективность разработанного метода в естественных условиях. Однако было выявлено, что для ЦМР, представляющих рельеф с небольшой вертикальной выраженностью, простой морфометрический подход к выделению дренажных ячеек ЦМР может быть чувствителен к передискретизации ЦМР. В дальнейших исследованиях необходимо повысить робастность этого этапа вычислений.
- Оценки типичной гранулярности ЦМР в данной работе основаны на эталонных ЦМР, реконструированных по топографическим картам, составленным вручную, с горизонтальным разрешением 0.25 мм в картографическом масштабе. Эксперименты показали, что ту же гранулярность можно получить и при гораздо более грубом разрешении ЦМР, однако оптимальный уровень детализации для качественного мелкомасштабного представления рельефа на картах требует высоко-

го разрешения ЦМР в сочетании с оптимальной гранулярностью ЦМР.

Соблюдение этих соображений должно способствовать более осмысленному использованию и совершенствованию разработанного метода.

### 3.2.5 Итоги

Гранулярность характеризует типичный размер форм рельефа, представленных в ЦМР. Разработка данного показателя обусловлена тем, что существующие меры качества ЦМР, такие как разрешение и точность, не способны охарактеризовать детальность ЦМР для целей картографирования. До настоящего момента было проблематично ответить, правильно ли выполнена генерализация ЦМР, чтобы соответствовать масштабу визуализации.

Оценка гранулярности осуществляется с помощью простой, но эффективной метрики, называемой шириной рельефа, которая основана на окружностях, вписанных в набор характерных точек поверхности. Тот факт, что ширина рельефа не требует определения точных границ рельефа (которые весьма спорны), делает ее простой в реализации и использовании. Также показано, что весь процесс оценки гранулярности может быть выполнен без каких-либо параметров, что является очень удобным свойством с практической точки зрения.

Экспериментальная работа, проведенная на многочисленных примерах эталонных ЦМР, восстановленных по генерализованным вручную топографическим картам, позволила вывести эмпирическое правило оптимальной относительной гранулярности для мелкомасштабного картографирования, которая составляет 5 — 6 мм в масштабе карты, однако может варьироваться в зависимости от расчлененности рельефа. Другим важным практическим результатом проведенного исследования является экспериментальное подтверждение того, что отбор форм рельефа при генерализации происходит по радикальному закону Топфера.

Два частных эксперимента на ЦМР, охватывающих однородные и неоднородные по генезису рельефа территории (Бодайбо и Грозный), показали, что гранулярность эффективно работает как ограничительная мера при обобщении ЦМР, в то время как одной только передискретизации для этого недостаточно. В обоих случаях совместное использование разрешения и гранулярности в качестве ограничительных метрик привело к одинаковому уровню детализации генерализованных вручную и автоматически ЦМР.

В будущих исследованиях желательно провести опрос экспертов для дополнительной оценки надежности гранулярности в качестве метрики, характеризующей уровень детализации ЦМР для целей картографирования. Аналитические приложения гранулярности в геоморфометрии также имеют высокий потенциал для будущих исследований.

### 3.3 Контроль детальности линейных пространственных объектов при обработке разными алгоритмами геометрического упрощения<sup>3</sup>

**Управление** детальностью может осуществляться на геометрическом уровне (при обобщении пространственных данных) и на уровне изображения (при визуализации пространственных данных). *Геометрический уровень* управления детальностью исследован на примерах генерализации цифровых моделей рельефа (см Разделе 3.2.3.2) и геометрического упрощения линий.

Ранее в настоящей работе было показано, что для линий разной формы следует использовать разные алгоритмы геометрического упрощения [Samsonov, Yakimova, 2017]. Как в таком случае обеспечить сопоставимое уменьшение детальности? Для решения этой проблемы был разработан новый метод, обеспечивающий сопоставимое уменьшение детальности линий при использовании разных алгоритмов геометрического упрощения [Samsonov, 2020a; Samsonov, 2021]. В основе метода лежит построение регрессионной зависимости между параметрами нескольких алгоритмов относительно заданной метрики. В качестве таких метрик предложено использовать изменение количества узлов линии, длины линии, количества изгибов линии, а также Модифицированное расстояние Хаусдорфа [Dubuisson, Jain, 1994] между результирующей и исходной линией.

#### 3.3.1 Методика

Целью излагаемой далее методики является управляемая генерализация линий, при которой появляется возможность изменять геометрические характеристики линий сходным образом при их обработке разными алгоритмами геометрического упрощения. Отслеживаемые изменения могут касаться как абсолютных значений метрик (например, количества точек), так и *коэффициентов сжатия*, которые отражают, во сколько раз абсолютная метрика изменяется после упрощения:

$$y = v_1/v_2 - 1, \quad (3.14)$$

где  $y$  — коэффициент сжатия,  $v_1$  — значение метрики до упрощения,  $v_2$  — значение метрики после упрощения. Поправка  $-1$  используется для того, чтобы коэффициент сжатия исходной линии был равен 0, что удобно для моделирования.

Разработанный подход основан на следующей идее: если установлена регрессионная модель зависимости между значением допуска алгоритма  $t$  и значением или коэф-

<sup>3</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov, Yakimova, 2020]

коэффициентом сжатия выбранной геометрической меры  $y = f(t)$ , то можно приравнять такие регрессии для двух разных алгоритмов  $A$  и  $B$  и получить *регрессионную модель между их параметрами*  $t_B = h(t_A)$ . Используя эту модель, можем упростить линию алгоритмом  $A$  с использованием произвольного значения допуска  $t_A$  и получить еще одно упрощение той же линии алгоритмом  $B$  с использованием автоматически вычисленного допуска  $t_B$ . Эти два упрощения будут аналогичны с точки зрения метрики, для которой были построены начальные регрессии. В терминах классификации [Mackaness, Ruas, 2007b] этот подход направлен на выполнение *оценки в целях настройки системы генерализации*.

Имея две линии, метрика которых изменена одинаково разными алгоритмами, получаем возможность сравнивать их по значениям других метрик и критериев. Например, чтобы визуально сопоставить характер упрощенных линий с целью выбора наилучшего алгоритма, можно потребовать, чтобы эти линии имели одинаковый коэффициент сжатия количества точек. Подобное требование имеет решающее значение, потому что если одна линия значительно более упрощена, чем другая, то и ее характер также искажается в большей степени. Следовательно, сравнение данных линий по характеру неправомерно.

Подход на основе моделирования имеет дополнительное преимущество: модели, построенные для разных метрик, могут быть комплексированы, чтобы получить единую модель, которая балансирует между ними. Далее будет показано, как этот подход может быть использован для получения модели, оценивающей изменения в уровне детализации линии. Поскольку эти изменения могут быть неодинаковы для линий разного характера, важно также исследовать, как это отражается на коэффициентах регрессионной зависимости.

### 3.3.1.1 Алгоритмы и общая гипотеза

Методика разрабатывалась на примере алгоритмов геометрического упрощения линий *Дугласа-Пейкера* [Douglas, Peucker, 1973], *Ли-Оупеншоу* [Li, Openshaw, 1992] и *Висвалингам-Уайатта* [Visvalingam, Whyatt, 1993]. Это хорошо известные процедуры геометрического упрощения, которые основаны на различных параметрах. Алгоритм *Дугласа-Пейкера (DP)* — итерационная процедура, которая работает на основе критерия кратчайшего расстояния (полосы пропускания). Это подход упрощения линии «сверху-вниз», который восстанавливает форму линии, начиная с глобально важных точек, и заканчивая точками, определяющими форму линии в локальной окрестности. Алгоритм *Висвалингам-Уайатта (VW)* также является итеративным, но он основан на критерии эффективной площади. Его принцип противоположный, «снизу-вверх», он не восстанавливает форму линии от грубой к детальной, а наоборот устраняет точки в порядке увеличения эффективной площади. Вследствие этого, область действия данной

процедуры локальна. Наконец, алгоритм *Ли-Оупеншоу* (*LO*) основан на концепции наименьшего видимого объекта (*SVO* — *smallest visible object*). Данный алгоритм не оценивает важность точек. Условно его можно классифицировать как последовательный, так как он обрабатывает каждый элемент *SVO* в линейной структуре и заменяет его, но Ли [Li, 2006] помещает эту процедуру в отдельный класс алгоритмов геометрического упрощения, который он характеризует как «алгоритмы масштабно-зависимой генерализации».

Имея в распоряжении три алгоритма, необходимо сначала выбрать основную процедуру упрощения, с параметром которой будут статистически связываться параметры оставшихся алгоритмов. Целесообразно выбрать алгоритм DP, поскольку он устанавливает стандарт де-факто в задачах компрессии пространственных данных и превосходит другие процедуры геометрического упрощения с точки зрения точности [White, 1985; Samsonov, Yakimova, 2017].

После выбора базовой процедуры возникает необходимость в гипотезе, которая бы помогла связать значения параметров алгоритмов математическим путем. Экспериментальная работа, представленная далее, показала, что метрики, описывающие изменения в количестве деталей (точек, изгибов и т. д.) линии, проявляют *степенную* зависимость от значения допуска алгоритма, независимо от алгоритма. Общая форма степенного закона такова:

$$y = \alpha x^\beta + \gamma. \quad (3.15)$$

Степенной закон является гибким, позволяет моделировать ускоряющийся, линейный и замедляющийся характер роста кривой зависимости (эти случаи соответствуют положительному, нулевому и отрицательному значению второй производной функции). Кроме того, можно показать, что если две функции  $y = f_A(d)$  и  $y = f_B(t)$  являются степенными, то значения их независимых аргументов также могут быть связаны степенной функцией  $t = f(d)$ .

Пусть  $d$  — значение допуска алгоритма DP. Тогда гипотеза о функциональной зависимости между  $d$  и допуском другого алгоритма  $t$  принимает следующий вид:

$$t = \alpha d^\beta \quad (3.16)$$

со свободным членом уравнения  $\gamma$  равным нулю, поскольку при нулевом значении параметра любого алгоритма геометрического упрощения генерализации быть не должно.

Параметры  $\alpha$  и  $\beta$  должны быть откалиброваны для каждой линии и каждой характеристики, которая измеряется после упрощения для оценки результатов генерализации. Калибровка начинается с некоторых априорных значений  $\alpha = \alpha_0$  и  $\beta = \beta_0$ , выбирае-

мых с учетом особенностей алгоритмов. В частности, можно предположить, что значение допуска  $VW$  имеет квадратичную зависимость от параметра  $DP$ , поскольку это соотношение площадной и линейной меры. Для данной пары алгоритмов было установлено  $\alpha_0 = 1$  и  $\beta_0 = 2$  в качестве стандартной параметризации, что соответствует экспериментальным данным в ранее опубликованных исследованиях (например, [Song, Miao, 2016]). Также предположим, что значение допуска алгоритма  $LO$ , которое устанавливает линейный размер минимально видимого объекта, имеет линейную зависимость от значения допуска  $DP$  и поэтому первоначально параметризуется с  $\alpha_0 = 2$  и  $\beta_0 = 1$ . Значение  $\alpha_0 = 2$  выбрано эмпирически на основе экспериментального анализа множества кривых.

Используя эти начальные параметризации, необходимо произвести серию геометрических упрощений каждым алгоритмом. Генерируется последовательность значений допуска  $DP$   $d = \{d_1, \dots, d_n\}$ , и ряд соответствующих допусков  $VW$  и  $LO$  выводится на основе  $d$ . Далее для каждого алгоритма строится эмпирическая кривая, показывающая зависимость между его параметром и величиной метрики, рассчитанной для генерализованной линии. Для этой зависимости методом нелинейной регрессии подбираются параметры степенной функции.

Предположим, что регрессия между значением допуска  $DP$  и выбранной геометрической метрикой равна  $y = kd^m$ , а аналогичная регрессия для другого алгоритма равна  $y = ax^b$ , где  $x = (t/\alpha_0)^{1/\beta_0}$ . Исходя из этих данных начальные параметры можно откалибровать по следующему уравнению:

$$kd^m = ax^b, \quad (3.17)$$

что преобразуется в:

$$\alpha = \alpha_0 \left[ \frac{k}{a} \right]^{\frac{\beta_0}{b}}, \quad \beta = \frac{m\beta_0}{b}. \quad (3.18)$$

Эти значения  $\alpha$  и  $\beta$ , параметризованные для конкретной метрики  $y$ , позволяют получить линии, имеющие статистически (с погрешностью аппроксимации) одинаковое значение  $y$ .

Предлагаемый подход имеет особенность: когда регрессионные модели получены для нескольких метрик, их можно объединить в одну модель, которая балансирует ними. Например, путем взвешенного осреднения параметров:

$$t = \left[ \sum_i w_i \alpha_i \right] d \left[ \sum_i w_i \beta_i \right], \quad (3.19)$$

где  $w_i$  — веса метрик, индексированных по  $i$ , и  $\sum_i w_i = 1$ . При выполнении серии упрощений, данный подход может быть использован для получения *последовательной*

серии уровней детализации. Билецки с соавторами [Biljecki et al., 2014] отмечает, что «любая комбинация значений метрик может быть уровнем детализации, и любая комбинация функций метрик также может представлять собой серию уровней детализации». Таким образом, последовательный ряд УД может быть построен в случае когда изменение метрики хорошо формализуется какой-либо монотонной функцией. Поскольку предложенный регрессионный подход также основан на получении монотонных функций по каждой метрике, появляется возможность получить последовательный ряд уровней детализации, используя значения допуска, вычисленные из комбинированной регрессионной модели.

### 3.3.1.2 Метрики

В качестве метрик были использованы две группы характеристик: **метрики точности** и **метрики геометрической структуры**. Обе группы исследованы с точки зрения того, насколько они показательны как характеристики изменения детальности линий при геометрическом упрощении.

В качестве метрики точности было использовано *Модифицированное расстояние Хаусдорфа, MHD* (см. Раздел 2.3.4). Отслеживалось изменение абсолютной величины данного показателя.

Для оценки изменения количества видимых деталей была применена иерархическая декомпозиция линии на следующие структурно-геометрические элементы:

1. Точки являются базовыми геометрическими элементами линии. Могут быть охарактеризованы расположением и плотностью.
2. Две последовательные точки образуют ребро, которое может быть охарактеризовано длиной и направлением (ориентацией).
3. Три последовательные точки образуют *триплеты*. В случае неколлинеарности точек триплеты формируют площадные объекты — треугольники. Их уже можно охарактеризовать показателями формы, в частности пропорциями и углом поворота в центральной точке.
4.  $N$  последовательных точек, имеющих постоянный знак угла поворота, образуют *изгиб* — наиболее заметную деталь в геометрии линии. Данное определение изгиба соответствует определению, использованному в работах [McMaster, 1986; Wang, Müller, 1998].
5. Собственно объект линии можно рассматривать как наивысший уровень рассмотрения.

Для того чтобы определить геометрические характеристики линий, которые можно использовать для характеристики изменения их детальности, было проведено исследование [Якимова et al., 2021], в котором произведен сравнительный анализ геометрии объектов с цифровых общегеографических карт предназначенных для масштабов 1 : 200 000,

1 : 500 000, 1 : 1 000 000, 1 : 2 500 000 и 1 : 8 000 000. Проведенные эксперименты показали что осредненные геометрические характеристики объектов при уменьшении масштаба исходного картографического материала могут меняться *монотонно* (возрастать или убывать), изменяться в тренде или оставаться стабильными. Для оценки детальности линий и управления ею подходят только характеристики первой группы (монотонные), к которым относятся:

- общее количество точек;
- количество изгибов;
- общая площадь изгибов;
- длина линии;
- показатели формы изгибов: длина, длина базовой линии, высота и площадь

Для полигональных объектов также было установлена монотонность таких показателей как общая площадь, общий периметр, средняя площадь и средний периметр.

Изменения геометрической структуры линии могут быть описаны в терминах как абсолютного количества (числа или длины) вышеперечисленных элементов, так и коэффициентов сжатия (отношения) этих величин. В настоящем исследовании использовались коэффициенты сжатия (Уравнение 3.14) применительно к количеству точек, количеству изгибов и длины линии как монотонным метрикам, надежно описывающим изменение детальности линий.

В дополнение к коэффициентам сжатия, была рассчитана *относительная извилистость*<sup>4</sup> для каждой стадии упрощения линии. Опираясь на то, что каждый изгиб может быть охарактеризован его длиной и длиной его базовой линии (линии, соединяющей начальную и конечную точки изгиба), относительная извилистость  $\tilde{K}$  может быть вычислена как среднее отношение между этими величинами:

$$\tilde{K} = \frac{1}{k} \sum_{j=1}^k \frac{l_j}{b_j}, \quad (3.20)$$

где  $k$  — количество изгибов,  $l_j$  — длина  $j$ -го изгиба,  $b_j$  — длина его базовой линии.

Изменение извилистости при геометрическом упрощении является показателем изменения формы линии.

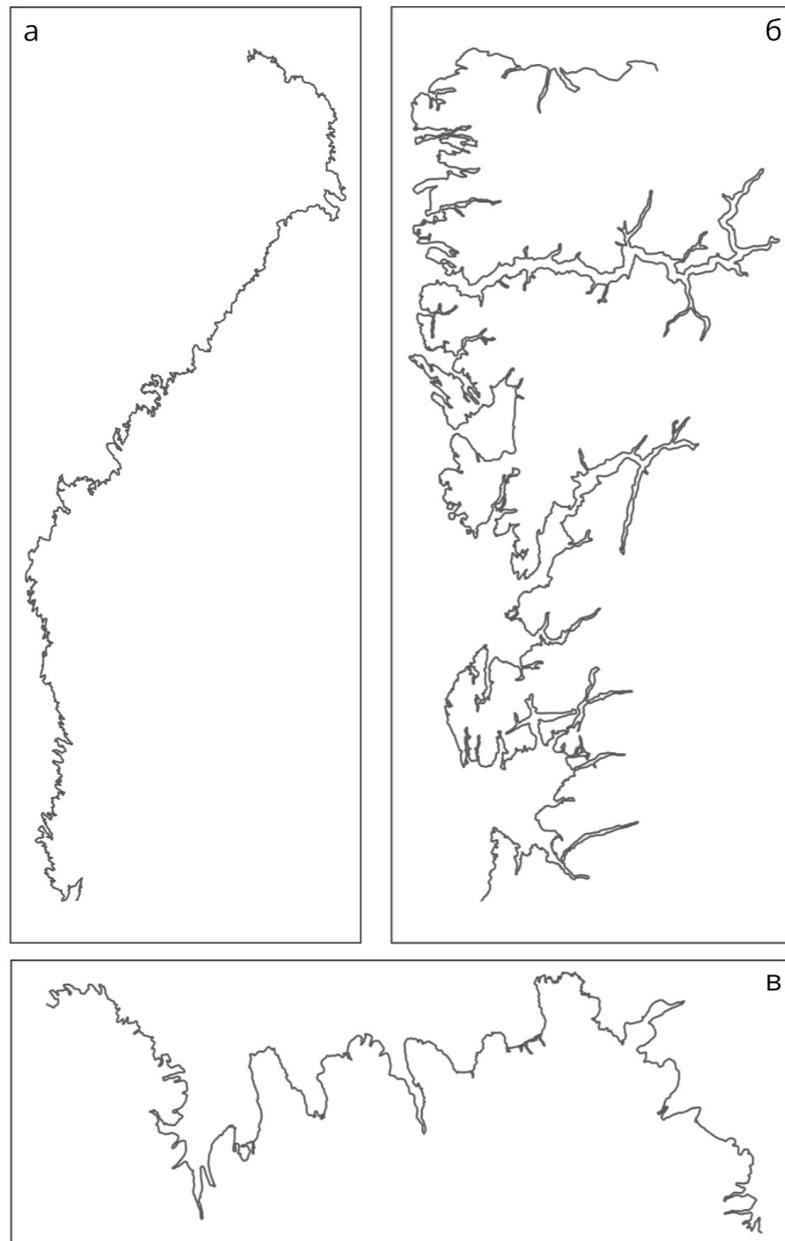
<sup>4</sup>Напомним, что общая извилистость или просто извилистость рассчитывается как отношение длины линии к длине отрезка, соединяющего ее начальную и конечную точки. Не определена для замкнутых линий.

### 3.3.2 Результаты и их анализ

#### 3.3.2.1 Исходные данные и постановка эксперимента

Для выполнения эксперимента были выбраны три морфологически разные береговые линии: западный берег Финляндии, северный берег Исландии и западный берег Норвегии (Рисунок 3.28). Линии были извлечены из набора данных *цифровой карты мира* DCW, соответствующего масштабу 1 : 1 000 000 [DCW, 1992].

Далее по тексту эти фрагменты береговых линий будут именоваться Финляндия, Норвегия и Исландия.



**Рисунок 3.28** – Исходные линии: (а) Финляндия; (б) Норвегия; (в) Исландия

Финляндия является примером линии с компактно уложенными изгибами небольшого размера, которые не выделяются в общей форме линии и имеют неглубокую иерар-

хию. Норвегия является примером линии с изгибами среднего размера, но плотно упакованными и образующими глубокую иерархию. Исландия является примером линии с большими и открытыми изгибами, которые явно выделяются в общей форме линии и имеют неглубокую иерархию.

Была выполнена серия упрощений каждой линии с использованием алгоритмов *DP*, *LO* и *VW*. Перед упрощением координаты линий были отмасштабированы таким образом, чтобы ограничивающий прямоугольник каждого линейного объекта имел длинную сторону, равную 1000, и был центрирован в точке  $(0, 0)$ . Это позволило нам привести линии к одному масштабу и сконцентрироваться на анализе влияния их морфологии. Значения допуска для *DP* были установлены в диапазоне от  $d = 0$  до  $d = 30$  с шагом приращения, равным 1. Соответствующие значения допуска для алгоритмов *LO* и *VW* были рассчитаны на основе значений  $d$  и с использованием стандартных параметризаций, описанных выше.

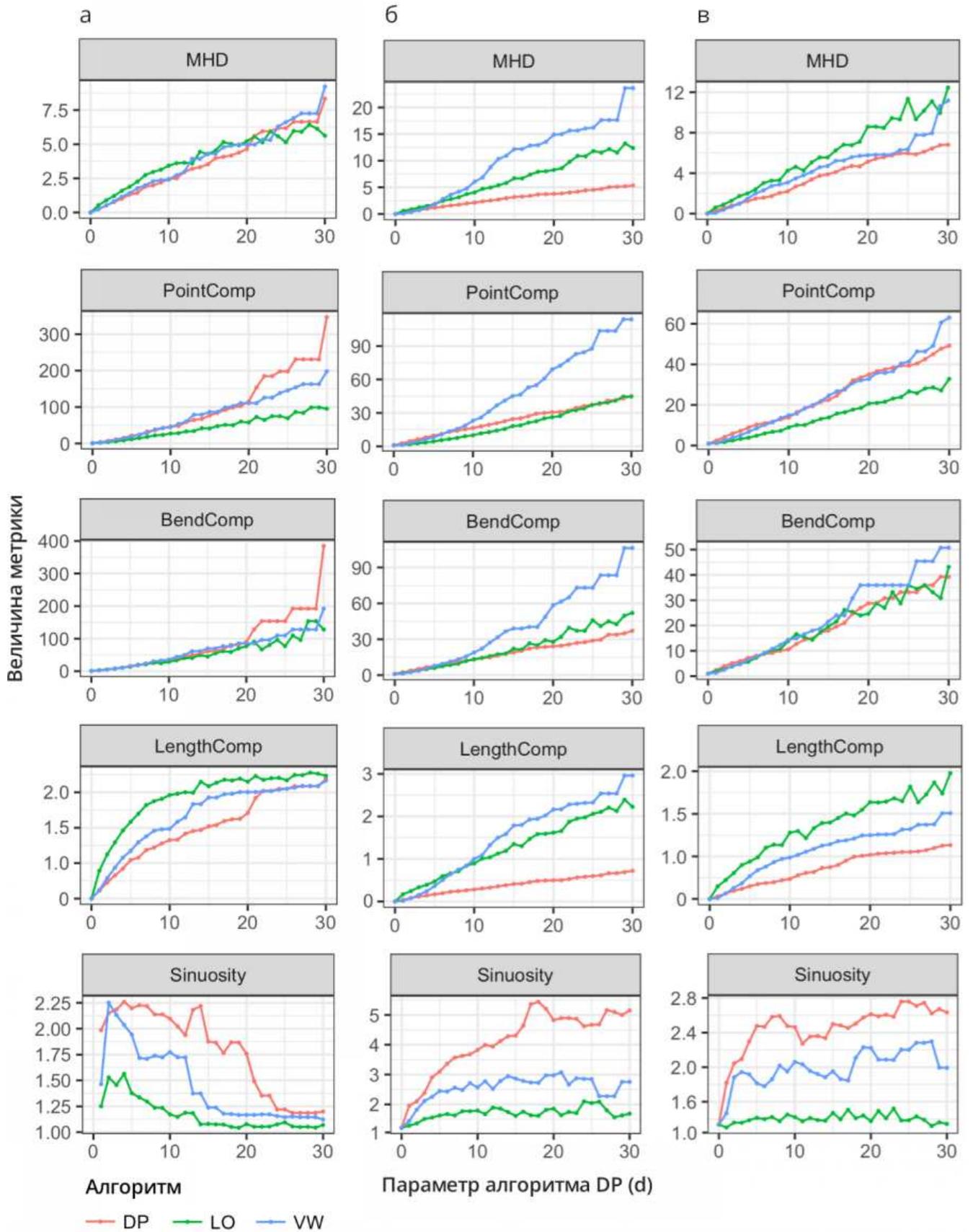
### 3.3.2.2 Анализ и интерпретация регрессионных моделей

Полученные значения вычисленных показателей представлены на Рисунке 3.29 в виде графиков функций от параметра  $d$ . Видно, что все показатели имеют тренд на увеличение в процессе генерализации (за исключением извилистости Финляндии, которая снижается), поэтому их значения можно использовать в регрессионном анализе. Поскольку некоторые характеристики демонстрируют нестабильное поведение для больших допусков, диапазон анализируемых значений допусков *DP* был ограничен до  $[0, 20]$ . Весь последующий анализ будет опираться только на зависимости, выявленные в этом диапазоне значений.

В Таблице 3.6 представлены значения параметров регрессии, полученные путем калибровки для каждой из выбранных линий с использованием степенной модели по методу наименьших квадратов (отметим, что  $k = a$  и  $m = b$  для алгоритма *DP*). Точность регрессии оценивается величиной  $\sigma$  — стандартного отклонения остатков регрессии.

Параметры степенного закона  $\alpha$  и  $\beta$  равны 1 для алгоритма *DP*, так как он используется в качестве эталонной процедуры. Учитывая это, таблицу следует читать следующим образом. Возьмем, к примеру, линию Финляндии, алгоритм *VW* и ограничение *MHD* (9-я строка в таблице). Степенной закон имеет вид  $t = 0.909d^{1.943}$ . Для заданного значения допуска *DP*  $d$  это уравнение обеспечит значение допуска *VW*  $t$ , такое, что исходная линия, обработанная двумя алгоритмами, будет иметь сходное значение метрики *MHD*. Точность этой модели определяется сочетанием  $\sigma = 0.080$  для параметризации *DP* и  $\sigma = 0.186$  для параметризации *VW* относительно ограничения *MHD* (строки 1 и 9 в Таблице 3.6 соответственно).

Приведенная таблица ясно показывает, что параметры регрессий зависят как от ограничивающей метрики, так и от характера обрабатываемой линии. Для содержатель-

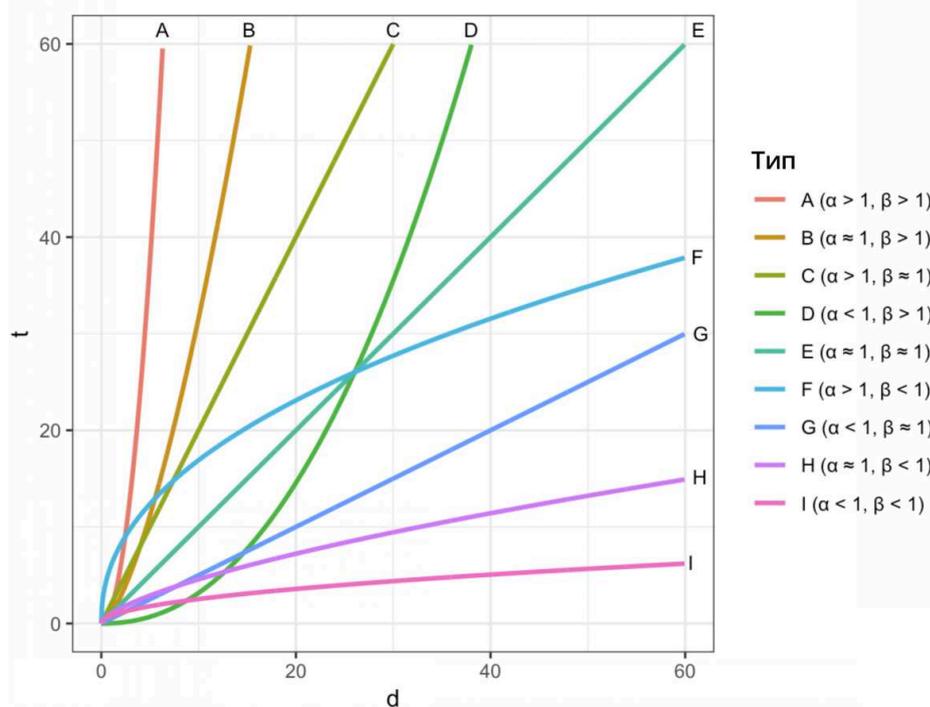


**Рисунок 3.29** – Расчетные метрики: (а) Финляндия; (б) Норвегия; (в) Исландия. Условные обозначения: *MHD* – модифицированное расстояние Хаусдорфа, *PointComp* – сжатие количества точек, *BendComp* – сжатие числа изгибов, *LengthComp* – сжатие длины

**Таблица 3.6** – Коэффициенты регрессии между параметрами алгоритмов Дугласа-Пейкера (DP), Ли-Оупеншоу (LO) и Висвалингам-Уайатта (VW) в соответствии с Уравнением 3.18

Алгоритм	Метрика	a	b	$\sigma$	$\alpha$	$\beta$
Финляндия						
DP	МНД	0.284	0.935	0.080	1.000	1.000
DP	Длина	0.188	0.624	0.036	1.000	1.000
DP	Точки	2.340	1.289	1.930	1.000	1.000
DP	Изгибы	1.382	1.383	2.223	1.000	1.000
LO	МНД	0.627	0.717	0.199	0.662	1.304
LO	Длина	0.569	0.384	0.085	0.111	1.625
LO	Точки	1.579	1.210	1.651	2.768	1.065
LO	Изгибы	1.582	1.274	2.511	1.798	1.085
VW	МНД	0.297	0.963	0.186	0.909	1.943
VW	Длина	0.238	0.639	0.062	0.473	1.952
VW	Точки	2.510	1.283	3.788	0.896	2.009
VW	Изгибы	2.120	1.253	3.201	0.505	2.208
Исландия						
DP	МНД	0.202	1.083	0.133	1.000	1.000
DP	Длина	0.026	0.996	0.016	1.000	1.000
DP	Точки	0.889	1.207	1.288	1.000	1.000
DP	Изгибы	0.616	1.250	1.190	1.000	1.000
LO	МНД	0.416	0.979	0.278	0.954	1.106
LO	Длина	0.160	0.644	0.035	0.120	1.547
LO	Точки	0.349	1.340	0.300	4.022	0.901
LO	Изгибы	0.890	1.124	1.426	1.440	1.113
VW	МНД	0.325	0.982	0.193	0.377	2.207
VW	Длина	0.078	0.774	0.036	0.059	2.574
VW	Точки	0.776	1.252	0.513	1.242	1.929
VW	Изгибы	0.413	1.470	1.553	1.722	1.701
Норвегия						
DP	МНД	0.303	0.856	0.054	1.000	1.000
DP	Длина	0.045	0.809	0.007	1.000	1.000
DP	Точки	1.885	0.925	0.388	1.000	1.000
DP	Изгибы	1.487	0.923	0.444	1.000	1.000
LO	МНД	0.376	1.044	0.159	1.625	0.820
LO	Длина	0.121	0.874	0.035	0.646	0.926
LO	Точки	0.310	1.470	0.273	6.826	0.629
LO	Изгибы	0.681	1.245	0.941	3.745	0.742
VW	МНД	0.305	1.315	0.767	0.989	1.302
VW	Длина	0.064	1.190	0.084	0.549	1.360
VW	Точки	0.585	1.580	1.243	4.398	1.171
VW	Изгибы	0.592	1.503	2.408	3.408	1.229

ного анализа отметим, что  $\beta = 1$  дает линейную функцию,  $\beta < 1$  дает замедляющийся рост, а  $\beta > 1$  дает ускоряющийся рост кривой. Значение  $\alpha$  линейно влияет на результат. Условная классификация кривых степенной функции, облегчающая интерпретацию таблицы, представлена на Рисунке 3.30. Будем использовать обозначенные в ней типы при последующем анализе.



**Рисунок 3.30** – Типы степенной функции при различных коэффициентах

Тип кривой степенной функции отражает специфический характер взаимодействия между значениями допусков двух алгоритмов для конкретной линии. В частности, если заданы два уравнения  $y_1 = ax^b$  и  $y_2 = kx^m$ , то соотношение  $y_1 = y_2$  будет выполняться только при условии:

$$x = \exp \left\{ \frac{\log(k) - \log(a)}{b - m} \right\}. \quad (3.21)$$

Перед этой критической точкой одна из функций всегда дает меньшие значения, чем другая. На Рисунке 3.30 такие точки представляют собой пересечения кривых. Например, комбинация  $\alpha > 1$  с  $\beta < 1$  (тип F) приводит к интенсивному росту  $t$  при малых значениях  $d$ , однако функция неизбежно замедляет свой рост, и её значения станут меньше линейной функции (тип E) при  $d > \alpha^{\frac{1}{1-\beta}}$ . На Рисунке 3.30 это происходит, когда  $d > 23$ . Такая ситуация отражает случай, когда требуется быстро увеличить значение допуска в течение первых итераций упрощения, но замедлить рост допуска при более интенсивных упрощениях. Аналогичная логика справедлива для противоположной ситуации, когда  $\alpha < 1$  и  $\beta > 1$  (тип D). Исходя из этих соображений, дадим содержательную интерпретацию параметров регрессий в Таблице 3.6.

Рост MHD-калиброванного допуска для  $LO$  медленно ускоряется для Финляндии (тип  $D$ ), почти линейен для Исландии (тип  $E$ ) и медленно затухает для Норвегии (тип  $F$ ). MHD-калиброванный допуск для  $VW$  демонстрирует почти квадратичное поведение для Финляндии (тип  $B$ ), более ускоренное возрастание для Исландии (тип  $D$ , но для текущего диапазона значений они меньше, чем для Финляндии) и более медленное возрастание для Норвегии (тип  $D$ ). Наиболее важным моментом здесь является то, что относительно метрики MHD в случае Норвегии параметры обеих процедур  $LO$  и  $VW$  ведут себя куда более затухающим образом по отношению к параметру процедуры  $DP$ , нежели в случае с Финляндией и Исландией. Отметим, что изгибы имеют наибольшую извилистость именно для Норвегии, особенно для больших значений допуска, принадлежащих интервалу  $(10, 20)$ . Большая извилистость означает, что удаление изгиба с помощью любой метрики приведет к большому горизонтальному сдвигу, и следовательно большому значению MHD. Следовательно, менять допуск нужно более деликатно.

Калиброванные по сжатию длины (*LengthComp*) допуски ведут себя похоже на калиброванные по MHD, но параметры приводят к менее интенсивному росту функции для малых значений ( $\alpha$  меньше) и более интенсивному росту для больших значений ( $\beta$  больше). Сходство значений регрессии для параметров MHD и *LengthComp* также может быть обосновано их высокой корреляцией, которая колеблется от  $-0.81$  ( $LO$ -Финляндия) до  $-0.99$  ( $DP$ -Норвегия) для всех линий и алгоритмов при уровне значимости менее  $0.05$  (альтернативная гипотеза заключается в отсутствии корреляции). Таким образом, *сжатие длины линии может быть использовано в качестве косвенной метрики точности упрощения*.

Калиброванный по сжатию количества вершин (*PointComp*) допуск для алгоритма  $LO$  является линейным для Финляндии (тип  $C$ ) и замедляющимся для Исландии и Норвегии (тип  $F$ ). Тот же допуск для алгоритма  $VW$  почти квадратичен в случае Финляндии и Исландии (тип  $B$ ) и близок к линейному для Норвегии (тип  $C$ ). Калиброванные по сжатию количества изгибов (*BendComp*) допуски ведут себя похожим образом, но их параметры в среднем приводят к менее интенсивному росту функции в диапазоне малых значений ( $\alpha$  меньше) и более интенсивному росту в диапазоне больших значений ( $\beta$  больше). Кроме того, калиброванный по сжатию количества изгибов допуск для Исландии характеризуется медленным ускорением вместо медленного затухания роста. Сходство коэффициентов регрессии для этих метрик также может быть обосновано их высокой корреляцией, которая колеблется от  $0.97$  ( $LO$ -Исландия) до  $0.99$  ( $DP$ -Исландия) для всех линий и алгоритмов при уровне значимости менее  $0.05$ . Таким образом, обе метрики аналогично реагируют на упрощение.

Допуски, калиброванные по вершинам и изгибам, обычно имеют большие значения, чем допуски, калиброванные по MHD и длине, поскольку их  $\alpha$  больше, а  $\beta$  близки по значению или меньше. Это означает, что метрика, используемая в алгоритме  $DP$ , приводит к очень быстрому сжатию количества деталей линии. Кроме того, полученные па-

раметризации показывают, что для более сложных линий с глубокой иерархией изгибов (Норвегия) регрессионные кривые как  $LO$ , так и  $VW$  имеют приоритет малых значений допуска, где изменения в деталях происходят быстро. Чем глубже иерархия, тем больше изгибов удаляется при первых итерациях упрощения. Алгоритмы  $LO$  и  $VW$  нуждаются в более интенсивном росте параметра, чтобы конкурировать с процедурой  $DP$  в диапазоне малых значений допуска.

Сравнивая алгоритмы  $LO$  и  $VW$ , можно сказать, что их реакция на характер линии, как правило, очень похожа. Оба алгоритма требуют быстрого роста значения допуска в малом диапазоне значений и замедления роста допуска в диапазоне больших значений (кривая медленно затухает или близка к линейной функции). В то время как в случае простой линии соотношение обратное: значение допуска алгоритмов  $LO$  и  $VW$  должно медленно расти относительно значения допуска  $DP$  в диапазоне малых допусков, и ускоряться в диапазоне больших значений допуска (кривая близка к квадратичной).

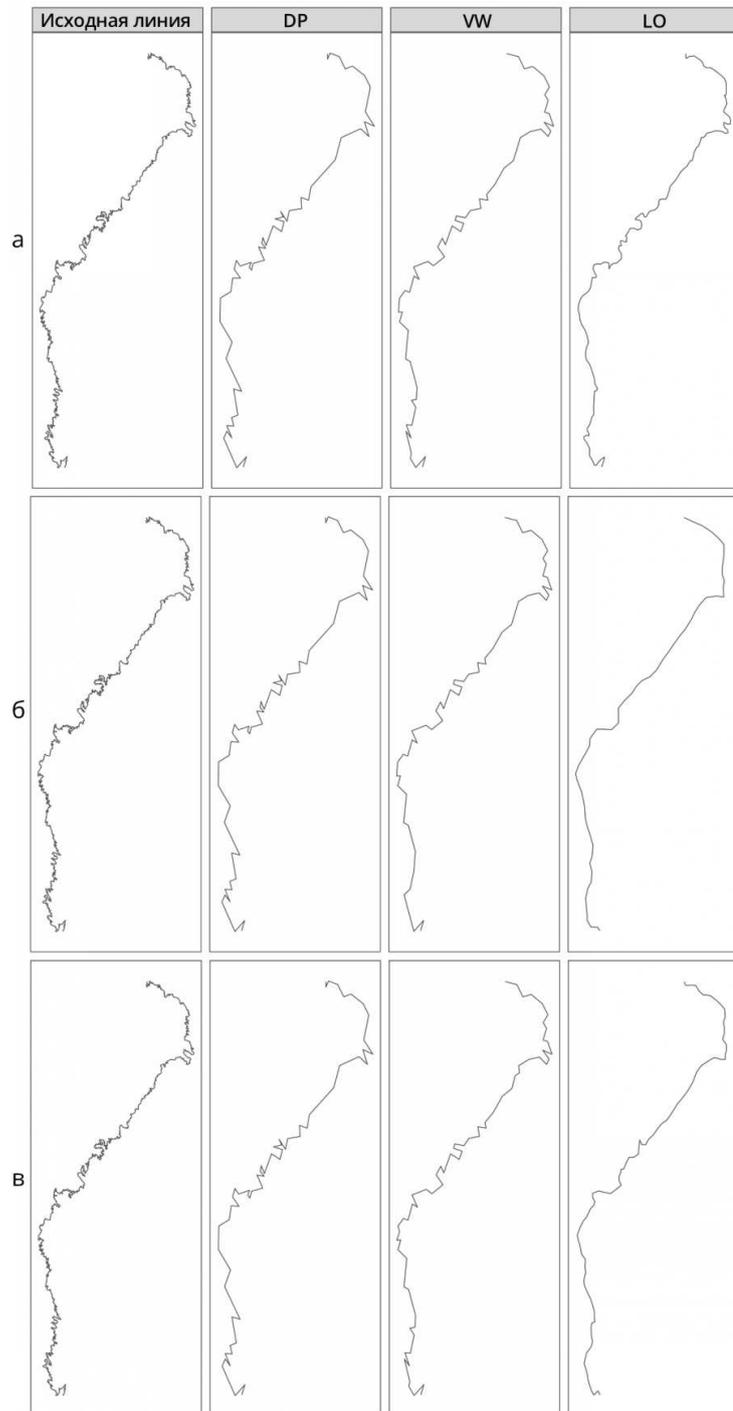
На основе проведенных рассуждений можно сделать вывод, что предлагаемая гипотеза степенного закона наглядно характеризует соотношения параметров нескольких алгоритмов в случае, когда изменения геометрической точности и количества деталей должны быть сопоставимы, а характер линий отличается.

### 3.3.2.3 Анализ и моделирование изменений уровня детализации

Проанализировав поведение уравнений регрессии для линий разной формы, перейдем к визуальному анализу результатов с целью выяснить, можно ли добиться с их помощью сопоставимого уменьшения детальности линий, используя предложенные метрики и их комбинации. На Рисунке 3.31, Рисунке 3.32 и Рисунке 3.33 показаны результаты генерализации, полученные относительно метрик  $MHD$ , сжатия количества вершин и изгибов. Значение допуска  $DP$  было установлено на 10, 15 и 20 для Финляндии, Исландии и Норвегии соответственно. Значения допуска  $VW$  и  $LO$  рассчитывались с использованием полученных параметризаций.

Анализируя данные изображения, можно сделать следующие выводы:

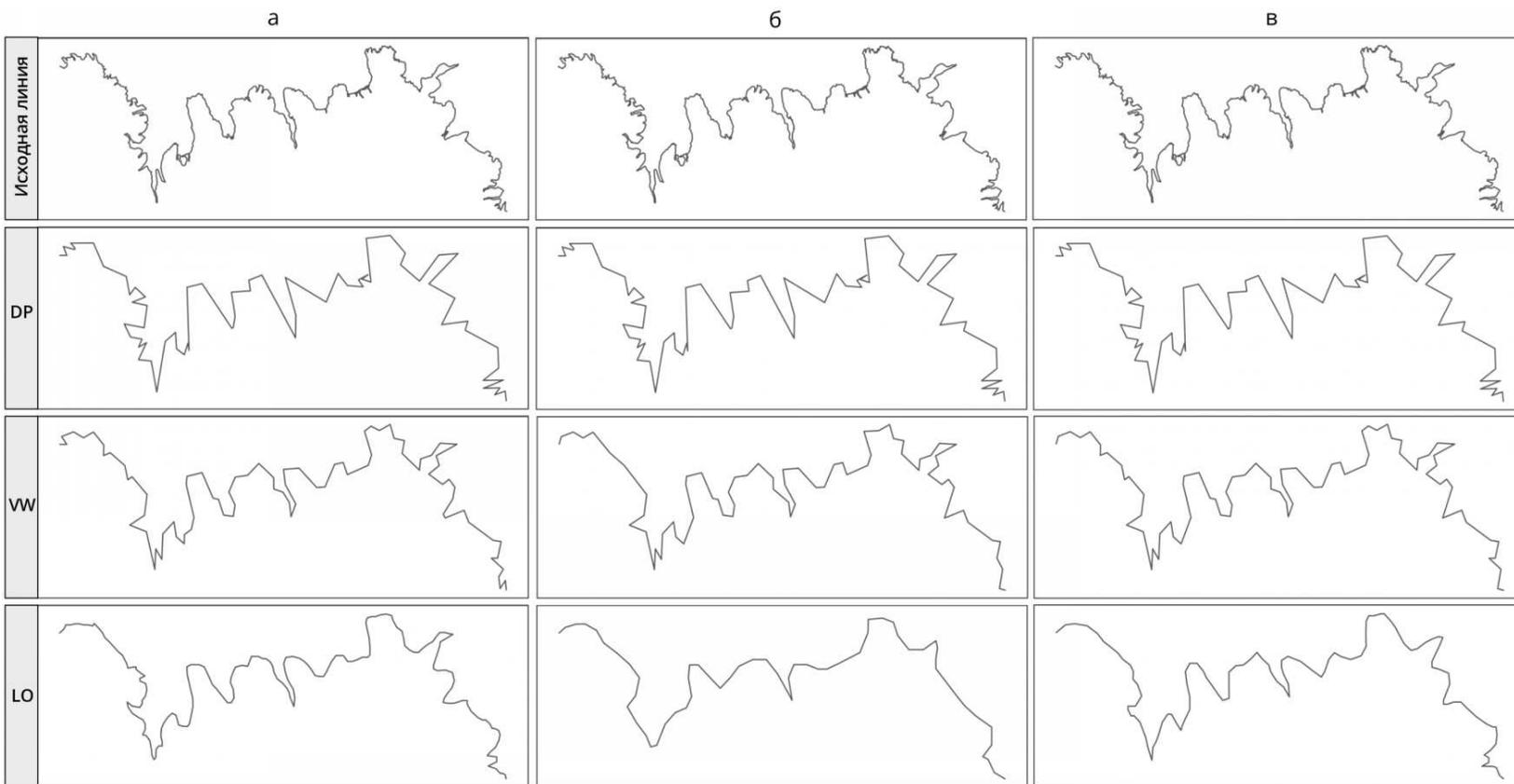
- $MHD$  в сравнении с другими метриками дает визуально тот же уровень детализации для Финляндии, немного более подробный для Исландии и значительно более подробный для Норвегии, особенно для алгоритма  $VW$ .
- Сжатие количества вершин дает визуально тот же уровень детализации по алгоритмам  $DP$  и  $VW$  и значительно менее детализированную форму по алгоритму  $LO$  для каждой из трех линий. Это означает, что алгоритм  $LO$  нуждается в большем количестве точек для достижения того же уровня детализации, независимо от характера линии.



**Рисунок 3.31** – Геометрическое упрощение береговой линии Финляндии при одинаковой величине: (а) МНД, (б) сжатия количества вершин и (в) сжатия количества изгибов. DP – результат работы алгоритма Дугласа-Пейкера; VW – результат работы алгоритма Висвалингам-Уайатта; LO – результат работы алгоритма Ли-Оупеншоу



**Рисунок 3.32** – Геометрическое упрощение береговой линии Норвегии при одинаковой величине: (а) МНД, (б) сжатия количества вершин и (в) сжатия количества изгибов. DP – результат работы алгоритма Дугласа-Пейкера; VW – результат работы алгоритма Висвалингам-Уайатта; LO – результат работы алгоритма Ли-Оупеншоу



**Рисунок 3.33** – Геометрическое упрощение линии Исландии при одинаковой величине (а) MHD; (б) сжатия количества вершин и (в) сжатия количества изгибов. DP – результат работы алгоритма Дугласа-Пейкера; VW – результат работы алгоритма Висвалингам-Уайатта; LO – результат работы алгоритма Ли-Оупеншоу

- Сжатие количества изгибов дает одно то же или немного более подробное изображение по алгоритму VW для каждой линии. Оно также дает менее детализированное, чуть менее детализированное и чуть более детализированное изображение при использовании алгоритма LO для Финляндии, Исландии и Норвегии соответственно.

Обобщив эту информацию, можно высказать предположение, что изменения в MHD должны работать лучше как метрика изменения детальности для Финляндии, сжатие количества изгибов является лучшей метрикой изменения детальности для Норвегии, а Исландия занимает промежуточное положение. Учесть все три варианта можно путем построения комплексной параметризации, моделирующей уменьшение детальности на основе взвешенного осреднения параметров MHD и сжатия количества изгибов в соответствии с формулой:

$$t_{12} = (w_1\alpha_1 + w_2\alpha_2)d^{w_1\beta_1+w_2\beta_2}, \quad (3.22)$$

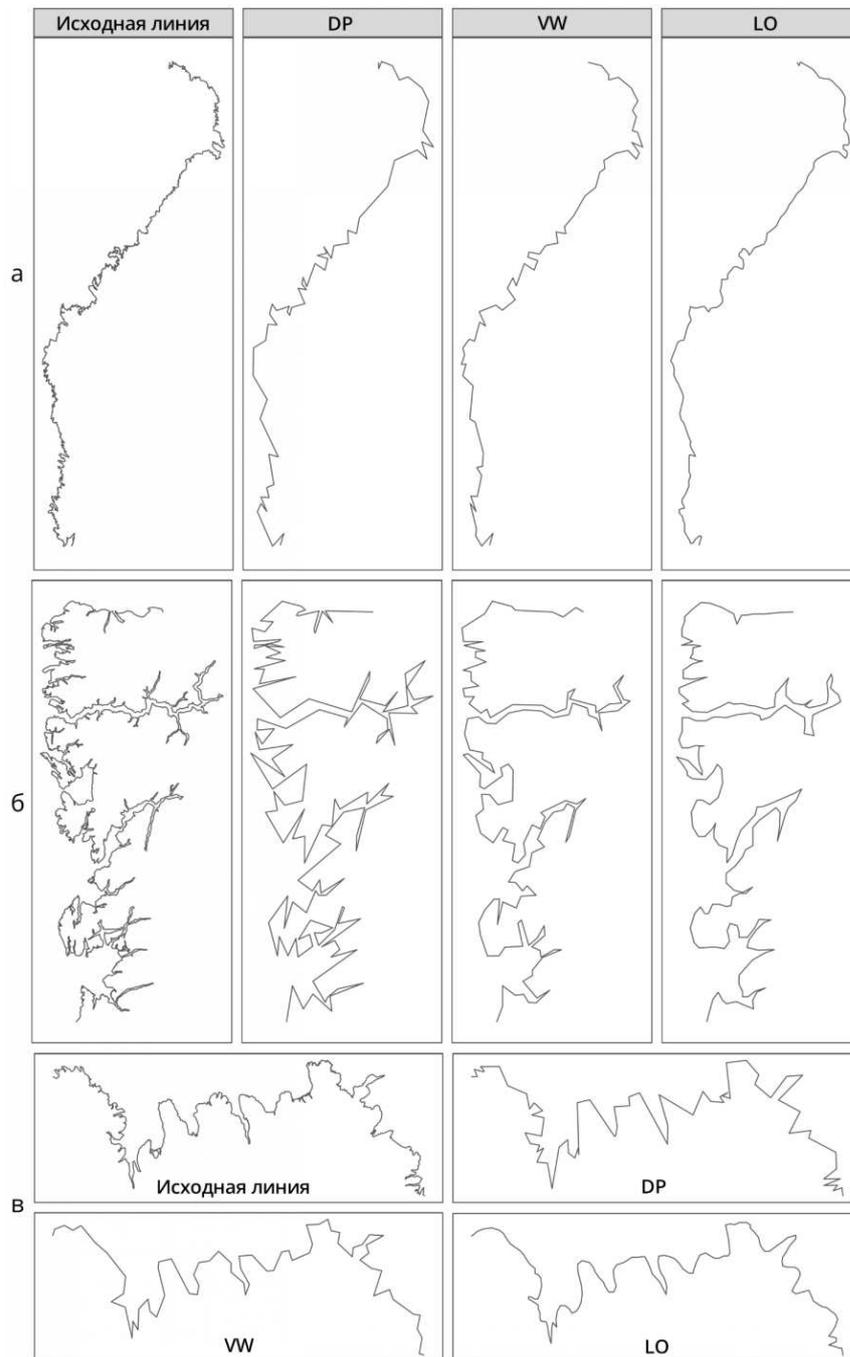
где индексы 1 и 2 соответствуют значениям параметров регрессий по MHD и сжатию количества изгибов соответственно и  $w_1 \in [0, 1]$ ,  $w_2 \in [0, 1]$  — их веса, зависящие от характера линии и удовлетворяющие условию  $w_1 + w_2 = 1$ .

На Рисунке 3.34 показаны результаты упрощения, полученные с помощью пар весовых коэффициентов  $(w_1, w_2)$ , равных  $(1, 0)$ ,  $(0.7, 0.3)$  и  $(0, 1)$  для Финляндии, Исландии и Норвегии соответственно. Визуальный анализ показывает, что эта комбинированная параметризация дает более последовательное изменение детальности линий. Важно отметить, что этот вид параметризации направлен на моделирование *снижения уровня детализации*, а не на оценку его абсолютного значения. Отметим также, что исследование [Самсонов, Якимова, 2021], в котором ряд респондентов попросили оценить соотношение детальности линий в целом подтвердило выводы настоящей работы и применимость выбранных метрик и их взвешенной комбинации.

### 3.3.3 Обсуждение

В настоящем разделе исследованы функциональные зависимости значений допусков нескольких алгоритмов геометрического упрощения, которые могут быть использованы для получения линий сходной детальности.

Пусть дано  $N$  линий,  $M$  метрик, которые характеризуют изменения в геометрии линии после упрощения, и  $K$  алгоритмов, которые могут быть использованы для упрощения. Используя предложенную методику, можем получить  $L = N \times M \times K$  регрессионных моделей, соответствующих всем возможным сочетаниям алгоритмов, метрик и линий. Если степенной закон используется в качестве модели для каждой метрики, то получаем  $2M \times K$  коэффициентов регрессии, которые могут быть связаны с каждой



**Рисунок 3.34** – Упрощение всех линий для достижения гомогенного уровня детализации при комплексировании между метриками МНД и сжатия количества изгибов для линий (а) Финландии; (б) Норвегии и (в) Исландии. DP – результат работы алгоритма Дугласа-Пейкера; VW – результат работы алгоритма Висвалингам-Уайатта; LO – результат работы алгоритма Ли-Оупеншоу

линией и сохранены в виде таблицы.

Эти коэффициенты могут быть использованы для:

- Генерализации линий разными алгоритмами при достижении схожих значений целевых метрик. Результаты такого упрощения могут быть использованы для сравнения линий с точки зрения других характеристик, которые трудно формализуемы. Например, чтобы выбрать алгоритм, который сохраняет характер линии.
- Генерализации множества линейных объектов (слоя), в которой изменения выбранной метрики (в том числе детализации) схожи для всех линий, независимо от алгоритма, выбранного для каждой линии. Это случай совместной (коллаборативной) генерализации.
- Расширения возможностей моделей генерализации (например, веб-фреймворка, предложенного в работе [Foerster, Stoter, Lemmens, 2007]) возможностью синхронизации степени обобщения разных линий при их обработке различными алгоритмами.

Заметим, что нет необходимости хранить калиброванные коэффициенты  $\alpha$  и  $\beta$ , которые связывают значения допусков алгоритмов, так как они могут быть легко получены на лету (Уравнение 3.18).

### 3.3.4 Итоги

В настоящем разделе диссертации предложен подход, позволяющий получать упрощенные линии, имеющие близкие значения метрики изменения детальности при обработке различными алгоритмами. Основные результаты представленного блока исследований можно обобщить следующим образом:

- Предложен формальный способ параметризации алгоритмов геометрического упрощения на основе регрессионного моделирования между метриками изменения детальности и значением допуска каждого алгоритма.
- Показано, что степенная функция  $t = \alpha t^\beta$  является эффективной моделью описания регрессионной зависимости между значениями параметров алгоритмов Дугласа-Пейкера, Ли-Опеншоу и Висвалингам-Уайатта друг с другом и по отношению к вычисляемым метрикам.
- Разработана процедура калибровки параметров регрессии между значениями допусков различных алгоритмов, позволяющая получить линии, имеющие схожее значение метрики изменения детальности. Коэффициенты регрессий отражают различия в характере линий и специфичны для комбинации алгоритмов.
- Продемонстрировано, что несколько метрик могут быть объединены методом взвешенного осреднения коэффициентов регрессии, чтобы получить модель, ба-

лансирующую между ними и позволяющую получить аналогичное снижение детальности различными алгоритмами.

Будущие исследования в этом направлении могут включать разработку новых регрессий, учитывающих морфологию элементов деталей линий, таких как изгибы. Это поможет углубить понимание того, как меняется детальность представления географических линий в базах пространственных данных. Задача выбора весовых коэффициентов для комбинирования регрессионных моделей по различным метрикам требует отдельного исследования. Несомненно, одним из возможных направлений будущих разработок является расширение перечня анализируемых алгоритмов геометрического упрощения. Наконец, следует рассмотреть возможности применения предложенного подхода к параллельному (коллаборативному) обобщению элементов множества линий различными алгоритмами.

### 3.4 Повышение детальности изображения при визуализации цифровых моделей рельефа с помощью дополнительных горизонталей<sup>5</sup>

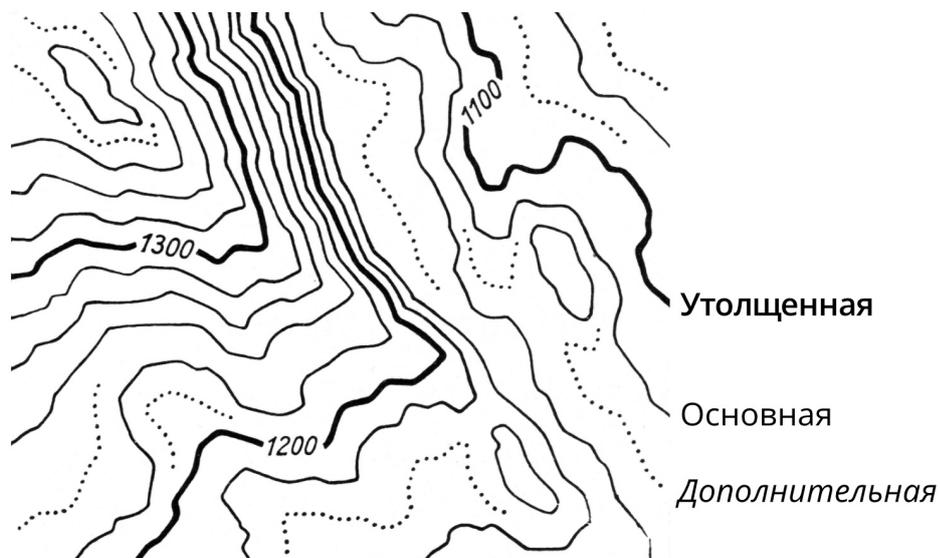
**Управление** детальностью на уровне изображения заключается в изменении параметров визуализации пространственных данных, которое приводит к изменению детальности картографического изображения. Наиболее сложно этот процесс устроен для цифровых моделей рельефа и других геополей, при визуализации которых в ряде сценариев происходит не просто отображение ячеек, а преобразование модели данных. В частности, горизонталы (изолинии) и штрихи представляют собой векторные объекты, извлеченные из цифровой модели рельефа с определенными параметрами, которые определяют плотность их размещения. В зависимости от параметров их извлечения можно получать изображение рельефа разной детальности.

При использовании горизонталей одним из методов повышения детальности изображения является построение *дополнительных горизонталей*, которые проводятся на половине высоты сечения в местах, где заложение велико и оказываются скрыты малые формы рельефа [Imhof, 1982]. Несмотря на распространенность этого приема, до момента выполнения данного исследования не было разработано метода его автоматизации. В рамках настоящего исследования был впервые разработан метод автоматизированного построения дополнительных горизонталей, в основе которого лежит построение двух растров: заложений и центральности [Samsonov et al., 2019].

<sup>5</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov et al., 2019]

### 3.4.1 Терминология

*Дополнительные горизонталы* — термин, общепринятый в отечественной литературе для обозначения горизонталей, выборочно проводимых на половине высоты сечения [Заруцкая, Сваткова, 1982; Верещака, 2002]. Аналогичным термином «*supplementary contour*» такие горизонталы называются и в ряде зарубежных трудов [Keates, 1989; Robinson *et al.*, 1995; USGS, 2005; Kimerling, 2016; Kenneth Field, 2018]). Иногда можно встретить название *промежуточные горизонталы* — в частности, у Эдуарда Имгофа в его основополагающей работе по картографическому изображению рельефа [Imhof, 1982]. Здесь есть некоторая путаница, т.к. другие исследователи, такие как Mackaness, Steven [2006], Kimerling [2016], а также USGS [2005], используют термин «*промежуточные горизонталы*» для обозначения обычных горизонталей, чтобы отличить их от *утолщенных* (которые обычно подписываются и рисуются более толстой линией). В данной части работы используются термины *утолщенная (index)*, *основная (regular)* и *дополнительная (supplementary) горизонталь*, как показано на Рисунке 3.35. Также обратим внимание читателя на то, что в данном разделе диссертации повсеместно используется швейцарский стиль отображения дополнительных горизонталей — пунктиром.



**Рисунок 3.35** — Терминология, используемая в отношении горизонталей. Изображение по Imhof [1982] выполнено в швейцарском стиле (дополнительная горизонталь показана пунктиром, а не штриховой линией)

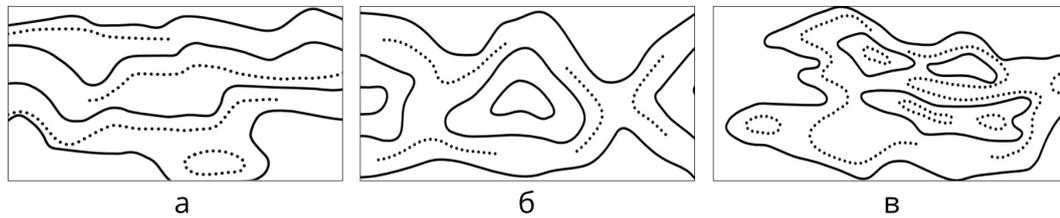
### 3.4.2 Методика

#### 3.4.2.1 Принципы размещения

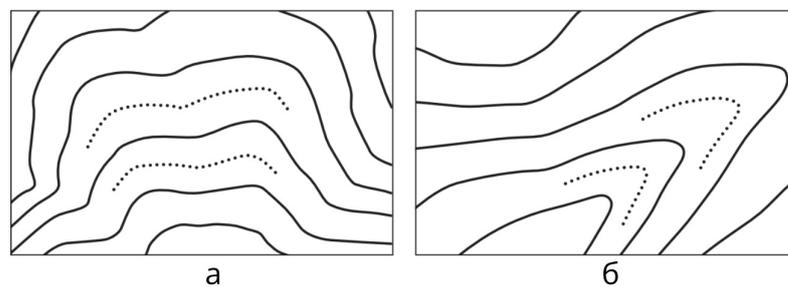
Анализ литературы показывает, что при проведении дополнительных горизонталей картографы руководствуются следующими принципами.

1. *Интервал и базовый уровень.* Как показал Имгоф, размещение вспомогательных горизонталей, т.е. имеющих произвольную высоту, не кратную половине основного сечения, создает запутанные карты, которые невозможно правильно интерпретировать [Imhof, 1982]. По этой же причине Имгоф также выступает против размещения нескольких вспомогательных горизонталей между соседними основными. Тем не менее, существуют редкие примеры карт со вспомогательными горизонталями. В частности, их применение допускается как отечественными [Руководство, 1978], так и, например, американскими [Hypsography, 2001] руководствами и наставлениями по составлению топографических карт.
2. *Центральность.* Линия дополнительной горизонтали проводится, если ее расстояние до следующей горизонтали с одной стороны значительно отличается от расстояния до горизонтали с другой стороны. Это особенно важно для террас (слева на Рисунке 3.36), обрывов плосковершинных хребтов и плато. Также если следующая выше или ниже по склону горизонтали находятся примерно на одинаковом расстоянии, дополнительная горизонталь не проводится, поскольку она не несет полезной для чтения рельефа информации (Рисунок 3.37, [Imhof, 1982; Hypsography, 2001]).
3. *Плотность.* Линии дополнительных горизонталей должны находиться не ближе определенного расстояния от горизонталей основного сечения. По этой причине их не наносят на крутых склонах, где они будут располагаться очень близко или даже сливаться с основными горизонталями и тем самым затруднять интерпретацию [Imhof, 1982].
4. *Плоские участки.* Дополнительные горизонтали рисуются вне зависимости от их центральности на плоских участках, где заложение между основными горизонталями велико. Имгоф рекомендует порог расстояния в 10-20 миллиметров в масштабе карты [Imhof, 1982], в то время как стандарт [Hypsography, 2001] использует порог площади  $33 \times 33$  миллиметра (или  $0.5 \times 0.5$  мили в масштабе 1 : 24 000).
5. *Малые формы.* Холмы, впадины, седловины, дюны, оползневые тела, боковые морены и другие малые геоморфологические формы часто требуют дополнительных горизонталей Рисунок 3.36. В руководстве по картографическим работам СССР [Руководство, 1978] для седловин приводится конкретный пример: дополнительные горизонтали должны быть нанесены по обе стороны седловины.
6. *Протяженность.* Как правило, линии дополнительных горизонталей продлеваются до участков, где они занимают примерно центральное положение между соседними горизонталями основного сечения. Мнения о необходимой центральности при этом разнятся. Американский стандарт USGS [Hypsography, 2001] требует величины 0.6, а отечественный [Руководство, 1978] — 1.0, т.е. ровно посередине. Швейцарское же картографическое агентство *Swisstopo* предписывает не только

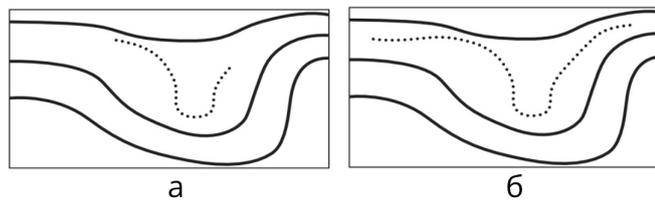
доводить дополнительные горизонтали до среднего положения, но и немного их продлевать дальше. Это повышает читаемость и позволяет избежать излишне коротких линий (Рисунок 3.38, *Geländedarstellung* [2008]).



**Рисунок 3.36** – Дополнительные горизонтали уточняют изображение (а) рельефа террас, (б) перевалов и (в) небольших холмов [*Cartographic generalization*, 1987]



**Рисунок 3.37** – Избыточные дополнительные горизонтали, которых следует избегать: (а) на склоне и (б) в районе тальвега. Они не несут полезной информации, поскольку их расстояния до соседних горизонталей выше и ниже по склону одинаковы [*Imhof*, 1982]



**Рисунок 3.38** – Продление дополнительных горизонталей до центрального положения между соседними горизонталями в соответствии с рекомендациями *Geländedarstellung* [2008]: (а) исходная линия; (б) результат продления

#### 3.4.2.2 Общая последовательность действий

Согласно предложенной методике, выполняется следующая последовательность действий:

1. Вычисляются горизонтали-кандидаты на половине высоты основного сечения.
2. Обрабатываются кандидаты, оконтуривающие вершины и понижения. Они либо рисуются, либо удаляются целиком, исходя из их длины и средней ширины замкнутой ими области.

3. Остальные горизонтали-кандидаты рассматриваются на всем протяжении, чтобы определить их участки, которые должны быть показаны на карте. Для этого анализируется релевантность каждой вершины горизонтали. Вершина считается релевантной, если она удовлетворяет ограничениям (а) заложения и (б) центральности.

- *Ограничение заложения* гарантирует наличие достаточного пространства между основными горизонталями для размещения дополнительных.
- *Ограничение центральности* гарантирует, что дополнительная горизонталь не будет проходить по геометрическому центру между соседними горизонталями, как показано на Рисунке 3.37.

Вершины, удовлетворяющие обоим ограничениям, соединяются отрезками.

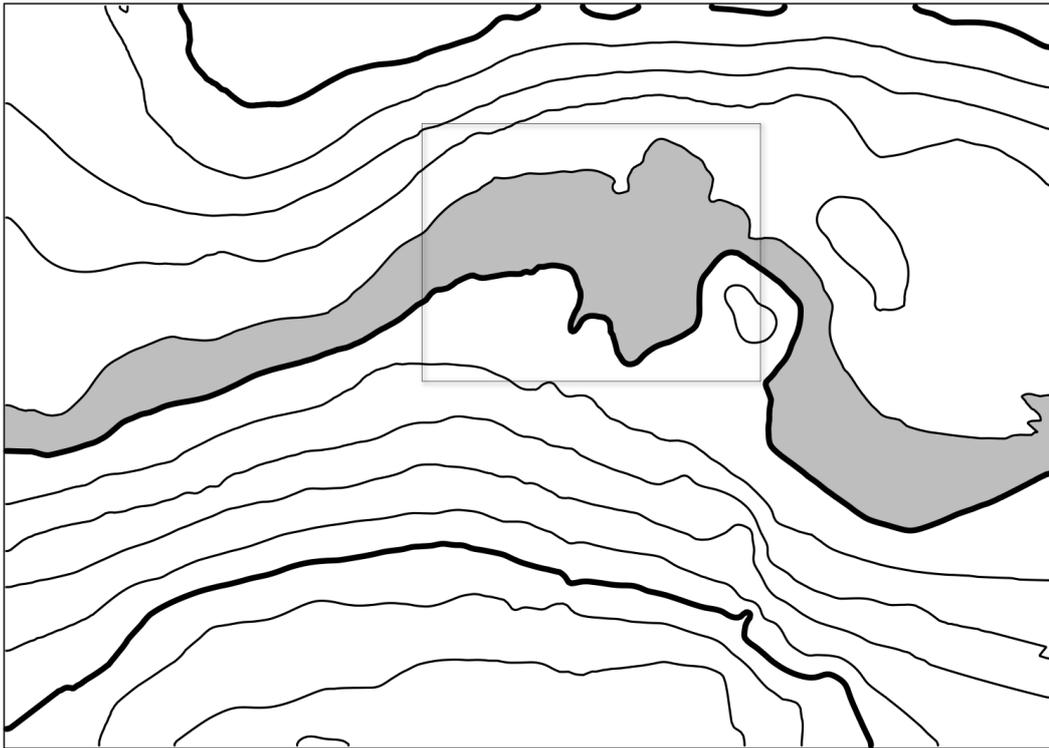
4. Проверяется, что полученные отрезки достаточно длинные и не разделены короткими промежутками. Короткие промежутки заполняются, короткие отрезки — удаляются.
5. При необходимости дополнительные горизонтали удлиняются так, чтобы они достигали центрального положения между соседними горизонталями, как показано на Рисунке 3.38.

Далее рассмотрим методические подробности, которые лежат в основе данной последовательности действий. Все этапы обработки проиллюстрированы на примере цифровой модели рельефа, представленной на Рисунке 3.39. Данная ЦМР была получена в результате геодезической съемки прибрежной территории в районе Беломорской биологической станции имени Н.А. Перцова МГУ имени М.В. Ломоносова, Карелия, Россия ( $66^{\circ}34'$  с.ш.,  $33^{\circ}08'$  в.д.) во время студенческой полевой практики в 2015 году.

### 3.4.2.3 Оценка ширины

Горизонтали основного сечения и граница карты делят картографируемую область на множество областей. Каждая такая область потенциально может содержать одну и более дополнительных горизонталей. Пример области заштрихован на Рисунке 3.39.

*Ширина области* используется для того, чтобы (а) убедиться, что достаточно места для размещения дополнительной горизонтали, и (б) выявить избыточно широкие области на плоских участках местности, где дополнительные горизонтали должны размещаться без учета их центральности. При этом области, ограниченные горизонталями основного сечения, в общем случае имеют произвольную форму и площадь, поэтому необходим метод, который позволил бы оценивать их ширину в произвольной точке. В дальнейших рассуждениях будем использовать именно термин «ширина области», а не заложение, поскольку заложение не определено для замкнутых горизонталей, в случае если ограничиваемые ими области не содержат других горизонталей основного сечения.



**Рисунок 3.39** — Две соседние горизонтали и граница карты определяют область (заштрихована серым), которая может содержать дополнительные горизонтали. Прямоугольник обозначает область на Рисунке 3.40

**Ширину области**  $W(p, r_i)$  в точке  $p$ , расположенной внутри области  $r_i$  определим как диаметр наибольшего круга  $c$  который содержит  $p$  и находится внутри  $r_i$ :

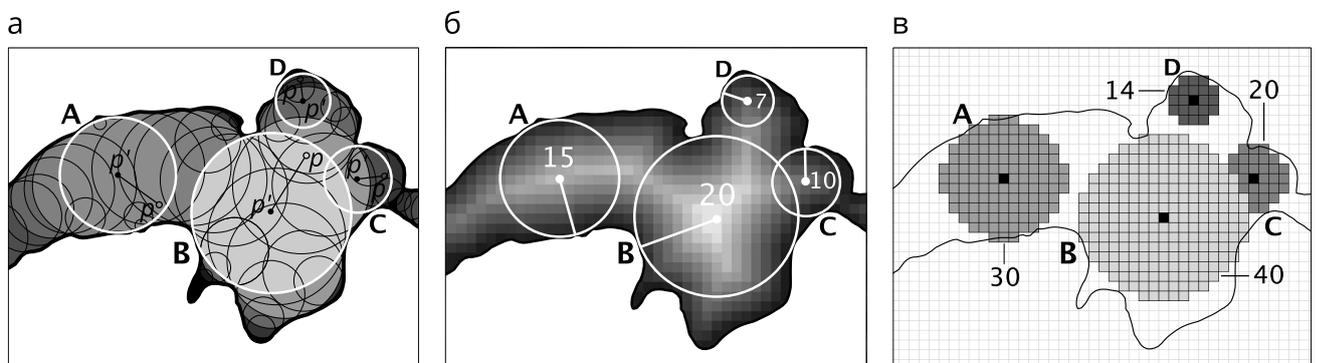
$$W(p, r_i) = \max_{p \in c \subseteq r_i} d(c), \quad (3.23)$$

где  $d(\cdot)$  — функция, вычисляющая диаметр круга. Этот круг назовем *доминантным кругом* точки  $p$ , а его центр (обозначаемый как  $p'$ ) — *доминантным соседом* точки  $p$ . Рисунок 3.40а иллюстрирует определение ширины области с использованием этих понятий. Для наглядности на нем показано только ограниченное число точек  $p$  и соответствующих им доминантных сущностей.

Вычисление ширины области можно эффективно автоматизировать с помощью алгебры карт в растровом режиме. Поиск точных доминантных кругов внутри области в этом случае заменяется приближенными вычислениями по ячейкам регулярной сетки. Алгоритм выглядит следующим образом:

1. Генерируется растровое покрытие, которое охватывает область моделирования с необходимым пространственным разрешением (вопрос выбора оптимального разрешения будет рассмотрен далее).
2. Для каждой ячейки растрового покрытия вычисляется и сохраняется евклидово расстояние до ближайшей основной горизонтали (Рисунок 3.40б). Это соответствует операции алгебры карт *FocalProximity* [Tomlin, 2012].

3. Создается выходное растровое покрытие с тем же пространственным разрешением и охватом и инициализируется нулевыми значениями.
4. Удвоенное значение каждой ячейки растра расстояний распространяется на выходные ячейки, которые покрываются окрестностью круга соответствующего диаметра (Рисунок 3.40в).
5. Результирующее значение определяется по следующему правилу: если ячейка пуста или имеет значение меньше, чем удвоенное значение растра расстояний, то ее значение заменяется удвоенным значением растра расстояний; в противном случае оно остается неизменным. Это показано на Рисунке 3.40в, где ячейки нижней левой части круга  $C$  перекрываются ячейками круга  $B$ , который имеет больший радиус.



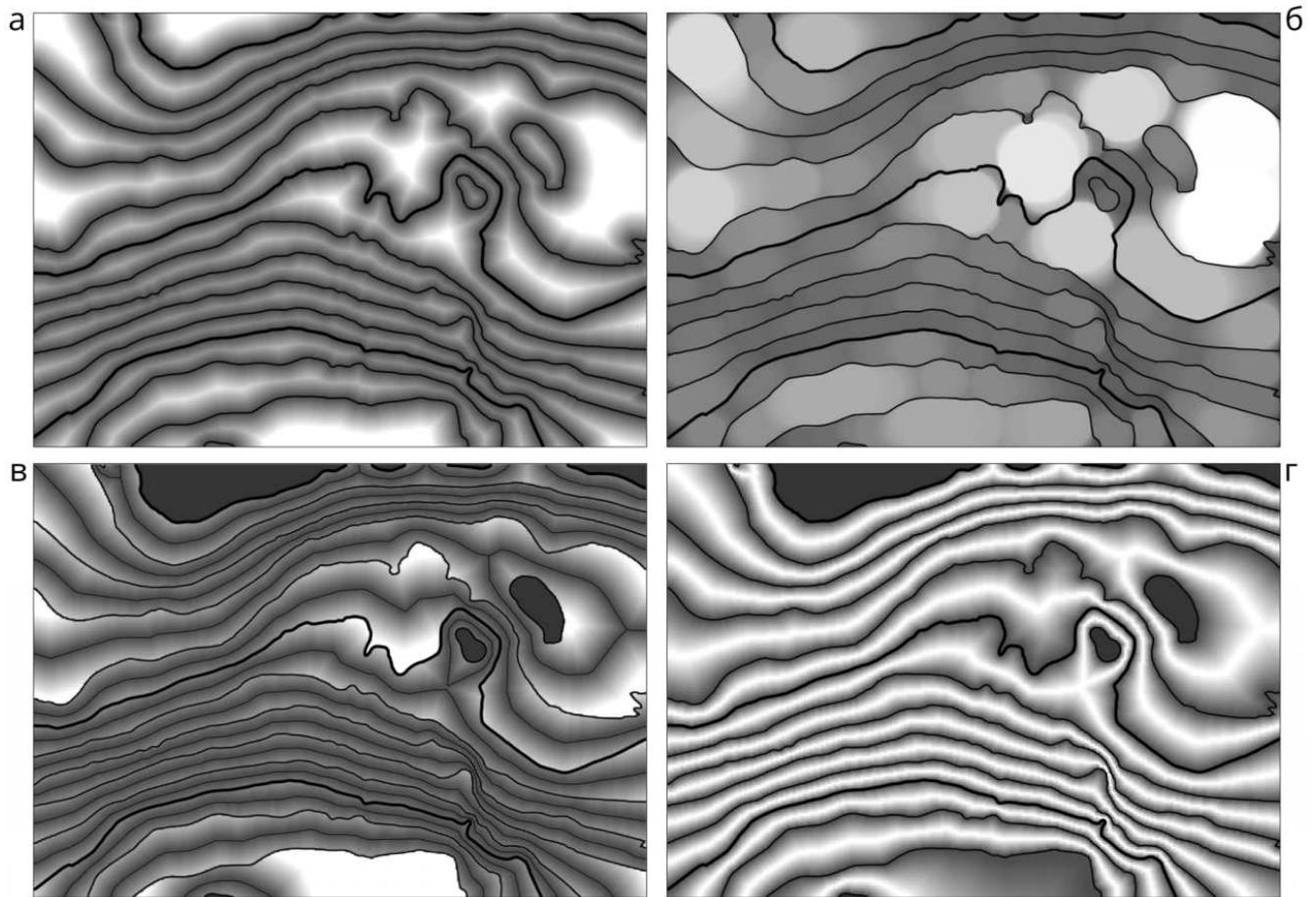
**Рисунок 3.40** – Вычисление ширины области: (а) Определение ширины на основе кругов (более яркие круги обозначают большую ширину области). Точки  $p'$  являются доминантными соседями точек  $p$ . Их окружности  $A$ ,  $B$ ,  $C$  и  $D$  нарисованы белым цветом; (б) Скалярное поле с расстоянием до ближайшей горизонтали; (в) Распространение удвоенного расстояния до пиксельной окрестности соответствующего размера (более яркие ячейки обозначают большие диаметры круга). Размер ячейки – 2.5 м

Полный пример вычисления ширины области показан на Рисунке 3.41. На Рисунке 3.41а показано растровое поле с евклидовым расстоянием до ближайшей горизонтали, а на Рисунке 3.41б представлен полученный растр ширины области.

#### 3.4.2.4 Оценка центральности

Как было сказано ранее, дополнительная горизонталь считается информативной, если она не равноудалена от горизонталей основного сечения. Для оценки этого свойства используем понятие *центральности*. Центральность точки, которая находится точно между двумя основными горизонталями, имеет максимальное значение 1; центральность точки, которая находится близко к одной из основных горизонталей, близка к минимальному значению 0.

Центральность, в отличие от расстояния до ближайшей горизонтали, нельзя измерить вдоль какой-то прямой. Чтобы вывести расчет центральности, анализируем раз-

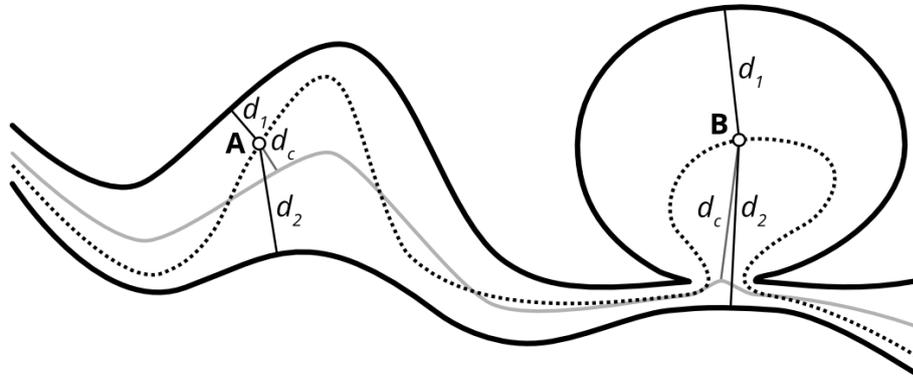


**Рисунок 3.41** – Вычисление ширины и центральности области: (а) расстояние до ближайшей горизонтали, (б) ширина области, (в) расстояние до центральной линии каждой области, (г) центральность. Яркие оттенки соответствуют большим значениям. Самые темные оттенки на (в) и (г) соответствуют областям, где центральность не определена

личные ситуации, которые представлены на Рисунке 3.42. Наивный способ вычисления центральности заключается в получении отношения расстояний до первой  $d_1$  и второй  $d_2$  ближайшей горизонтали:

$$C_{12} = d_1/d_2. \quad (3.24)$$

Эта формула эффективно работает в простых случаях, таких как точка  $A$  на Рисунке 3.42, где горизонтали следуют друг за другом ( $C_{12} \approx 0.4$ ), но она не работает при более сложных конфигурациях горизонталей. Например, точка  $B$  расположена в широкой изолированной части области, где дополнительная горизонталь была бы очень информативна. Однако точка  $B$  имеет высокое значение  $C_{12} \approx 0.8$  в соответствии с Формулой 3.24, поэтому при дальнейшей обработке она будет отфильтрована.



**Рисунок 3.42** – Вычисление центральности:  $d_1$  расстояние до первой ближайшей горизонтали,  $d_2$  расстояние до второй ближайшей горизонтали,  $d_c$  расстояние до центральной линии между первой и второй ближайшей горизонталями (изображена серой линией)

Более надежная мера центральности может быть получена на основе центральной линии между двумя горизонталями, т.е. линии, равноудаленной от обеих горизонталей. В векторном режиме эту линию можно получить, если построить диаграмму Вороного для основных горизонталей и извлечь границу между ее ячейками. С расстоянием  $d_c$  до центральной линии, центральность может быть рассчитана как:

$$C = \frac{d_1}{d_1 + d_c}. \quad (3.25)$$

Эта формула по-прежнему работает для точки  $A$  и при этом эффективно справляется со случаем точки  $B$  ( $C \approx 0.3$ ).

Как и в случае оценки ширины области, центральность вычисляется с помощью алгебры карт в растровом режиме. Расстояние до ближайшей горизонтали  $d_1$  есть евклидово расстояние, полученное ранее при оценке ширины области (Рисунок 3.41а). Для расчета  $d_c$  центральные линии получаются с помощью операции *евклидова распределения* (также называемой *FocalNeighbor* [Tomlin, 2012]), применяемой к основным горизонталям. Получаемый в результате этой операции категориальный растр содержит в каждой

ячейке идентификатор ближайшей горизонтали и является растровой аппроксимацией искомой диаграммы Вороного. Границы между полученными растровыми зонами являются искомыми центральными линиями, показанными на Рисунке 3.41в в виде тонких линий между горизонталями.

В отличие от  $d_1$ , которое всегда является расстоянием по прямой, значение  $d_c$  может быть криволинейным, если прямая видимость между точкой и центральной линией перекрыта одной из горизонталей, формирующих область. Поэтому, чтобы найти  $d_c$  (а) строится дополнительная *стоимостная поверхность*, которая запрещает пересечение основных горизонталей во время вычисления кратчайшего пути, и (б) применяется операция растровой алгебры *стоимостное расстояние*, также называемая *FocalProximity in FRICTIONLAYER* [Tomlin, 2012]), которая использует стоимостную поверхность в качестве параметра трения. Стоимостная поверхность строится путем стирания ячеек, пересекающихся с регулярными горизонталями, и присвоения всем остальным ячейкам значения 1. Результирующая поверхность стоимостного расстояния показана на Рисунке 3.41в.

Следует обратить внимание на то, что центральность не определена для областей, на границе которых имеется только одна горизонталь. Эти области показаны самым темным оттенком на Рисунке 3.41в. Для них центральность принудительно устанавливается равной 0, чтобы позволить размещать в них дополнительные горизонтали исходя только из ширины.

После нахождения расстояния до ближайшей горизонтали  $d_1$  и расстояния до центральной линии  $d_c$  вычисляется центральность каждой ячейки растра согласно Формуле 3.25. На Рисунке 3.41г показана результирующая растровая поверхность.

### 3.4.2.5 Фильтрация

Каждая вершина в составе дополнительной горизонтали должна пройти два ограничения фильтрации: вершина удаляется, если (а) область в месте расположения вершины слишком узкая или (б) вершина находится близко к центральной линии. Удаление вершины приводит к разбиению линии-кандидата на два новых участка, между которыми остается промежуток.

*Замкнутые* дополнительные горизонтали, оконтуривающие впадины и пики, и при этом не охватывающие другой горизонтали, сохраняются целиком, если их длина превышает  $L_{\min}$  и если средняя ширина оконтуриваемой горизонтальной области больше заданного параметра  $W_e$ . При этом фильтрация по площади не используется, так как большой площадью может обладать излишне узкая, но при этом длинная область.

При фильтрации всех остальных дополнительных горизонталей необходимо различать широкие, узкие и средние по ширине участки областей:

- *Широкие участки*: не требуют учета центральности при проведении дополнитель-

ных горизонталей. Для их распознавания вводится параметр  $W_{\max}$ . Если в вершине горизонтали  $W > W_{\max}$ , то линия будет продолжена.

- *Узкие области:* исключают проведение дополнительных горизонталей. Для их распознавания вводится параметр  $W_{\min}$ . Если в вершине горизонтали  $W < W_{\min}$ , то линия будет прервана.
- *Средние области:* если вершина находится в участке области с шириной между  $W_{\min}$  и  $W_{\max}$ , то дополнительная горизонталь будет прервана, как только приблизится слишком к центральной линии. Для этого необходимо установить величину порога центральности  $\widehat{C}$  и проверять условие  $C < \widehat{C}$ , которое обсуждается ниже.

Обратим внимание на то, что случай широких участков позволяет проводить дополнительные горизонталы с произвольной центральностью, что эквивалентно  $\widehat{C} = 1$ . Логично предположить, что в случае узких участков должно быть  $\widehat{C} = 0$ , однако пороговое значение должно быть все же ненулевым, чтобы исключить дополнительные горизонталы слишком близко к основным.

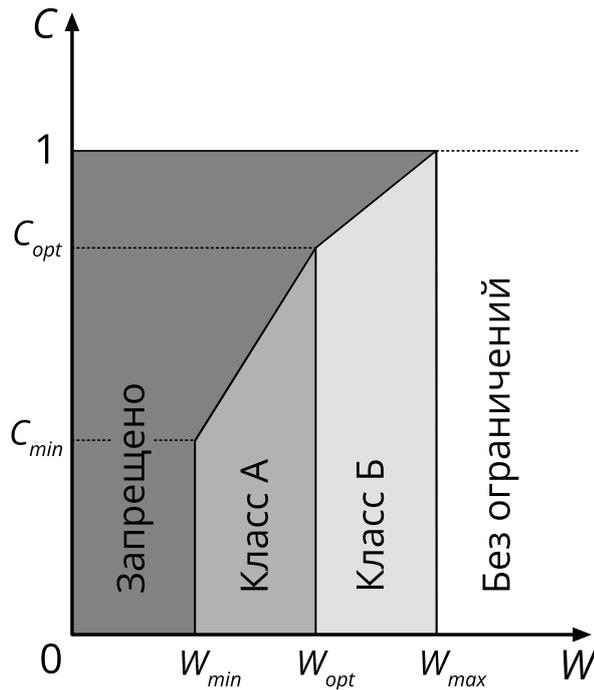
Введем более общие обозначения. Пусть  $\widehat{C}(W_{\min}) = C_{\min}$ , а  $\widehat{C}(W_{\max}) = C_{\max}$ . В этом случае порог центральности при изменении  $W$  может линейно изменяться между  $C_{\min}$  и  $C_{\max}$ :

$$\widehat{C}(W) = C_{\min} + (C_{\max} - C_{\min}) \frac{W - W_{\min}}{W_{\max} - W_{\min}}. \quad (3.26)$$

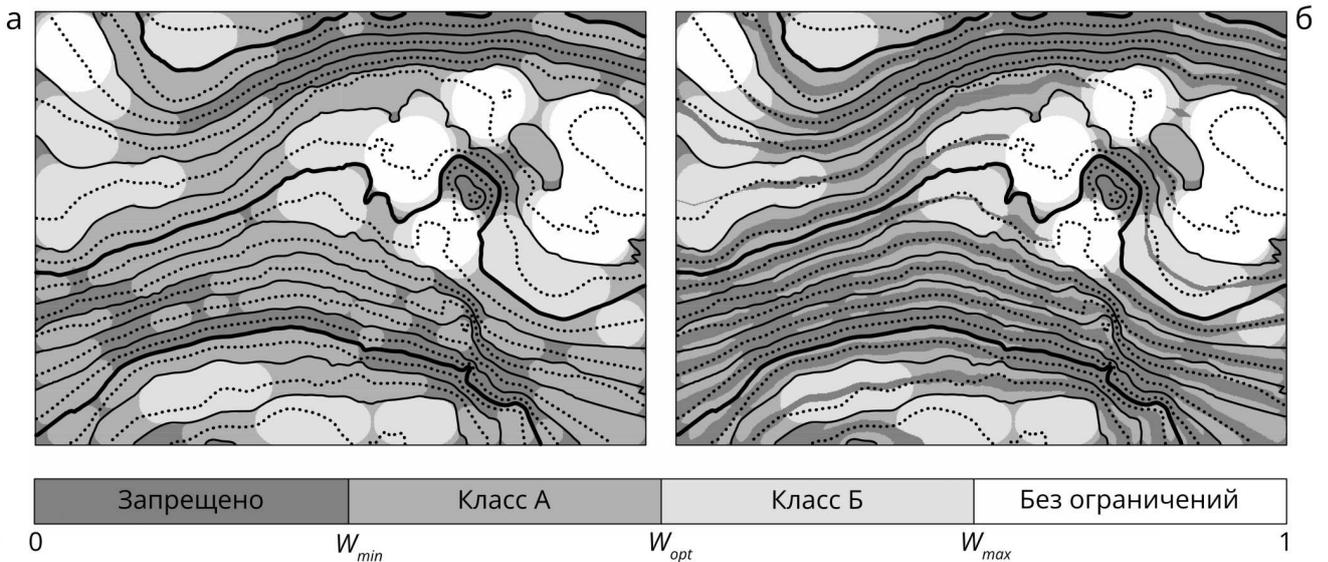
Экспериментальная работа, однако, показала, что простая линейная интерполяция не всегда эффективна. Для улучшения точности контроля было решено разбить диапазон ширины области между  $W_{\min}$  и  $W_{\max}$  оптимальным параметром центральности  $W_{opt}$  на два класса. Как показано на Рисунке 3.43, класс А отображает ширину области между  $W_{\min}$  и  $W_{opt}$  на  $[C_{\min}, C_{opt})$ , а класс Б отображает ширину области между  $W_{opt}$  и  $W_{\max}$  до  $[C_{opt}, C_{\max}]$ . Таким образом, на каждом участке используется линейная интерполяция в соответствии с Формулой 3.26 в качестве примера на Рисунке 3.43 значение  $C_{\min} = 0.4$ ,  $C_{opt} = 0.8$  и  $C_{\max} = 1$ . Большинство дополнительных горизонталей генерируется в областях класса В. Ограничение центральности не применяется, если ширина области больше  $W_{\max}$ .

Три серых оттенка на Рисунке 3.44а обозначают классы ширины области, соответствующие  $W_{\min}$ ,  $W_{opt}$  и  $W_{\max}$ . В данном случае они заданы на уровне 25%, 50% и 75% от максимальной ширины области в пределах ЦМР.

Три серых оттенка на Рисунке 3.44б обозначают зоны подходящей центральности в соответствии с зависимостью на Рисунке 3.43. Дополнительные горизонталы могут быть размещены на любых светлых участках на Рисунке 3.44б. В пределах темных участков их отображение блокируется.

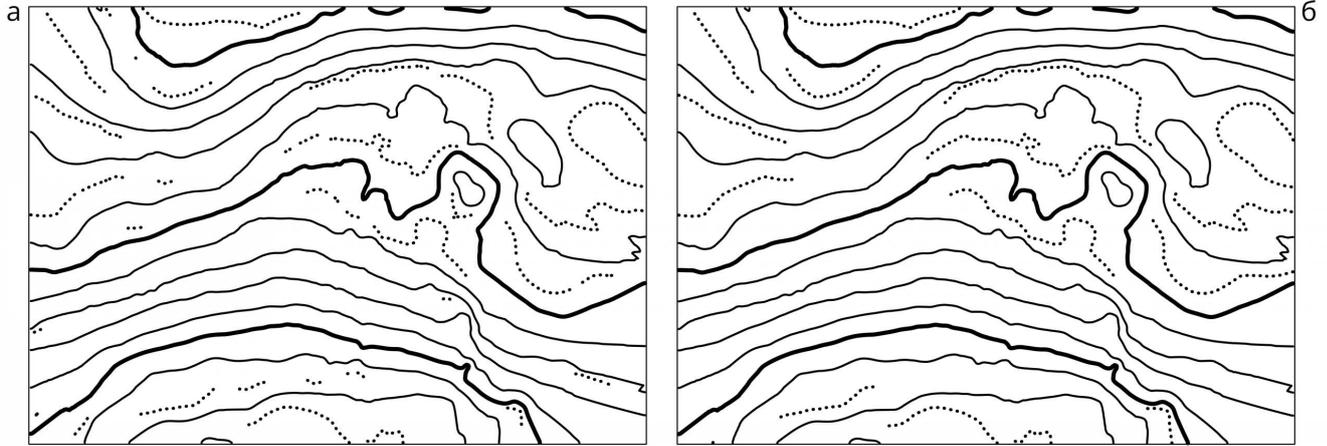


**Рисунок 3.43** – Сопоставление ширины области с порогом центральности. В самой темной области размещение дополнительных горизонталей запрещено; вершины фильтруются по центральности, если они относятся к классу А или Б; к вершинам в белой области фильтрация по центральности не применяется



**Рисунок 3.44** – Связь ширины и центральности: (а) ширина области, (б) порог центральности, полученный из (а) с помощью функции на Рисунке 3.43

На Рисунке 3.45а показан результат фильтрации дополнительных горизонталей по ширине и центральности. Видно, что линии прерываются многочисленными разрывами, а некоторые линии излишне короткие. Для устранения этих недостатков применяется алгоритм, рассмотренный в следующем разделе.



**Рисунок 3.45** – Дополнительные горизонталей: (а) после фильтрации вершин по ширине области и порогу центральности, (б) после заполнения коротких промежутков, удаления коротких участков и продления линий

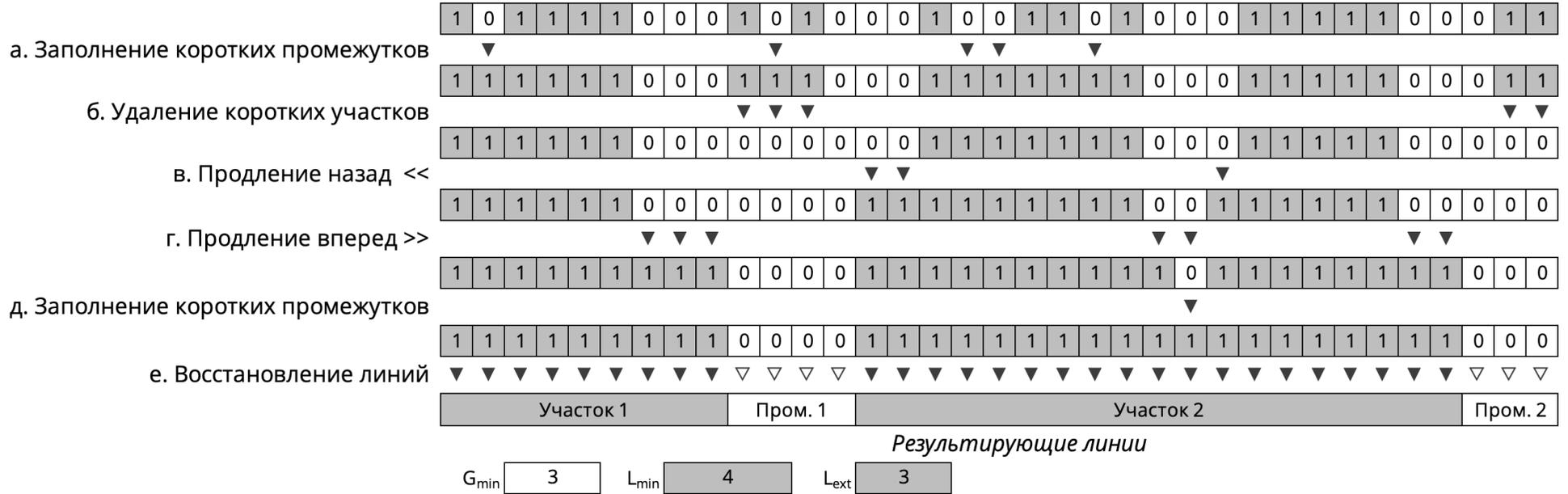
#### 3.4.2.6 Обработка промежутков и коротких участков

Визуальная непрерывность дополнительных горизонталей обеспечивается путем заполнения коротких промежутков между ними и удаления коротких участков. Для этого каждая вершина, которая была удалена на предыдущем шаге, помечается флагом  $F = 0$ , а оставшиеся помечаются флагом  $F = 1$ . Эти флаги меняются на противоположные согласно следующим критериям:

- (1) Вершины помечаются флагом  $F = 1$ , если они являются частью промежутка  $F = 0$ , длина которого меньше  $G_{\min}$  (Рисунок 3.46а).
- (2) Вершины помечаются флагом  $F = 0$ , если они являются частью промежутка  $F = 1$ , длина которого меньше  $G_{\min}$  (Рисунок 3.46б).

Важно, чтобы операции проводились именно в таком порядке, как указано. В этом случае последовательность коротких участков, разделенных короткими промежутками, будет объединена в непрерывную дополнительную горизонталь. Если же сначала удалить короткие участки, такая горизонталь наоборот будет удалена. Таким образом, первичное заполнение промежутков повышает вероятность «выживания» дополнительной горизонтали.

Исходные вершины



**Рисунок 3.46** – Создание непрерывных дополнительных горизонталей путем заполнения промежутков, удаления коротких участков линий и продления линий (для простоты предполагается постоянное расстояние между вершинами)

### 3.4.2.7 Продление линий

Этот, заключительный, этап обработки данных направлен на соблюдение принципа, согласно которому дополнительные горизонталы должны достигать центрального положения между двумя соседними горизонталями основного сечения. В целом, нет гарантии, что линия дополнительной горизонтали когда-либо достигнет центрального положения, поэтому необходимо ввести некий порог продления, за пределами которого оно не будет выполняться.

Для этого вводится *центральность продления*  $C_{ext}$  и *длина продления*  $L_{ext}$ . Обработка выполняется отдельно для первой и последней вершины каждого оставшегося участка дополнительной горизонтали. Начиная с первой вершины, алгоритм сдвигается в сторону промежутка до тех пор, пока пройденное вдоль линии расстояние не станет больше, чем  $L_{ext}$ . Затем выбирается вершина  $p_l$  с наибольшей центральностью  $C_l$  среди пройденных вершин. Если  $C_l \geq C_{ext}$ , то все пройденные вершины между первой вершиной и  $p_l$  помечаются флагом  $F = 1$  (Рисунок 3.46в). В противном случае флаги не изменяются. Аналогичная процедура выполняется с последней вершины в противоположную сторону линии (Рисунок 3.46г).

Побочным эффектом процедуры удлинения является сокращение промежутков между участками дополнительной горизонтали, что требует еще одной итерации заполнения промежутков (Рисунок 3.46д). На последнем этапе последовательности вершин с  $F = 1$  преобразуются уже в векторные линии (Рисунок 3.46е).

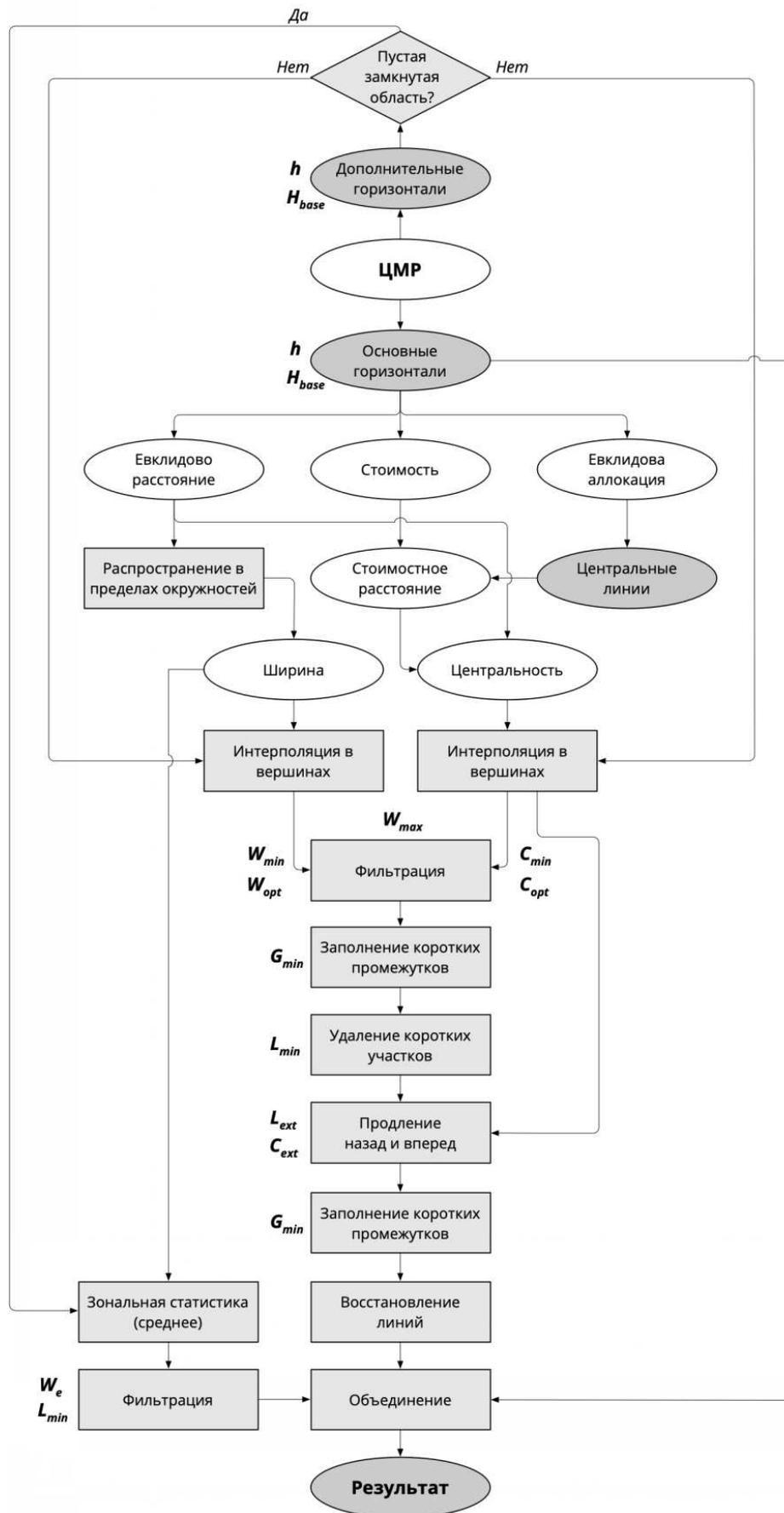
Рисунок 3.45б иллюстрирует комбинированный эффект заполнения промежутков, удаления коротких участков и продления линий до центрального положения. По сравнению с Рисунком 3.45а, можно заметить, что на Рисунке 3.45б небольшие участки дополнительных горизонталей были удалены или объединены в более длинные непрерывные линии, при этом многие из них достигают положения, близкого к центральному.

Таким образом, после выполнения перечисленных действий, обеспечивается реализация всех основных принципов проведения дополнительных горизонталей при изображении рельефа.

### 3.4.3 Итоговый алгоритм и его реализация

Итоговый алгоритм, объединяющий пройденные шаги, представлен на Рисунке 3.47. Конечный результат представляет собой комбинацию основных горизонталей, замкнутых дополнительных горизонталей для впадин и пиков, которые не охватывают никакие другие горизонталы, и дополнительных горизонталей, созданных описанным методом.

Набор параметров, управляющих процессом обработки, представлен в Таблице 3.7. Пояснения к каждому параметру приведены в столбце «Описание». В столбце «Диапазон



**Рисунок 3.47** – Алгоритм генерации горизонталей для изображения рельефа на карте. Растровые покрытия представлены белыми овалами (овал DEM является началом рабочего процесса), векторные объекты представлены серыми овалами, а процедуры обработки – серыми прямоугольниками. Обозначения на схеме приведены согласно Таблице 3.7

Таблица 3.7 – Параметры инструмента генерации дополнительных горизонталей

Символ	Название	Описание	Значение по умолчанию	Диапазон значений
$h$	Сечение рельефа	Вертикальный интервал между уровнями основных горизонталей.	–	$(0, +\infty)$
$H_{base}$	Базовый уровень	Уровень, от которого отсчитывается сечение рельефа.	0	$(-\infty, +\infty)$
$W_{min}$	Минимальная ширина области	Если ширина области меньше, чем $W_{min}$ , то дополнительные горизонталы не наносятся.	$0.5 \cdot W_{opt}$	$[0, W_{opt}]$
$W_{opt}$	Оптимальная ширина области	Оптимальная ширина области, в которой необходимо построить дополнительные горизонталы.	–	$[0, +\infty)$
$W_{max}$	Максимальная ширина области	Если ширина области больше, чем $W_{max}$ , то дополнительные горизонталы фильтруются по порогу центральности.	$1.5 \cdot W_{opt}$	$[W_{opt}, +\infty)$
$W_e$	Минимальная средняя ширина области для замкнутых доп. горизонталей	Замкнутая дополнительная горизонталь наносится только в том случае, если средняя ширина ограниченной ею области больше $W_e$ .	$0.25 \cdot W_{opt}$	$[0, W_{opt}]$
$C_{min}$	Минимальная центральность	Максимально допустимая центральность для $W_{min}$ .	0.4	$[0, C_{opt}]$
$C_{opt}$	Оптимальная центральность	Максимально допустимая центральность для $W_{opt}$ .	0.8	$[0, 1]$
$C_{ext}$	Центральность продления	Минимальная центральность, которая должна быть достигнута при продлении дополнительной горизонтали.	0.8	$[C_{opt}, 1]$
$L_{min}$	Минимальная длина участка	Дополнительные горизонталы короче $L_{min}$ удаляются.	$W_{opt}$	$[0, +\infty)$
$L_{ext}$	Длина продления	Участки дополнительных горизонталей расширяются в обе стороны на $L_{ext}$ .	$W_{opt}$	$[0, L_{min}]$
$G_{min}$	Минимальная длина промежутка	Промежуток между дополнительными горизонталями короче чем $G_{min}$ заполняется.	$W_{opt}$	$[0, L_{min}]$

значений» представлен возможный диапазон значений.

Оптимальная центральность  $C_{opt}$  является единственным параметром, который не зависит от параметров ЦМР. Сечение рельефа  $h$ , базовый уровень  $H_{base}$ , оптимальная ширина  $W_{opt}$  и минимальная длина  $L_{min}$  зависят от размаха высот, масштаба карты и пространственного охвата ЦМР. Все остальные параметры ограничены значениями  $W_{opt}$ ,  $L_{min}$  и  $C_{opt}$ . Значения по умолчанию получены в результате экспериментальной работы, описанной в следующем разделе.

Алгоритм был реализован на языке *Python* в виде свободно распространяемого дополнительного модуля для *Esri ArcGIS for Desktop / ArcGIS Pro*, который можно загрузить из репозитория *GitHub* [Samsonov, Walther, 2019]. Инструмент предоставляет графиче-

ский пользовательский интерфейс для выбора параметров и создает как обычные, так и дополнительные горизонталы. Цифровые модели рельефа, представленные в следующем разделе, также включены в репозиторий для обеспечения воспроизводимости полученных результатов.

### 3.4.4 Результаты

На Рисунках 3.48—3.53 представлены дополнительные горизонталы, созданные на основе шести цифровых моделей рельефа. На всех шести рисунках карта, обозначенная (а), показывает дополнительные горизонталы после фильтрации по ширине и центральности. Серым цветом отмечены области, в которых дополнительные горизонталы не отображаются. Карта, обозначенная (б) на Рисунках 3.48—3.53, показывает итоговые линии после всех этапов обработки. Для всех шести примеров использованы значения параметров по умолчанию, перечисленные в Таблице 3.7.  $W_{opt}$  было установлено равным половине максимальной ширины области для каждой модели.

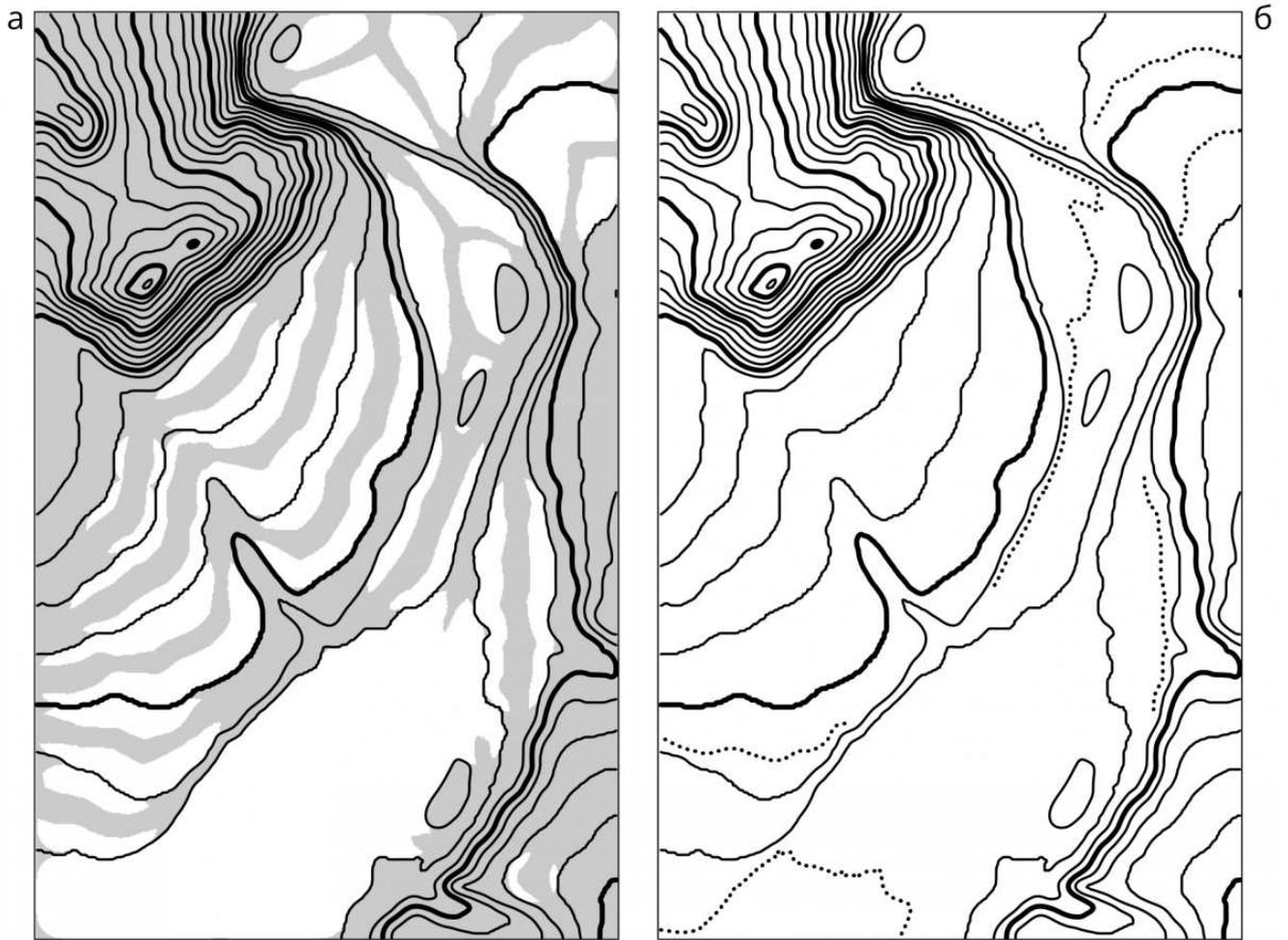
Модели 1, 2 и 3 взяты из альбома примеров изображения рельефа на топографических картах ЦНИИГАиК [Гольдман, 1968].

**Модель 1** представляет низкоргорный рельеф (Рисунок 3.48). Широкие области между горизонталями основного сечения соответствуют речной пойме, занимающей самые низкие положения в рельефе. Центральная дополнительная горизонталь в направлении юг-юго-запад наглядно показывает преимущества предлагаемого подхода с точки зрения рисовки рельефа. На карте слева (Рисунок 3.48а) показаны шесть участков дополнительной горизонтали различной длины; два самых коротких участка обозначены одиночными точками. После заполнения промежутков между участками горизонталь (Рисунок 3.48б) представляет собой одну непрерывную линию. Аналогичным образом удаляются короткие участки, между которыми получились значительные промежутки (см. четыре коротких участка слева от центра карты).

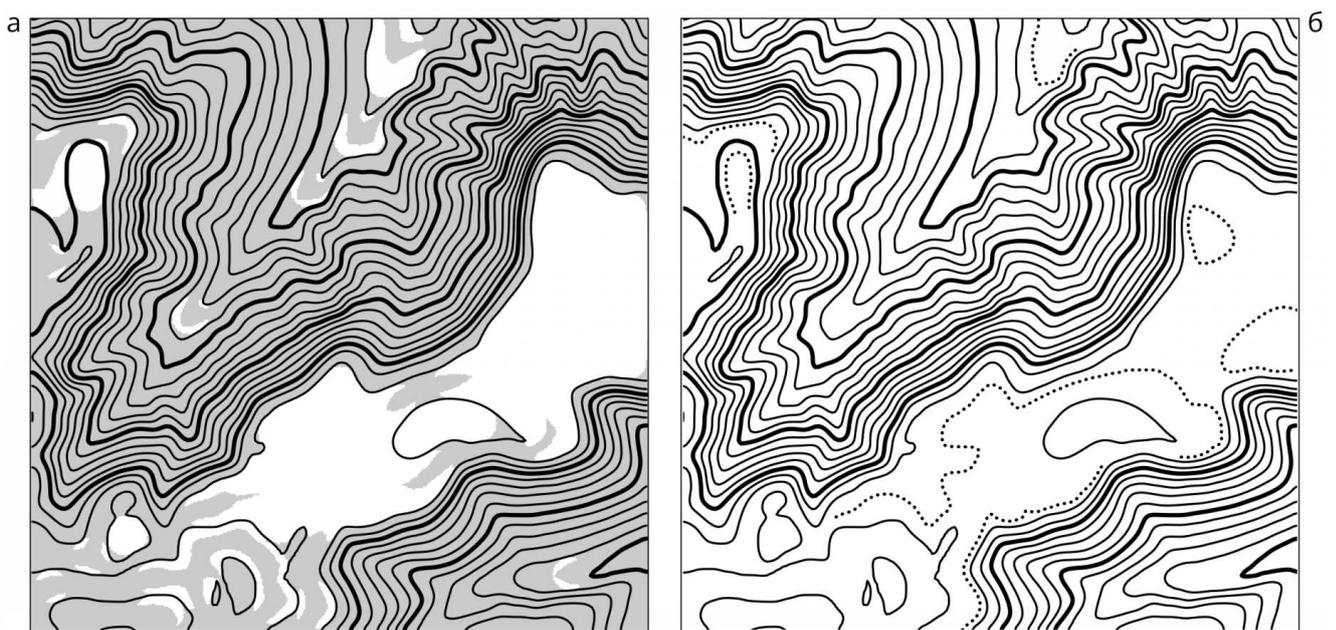
**Модель 2** представляет ледниковый рельеф с множеством вытянутых мезоформ (Рисунок 3.49). Выявлены многочисленные замкнутые привершинные дополнительные горизонталы, подчеркивающие их круглую форму. Также изображение Модели 2 содержит множество дополнительных горизонталей, которые были продлены до центрального положения между соседними горизонталями. Особенно заметно это в северо-западной части карты. Продление позволяет подчеркнуть место, в котором поперечный профиль склона становится прямым.

**Модель 3** представляет вулканическое трапповое плато с характерной широкой горизонтальной поверхностью на самых больших высотах (Рисунок 3.50). В данном случае дополнительные горизонталы успешно показывают затухание роста высоты и пологовыпуклый характер бровки плато.

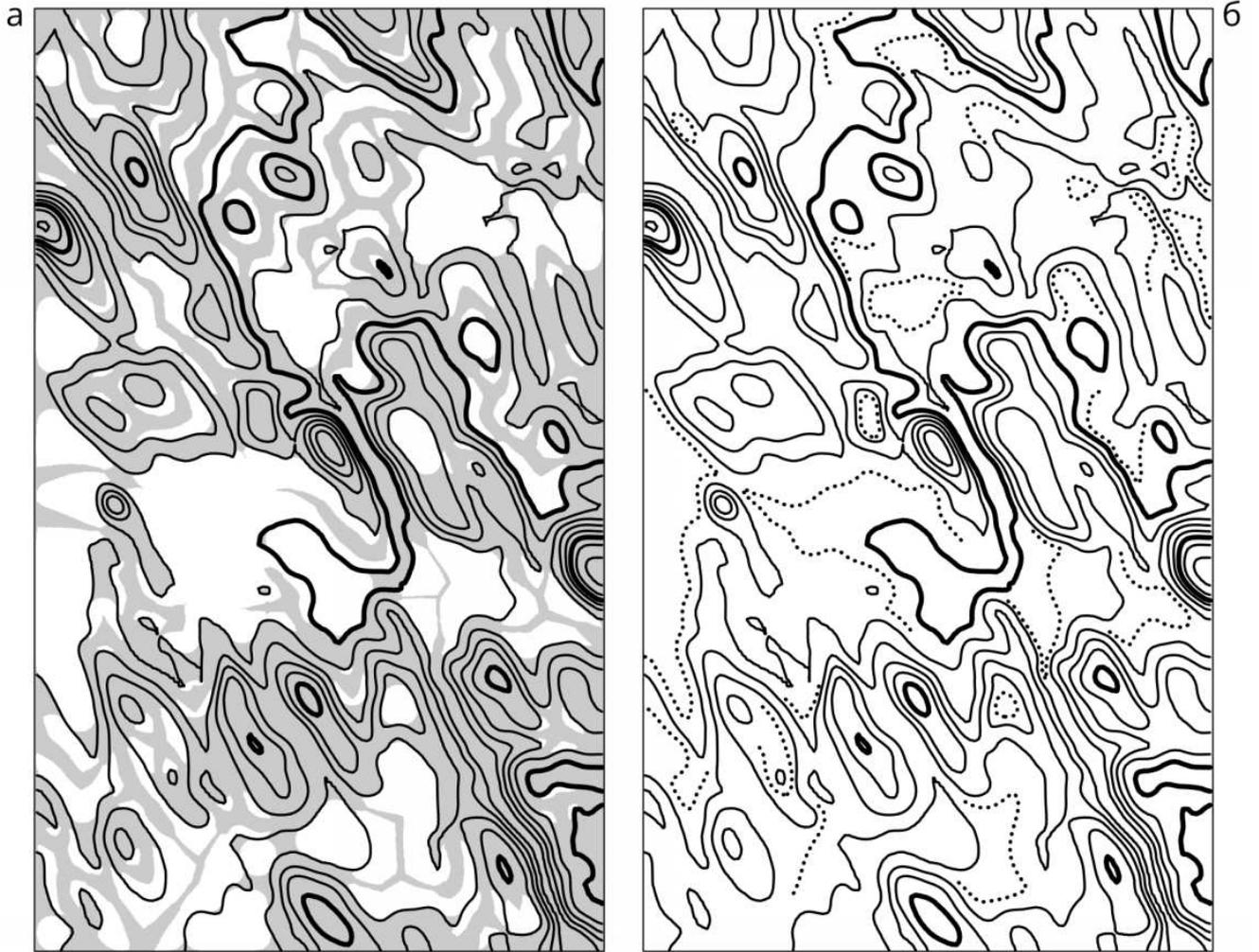
**Модель 4** представляет собой фрагмент цифровой модели рельефа SRTM на неболь-



**Рисунок 3.48** — Дополнительные горизонтали для поймы на низкогорном участке (модель 1): (а) допустимые области построения дополнительных горизонталей (отмечены белым цветом); (б) дополнительные горизонтали после заполнения промежутков, удаления коротких участков и продления. Сечение рельефа — 10 метров



**Рисунок 3.49** — Дополнительные горизонтали при изображении ледникового рельефа (модель 2). Основное сечение рельефа 2.5 метра. Также см. подпись к Рисунку 3.48



**Рисунок 3.50** – Дополнительные горизонталы при изображении вулканического траппового плато (модель 3). Основное сечение рельефа 10 метров. Также см. подпись к Рисунку [3.48](#)

шой участок северного макросклона Кавказских гор (Рисунок 3.51). Это предгорный регион, при этом главный хребет расположен в юго-западном углу карты. Широкие области между горизонталями, нуждающиеся в повышении детальности изображения, в данном случае ожидаемо сосредоточены в предгорной части. Наличие дополнительных горизонталей существенно обогащает рисовку такого сложного рельефа, но при этом есть много относительно коротких участков горизонталей, что говорит о потенциале для дальнейшего совершенствования предложенной методики.

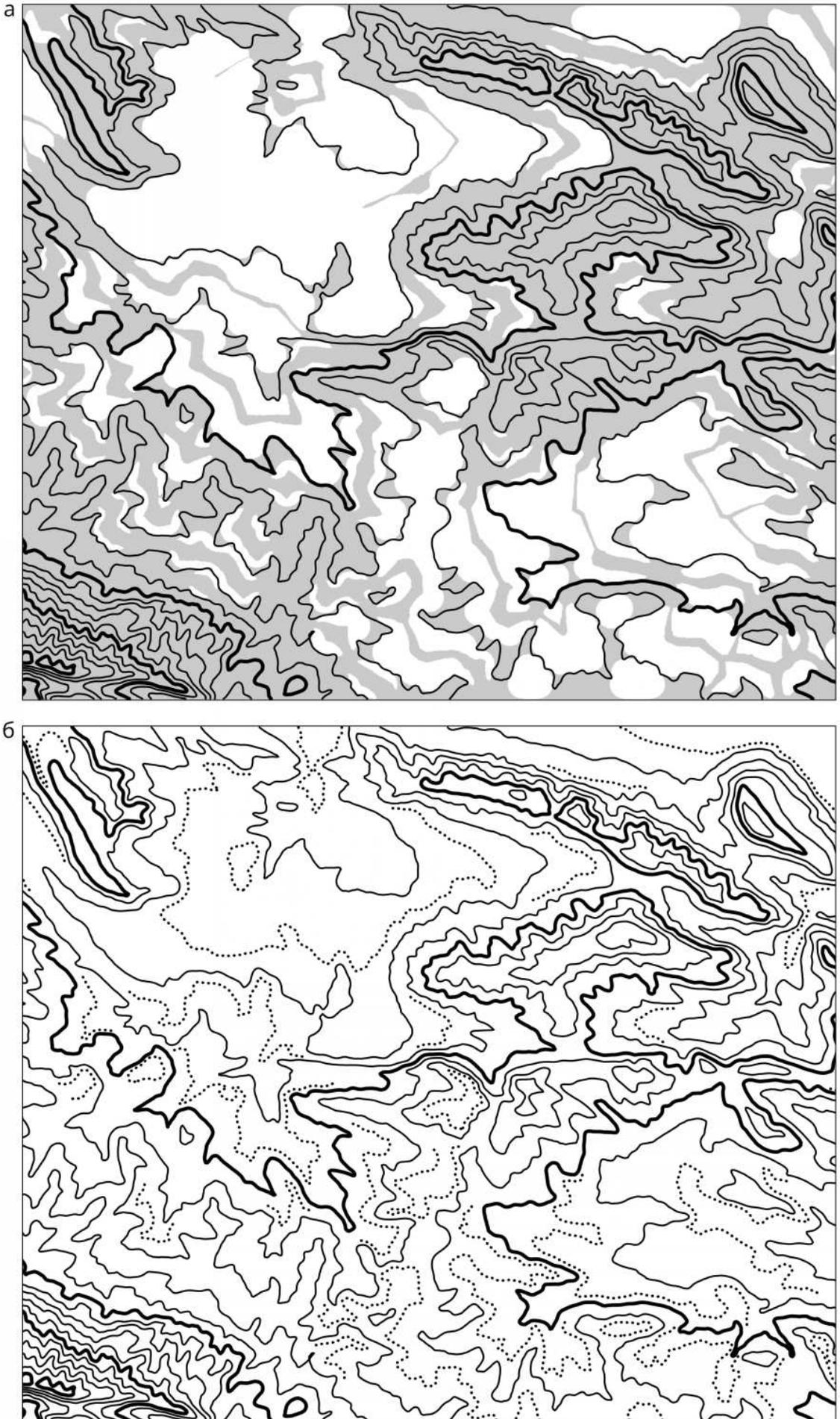
**Модель 5** — это участок цифровой модели рельефа SRTM на территорию карстового массива Дивача в Словении с диапазоном высот от 325 до 500 м (Рисунок 3.52). Дополнительные горизонтали особенно полезны для такого сложного типа рельефа — они четко иллюстрируют карстовые воронки, не выражающиеся основным сечением, а также показывают прочие незаметные в основном сечении вариации формы поверхности. Бергштрихи отрисованы вручную, чтобы указать направление понижения высот.

**Модель 6** представляет фрагмент долины реки Протвы в районе Сатинской учебно-научной базы Географического факультете МГУ (Рисунок 3.53). ЦМР имеет горизонтальное разрешение 5 м и создана по данным аэрофотосъемки. Размах высот от 140 до 180 м. По картографическому изображению видно, что пара дополнительных горизонталей оказалась очень полезна для рисовки рельефа самых низких участков поймы реки, которые не выражены в горизонталях основного сечения. Несколько коротких дополнительных горизонталей, указывающих на микроформы пойменного рельефа, не были включены в окончательный вариант карты, поскольку не удовлетворяют критериям длины и/или средней ширины.

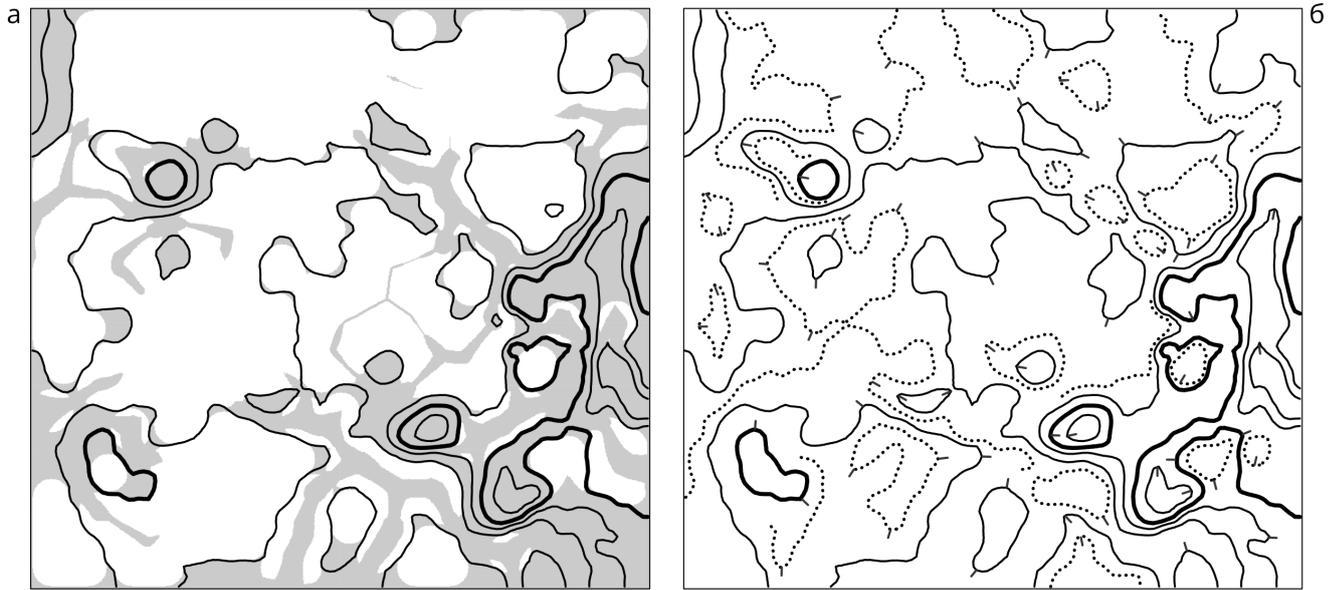
### 3.4.5 Обсуждение

Дополнительные горизонтали, созданные с помощью разработанной методики и представленные на Рисунках 3.48—3.53 имеют качество, пригодное для картографического производства. На данные изображения были получены отзывы и комментарии двух старших экспертов-картографов, специализирующихся на топографическом картографировании в швейцарском национальном картографическом агентстве *Swisstopo* и *Институте картографии и геоинформации университета ETH Zurich* — двух учреждениях, которые известны во всем мире своими школами топографического картографирования ([Tufte, 2001; Kenneth Field, 2018]). Экспертов попросили прокомментировать размещение дополнительных горизонталей. Полученные отзывы были положительными, при этом эксперты отметили, что ручное редактирование не смогло бы существенно улучшить автоматически сгенерированные дополнительные горизонтали.

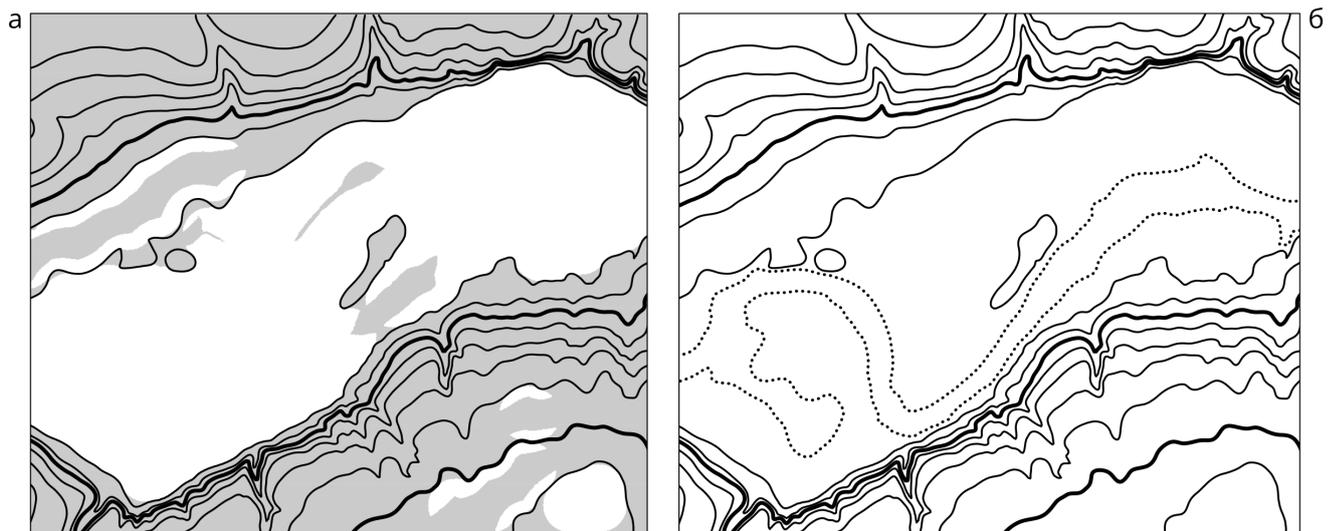
Отметим, что значения по умолчанию почти для всех параметров алгоритма в Таблице 3.7 могут быть либо заданы заранее, либо получены из параметра оптимальной ширины области  $W_{opt}$ . Этот параметр используется для расчета значений параметров



**Рисунок 3.51** – Дополнительные горизонталы при изображении предгорной части Кавказских гор (модель 4). Основное сечение рельефа 100 метров. Также см. подпись к Рисунку [3.48](#)



**Рисунок 3.52** – Дополнительные горизонталы при изображении карстового массива Дивача в Словении (модель 5). Основное сечение рельефа 25 метров. Также см. подпись к Рисунку [3.48](#)



**Рисунок 3.53** – Дополнительные горизонталы при изображении долины реки Протвы в районе Сатинской учебно-научной базы Географического факультете МГУ (модель 6). Основное сечение рельефа 5 метров. См. также подпись к Рисунку [3.48](#)

по умолчанию для ширины области, заполнения промежутков, удаления коротких участков, продления до центрального положения и фильтрации замкнутых дополнительных горизонталей. При этом значения параметров по умолчанию не гарантируют идеальных результатов для всех типов рельефа, но служат хорошей отправной точкой для более точной настройки, если это необходимо.

Также отметим, что вычисление ширины области и центральности — относительно дорогие вычислительные операции. В частности, вычисление растра ширины в растровом режиме имеет алгоритмическую сложность  $O(n^2)$ , где  $n$  — ширина или высота модели рельефа в ячейках. Однопоточная реализация на языке Python создает растр ширины размером  $500 \times 500$  ячеек примерно за 20 секунд на ноутбуке *Intel Core i7 2,6 ГГц*. Такие вычислительные затраты не позволяют визуализировать поверхности в режиме реального времени, но подходят для создания топографических карт в производственном режиме. Для повышения скорости обработки многие этапы рабочего процесса могут быть реализованы в многопоточном режиме и/или на высокопроизводительном компилируемом языке типа *Rust* или *C++*.

### 3.4.6 Итоги

Изображение рельефа с помощью горизонталей — это всегда баланс между детальностью и графической нагрузкой изображения. Более частое сечение рельефа позволяет повысить детальность изображения, но может привести к чрезмерной графической нагрузке, повышенной плотности горизонталей на единицу площади карты. Это снижает читаемость карты и, как показывает практика, мешает трехмерному эффекту отмывки, которая часто используется в сочетании с горизонталями. Увеличение высоты сечения позволяет избежать подобных недостатков, однако приводит к уменьшению детальности изображения, в том числе полному исчезновению на изображении отдельных форм рельефа и их элементов. Хорошо зарекомендовавшее себя в ручной картографии компромиссное решение — сочетание горизонталей умеренно плотного сечения с прицельно расположенными дополнительными горизонталями, которые выявляют подобные «скрытые нюансы» рельефа и позволяют локально повысить детальность его изображения. Разработанный метод впервые позволил решить эту задачу автоматизированным путем. Таким образом, был пройден важный этап управления детальностью картографического изображения рельефа путем подходящего отбора (генерализации) дополнительных горизонталей.

### 3.5 Картографическая визуализация цифровых моделей рельефа с помощью штрихов и управление его детальностью<sup>6</sup>

Представленный в настоящем разделе диссертации метод [Samsonov, 2014] позволил не только впервые полноценно автоматизировать построение штрихов крутизны и теневых штрихов, но также алгоритмизировать получение картографических изображений различной детальности. Цифровая модель рельефа сглаживается с помощью гауссова фильтра с заданными параметрами (размер, количество итераций). После этого выполняется построение изолиний с заданным сечением, и вдоль изолиний с равным шагом выпускаются линии тока. Более высокая детальность изображения достигается при использовании более частотного сечения рельефа, малого интервала расстановки штрихов, а также размера сглаживающего фильтра и количества итераций сглаживания. При этом наглядно показано, что без использования сглаживания изображение не обладает необходимой степенью упорядоченности.

В данном разделе преследуются три основные цели:

1. Разработать алгоритм построения штрихов, полностью удовлетворяющий пяти оригинальным требованиям Имгофа.
2. Исследовать возможности управления детальностью изображения рельефа при использовании штрихов.
3. Исследовать возможности применения штрихов в морфометрическом картографировании.

#### 3.5.1 Алгоритм

Разработанный алгоритм построения штрихов соответствует следующим требованиям:

1. Горизонтали используются в качестве направляющих для размещения штрихов.
2. Штрихи получаются как линии тока между горизонталями, а не прямые отрезки.
3. Расхождение (дивергенция) линий тока компенсируется вставкой дополнительных штрихов.
4. Штрихи вверх по склону генерируются наряду со штрихами вниз по склону.

Алгоритм состоит из четырех этапов и управляется семью параметрами, которые описаны далее. Общая схема алгоритма представлена на Рисунке 3.54. Все метрические

<sup>6</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov, 2014]

параметры — в единицах, угловые — в градусах, а глубина рекурсии — безразмерная величина. В этом параграфе коснемся технических аспектов алгоритма, а выбор и адаптация его основных параметров, таких как сечение рельефа и шаг штриховки, будут рассмотрены в следующем разделе, наряду с оценкой их влияния на детальность изображения.

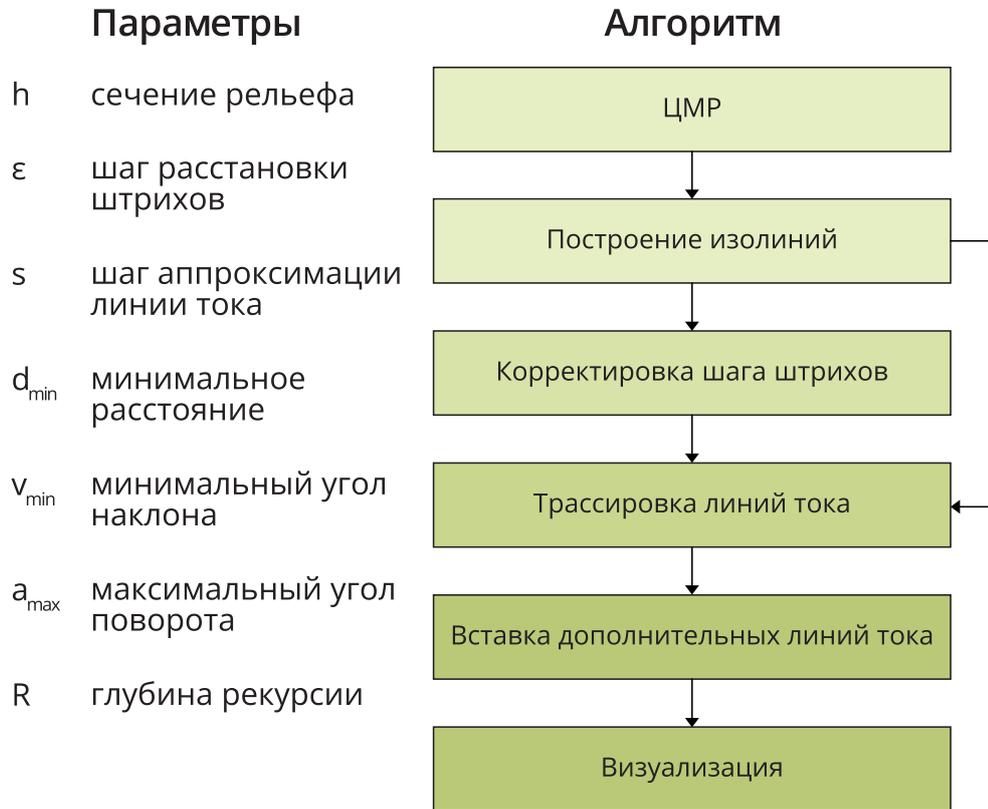


Рисунок 3.54 — Алгоритм построения штрихов и его параметры

### 3.5.1.1 Построение горизонталей

Горизонталы строятся с заданным сечением рельефа  $h$ . ЦМР рассматривается как состоящая из квадратных ячеек с ребрами, соединяющими узлы. На первом этапе отмечаются все ячейки ЦМР и их ребра, пересекающиеся с текущим уровнем горизонтали  $H$  (Рисунок 3.55). Затем трассировка начинается с первой найденной ячейки, пересекаемой текущим уровнем горизонтали, и продолжается в обоих направлениях, пока линия не замкнется или не достигнет границы ЦМР. Используется билинейная функция для вычисления высоты в центре ячейки, чтобы разрешить неоднозначность в случае, когда ячейка пересекается двумя линиями и не ясно, какие точки должны быть соединены (центральная ячейка на Рисунке 3.55).

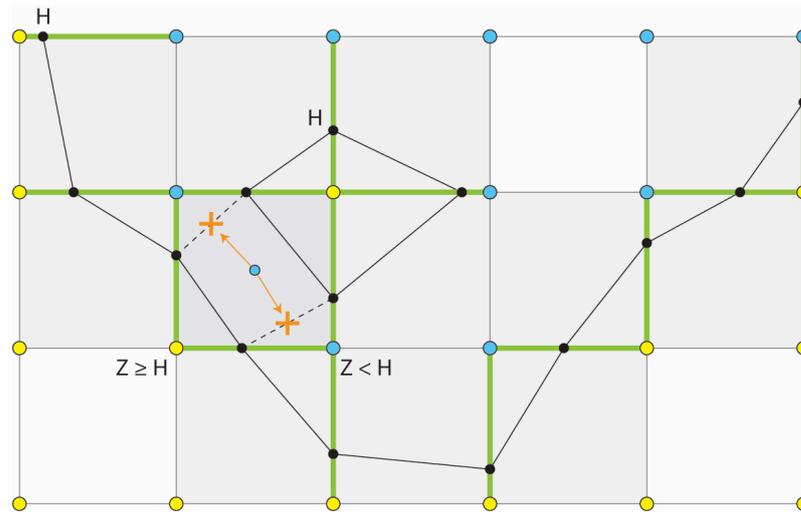


Рисунок 3.55 – Построение горизонталей

### 3.5.1.2 Корректировка интервала штрихов

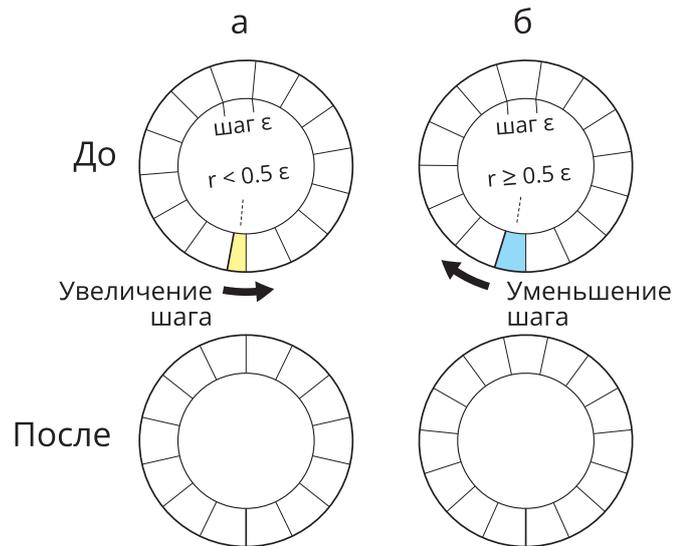
Начальные точки, с которых начинается построение штрихов, распределяются вдоль каждой горизонтали с постоянным расстоянием — шаг штриховки  $\epsilon$ . Значение  $\epsilon$  задается пользователем. Однако оно должно быть настроено для каждой горизонтали индивидуально, так как длина горизонтали обычно не кратна шагу расстановки штрихов. Шаг между последним и первым штрихом (остаток  $r$ ) для замкнутых горизонталей может быть короче, чем у других, что сильно заметно и нарушает единство изображения (Рисунок 3.56). Целесообразно выделить два случая: а)  $r < \epsilon/2$  и б)  $r \geq \epsilon/2$ . Стратегия заключается в минимизации изменений. В случае (а) на Рисунке 3.56 интервал немного увеличивается, для случая (б) — уменьшается. Математически это делается путем округления числа шагов до ближайшего целого числа и последующей подгонки заданного шага под новое число:

$$\epsilon' = \begin{cases} w < 0 : \frac{L}{\lceil n \rceil}, \\ w > 0 : \frac{L}{\lfloor n \rfloor}, \\ w = 0 : \epsilon, \end{cases} \quad w = (\lceil n \rceil - n) - (n - \lfloor n \rfloor), \quad (3.27)$$

где  $\epsilon$  — заданный шаг,  $\epsilon'$  — корректируемый шаг,  $n = l/\epsilon$  — количество шагов, которое укладывается в длину горизонтали (дробное!),  $L$  — длина горизонтали,  $w$  — параметр близости, операторы  $\lfloor \cdot \rfloor$  и  $\lceil \cdot \rceil$  означают наибольшее целое число снизу и сверху соответственно.

### 3.5.1.3 Трассировка линий тока

При построении линий тока алгоритм движется по каждой горизонтали от ее первой вершины в сторону последней в соответствии с скорректированным шагом  $\epsilon'$ . Исполь-



**Рисунок 3.56** – Регулировка шага расстановки штрихов: (а) узкий остаток; (б) широкий остаток

зуются векторный метод трассировки линий тока. Их расчет на основе раstra направлений стока [O’Callaghan, Mark, 1984] довольно груб для построения штрихов, так как линии тока получаются в виде ступенчатых кривых, зависящих от размера ячеек. Аккуратные изгибы линий тока, очерчивающие земную поверхность можно получить только в векторном режиме.

Линия тока  $(x(t), y(t))$ , исходящая вниз по склону из точки  $p_0 = (x_0, y_0)$ , может быть найдена как решение задачи Коши для системы дифференциальных уравнений первого порядка:

$$\begin{cases} x'(t) = -f_x(x, y) \\ y'(t) = -f_y(x, y) \end{cases}, \quad x(0) = x_0, \quad y(0) = y_0, \quad (3.28)$$

где  $t > 0$  – параметр,  $f$  – функция, используемая для моделирования поверхности,  $f_x = \frac{\partial f}{\partial x}$  и  $f_y = \frac{\partial f}{\partial y}$  – ее частные производные, составляющие вектор градиента  $\nabla f = \left\{ \frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right\}$ . Знаки в правой части уравнений меняются на противоположные, если направление трассировки линии тока вверх по склону.

Поле высот в каждой ячейке ЦМР аппроксимируется с помощью билинейной функции для расчета высоты и градиента в произвольной точке:

$$f(x, y) = a_{00} + a_{10}x + a_{01}y + a_{11}xy, \quad (3.29)$$

где  $a_{ij}$  – коэффициенты, определяемые углами ячейки ЦМР, а  $0 \leq x, y \leq 1$  – локальные координаты внутри ячейки, определяемые относительно одного из ее углов. Производные билинейной функции определяются как:

$$\begin{cases} f_x = a_{10} + a_{11}y \\ f_y = a_{01} + a_{11}x. \end{cases} \quad (3.30)$$

Линия тока аппроксимируется отрезками конечной длины  $s$ , которая является одним из параметров алгоритма. Это означает, что каждая следующая точка находится на расстоянии  $s$  от предыдущей. Величина  $s$  должна быть достаточно мала, чтобы обеспечить визуальную гладкость штрихи в масштабе карты, и задается пользователем.

Линия тока вниз по склону трассируется путем итеративного смещения с шагом  $s$  от начальной точки, расположенной на горизонтали, в направлении, *противоположном* градиенту. Трассировка заканчивается, когда выполнится одно из следующих условий:

1. Высота  $z_p$  в текущей точке  $p(x, y)$  меньше или равна высоте нижестоящего уровня горизонталей.
2. Угол наклона  $v_p$  в текущей точке  $p(x, y)$  равен или меньше наименьшего заданного угла наклона  $v_{min}$  (например,  $1^\circ$ ).
3. Расстояние  $d$  до одного из ранее прослеженных штрихов меньше  $d_{min}$ .

Для участков поверхности, которые не могут быть охвачены нисходящими линиями тока (например, привершинных поверхностей), необходимо также трассировать линии тока вверх по склону. Алгоритм их трассировки аналогичен нисходящим линиям тока, за исключением того, что:

- трассировка идет в направлении  $\nabla f$ , а не противоположно ему.
- первое условие окончания трассировки должно быть переформулировано как «высота  $z_p$  в текущей точке  $p(x, y)$  больше или равна высоты вышестоящего уровня горизонталей».

*Нисходящие линии тока имеют больший вес, чем восходящие.* Восходящая линия будет сохранена, если ее трассировка прекращена вторым условием (малый угол наклона). Если же первое условие наступает раньше, то такая линия отклоняется, поскольку достигнут вышестоящий уровень горизонталей, а относительно него будут обязательно построены нисходящие линии тока.

Еще один нюанс построения штрихов заключается в том, что когда линия тока аппроксимируется конечными элементами, она может перескочить через тальвег, и затем начать спускаться вниз по склону, зигзагообразно переходя с борта на борт долины. При этом обычно, если не контролировать такое явление, в эрозионной форме рельефа создается только одна линия тока с подобным поведением, поскольку остальные обрезаются рядом с ней по критерию расстояния (критерий №3). Чтобы предотвратить этот эффект, в процессе трассировки контролируется угол поворота между двумя последовательными отрезками линии тока. Если угол поворота больше заданного порога (например,  $45^\circ$ ), то трассировка линии тока прерывается.

### 3.5.1.4 Вставка дополнительных штрихов на выпуклых склонах

Выпуклые склоны вызывают расходящееся поведение линий потока. Это означает, что соседние линии тока становятся более удаленными друг от друга при движении вниз по склону. Поскольку шаг штриховки должен быть одинаковым по всей карте (пятое правило Имгофа), необходимо вставлять дополнительные штрихи, когда расстояние между существующими штрихами превышает заданный порог (Рисунок 3.57).

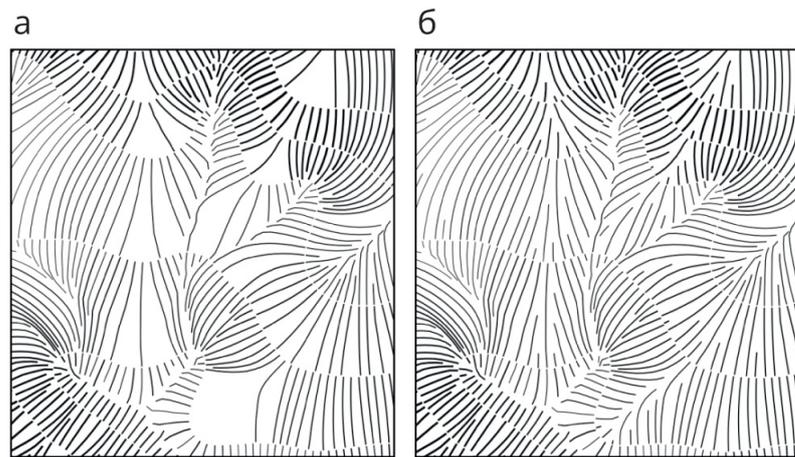


Рисунок 3.57 – Вставка дополнительных штрихов.(а) до вставки, (б) после вставки

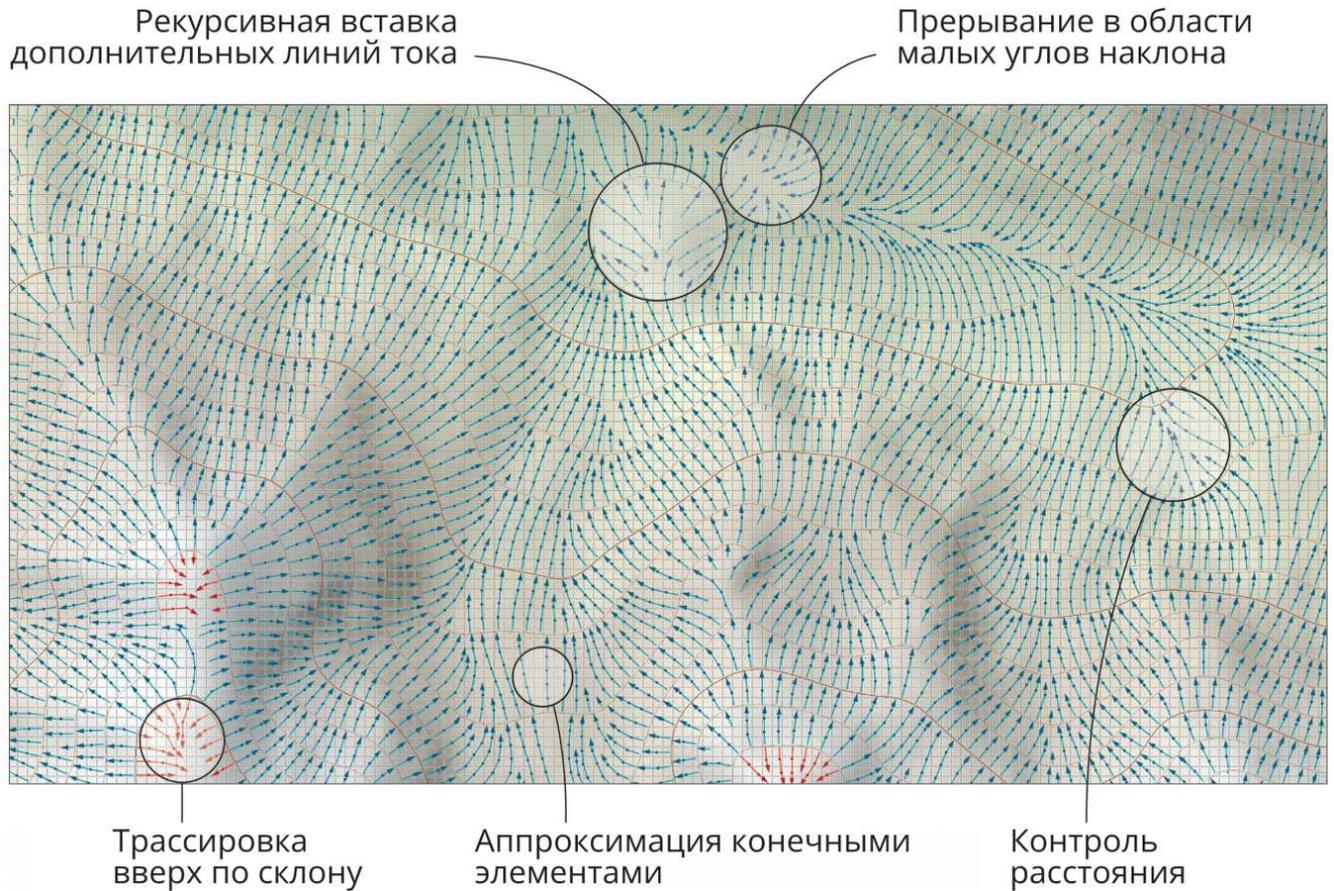
Для этого анализируются все пары соседних штрихов одного уровня. Если расстояние между ними становится равным или больше  $2\epsilon'$  (удвоенный скорректированный шаг), то посередине между ними вставляется и прослеживается вниз по склону дополнительная линия тока. Условия останковки трассировки аналогичны основным штрихам. Операция повторяется рекурсивно для двух новых пар штрихов: левый-средний и правый-средний, пока не отпадет необходимость вставлять дополнительные штрихи (т.е. расстояние между соседними штрихами везде будет меньше  $2\epsilon'$ ). Также задается допустимый уровень рекурсии  $R$ , который ограничивает максимальную глубину вставки штрихов.

Основные особенности нового алгоритма проиллюстрированы на Рисунке 3.58.

## 3.5.2 Результаты

### 3.5.2.1 Реализация

Алгоритм был реализован в виде специализированного программного обеспечения на языке программирования C++. В качестве входного формата данных выбран USGS DEM. Полученные линии штрихов были экспортированы в формат Esri File Geodatabase и визуализированы в ArcGIS Desktop 10.



**Рисунок 3.58** – Основные особенности авторского алгоритма построения штрихов

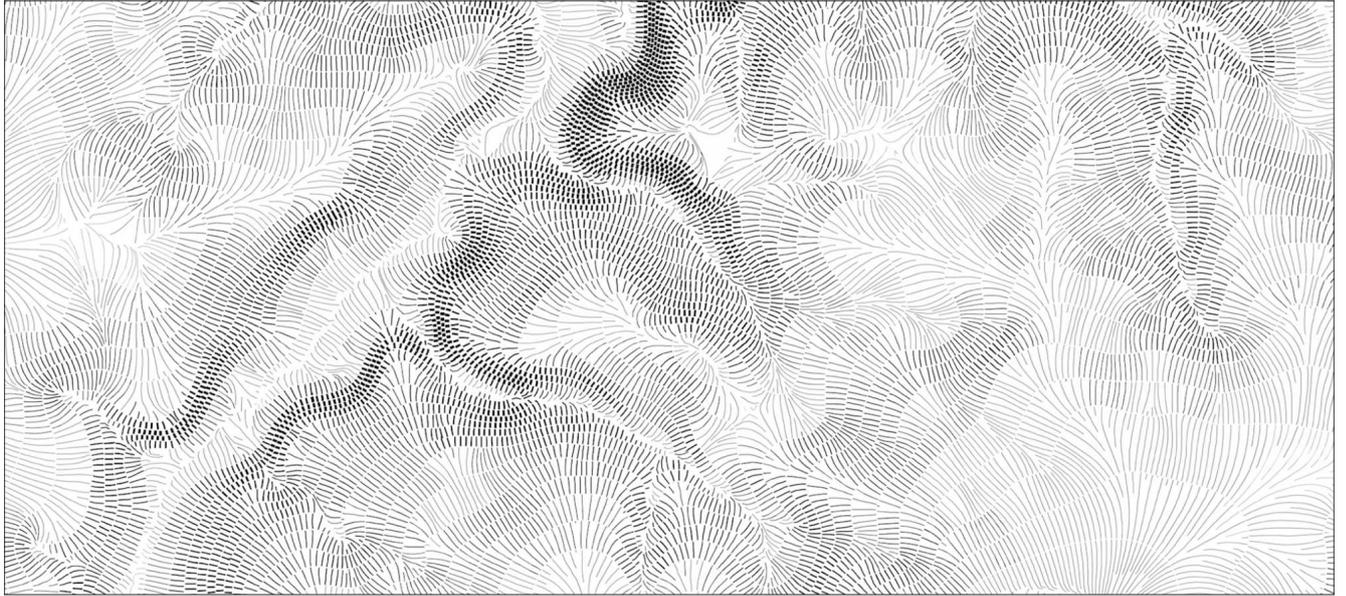
### 3.5.2.2 Штрихи крутизны и теневые штрихи

Первые два примера представляют традиционные штрихи крутизны Рисунок 3.59 и теневые штрихи (Рисунок 3.60), построенные с использованием предложенного подхода. Изображения демонстрируют способность нового алгоритма реализовать все основные требования к построению штрихов: расположение штрихов горизонтальными рядами (горизонталы используются как направляющие), представление в виде линий тока, равномерная плотность штрихов за счет вставки дополнительных линий тока. Ширина штрихов определяется на этапе визуализации в зависимости от угла наклона или освещенности.

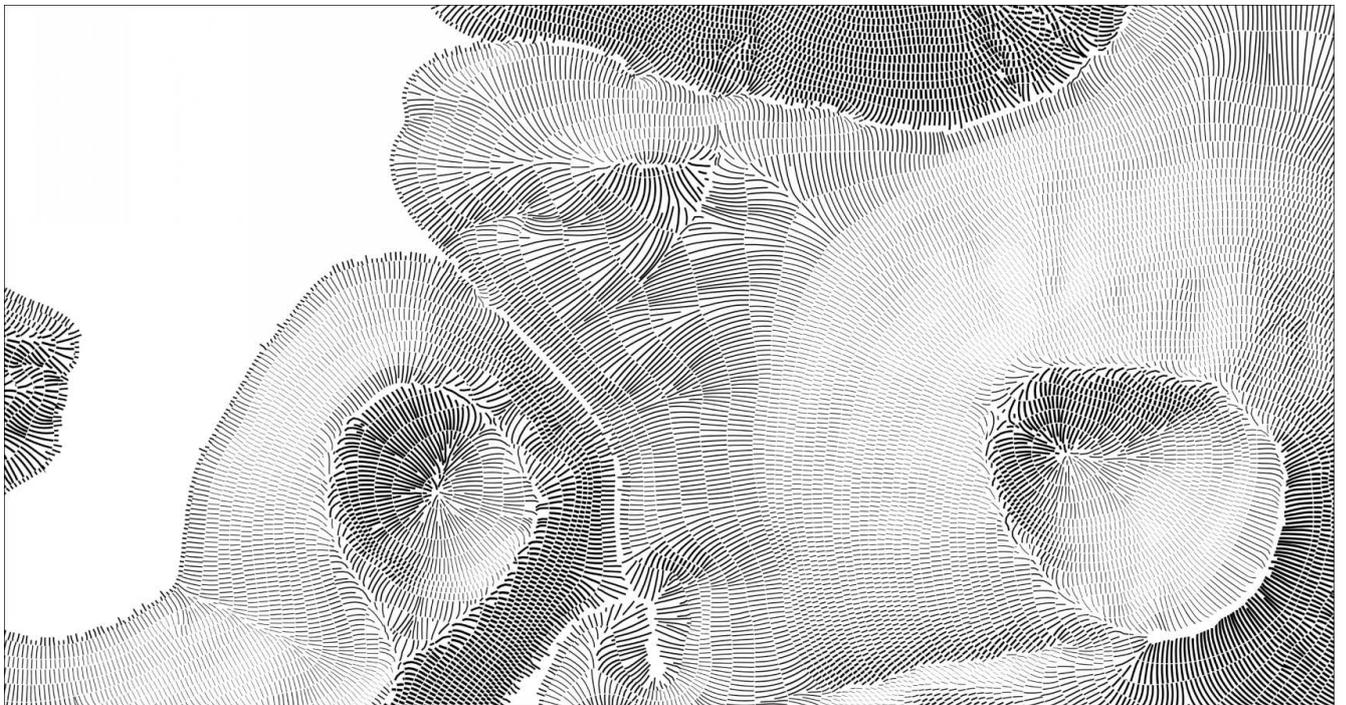
### 3.5.2.3 Детальность изображения

Детальность изображения штрихами определяется тремя ключевыми факторами: сечением рельефа, шагом расстановки штрихов, а также степенью сглаженности ЦМР.

Сечение рельефа для штриховки оптимизируется исходя из масштаба картографирования и крутизны склонов. Оно должен быть достаточно большим, чтобы избежать очень коротких штрихов на участках со значительными углами наклона. Имгоф рекомендует, чтобы длина штриха  $l_{\min}$  составляла не менее 0.2 мм, но предпочтительнее 0.3



**Рисунок 3.59** – Эрозионный рельеф (долина р. Раменки в Москве), изображенный штрихами крутизны с помощью авторского алгоритма



**Рисунок 3.60** – Вулканический рельеф, изображенный теньвыми штрихами с помощью авторского алгоритма

мм [Imhof, 1982]. Однако избыточно крупное сечение тоже может принести проблемы. Верхняя граница оптимальной величины сечения определяется исходя из трех факторов. Первый — ограничить дивергенцию и конвергенцию штрихов в разумных пределах и тем самым избежать широкого использования вставок и прерываний штрихов. Вторая проблема связана с длинными штрихами, которые имеют тенденцию менять свое направление более одного раза в соответствии с кривизной поверхности. При этом чрезмерно короткие штрихи теряют способность выделять направление склона.

Опыт применения разработанного алгоритма показал, что наилучшее представление о форме поверхности дают умеренно изогнутые штрихи средней длины, для которых четко видны ограничивающие их сверху и снизу горизонталы. Такие штрихи представляют собой жесткий и хорошо воспринимаемый каркас поверхности с хорошим акцентом на экспозицию склона. О влиянии сечения рельефа на детальность и упорядоченность изображения можно судить по картографическим изображениям (а) — (в) на Рисунке 3.61. Можно отметить, что изображение (а) слабо отражает экспозицию склона, изображение (в) страдает от длинных волнистых штрихов и чрезмерного использования вставок, а изображение (б) наиболее сбалансированно из трех.

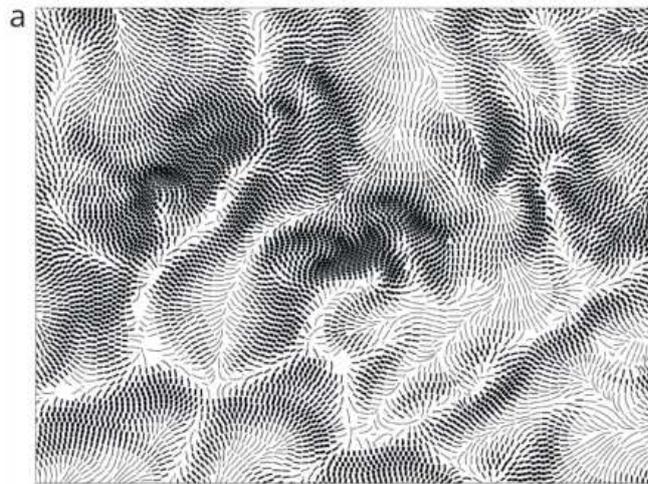
Шаг расстановки штрихов  $\epsilon$ , как и сечение рельефа, значительно влияет на детальность изображения и должен выбираться с учетом характера рельефа и масштаба карты. И здесь так же необходимо найти оптимум. Большой шаг оставляет много пространства для широкой вариации толщины штрихов, обеспечивает легкое восприятие каждого штриха. Однако чрезмерно разреженная штриховка не способна адекватно отобразить все изгибы поверхности, поскольку они теряются в промежутках между штрихами. Имгоф отмечал, что:

*Количество штрихов на сантиметр длины карты во многом определяет ее графический характер. Очень тонкие линии имеют сглаживающий, локализованный, уплотняющий эффект, в то время как толстые линии имеют эффект грубого растра, кажутся более резкими, менее детальными, больше похожими на гравюру на дереве, но при этом четкими и более прозрачными* [Imhof, 1982, стр. 218].

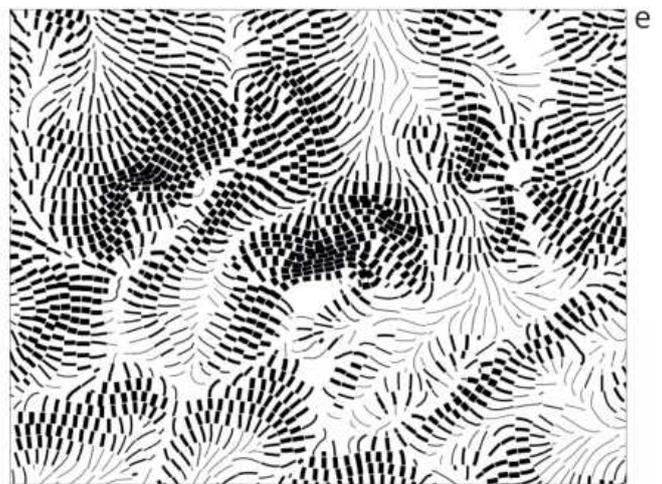
Влияние шага штриховки на детальность и структуру изображения продемонстрировано на изображениях (г), (д) и (е) на Рисунке 3.61. Штрихи на изображении (г) трудно воспринимать по отдельности из-за чрезмерно малого шага, а разнообразие градаций ширины сильно ограничено недостатком свободного пространства. Карта на рисунке (е) выглядит рыхло: индивидуальность каждого штриха преобладает над общим рисунком изображения рельефа из-за слишком большого интервала их расстановки. Наконец, изображение (д) снова хорошо сбалансировано.

Можно ли автоматически рассчитать оптимальное сечение рельефа и шаг штриховки? Хорошей отправной точкой является расчет наибольшего угла наклона  $\nu_{\max}$  по всей ЦМР и сопоставление его с наименьшим допустимым заложением горизонталей  $l_{\min}$ . Тогда, если масштаб карты равен  $1 : M$ , сечение рельефа  $h$  может быть рассчитано как:

## Изменение высоты сечения рельефа



## Изменение шага штрихов



**Рисунок 3.61** — Влияние сечения рельефа (а, б, в) и шага расстановки линий тока (г, д, е) на детальность штриховки.

$$h = M \cdot l_{\min} \cdot \tan \nu_{\max}. \quad (3.31)$$

Чтобы вывести правило оптимального выбора шага штриховки, следует вспомнить, что каждый штрих ориентирован в соответствии с экспозицией склона. Это направление будет визуально восприниматься только если длина штриха больше его толщины. В этом случае главная (самая длинная) ось штриха как полигонального объекта будет совпадать с линией тока. Сохранение этого визуального эффекта может быть обеспечено, если шаг штриховки равен  $l_{\min}$ . В этом случае только на участках с максимальными углами наклона толщина штриха может быть равна его длине (если точнее, штрихи будут касаться). Во всех остальных местах длина штриха будет больше толщины. Таким образом, формула для выбора шага штриховки выглядит следующим образом:

$$\epsilon = M \cdot l_{\min}. \quad (3.32)$$

Вычисления по Формуле 3.31 и Формуле 3.32 не могут гарантировать оптимальную детальность изображения для любой ЦМР. Но с их помощью может быть начат процесс выбора параметров, которые далее адаптируются к особенностям рельефа территории.

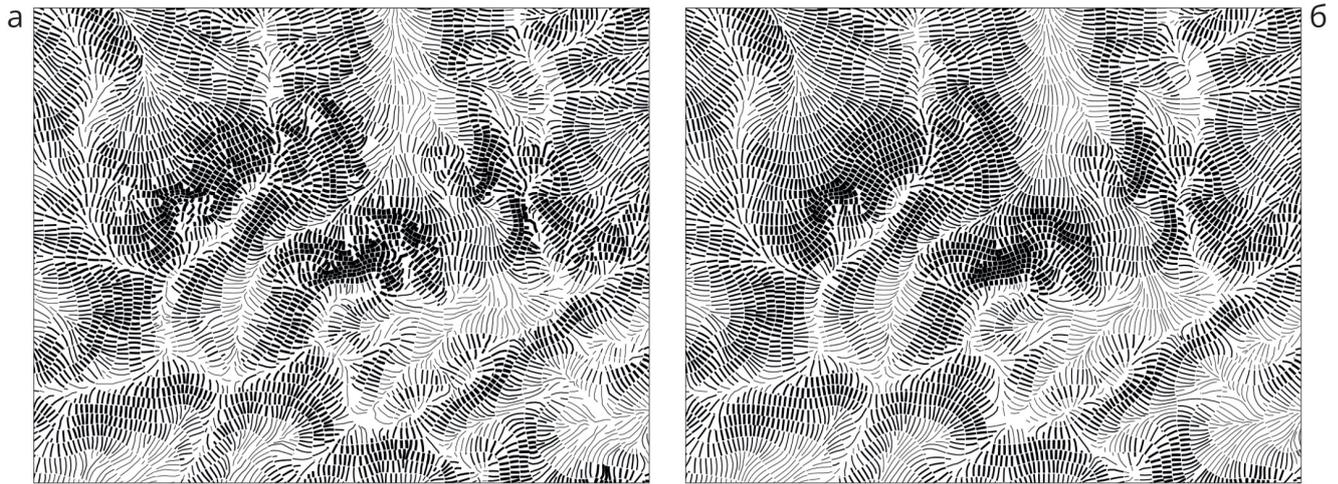
Экспериментальная работа на сильно расчлененном рельефе показывает, что сбалансированная детальность штриховки не может быть обеспечена только путем корректировки сечения рельефа и шага расстановки штрихов. Также необходимо обобщение и сглаживание ЦМР для предотвращения плохой читаемости изображения вследствие его избыточной детальности.

Здесь уместно снова процитировать Имгофа:

*‘...изобилующие элементами сложные формы [рельефа], которые нередко встречаются и могут быть описаны с помощью современных точных съемок, не могут быть выражены с помощью четко определенного, жесткого геометрического рисунка штрихов. Поскольку каждый отдельный штрих должен быть перпендикулярен соответствующему участку горизонтали, очень сложные, плотные участки горизонталей приводят к весьма запутанному наложению штриховых паттернов. Упорядоченное расположение штрихов предполагает весьма простое и уплощенное изображение форм рельефа. Степень генерализации карт со штриховкой рельефа значительно превышает степень генерализации карты в горизонталях того же масштаба’.* [Imhof, 1982, стр. 222].

Таким образом, в общем случае детальность ЦМР для штриховки должна быть меньше, чем для горизонталей или отмывки в том же масштабе. Сглаживание обычно критикуют как метод обобщения рельефа, поскольку оно имеет тенденцию к уменьшению углов наклона и амплитуды высот. Однако экспериментальная работа показывает, что

даже незначительное сглаживание, которое не вносит существенных изменений в распределение высот и углов наклона, устраняет небольшие колебания высот, которые существенно осложняли бы рисунок штриховки. Изображение (а) в левой части на Рисунке 3.62 получено из исходной ЦМР, а (б) — после двух итераций фильтром среднего значения с размером окна  $3 \times 3$ . Видно, что хаотичная компонента направлений штрихов ликвидирована, а сглаженное изображение выглядит гораздо более упорядоченным и аккуратным.

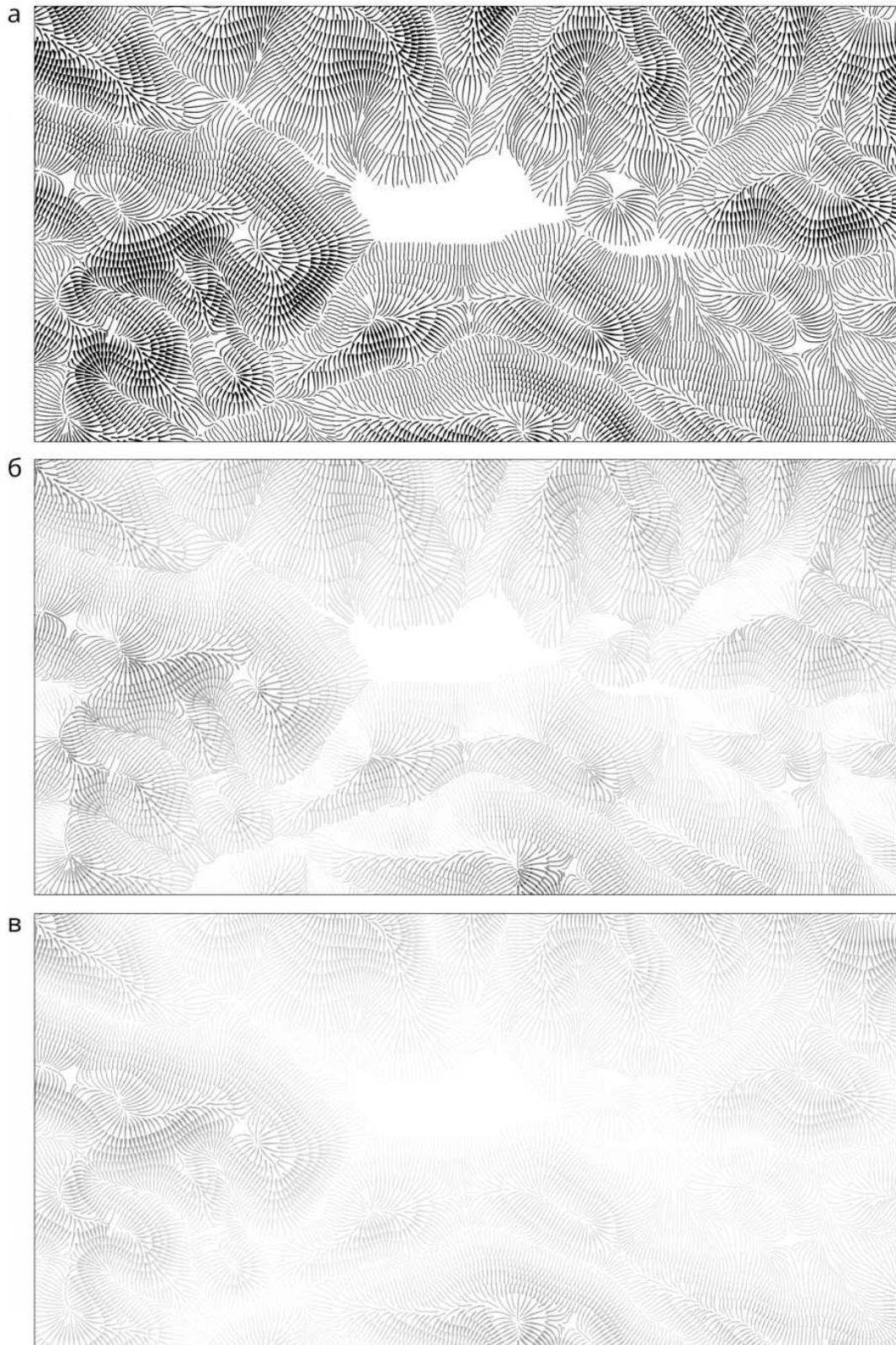


**Рисунок 3.62** — Влияние сглаживания ЦМР на упорядоченность штрихового изображения: (а) исходная ЦМР; (б) ЦМР, обработанная сглаживающим фильтром

При значительном уменьшении масштаба необходимо применять полноценные методы генерализации, и в этом случае распределение высот уже не может быть в точности сохранено. Однако сглаживание склонов — неизбежное следствие генерализации ЦМР. Это означает, что штриховка как строгий метод представления рельефа подходит только для крупных и средних масштабов картографирования.

#### 3.5.2.4 Сужение и атмосферная перспектива

**Сужение** (англ. *tapering*) — графический эффект, который заключается в уменьшении толщины штриха от его начала к его концу. Сужение позволяет сделать штриховое изображение более изящным. Сужение прежде всего необходимо для штрихов, построение которых было прервано по критериям 2 и 3, т.е. по достижению пологого склона или касания с ранее построенным штрихом. **Атмосферная перспектива** — графический эффект, при котором более удаленные объекты выглядят размытыми, менее четкими. Примеры использования данных эффектов показаны на Рисунке 3.63. Как можно видеть, данные эффекты делают изображение более естественным и изящным, а атмосферная перспектива дает также эффект более высокой детальности изображения привершинных поверхностей, и меньшей детальности изображения днщ долин.



**Рисунок 3.63** — Дополнительные эффекты при отображении штрихов: (а) сужение; (б) атмосферная перспектива; (в) комбинация сужения и атмосферной перспективы

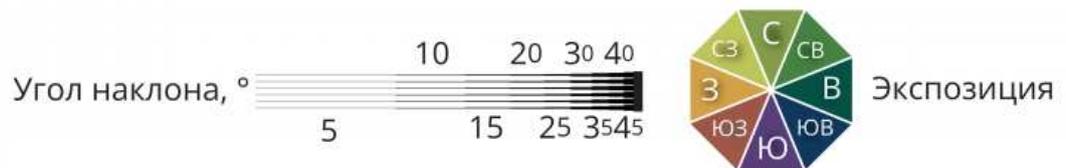
### 3.5.2.5 Штрихи крутизны и экспозиции

Создание карт углов наклона и экспозиций склонов способом количественного фона является общей практикой в геоморфометрии. Эти характеристики могут быть объединены на одном изображении. В этом случае за экспозицию отвечает цветовой тон и его светлота, а увеличение насыщенности применяется при отображении классов углов наклона от малых к большим. Цветовой контраст между более светлыми и теплыми северо-западными и более темными и холодными юго-восточными классами экспозиции позволяет ввести выразительный трехмерный эффект в изображение [Brewer, Marlow, 1993].

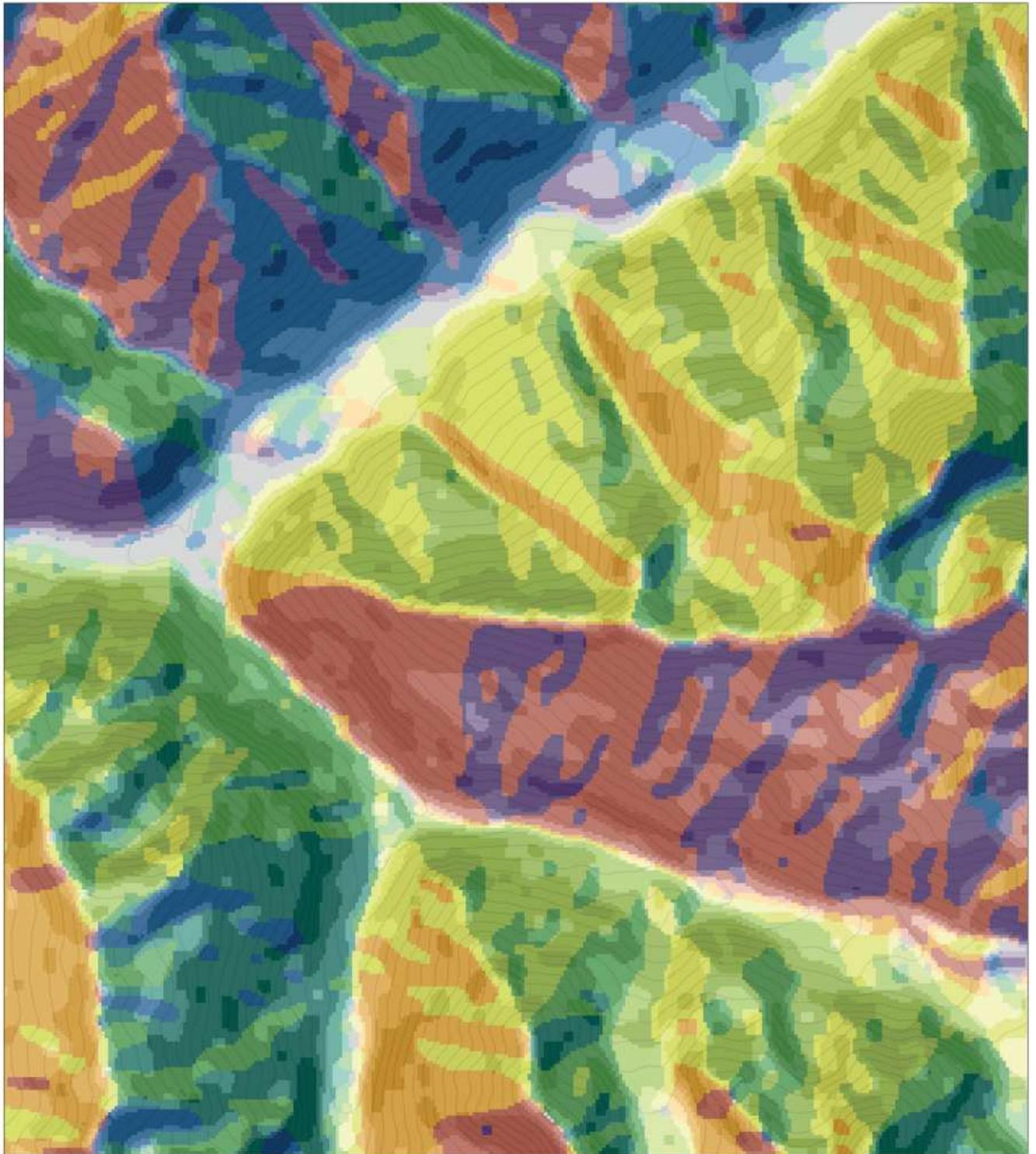
На Рисунке 3.64 представлен экспериментальный метод штрихов крутизны и экспозиции, предназначенный для морфометрического картографирования. Углы наклона кодируются толщиной штриха, а экспозиция — цветом. Углы наклона классифицированы с равным интервалом в  $5^\circ$ , а экспозиция разделена на 8 классов по сторонам света. Для сравнения на Рисунке 3.65 представлена карта крутизны и экспозиции склонов на ту же территорию, построенная традиционным методом качественно-количественного фона. Здесь применение той же классификации, что и для штрихов (8 экспозиций  $\times$  10 градаций углов наклона), привело бы к 80 цветам, которые практически невозможно различить. По этой причине пришлось сократить количество градаций углов наклона для площадного способа до 5 категорий: 0 — 10, 10 — 20, 20 — 30, 30 — 40 и  $> 40^\circ$ . Таким образом на этом этапе выявлено одно из ограничений традиционных карт углов наклона и экспозиций склона.

Вместе с тем, было необходимо разработать единую цветовую шкалу для двух видов карт, чтобы сделать их сравнение более объективным. Оригинальная цветовая шкала Brewer, Marlow [1993] оказалась слишком яркой для штрихов, которые становятся едва различимыми на белом фоне. Для решения проблемы была разработана более темная цветовая шкала, которая использует аналогичные принципы изменения оттенка, светлоты и насыщенности. Цвета, представляющие наибольшие углы наклона на традиционной карте, были использованы для штриховки. Необходимая для классификации штрихов атрибутивная информация о средней длине и экспозиции склона вдоль штриха была извлечена из ЦМР.

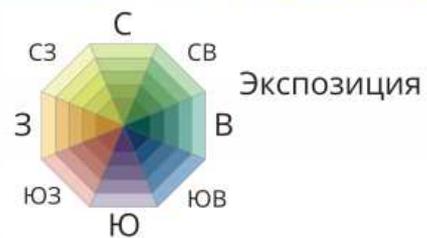
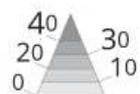
Анализ изображений, представленных на Рисунке 3.64 и Рисунке 3.65, выявляет одно существенное различие между двумя методами: фон дает строго зональную характеристику поверхности, в то время как штрихи дают сочетание зональной и локальной характеристики. Экспозиции на карте фона легче соотносить с градациями легенды, чем на карте штрихов, поскольку участки одной и той же экспозиции имеют фиксированный цвет. Это также дает возможность более четко выделять склоны разных экспозиций. Но вместе с тем, на карте штрихов экспозиция не только кодируется цветом, но и подчеркивается направлением линии тока. Таким образом, штриховка позволяет более точно



**Рисунок 3.64** – Штрихи крутизны и экспозиции, полученные с помощью авторского алгоритма. Долина ледника Джанкуат, Северный Кавказ.



Плоские участки  
показаны серым цветом  
Угол наклона, °



**Рисунок 3.65** — Карта крутизны и экспозиции склонов методом качественно-количественного фона. Долина ледника Джанкуат, Северный Кавказ.

оценивать локальные особенности экспозиции склонов. Она также отображает распределение градиента поверхности, благодаря чему пользователь карты может визуально проследить линии тока, глядя на изображение.

С углами наклона ситуация аналогичная: фон дает более удобное изображение для определения областей с заданным интервалом углов наклона, а штриховка дает больше информации о локальных особенностях за счет индивидуальной длины каждого штрих. Более того, штриховка позволяет выделить в два раза больше классов по углам наклона, чем фон. В отличие от экспозиций соотношение длин штрихов с легендой углов наклона выполняется более просто: достаточно найти в ней штрих аналогичной длины. Для ряда пользователей, особенно с нарушением цветового зрения, определение угла склона по длине может быть гораздо проще, чем по цвету.

### 3.5.3 Итоги

Штрихи представляют собой уникальный способ изображения рельефа, который комбинирует в себе свойства светотеневых и морфометрических подходов. Новый алгоритм, представленный в данном разделе диссертации, реализует все традиционные требования к построению штрихов. Он преодолевает основные проблемы предыдущих алгоритмов, такие как использование прямых отрезков, недостаточный учет выпуклых склонов и привершинных поверхностей.

Особое внимание уделено управлению детальностью штриховки. В процессе методических разработок и экспериментов удалось наглядно продемонстрировать как сечение рельефа, шаг расстановки штрихов и степень сглаженности ЦМР влияют на детальность изображения и могут быть подобраны для ее оптимизации. Разработанный подход позволяет получать высококачественные штрихи, близкие по внешнему виду к ручной работе.

Предложенный метод штрихов крутизны и экспозиции открывает новые возможности для применения штрихов в морфометрическом картографировании. Представляется перспективным применение штрихов для отображения геофизических векторных полей при этом они могут быть стилизованы стрелками, чтобы подчеркнуть их направление. Важным направлением применения штрихов является также археологическое и историческое картографирование, где этот способ применяется для стилизации изображения.

Основным недостатком штриховки является ее графическая сложность, которая ограничивает возможность комбинирования штрихов с другими слоями. Штриховка, однако, не заменяет способ качественно-количественного фона, но дает другую, более локализованную, характеристику поверхности. Оба метода можно использовать в дополнение друг к другу, чтобы получить наиболее полное представление поверхности.

Наш опыт работы со штрихами показал, что текущие результаты могут быть улуч-

шены за счет дальнейшего развития в следующих направлениях:

1. **Выравнивание длины** штрихов. В текущей версии алгоритма контролируется только расстояние между штрихами. Однако длина соседних штрихов на пологих склонах может резко отличаться, так как положение конечной точки зависит от локального угла наклона в конкретной ячейке сетки, куда пришла линия тока в процессе трассировки. Длину соседних штрихов следует выравнивать, обрезая или удлиняя некоторые из них в зависимости от ситуации.
2. **Разработка дополнительного модуля** для QGIS, ArcGIS Desktop или другого популярного программного обеспечения ГИС, чтобы сделать использование штриховки более распространенным, простым и удобным.

Было бы также полезно опробовать штриховку в различных исследованиях, связанных с визуализацией и анализом поверхностей.

### 3.6 Конфляция гидрографических данных и цифровых моделей рельефа разной детальности<sup>7</sup>

**Интеграция** разнородных пространственных данных предполагает их совмещение в пространстве для проведения дальнейшего совместного анализа и картографирования. При этом задача усложняется в случае, когда данные обладают разной детальностью и обнаружение идентичных объектов становится затруднительным.

*Конфляция* как случай интеграции связана с высокоточным (пообъектным) совмещением данных, полученных из разных источников [Saalfeld, 1988; de Smith, Goodchild, Longley, 2018]. В настоящем исследовании впервые решена проблема автоматизированной конфляции данных о рельефе и гидрографии [Samsonov, 2022]. Линии водотоков упорядочиваются согласно методу Хака [Hack, 1957] (аналогичный каталогам рек СССР), где основной водоток от истока до устья имеет порядок 1, его непосредственные притоки имеют порядок 2 и т.д. В пределах заданного расстояния от каждого водотока ищется водоток-контрагент. Генерируются связи конфляции между контрагентами и соответствующими точками водотоков. Далее осуществляется смещение векторных объектов (узлов или структурных линий), представляющих ЦМР вдоль связей конфляции, по результатам которого линии контрагентов совмещаются с водотоками, а окружающие их объекты смещаются пропорциональным образом. По высотам смещенных объектов восстанавливается триангулированная поверхность, которая далее растеризуется на ту же сетку, что и исходная ЦМР.

<sup>7</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Samsonov, 2020b]

### 3.6.1 Методика

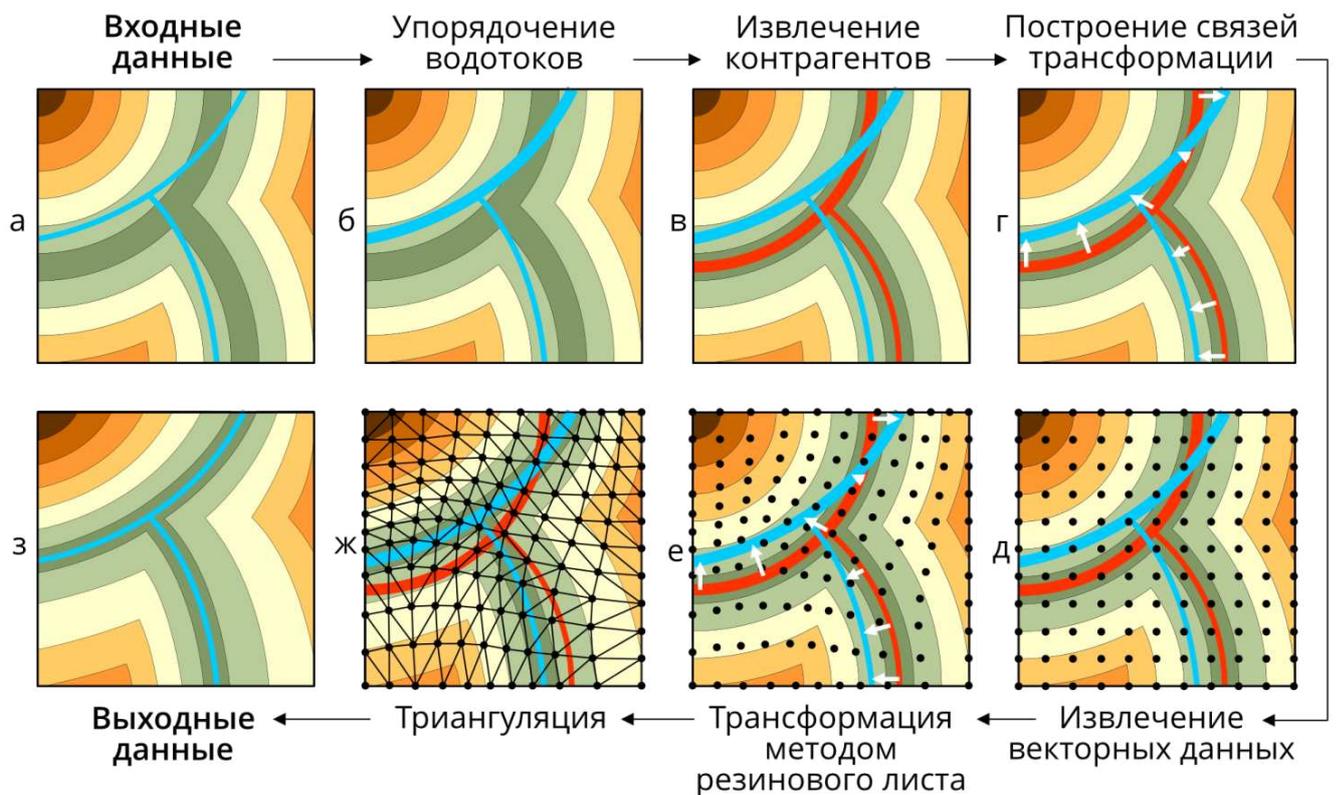
#### 3.6.1.1 Общие принципы

Разработана методика, которая позволяет совмещать ЦМР с опорными линиями гидрографии. Методика основана на следующих принципах:

1. Модель земной поверхности, представленная в ЦМР, в результате обработки должна быть пространственно совмещена с опорными линиями гидрографии (водотоками). Детальность ЦМР и линий гидрографии может быть произвольной (в том числе различной).
2. Конфляция выполняется путем непрерывного смещения местоположений, в которых известна отметка высоты. Новые формы рельефа не создаются на поверхности и не удаляются из нее. Опорные линии гидрографии по умолчанию имеют первостепенное значение в процессе совмещения и поэтому остаются на своих местах, однако процесс при необходимости может быть инвертирован.
3. Местоположения с отметками высот должны быть представлены векторными объектами (точками, линиями, полигонами). Как растровые, так и триангуляционные ЦМР могут быть представлены в виде набора векторных объектов (узлы и ребра интерполяционной сетки) без потери данных. Кроме того, объекты могут представлять собой структурный каркас поверхности, извлеченный из исходной ЦМР в процессе генерализации. Таким образом, векторное представление абстрагирует формат ввода данных о рельефе и может использоваться при различных сценариях конфляции — совместно с генерализацией ЦМР или без нее.
4. Трансформация высотных данных осуществляется путем смещения координат отметок высот согласно векторам (связям конфляции), направленным к опорным линиям гидрографии и исходящим из наиболее схожих траекторий на поверхности ЦМР — *водотоков-контрагентов* (или просто *контрагентов*). Каждая опорная линия гидрографии должна иметь своего контрагента.
5. Водотоки-контрагенты извлекаются из исходной ЦМР автоматизированным путем и должны представлять собой топологически корректную сеть, аналогичную сети опорных линий гидрографии. Метод выделения контрагентов должен быть устойчив к ошибкам в ЦМР и гидрографических линиях (таким как искусственные замкнутые локальные понижения, неправильные направления линий гидрографии), а также к нестандартным конфигурациям водотоков (разветвленное русло, дельты, каналы).
6. Результирующая ЦМР восстанавливается по трансформированным высотным данным.

Разработанный метод конфляции состоит из семи этапов, которые показаны на Рисунке 3.66:

1. Порядковая классификация опорных гидрографических линий (Рисунок 3.66б).
2. Трассировка контрагентов (Рисунок 3.66в).
3. Генерация связей конфляции (Рисунок 3.66г).
4. Извлечение высотных данных в виде векторных объектов (Рисунок 3.66д).
5. Смещение высотных данных по связям конфляции методом резинового листа (Рисунок 3.66е).
6. Создание триангуляционной ЦМР (TIN) из смещенных высотных данных (Рисунок 3.66ж).
7. Реконструкция согласованной с гидрографией ЦМР на основе триангуляционной модели (Рисунок 3.66з).



**Рисунок 3.66** – Этапы конфляции: (а) входная ЦМР и линии гидрографии, (б) порядковая классификация линий гидрографии, (в) выделение водотоков-контрагентов, (г) генерация связей конфляции, (д) извлечение высотных данных, (е) трансформация высотных данных, (ж) триангуляция трансформированных высотных данных, (з) восстановление результирующей ЦМР. Опорные гидрографические линии показаны светло-голубым цветом. Контрагенты показаны красным цветом.

В следующих разделах подробно описан каждый этап метода.

### 3.6.1.2 Порядковая классификация водотоков

Порядковая классификация — этап предварительной обработки, который выполняется для выявления существующих топологических связей в системе гидрографических линий и её организации в иерархическую структуру, минимизирующую количество элементов гидрографической сети и устанавливающую соподчиненность между ними. Классификация необходима для обеспечения однозначной последовательности, в которой будут трассироваться водотоки-контрагенты, а также для воспроизведения топологических отношений, обнаруженных в эталонных линиях, в виде аналогичных отношений в сети контрагентов.

Для установления иерархии водотоков был применен метод Хака [Hack, 1957], также известный как *естественное упорядочение водотоков*. Согласно этой процедуре, самый длинный путь вверх по течению от каждого замыкающего створа в речной сети получает порядок 1. Затем длиннейшие притоки водотоков порядка 1 прослеживаются вверх по течению, и полученным линиям присваивается порядок 2. Процесс продолжается до тех пор, пока всем гидрографическим линиям не будут присвоены порядки. Дополнительная информация, такая как площадь водосборного бассейна (или аккумуляция стока), может быть использована для взвешивания длины водотоков и повышения приоритетности рек с большими бассейнами. Основное преимущество метода упорядочивания Хака заключается в том, что оно устанавливает субординацию между водотоками в сети. Кроме того, он минимизирует количество водотоков с точки зрения их сетевой модели, что уменьшает количество контрагентов, которые необходимо найти.

Стандартная схема упорядочивания Хака нуждается в уточнении для участков реки с разветвленным руслом, поскольку все протоки внутри разветвленного участка будут иметь одинаковый порядок. В нашем случае самый длинный проток, проходящий через разветвленный участок, должен сохранять свой порядок, а его притоки будут классифицироваться так, если бы они не вытекали из основного протока. Таким образом, прямые притоки основного протока должны получать тот же порядок, увеличенный на единицу, и процесс будет продолжаться рекурсивно для притоков притоков, пока не будет обработан весь разветвленный участок. Уточненная таким образом схема порядковой классификации называется в данной работе *модифицированной схемой (методом) Хака*.

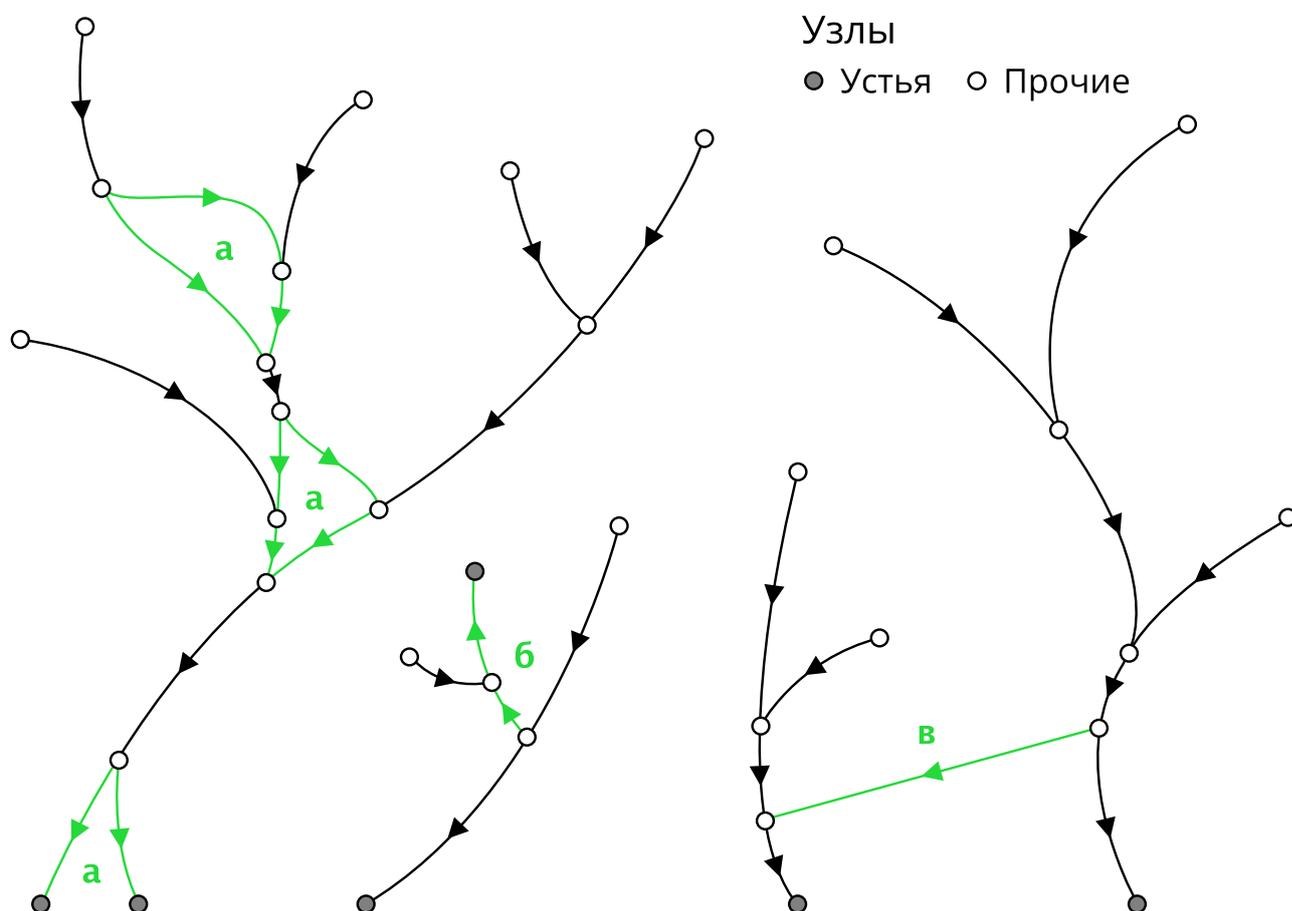
Порядковая классификация линий гидрографии требует знания их исходных топологических отношений. Поэтому перед применением модифицированной схемы Хака необходимо провести анализ исходных линий. Для получения желаемого результата применяется следующая последовательность действий:

1. Разрезать гидрографические линии в точках пересечения.
2. Определить для полученных линий соседей выше и ниже по течению.
3. Выявить устьевые участки (линии, у которых нет соседей ниже по течению).
4. Классифицировать все водотоки согласно модифицированной схеме Хака, начи-

ная с устьевых участков.

5. Представить топологическую структуру полученной сети в табличной форме.

Результат первого этапа показан на Рисунке 3.67. На этом рисунке выделены три особые конфигурации водотоков. На *разветвленное русло и дельты (A)*, а также *каналы (C)* при поиске контрагентов будет наложено дополнительное ограничение на сохранение бифуркации. Некорректно оцифрованные водотоки (*B*) рассматриваются как есть. Алгоритм предварительной обработки не пытается выявить и исправить такие случаи, поскольку надежно это сделать нельзя. В частности, несоответствие между гидрографией и ЦМР в силу различий в детальности не позволяет рассчитывать на высотную информацию для определения истинного направления водотока. Тем не менее, как будет показано далее, это не снижает качество совмещения, поскольку стратегия наименьшего стоимостного пути может быть эффективно использована для отслеживания контрагентов вне зависимости от корректности оцифровки эталонных водотоков.



**Рисунок 3.67** — Линии гидрографии, разделенные узлами в точках пересечений. Нестандартные конфигурации: (а) разветвленное русло и дельты, (б) некорректное направление оцифровки, (в) канал, соединяющий два естественных водотока.

Гидрографическая сеть, в которой установлены порядки согласно предложенной схеме, представлена на Рисунке 3.68. Эта сеть также отражена в Таблице 3.8, которая обобщает топологические отношения между участками сети и определяет последо-

вательность, в которой должны быть прослежены их контрагенты. Таблица содержит шесть переменных:

- ID (уникальный идентификатор водотока);
- CONFL (ID водотока, в который впадает данный водоток);
- BIFUR (ID водотока, из которого вытекает данный водотока);
- ITER (номер итерации, на которой должен быть извлечен контрагент для данного водотока);
- ORDER (порядок согласно модифицированной схеме Хака);
- TYPE (тип водотока по отношению к бифуркационному процессу).

Первые три переменные (ID, CONFL и BIFUR) вычисляются в процессе порядковой классификации. Значения CONFL и BIFUR задают топологические отношения между водотоками, которые в свою очередь определяют четыре типа узлов сети, указанные в легенде Рисунка 3.68:

- *устье* (конечный узел водотока с  $\text{CONFL} = -1$ );
- *исток* (начальный узел водотока с  $\text{BIFUR} = -1$ );
- *слияние* (конечный узел водотока с  $\text{CONFL} \neq -1$ );
- *разветвление* (начальный узел водотока с  $\text{BIFUR} \neq -1$ ),

где  $-1$  означает отсутствие такого водотока.

Также существует два типа водотоков по отношению к процессу разветвления (многорукости) русла:

- *основной рукав* (водоток с  $\text{BIFUR} = -1$ );
- *второстепенные рукава* (водотоки с  $\text{BIFUR} \neq -1$ ).

Эти типы закодированы в Таблице 3.8 в переменной TYPE, а также в легенде Рисунка 3.68.

Неотрицательные значения переменных CONFL и BIFUR устанавливают зависимости в сети гидрографических линий. В частности, если  $\text{CONFL} = j$  или  $\text{BIFUR} = k$  для некоторого водотока с  $\text{ID} = i$ , то водоток  $i$  считается *подчиненным* к  $j$  и  $k$ , а значит, расположение его начальной и конечной точки привязано к этим водотокам. Следовательно,  $j$  и  $k$  считаются *суперординатными* (т.е. имеющими более высокий порядок) по отношению к  $i$ . Сохранение топологических отношений в процессе конфляции требует, чтобы контрагенты суперординатных водотоков были выделены ранее субординатных.

Переменная ITER в Таблице 3.8 устанавливает последовательность, в которой должен быть организован этот процесс. Алгоритм ее генерации начинается с установки  $\text{ITER} = 1$  для водотоков с  $\text{CONFL} = -1$  и  $\text{BIFUR} = -1$ . Затем соответствующие водотоки исключаются из списка и начинаются итерации, начиная с  $i = 2$  в качестве переменной цикла. На каждой  $i$ -й итерации  $\text{ITER} = i$  устанавливается для водотоков с  $\text{CONFL} \notin \text{ID}$  и  $\text{BIFUR} \notin \text{ID}$ . Соответствующие водотоки исключаются из списка, и итерации продолжаются с  $i = i + 1$ , пока не будут обработаны все водотоки.



**Рисунок 3.68** — Линии водотоков, классифицированный в соответствии с модифицированным методом Хака

**Таблица 3.8** — Атрибуты водотоков по результатам порядковой классификации модифицированным методом Хака

ID	CONFL	BIFUR	ITER	ORDER	ТИП
1	-1	-1	1	1	Основной
2	-1	1	2	1	Второстепенный
3	1	-1	2	2	Основной
4	1	-1	2	2	Основной
5	1	-1	2	2	Основной
6	1	1	2	2	Второстепенный
7	4	-1	3	3	Основной
8	3	1	3	3	Второстепенный
9	-1	-1	1	1	Основной
10	-1	9	2	1	Второстепенный
11	10	-1	3	2	Основной
12	-1	-1	1	1	Основной
13	-1	12	2	1	Второстепенный
14	12	-1	2	2	Основной
15	12	-1	2	2	Основной
16	12	-1	2	2	Основной
17	14	-1	3	3	Основной

Как видно из Таблицы 3.8, количество итераций не обязательно равно модифицированному порядку Хака. Более того, оно не зависит от типа водотока. При этом порядок обработки водотоков с одинаковым ITER на каждой итерации не имеет значения, так как эти водотоки независимы друг от друга. Определив последовательность итераций, можем перейти к основному этапу разработанной методики — выделению водотоков-контрагентов.

### 3.6.1.3 Выделение водотоков-контрагентов

Термин **водоток-контрагент** используется в данной работе для обозначения пути на поверхности ЦМР, наиболее похожего на заданную линию гидрографии. Поскольку сходство является общим понятием, необходимо определить его таким образом, чтобы его можно было оценить автоматически. Это делается в терминах расстояний, определенных в следующем параграфе. Далее для краткости в данном разделе исходные линии гидрографии будут называться просто **водотоками**, а соответствующие им водотоки-контрагенты — **контрагентами**. Таким образом, для каждого водотока ищется его контрагент.

#### Расстояния

Ранее в Разделе 2.3.4 была введена метрика модифицированного расстояния Хаусдорфа. В рамках настоящего раздела необходимо ввести более широкую систему метрик, начиная с базового расстояния Хаусдорфа, поскольку далее они будут использоваться в качестве критериев точности извлечения контрагентов.

Если даны два подмножества  $\mathcal{A} = \{a\}$  и  $\mathcal{B} = \{b\}$  метрического пространства  $M$  с заданным расстоянием  $d$ , то расстояние **Хаусдорфа**  $d_H(\mathcal{A}, \mathcal{B})$  между ними вычисляется как [Hausdorff, 1914]:

$$d_H(\mathcal{A}, \mathcal{B}) = \max \left\{ \sup_{a \in \mathcal{A}} \inf_{b \in \mathcal{B}} d(a, b), \sup_{b \in \mathcal{B}} \inf_{a \in \mathcal{A}} d(a, b) \right\}. \quad (3.33)$$

При оценке  $d_H$  в двумерном пространстве используется только стандартное евклидово расстояние  $d(a, b) = \sqrt{(x_a - x_b)^2 + (y_a - y_b)^2}$ .

На практике, когда  $\mathcal{A}$  и  $\mathcal{B}$  являются конечными множествами, вычисляется дискретная версия  $d_H$ . Для каждой точки  $a$  в  $\mathcal{A}$  находится ближайшая точка  $b_a$  в  $\mathcal{B}$  и  $\vec{d}_H(\mathcal{A}, \mathcal{B}) = \max_a d(a, b_a)$  принимается за приближение  $\sup_{a \in \mathcal{A}} \inf_{b \in \mathcal{B}} d(a, b)$ . Аналогично,  $\vec{d}_H(\mathcal{B}, \mathcal{A}) = \max_b d(b, a_b)$  вычисляется как аппроксимация  $\sup_{b \in \mathcal{B}} \inf_{a \in \mathcal{A}} d(a, b)$ . Наконец, дискретное расстояние Хаусдорфа определяется как:

$$d_H(\mathcal{A}, \mathcal{B}) = \max \left\{ \vec{d}_H(\mathcal{A}, \mathcal{B}), \vec{d}_H(\mathcal{B}, \mathcal{A}) \right\}. \quad (3.34)$$

Расстояние  $\vec{d}_H$  называется **направленным расстоянием Хаусдорфа**. В общем случае  $\vec{d}_H(\mathcal{A}, \mathcal{B}) \neq \vec{d}_H(\mathcal{B}, \mathcal{A})$ . Поскольку и  $d_H$ , и  $\vec{d}_H$  являются минимаксными мерами, они чувствительны к выбросам из множества. Это свойство полезно для ограничения максимального отклонения контрагента и линии гидрографии друг от друга.

Введенное в [Dubuisson, Jain, 1994] для сопоставления изображений **Модифицированное расстояние Хаусдорфа**  $d_M$  основано на аналогичной идее, но заменяет максимум кратчайших расстояний их средним значением:

$$d_M = \max \left\{ \frac{1}{|\mathcal{A}|} \sum_a d(a, b_a), \frac{1}{|\mathcal{B}|} \sum_b d(b, a_b) \right\}, \quad (3.35)$$

где  $|\mathcal{A}|$  и  $|\mathcal{B}|$  означает количество элементов в  $\mathcal{A}$  и  $\mathcal{B}$  соответственно. По сравнению с  $d_H$  и  $\vec{d}_H$ ,  $d_M$  нечувствительно к выбросам и дает представление о средней близости между линиями. Это свойство полезно для выбора контрагента среди множества кандидатов.

Три расстояния, описанные выше, однако, не являются лучшими мерами близости между двумя путями на поверхности, поскольку они не учитывают непрерывность и направленность линий. Более сильной мерой близости для сравнения двух линейных объектов является расстояние Фреше, определяемое как [Fréchet, 1906]:

$$d_F(\mathcal{A}, \mathcal{B}) = \inf_{\alpha, \beta} \max_{t \in [0, 1]} \left\{ d(\mathcal{A}(\alpha(t)), \mathcal{B}(\beta(t))) \right\}, \quad (3.36)$$

где  $\mathcal{A}(u)$ ,  $\mathcal{B}(v)$ ,  $u, v \in [0, 1]$  — *параметрические описания* линий, а  $\alpha(t)$ ,  $\beta(t)$ ,  $t \in [0, 1]$  — так называемые *репараметризации*, которые оптимизируются для нахождения таких  $u = \alpha(t)$  и  $v = \beta(t)$ , которые минимизируют максимальное расстояние между  $\mathcal{A}(u)$  и  $\mathcal{B}(v)$  для общего  $t$ . Параметрическое описание подразумевает, что  $\mathcal{A}(0)$  — начальная точка  $\mathcal{A}$ ,  $\mathcal{A}(1)$  — конечная точка  $\mathcal{A}$ , а  $\mathcal{A}(u)$  — некоторая промежуточная точка на кривой, которая непрерывно движется вдоль  $\mathcal{A}$  от своего начала до конца, когда  $u$  меняется с 0 на 1.

Расстояние Фреше обычно объясняется как минимальная длина поводка, необходимая человеку, чтобы выгулять собаку, в то время как оба они идут своими путями, и эти пути должны быть пройдены от начала до конца (Рисунок 3.69).

Если дискретное и модифицированное расстояния Хаусдорфа можно вычислить напрямую, то вычисление расстояния Фреше значительно сложнее. Для получения  $d_F$  и его дискретной версии (которая применима к пространственным объектам) можно использовать методы, разработанные в [Alt, Godau, 1995] и [Eiter, Mannila, 1994]. Алгоритм из последней работы используется в настоящем исследовании.

Определив расстояния, перейдем к формализации условий, которые могут быть использованы для трассировки контрагента одного водотока. После того как эти условия

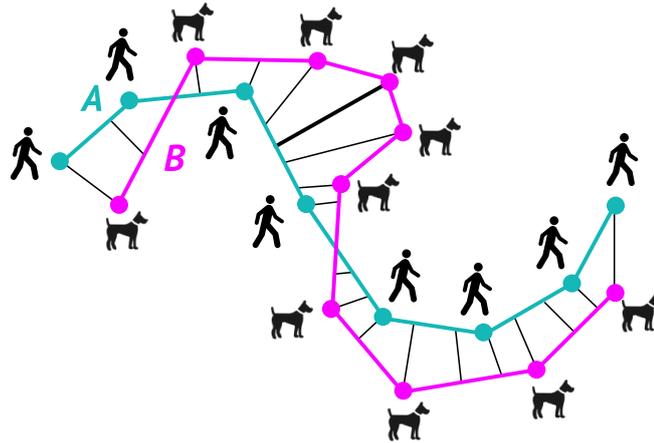


Рисунок 3.69 – Расстояние Фреше (выделено толстой черной линией)

будут определены, обсудим, как они должны быть расширены в случае нескольких гидрографических линий, составляющих единую сеть.

### Одна линия гидрографии

**Контрагент-кандидат**  $s$  для водотока  $h$  — это путь на поверхности ЦМР, который удовлетворяет следующим трем условиям:

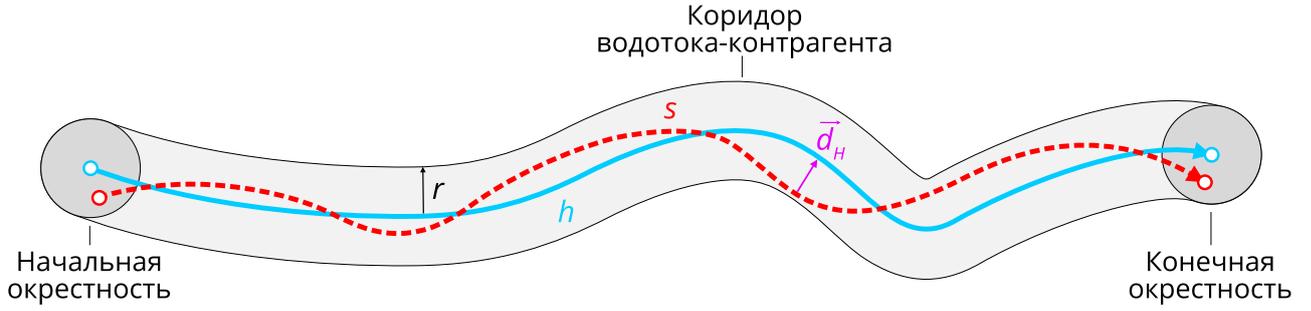
$$\begin{cases} d(s(0), h(0)) \leq r, \\ d(s(1), h(1)) \leq r, \\ \vec{d}_H(s, h) \leq r, \end{cases} \quad (3.37)$$

где  $r$  — пороговое расстояние, называемое **радиусом захвата**.

Первое условие в Формуле 3.37 означает, что  $s$  должен начинаться не дальше, чем  $r$  от первой точки  $h$ . Все местоположения, удовлетворяющие этому условию, находятся в буферной зоне  $h(0)$  в форме круга, которая называется **начальной окрестностью** и определяется смещением на расстояние не более  $r$  от  $h(0)$ . Второе условие в Формуле 3.37 задает аналогичное ограничение, связывающее конечные точки  $s$  и  $h$ . Соответственно, буферная зона  $h(1)$  в форме круга называется **конечной окрестностью**. Наконец, третье условие в Формуле 3.37 требует, чтобы  $s$  отклонялось от  $h$  не дальше, чем на  $r$ . Все местоположения, удовлетворяющие этому условию, находятся в буферной зоне  $h$ , называемой **коридором контрагента** и определяются смещением на расстояние не более  $r$  от  $h$ .

Определение контрагента-кандидата показано на Рисунке 3.70.

Поскольку  $\vec{d}_H(s, h)$  ограничивает только расстояние контрагента от водотока, но не наоборот, это самое слабое из всех возможных ограничений на расстояние. Качество кандидата может быть усилено с помощью расстояний Хаусдорфа и Фреше. В частности, можно выделить три следующих класса:



**Рисунок 3.70** — Определение контрагента-кандидата: водоток ( $h$ ), контрагент-кандидат ( $s$ ), радиус захвата ( $r$ ), направленное расстояние Хаусдорфа ( $\vec{d}_H$ ) от  $s$  до  $h$ .

- *сильный*:  $d_F(s, h) \leq r$ ;
- *стандартный*:  $d_H(s, h) \leq r$ ;
- *слабый*:  $\vec{d}_H(s, h) \leq r$ .

Важно отметить, что если применяется слабое ограничение, это не исключает, что кандидат будет удовлетворять также стандартному и сильному критерию, однако этого уже нельзя гарантировать. Введенные классы используются не только для отбора кандидатов, но и для оценки качества получаемых контрагентов (Таблица 3.9).

Пусть  $S = \{s\}$  — множество контрагентов-кандидатов. Для формализации условий выбора наилучшего контрагента из  $S$  представим каждого кандидата в виде параметрической кривой, индексируемой  $t$ :

$$s(t) = (x(t), y(t)), t \in [0, 1]. \quad (3.38)$$

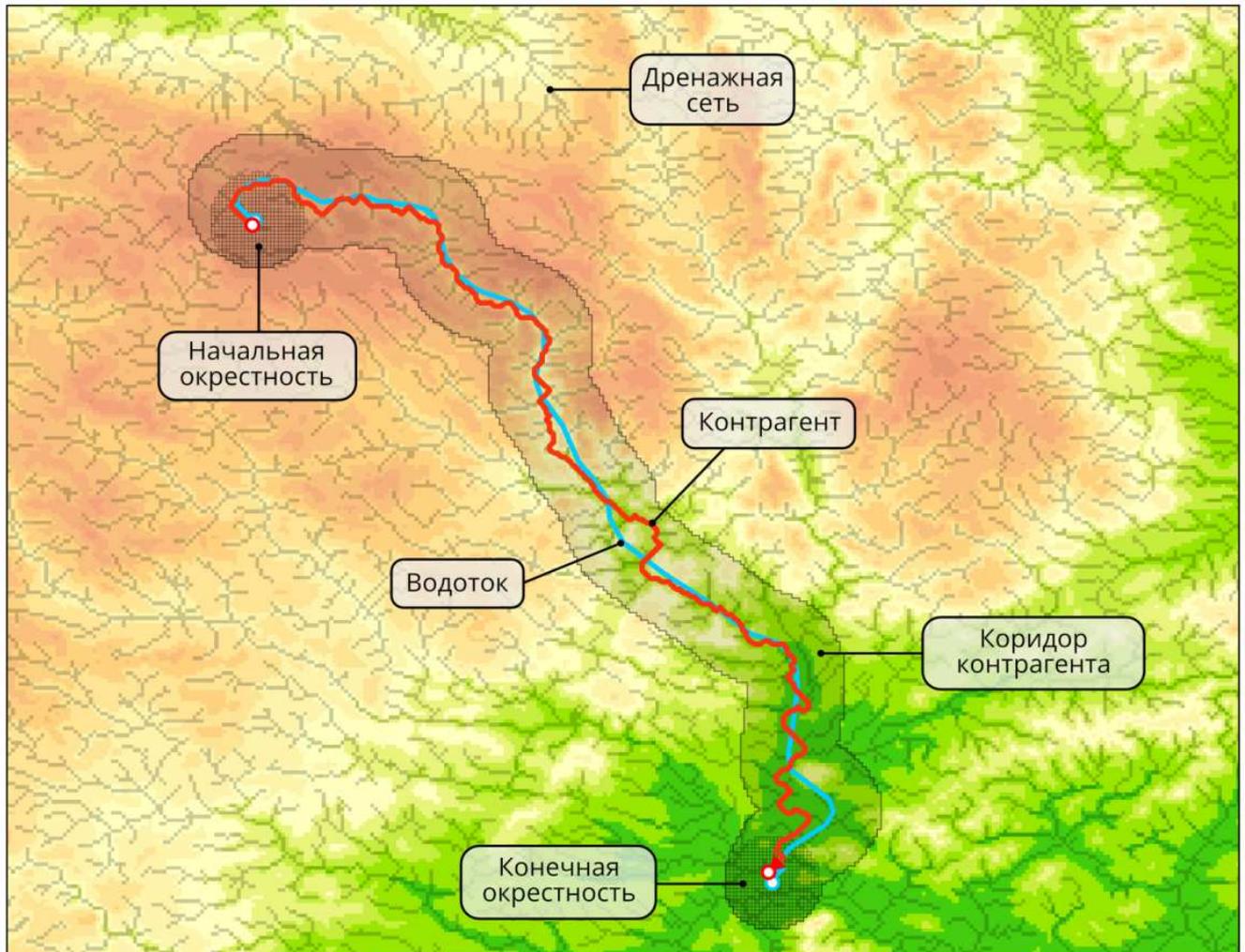
Пусть  $p = (x, y)$  — ячейка ЦМР, определяемая координатами ее центра, а  $F(p) = F(x, y)$  — поле площади водосбора (аккумуляции стока). Тогда контрагент  $S$  выбирается из  $S$  путем поиска такого его элемента  $s$ , который удовлетворяет следующим трем условиям:

$$\begin{cases} \min_{\forall p \in s} F(p) \geq a, \\ F(s(u)) \geq F(s(v)) \mid u > v, \\ d_M(s, h) = \min_{\forall s_j \in S} d_M(s_j, h). \end{cases} \quad (3.39)$$

Первые два условия в Формуле 3.39 заставляют  $s$  быть элементом сети потенциальных водотоков на поверхности ЦМР: он проходит через ячейки с площадью водосбора больше определенного порога ( $\min F(s) \geq a$ ) в направлении монотонного увеличения этой площади ( $F(s(u)) \geq F(s(v)) \mid u > v$ ). Третье условие в Формуле 3.39 принуждает  $s$  быть ближайшим к водотоку кандидатом с точки зрения модифицированного расстояния Хаусдорфа.

Пример одного извлеченного контрагента показан на Рисунке 3.71. Этот рисунок

дополнительно содержит сеть потенциальных водотоков, представленную серым цветом и соответствующую ячейкам раstra площади водосбора, которые удовлетворяют условию  $\min_{\forall p \in s} F(p) \geq a$ .



**Рисунок 3.71** – Водоток и его контрагент

Разнообразие пространственных данных приводит к ситуациям, в которых невозможно извлечь контрагенты с помощью базового подхода, описанного выше. Среди распространенных случаев следует отметить некорректное направление гидрографических линий, ошибки в ЦМР и получаемых на ее основе направлениях стока, существенные различия в детальности ЦМР и гидрографии. Кроме того, неудачный выбор параметров  $r$  и  $a$ , а также использование не слабого, а стандартного или сильного критерия близости может воспрепятствовать получению контрагента. Для преодоления подобных сложностей разработанная методика включает альтернативную стратегию *кратчайшего стоимостного пути*, описанную ниже.

Пусть  $E(p, h) = \min_{q \in h} d(p, q)$  – поле евклидовых расстояний до  $h$ , где  $p$  и  $q$  – ячейки ЦМР, определяемые координатами их центров, а  $h(t) = (X(t), Y(t))$ ,  $t \in [0, 1]$  – параметрическое описание водотока с  $h(0)$  в качестве начальной точки и  $h(1)$  в качестве

конечной точки. **Контрагент наименьшей стоимости** будет получен путем трассировки кандидата  $s$ , минимизирующего следующий функционал:

$$\int_0^1 W(s(t)) [E(s(t), h) + 1] dt \rightarrow \min \quad (3.40)$$

при следующих условиях:

$$\begin{cases} s(0) = h(0), \\ s(1) = h(1), \end{cases} \quad (3.41)$$

где  $W(p)$  – штрафная функция отклонения от водотока, определяемая как:

$$W(p) = \begin{cases} 1 & F(p) \geq a, \\ w\hat{Z}(p) & F(p) < a, \end{cases} \quad \hat{Z}(p) = Z(p) - \min(Z) + 1. \quad (3.42)$$

Константа  $w \geq 1$  в Формуле 3.42, определяющая масштаб штрафа, задается как один из параметров конфляции. Она увеличивает стоимость пути, когда он проходит через ячейки  $p$ , не принадлежащие сети потенциальных водотоков на ЦМР ( $F(p) < a$ ).

Заметим также, что штрафная функция повышает приоритет путей, стремящихся к более низким высотам, что достигается умножением  $w$  на **относительную высоту**  $\hat{Z}(p)$  в Формуле 3.42. Эта переменная рассчитывается как превышение над минимальной высотой по ЦМР плюс свободный член 1, который добавляется, чтобы точка с  $Z(p) = \min(Z)$  не получила автоматически нулевую стоимость.

По сравнению с обычными, контрагенты наименьшей стоимости начинаются и заканчиваются точно в заданных точках, но считаются менее качественными, так как могут проходить через ячейки ЦМР, не принадлежащие дренажной сети.

Пример контрагента наименьшей стоимости показан на Рисунке 3.72. Наряду с водотоком и его контрагентом, на этом рисунке цветовым градиентом показана стоимостная поверхность. Ячейки, принадлежащие дренажной сети, имеют стоимость, равную 1, и окрашены в самый темный цвет. Стоимость быстро растет с увеличением расстояния от водотока и постепенно уменьшается вместе с высотой поверхности. В совокупности эти два фактора заставляют контрагента быть как можно ближе к водотоку и следовать пути, который идет вниз по склону.

### Несколько гидрографических линий

Если в качестве исходных данных задано несколько опорных водотоков, необходимо предпринять дополнительные усилия для сохранения характера их топологических связей. В частности, каждое слияние и разветвление должны быть представлены как аналогичный **контрагентный узел** между результирующими водотоками-контрагентами. Пусть  $h_1$  и  $h_2$  – исходные водотоки, из которых  $h_1$  является основным, а  $h_2$  – второсте-



**Рисунок 3.72** – Контрагент наименьшей стоимости (обозначен красной линией). Прямоугольник с черной обводкой ограничивает область, представленную на Рисунке 3.74

пенным (впадает в  $h_1$  или вытекает из него). Также пусть  $c_1$  и  $c_2$  — их соответствующие контрагенты, а  $s_2$  — некоторый контрагент-кандидат для  $h_2$ . Тогда применяются следующие пять **топологических правил**, представленных на Рисунке 3.73:

**Топологическое правило 1.** Если  $h_2$  является *притоком*  $h_1$  в некотором узле  $\gamma$ , то конечная окрестность  $s_2$  должна иметь центром точку  $v$  на  $c_1$ , которая ближе всего к  $\gamma$ . Для кандидата в контрагенты это означает, что условие  $d(s_2(1), h_2(1)) \leq r$  преобразуется в  $d(s_2(1), v) \leq r$ . Если для трассировки используется метод стоимостного пути, то условие  $s_2(1) = h_2(1)$  преобразуется в  $s_1(1) = v$ . Получающийся в результате контрагентный узел обозначен как  $\gamma'$  на Рисунке 3.73а.

**Топологическое правило 2.** Если  $h_2$  *вытекает* из  $h_1$  на некотором узле разветвления  $\lambda$ , то начальная окрестность  $s_2$  должна иметь центром точку  $u$  на  $c_1$ , которая ближе всего к  $\lambda$ . Для кандидата в контрагенты это означает, что условие  $d(s_2(0), h_2(0)) \leq r$  преобразуется в  $d(s_2(1), u) \leq r$ . Если для трассировки используется метод стоимостного пути, то условие  $s_2(0) = h_2(0)$  преобразуется в  $s_2(0) = u$ . Получающийся в результате контрагентный узел обозначен как  $\lambda'$  на рисунке Рисунке 3.73б.

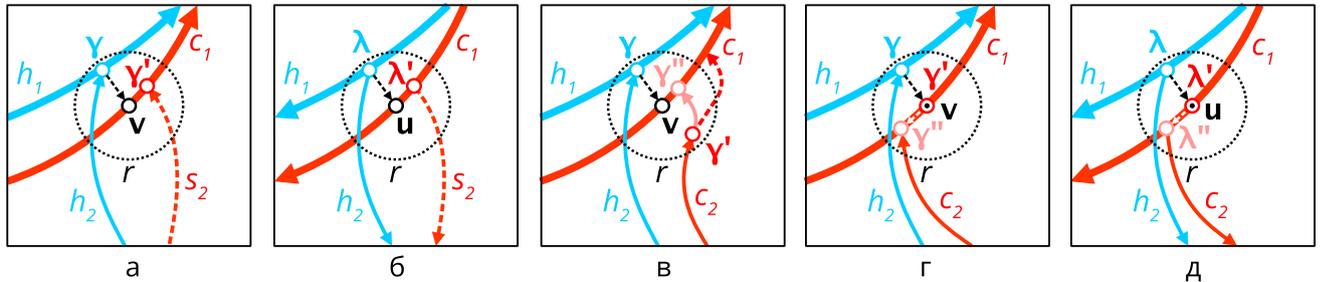
**Топологическое правило 3.** Если  $c_2$  является *обычным контрагентом* и не имеет общих точек с  $c_1$ , то он должен быть продлен, чтобы достичь  $c_1$  в точности, используя метод стоимостного пути. Рисунок 3.73 иллюстрирует случай, когда  $h_2$  является притоком  $h_1$ , но аналогичное правило применимо и в случае, когда  $h_2$  является рукавом, вытекающим из  $h_1$ . Получающийся в результате контрагентный узел обозначен как  $\gamma''$  на рисунке Рисунке 3.73в.

**Топологическое правило 4.** Если  $c_2 = \{p_k, k = 1, \dots, |c_2|\}$  является *притоком*  $c_1$  и имеет с ним более одной общей точки, то его следует заменить на  $c_2 = \{p_k \in c_2 \mid k = 1, \dots, k_{min}\}$ , где  $k_{min}$  — наименьший индекс общей точки. Получающийся контрагентный узел обозначен как  $\gamma''$  на рисунке Рисунке 3.73г.

**Топологическое правило 5.** Если  $c_2 = \{p_k, k = 1, \dots, |c_2|\}$  *вытекает* из  $c_1$  и имеет с ним более одной общей точки, то его следует заменить на  $c_2 = \{p_k \in c \mid k =$

$k_{max}, \dots, |c_2|$ }, где  $k_{max}$  — наибольший индекс общей точки. Получившийся контрагентный узел обозначен как  $\lambda''$  на Рисунке 3.73д.

Применяемые в описанной последовательности, эти правила гарантируют, что результирующие водотоки-контрагенты составят сеть с топологической структурой, аналогичной топологической структуре эталонных гидрографических линий.



**Рисунок 3.73** — Топологические правила для трассировки контрагентов нескольких водотоков: (а) слияние; (б) бифуркация; (в) слияние с продлением; (г) слияние с удалением перекрытия; (д) бифуркация с удалением перекрытия

Рисунок 3.74 иллюстрирует, как топологические правила 1 и 4 действуют на практике. Сначала прослеживается контрагент водотока более высокого порядка. Поскольку он не проходит через точку слияния водотоков, нет смысла искать его контрагент-приток, который заканчивается в соответствующей ячейке ЦМР. Вместо этого конечная окрестность этой линии перемещается в ближайшую точку на контрагенте более высокого порядка, а коридор поиска расширяется соответственно. Когда контрагент притока прослеживается методом стоимостного пути, он впадает в основного контрагента и проходит вдоль него короткий участок до достижения конечной ячейки. Этот общий участок обрезается для устранения перекрытия и получения итогового контрагентного узла.



**Рисунок 3.74** — Слияние водотоков и их контрагентов

### 3.6.1.4 Генерация связей конфляции

После того как для каждого водотока определен контрагент, между водотоками и контрагентами создаются связи конфляции. Поскольку контрагент извлекается из ЦМР,

его вершины разделены расстоянием не менее размера ячейки  $R$ . Для точной трансформации важно, чтобы плотность векторов была равномерной и соответствовала разрешению, поэтому вершины и водотоков и контрагентов предварительно сгущаются так, чтобы расстояние между ними не превышало  $R$ . Далее генерируются связи, соединяющие ближайшие вершины двух линий. На Рисунке 3.75 показаны связи, сгенерированные для фрагмента водотока-контрагента. Эти связи будут служить векторами, определяющими трансформацию ЦМР методом резинового листа.

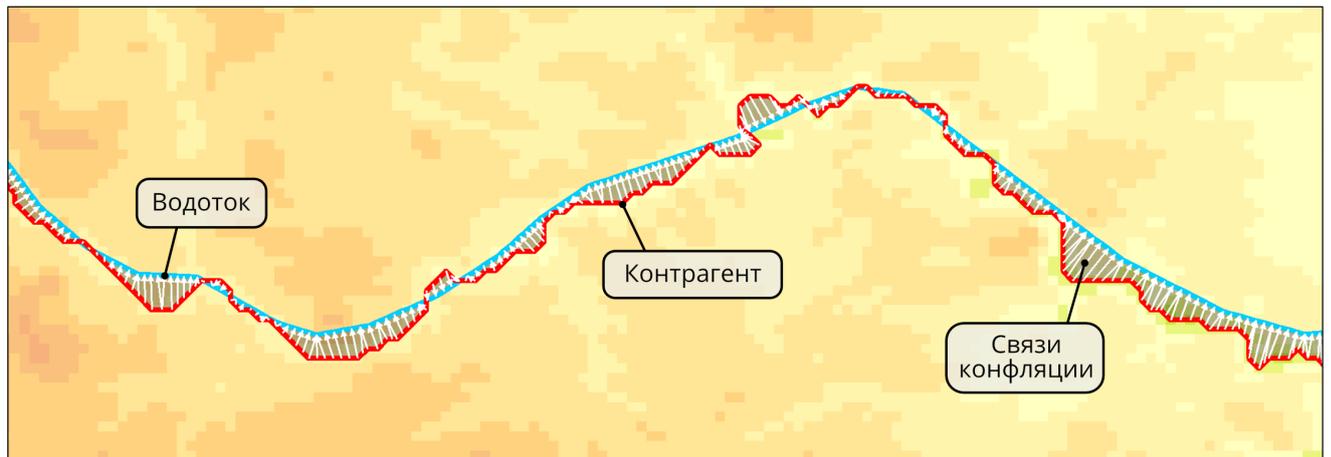


Рисунок 3.75 – Связи конфляции

### 3.6.1.5 Извлечение высотных данных

Этот и три последующих параграфа методического раздела дополнены Рисунками 3.76 - 3.77, которые иллюстрируют случаи конфляции с генерализацией и без. Пример фрагмента исходной растровой ЦМР с наложенным на нее водотоком представлен на рисунках Рисунке 3.76а и Рисунке 3.77а.

В случае, если конфляция выполняется без генерализации ЦМР, на первом этапе ячейки растровой ЦМР должны быть преобразованы в точки, а для триангуляционной ЦМР результирующими точками будут являться ее вершины. Точечное представление растровой ЦМР показано на Рисунке 3.76б.

Другой сценарий возникает, когда водотоки должны быть внедрены в структуру ЦМР в процессе ее генерализации — например, когда детальность ЦМР понижается для того, чтобы она соответствовала детальности гидрографии. Чтобы удовлетворить этому условию, необходимо использовать один из методов генерализации ЦМР, учитывающих гидрографическую сеть явным образом [Jordan, 2007; Li, Wang, Yang, 2008; Leonowicz, Jenny, Hurni, 2009; Ai, Li, 2010; Samsonov, 2011; Zhou, Chen, 2011; Chen et al., 2012]. Общим свойством этих методов является то, что в них гидрографическая сеть выделяется по ЦМР, а не подается из внешних данных, т.е. является скорее дренажной сетью. Это означает, что для того чтобы полученная ЦМР была совмещена с линиями реальных во-

дотоков, необходимо не просто выделение дренажной сети, а сети, включающей в обязательном порядке контрагенты, соответствующих водотокам.

Для решения этой задачи был адаптирован метод, разработанный в настоящей диссертации и представленный в Разделе 2.5. Соответствующие структурные линии рельефа (талвеги и водоразделы), извлеченные из фрагмента ЦМР, представлены на Рисунке 3.77б.

### 3.6.1.6 Трансформация высотных данных методом резинового листа

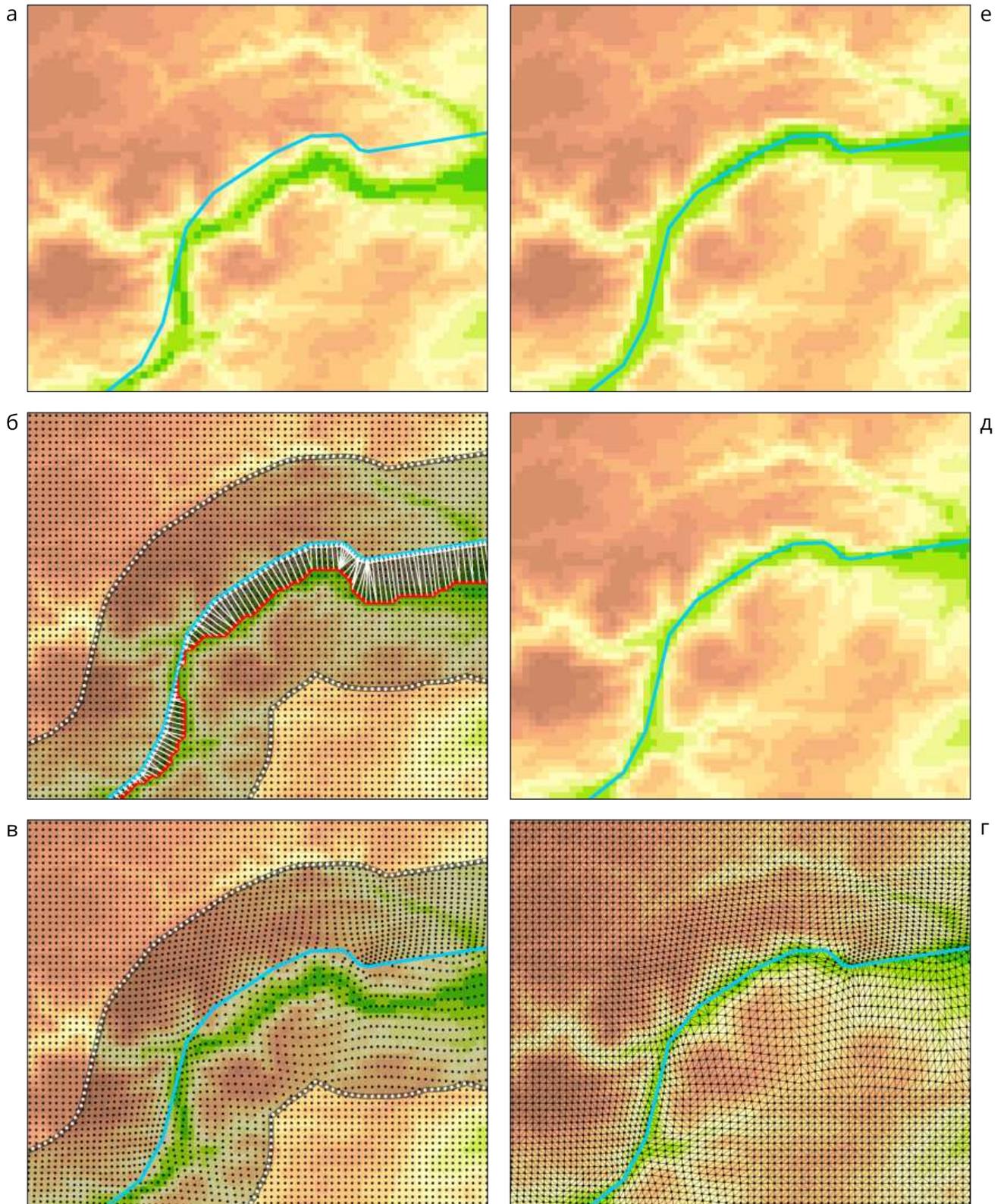
Полученные связи между водотоками и их контрагентами, а также высотные данные, представленные в виде векторных объектов, можно использовать для совмещения ЦМР с водотоками методом конфляции [Saalfeld, 1988]. Данный метод использует локализованные трансформации координат методом резинового листа [White, Griffin, 1985]. Поскольку он искривляет пространство не только в непосредственной близости совмещаемых точек, но и на отдалении от них, необходимо принудительно ограничить зону трансформации. Для этого создается **область конфляции**  $A$ , представляющая собой буфер замкнутой области, ограниченной водотоками, их контрагентами и связями конфляции. Радиус захвата  $r$  используется в качестве расстояния буферизации, но в общем случае может быть инициализирован независимым от него значением.

Область конфляции показана прозрачным серым цветом на Рисунках 3.76 — 3.77. Точки, расположенные вдоль границы этой области с шагом, равным разрешению ЦМР, определяют так называемые **тождественные связи**  $L_I$ . Они изображены белым цветом на Рисунках 3.76 — 3.77. Тожественные связи задают местоположения, которые остаются фиксированными в процессе конфляции и не позволяют преобразованию методом резинового листа распространяться за пределы созданной области. Примеры векторных объектов с высотными данными, трансформированных в пределах области конфляции, представлены на Рисунке 3.76в и Рисунке 3.77в соответственно. Важно отметить, что за пределами области конфляции высотные данные располагаются в исходных местоположениях.

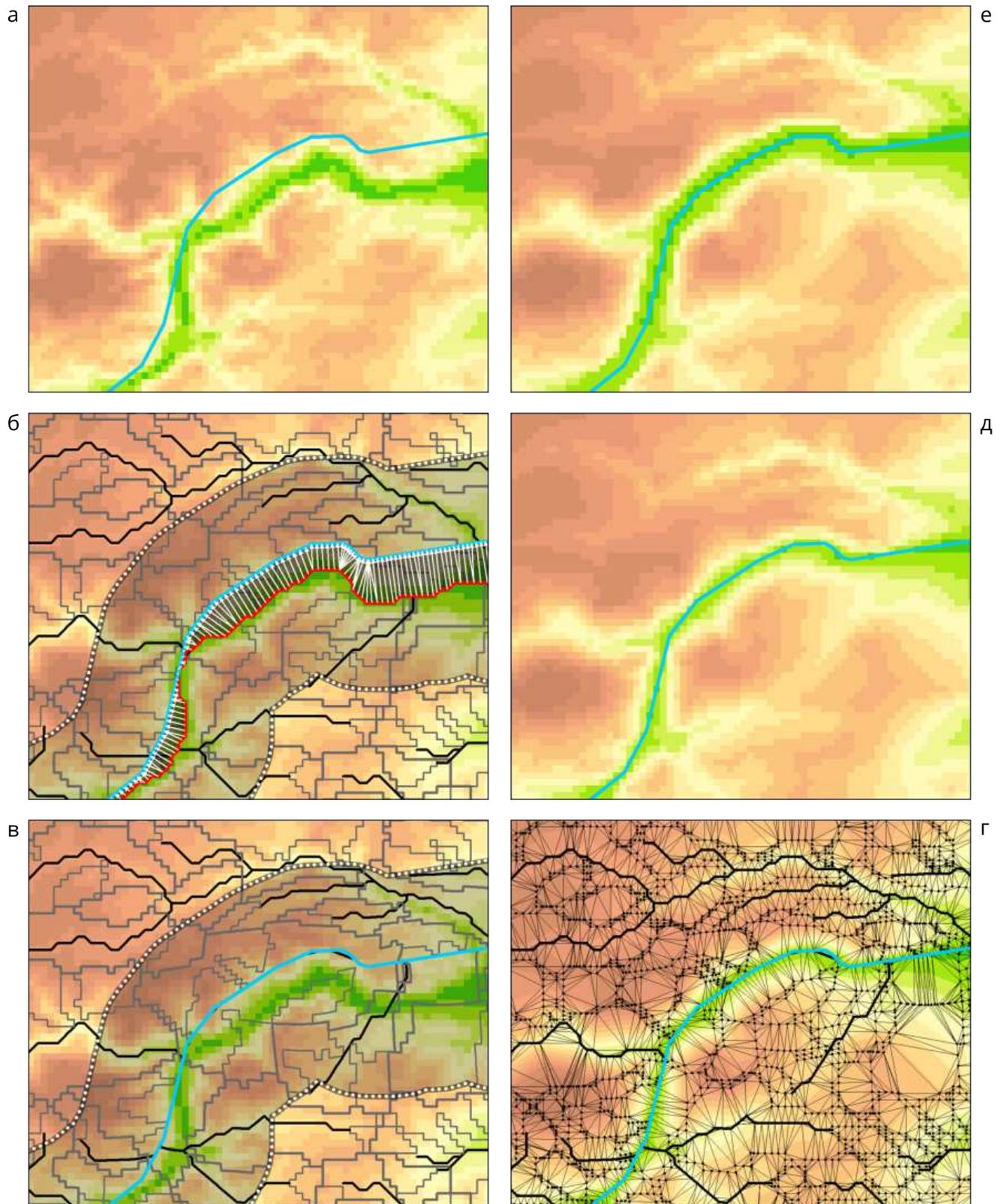
### 3.6.1.7 Восстановление ЦМР после трансформации

Восстановление ЦМР реализуется путем триангуляции трансформированных высотных данных и дальнейшей растеризации<sup>8</sup> полученной поверхности на ту же интерполяционную сетку, что и исходная ЦМР. В случае если конфляция реализовывалась по узлам ЦМР (без генерализации), то и линейный метод, и метод естественного соседа [Sibson, 1981] при растеризации дадут одинаковые высоты в ячейках ЦМР, расположенных вне области конфляции. Это следует из того, что центры этих ячеек совпадают с

<sup>8</sup>При необходимости, если исходная ЦМР — растровая



**Рисунок 3.76** – Конфляция растровой ЦМР с опорной гидрографической линией: (а) исходная ЦМР и водоток; (б) высотные данные (черные точки), водоток-контрагент (красная линия), связи конфляции (белые стрелки), область конфляции (прозрачный серый полигон) и тождественные связи (белые точки); (в) высотные данные после трансформации; (г) TIN; (д) растеризованная ЦМР после трансформации; (е) растеризованная ЦМР после трансформации и расширения долины водотока



**Рисунок 3.77** – Конфляция растровой ЦМР с опорной гидрографической линией в процессе генерализации: (а) исходная ЦМР и водоток; (б) потенциальные водотоки (черные линии), границы их бассейнов (серые линии), водоток-контрагент (красная линия), связи конфляции (белые стрелки), область конфляции (прозрачный серый полигон) и тождественные связи (белые точки); (в) высотные данные после трансформации; (г) TIN; (д) растеризованная ЦМР после трансформации; (е) растеризованная ЦМР после трансформации и расширения долины водотока

высотными данными, подаваемыми в функцию растеризации.

Триангуляции трансформированных точек и структурных линий, а также полученные растровые ЦМР представлены на рисунках Рисунке 3.76г-д и Рисунке 3.77г-д соответственно. Обратим внимание на то, что в обоих случаях устраняется плановое несоответствие между ЦМР и опорными линиями, а в случае генерализации дополнительно обеспечивается наличие долины контрагента в результирующей поверхности.

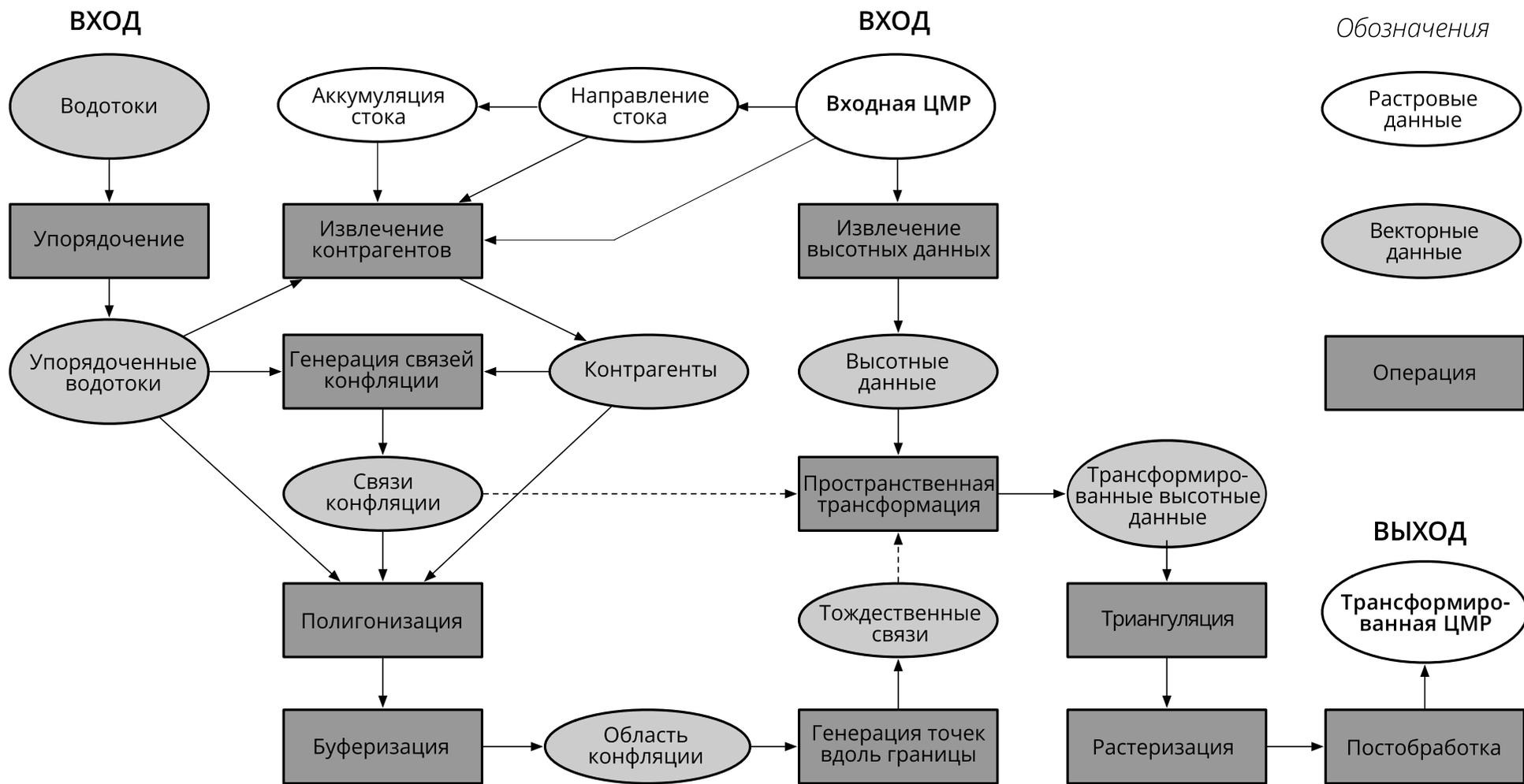
### 3.6.1.8 Постобработка

Предыдущий этап завершает процесс *горизонтального* совмещения. Однако как картографические, так и аналитические приложения ЦМР требуют, чтобы совмещение между ЦМР и объектами гидрографии было выполнено также и в *вертикальном* измерении. С картографической точки зрения это означает, что высоты вдоль каждого водотока должны монотонно уменьшаться, т.е. на его пути не должно быть замкнутых впадин или возвышений. В противном случае картографическое представление ЦМР будет географически недостоверным и эстетически непривлекательным. Чтобы предотвратить образование таких артефактов в результирующей ЦМР, применяется процедура, которая прорезает любые участки повышения высотных отметок на пути каждого водотока с помощью линейной интерполяции. Если конфляция выполняется для целей картографирования (визуализации), то полученная ЦМР может быть дополнительно подвергнута операции расширения вдоль водотоков для обеспечения видимости их долин на поверхности ЦМР согласно процедуре, описанной в Разделе 2.5.3. Итоговые результаты, полученный в результате процедур прорезания и расширения при конфляции в обычном режиме и с генерализацией, представлены на Рисунке 3.76е и Рисунке 3.77е.

Следует отметить, что если результирующую ЦМР планируется использовать для гидрологического анализа, то ограничения должны быть еще более жесткими. Помимо монотонного уменьшения высоты, также необходимо потребовать, чтобы водотоки были полностью представлены в дренажной сети ЦМР. Поскольку эта задача обычно решается с помощью операции прорезания (*carving, breaching*), ту же процедуру можно применить и к ЦМР, которая уже прошла процедуру конфляции. Преимущество такой ЦМР заключается в том, что прорезание не будет создавать новые долины на поверхности ЦМР, а вместо этого будет форсироваться корректное направление стока вдоль днищ уже существующих долин.

### 3.6.1.9 Общий алгоритм

Общий алгоритм для случая растровой ЦМР на входе представлен на Рисунке 3.78.



**Рисунок 3.78** – Общий рабочий процесс для совмещения растровой ЦМР с опорными водотоками

Алгоритм предусматривает оба варианта извлечения высотных данных — с генерализацией и без нее на этапе `Extract vector elevation data`. Для краткости в схеме опущены некоторые промежуточные операторы и данные. В частности, для получения растра аккумуляции стока из ЦМР необходимо устранить замкнутые локальные понижения.

### 3.6.2 Результаты

Разработанная методика была реализована на языке программирования Python с помощью модулей `numpy` [Oliphant, 2006] и `arcpy` [Esri, 2020]. Программное обеспечение находится в открытом доступе и может быть загружено с репозитория GitHub [Samsonov, 2020a]. Реализация была протестирована на примере двух широко используемых пространственных наборов данных, находящихся в открытом доступе: растровой ЦМР GEBCO\_2019 [GEBCO, 2019] и векторной базы данных Natural Earth [Patterson, Kelso, et al., 2020]. Разрешение GEBCO\_2019 в географических координатах составляет 15", что приблизительно равно 500 м вдоль экватора. Набор данных Natural Earth содержит пространственные данные на трех уровнях детализации 10, 50 и 110, предназначенных для картографирования в масштабах 1 : 10 000 000, 1 : 50 000 000 и 1 : 110 000 000 соответственно. В рамках настоящего эксперимента были использованы гидрографические данные с УД 10 (версия 4.1.0). Эти данные имеют существенно более низкую детальность, чем модель GEBCO, что приводит к их практически повсеместному несовпадению и дает возможность в полной мере оценить эффективность разработанной методики.

Оба набора данных были обрезаны в пределах территории в Центральной Европе, расположенной в основном в бассейне реки Дунай, и представлены на Рисунке 3.79. Размеры обрезанной по охвату территории ЦМР составляют  $1358 \times 1007$  ячеек, а обрезанная речная сеть содержит 28 объектов. Для обработки в метрических единицах оба набора данных были преобразованы в равнопромежуточную цилиндрическую проекцию (*Plate Carrée*), которая сохраняет геометрию исходного растра без деформаций. В качестве оптимальных значений параметров, найденных опытным путем, использовались радиус захвата  $r = 6000$ , минимальная аккумуляция стока  $a = 10$  и штрафной коэффициент  $w = 30$ . Растр направлений стока водотока был получен методом D8 после обработки исходной ЦМР с помощью процедуры прорыва замкнутых понижений (*depression breaching*) [Lindsay, 2016b]. По умолчанию использовался слабый критерий точности, определяемый направленным хаусдорфовым расстоянием  $\vec{d}_H$ .

Уникальные идентификаторы водотоков отмечены рядом с каждым объектом на Рисунке 3.79. Особый интерес представляют два случая: рукав с  $ID = 9$ , который участвует в формировании разветвленного русла Дуная у Будапешта, и водоток с  $ID = 36$ , который на самом деле представляет собой две реки из бассейнов Эльбы и Дуная, ошибочно объединенные в один пространственный объект. Данная ошибка намеренно не

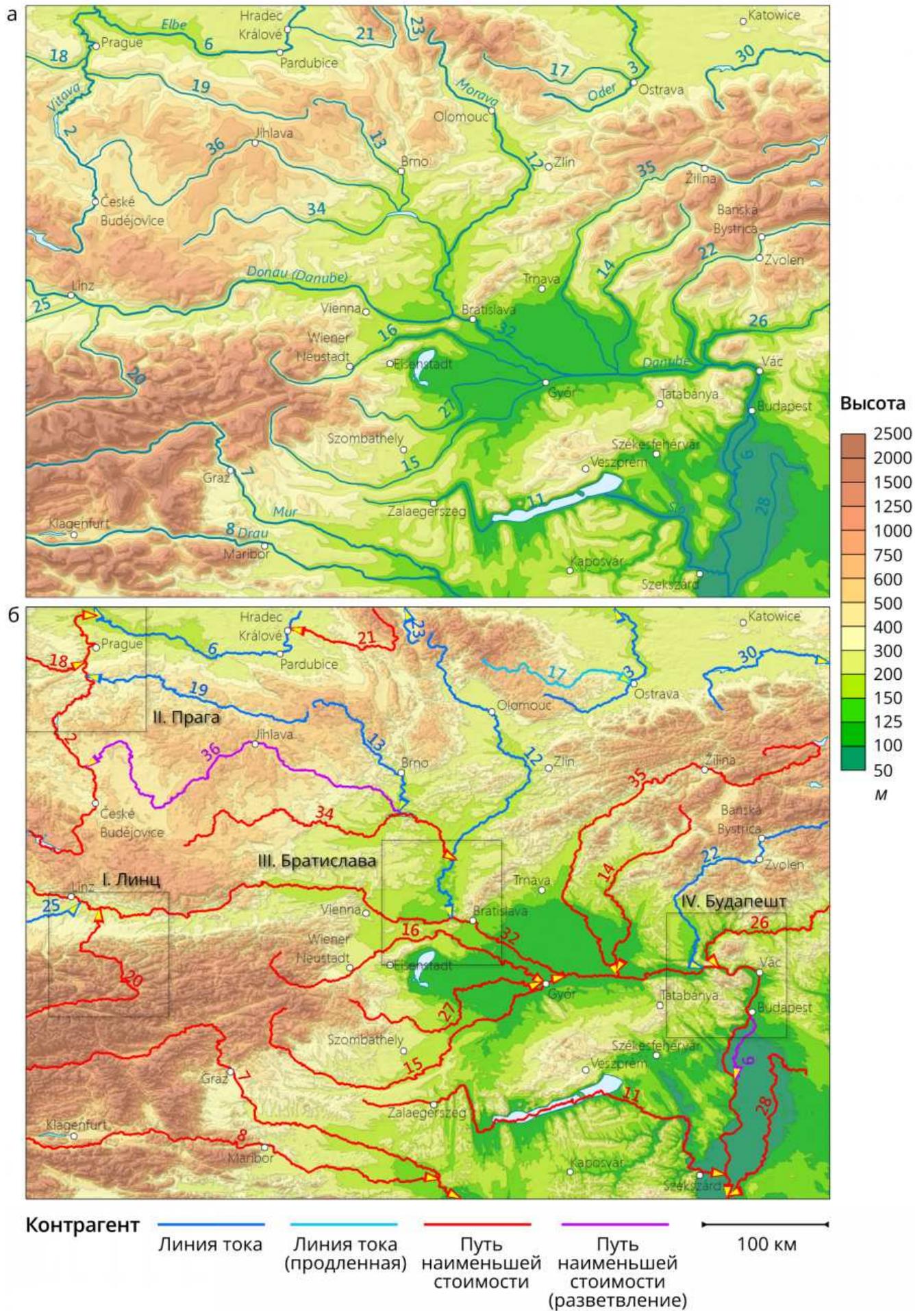
была исправлена в исходных данных, чтобы проверить устойчивость разработанного в нестандартных ситуациях.

Карта в нижней части Рисунка 3.79 представляет собой необработанную ЦМР GEBCO\_2019 с наложенными поверх нее контрагентами водотоков. Поскольку данные Natural Earth уже оцифрованы в соответствии с порядком Хака для картографических целей (основные водотоки являются непрерывными от истока до устья), исходная геометрия не изменялась, а вместо этого был рассчитан порядок водотоков и номера итераций в соответствии с алгоритмом предварительной обработки, описанным в Разделе 3.6.1.2. Полученные контрагенты представляют топологически корректную сеть со структурой, идентичной сети водотоков (конец каждого контрагента обозначен стрелкой). На Рисунке 3.79 специальным символом обозначены два типа контрагентов: рукава, полученные методом стоимостного пути ( $ID = 9, 34$ ) и обычный контрагент, который пришлось продлить до водотока более высокого порядка ( $ID = 17$ ). Алгоритм успешно справился с данными конфигурациями и форсировал соблюдение топологических правил 2 и 3.

Таблица 3.9 обобщает результаты выделения контрагентов. Первые шесть столбцов формируются на этапе порядковой классификации водотоков и соответствуют Таблице 3.8. Последние пять столбцов показывают тип и близость полученных контрагентов к соответствующим водотокам. Для каждого водотока вычислено направленное хаусдорфово расстояние  $\vec{d}_H$ , хаусдорфово расстояние  $d_H$  и расстояние Фреше  $d_F$ , а также выполнена оценка качества контрагента в терминах классов, введенных в Разделе 3.6.1.3. Несмотря на то, что для фильтрации кандидатов в контрагенты использовалось слабое ограничение, все водотоки, кроме одного, попадают в наиболее строгий класс, ограниченный расстоянием Фреше, а значения всех трех расстояний в основной массе довольно близки друг к другу.

Полученный результат обосновывает решение использовать слабое ограничение по умолчанию. Несмотря на то что меры сходства, отличные от расстояния Фреше, не учитывают непрерывность линий, их применение оправдано практическими соображениями: слабые ограничения могут быть удовлетворены более простыми и быстрыми вычислительными процедурами, что важно в случае необходимости конфликации данных большого объема. То, что контрагент с  $ID = 15$  не удовлетворяет обычному и сильному критерию, обусловлено применением топологического правила 4: этот контрагент достиг вышестоящего контрагента вне конечной окрестности, и его конечный участок был сокращен.

Верхнее изображение на Рисунке 3.79 представляет результат конфликации ЦМР с гидрографическими данными в режиме генерализации, при котором детальность ЦМР была приведена в соответствие с детальностью гидрографии. Минимальная длина основного потенциального водотока при генерализации была установлена в 45 ячеек ЦМР, а для второстепенных водотоков — в 15 ячеек.



**Рисунок 3.79** – Область и объекты конфляции: (а) водотоки и совмещенная с ними генерализованная ЦМР; (б) водотоки-контрагенты и исходная ЦМР

Таблица 3.9 – Статистика по результатам извлечения водотоков-контрагентов

ID	CONFL	BIFUR	ITER	ORDER	ТИП	КОНТРАГЕНТ	$d_F$ , м	$d_H$ , м	$\vec{d}_H$ , м	КАЧЕСТВО
2	6	-1	2	2	Основной	Стоимостной	4065	4065	4065	Сильный
3	-1	-1	1	1	Основной	Линия тока	3202	3202	3202	Сильный
6	-1	-1	1	1	Основной	Линия тока	4412	4384	4384	Сильный
7	-1	-1	1	1	Основной	Стоимостной	4705	4472	4472	Сильный
8	7	-1	2	2	Основной	Стоимостной	2809	2809	2809	Сильный
9	32	32	2	2	Второстепенн	Стоимостной	2357	2357	2356	Сильный
11	32	-1	2	2	Основной	Стоимостной	2375	2375	2375	Сильный
12	32	-1	2	2	Основной	Линия тока	5166	5136	5136	Сильный
13	36	-1	5	4	Основной	Линия тока	4586	4477	4477	Сильный
14	35	-1	3	3	Основной	Стоимостной	3307	3307	3307	Сильный
15	32	-1	2	2	Основной	Стоимостной	6264	6264	4277	Слабый
16	15	-1	3	3	Основной	Стоимостной	2884	2884	2884	Сильный
17	3	-1	2	2	Основной	Линия тока	4737	4737	4737	Сильный
18	2	-1	3	3	Основной	Стоимостной	2501	2501	1754	Сильный
19	2	-1	3	3	Основной	Линия тока	2510	2510	2510	Сильный
20	32	-1	2	2	Основной	Стоимостной	3341	3341	2581	Сильный
21	6	-1	2	2	Основной	Стоимостной	1987	1987	1987	Сильный
22	32	-1	2	2	Основной	Линия тока	4604	4604	4604	Сильный
23	-1	-1	1	1	Основной	Линия тока	1378	1378	1378	Сильный
25	32	-1	2	2	Основной	Линия тока	2519	2519	2320	Сильный
26	32	-1	2	2	Основной	Стоимостной	4185	4185	4185	Сильный
27	16	-1	4	4	Основной	Стоимостной	4408	4408	4408	Сильный
28	32	-1	2	2	Основной	Стоимостной	2998	2998	2998	Сильный
30	-1	-1	1	1	Основной	Линия тока	4059	4059	4059	Сильный
32	-1	-1	1	1	Основной	Стоимостной	4206	4206	4206	Сильный
34	12	-1	3	3	Основной	Стоимостной	4417	4417	4417	Сильный
35	32	-1	2	2	Основной	Стоимостной	3689	3689	3689	Сильный
36	2	34	4	3	Второстепенн	Стоимостной	4032	4032	4032	Сильный

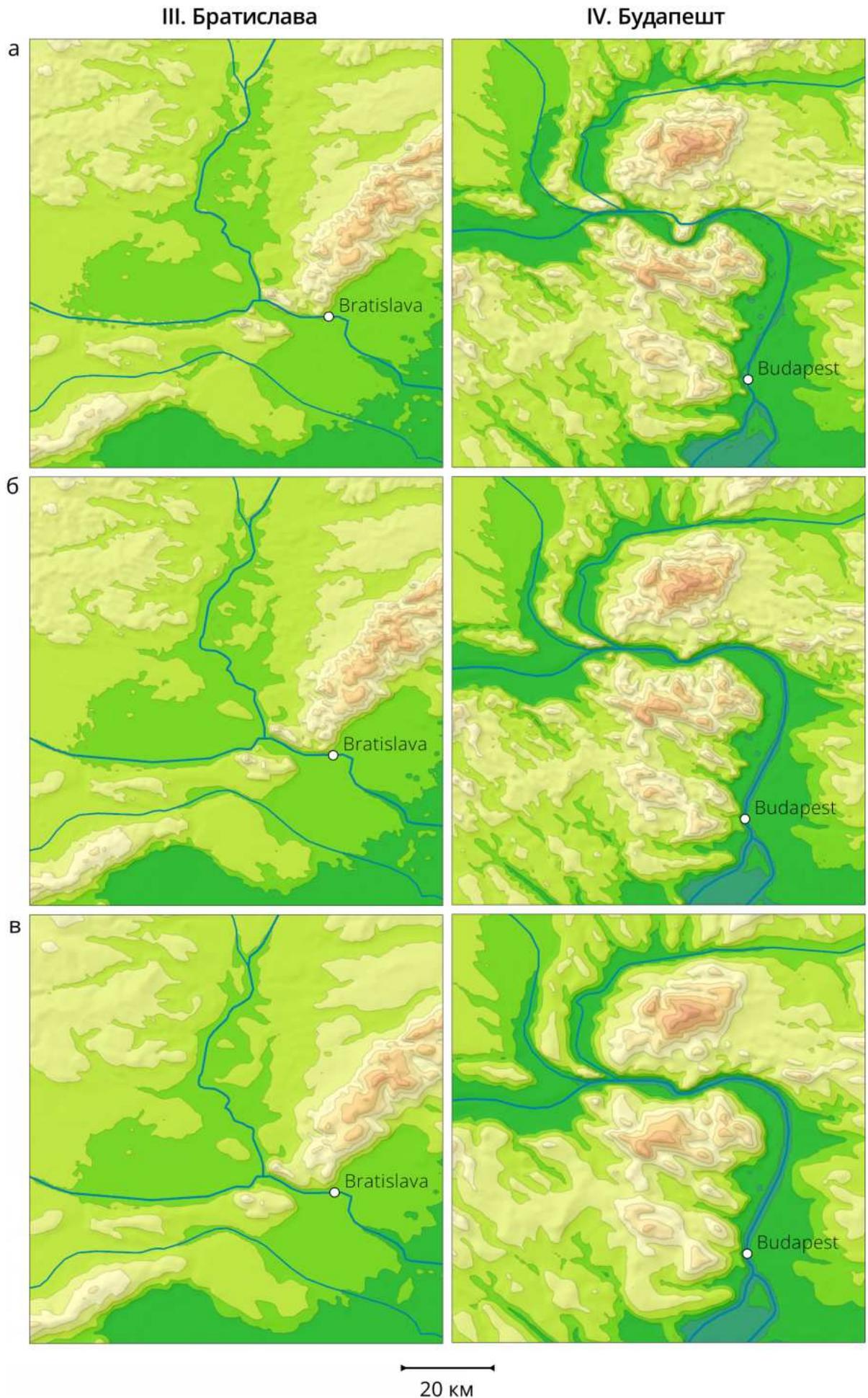
Карта наглядно демонстрирует, что ЦМР идеально согласована с водотоками, и с помощью разработанной методики можно получить высококачественный картографический результат. Карта также демонстрирует, что вышеупомянутый гидрографический объект с  $ID = 36$  не представляет проблемы для реализованного алгоритма, и полученная ЦМР совмещена с данным водотоком.

Для детальной иллюстрации эффективности разработанной методики было выбрано четыре участка, условно названные по имени крупнейшего города, расположенного внутри. Эти участки обозначены прямоугольниками на нижней карте на Рисунке 3.79. Первые два региона, *I. Линц* и *II. Прага*, представлены на Рисунке 3.80, а два оставшихся, *III. Братислава* и *IV. Будапешт*, показаны на Рисунке 3.81. Оба рисунка организованы по схожей схеме: верхнее изображение показывает исходную ЦМР, среднее — совмещенную, а нижнее — совмещенную с генерализацией. Минимальная длина основного потенциального водотока была установлена в 9 ячеек ЦМР, а второстепенного — в 3. На всех изображениях водотоки одинаковы. Полученные картографические представления доказывают эффективность разработанного метода как для негенерализованной, так и для генерализованной ЦМР. Если для исходной ЦМР очевидна несогласованность с водотоками, то видно, что эта проблема эффективно решена процедурой конфляции.

Чтобы подкрепить визуальные наблюдения статистически, была выполнена численная оценка согласованности между ЦМР и водотоками с использованием индекса согласия *Каппа Коэна* [Cohen, 1960]. Для этого была извлечена дренажная сеть (сеть потенциальных водотоков) из исходной и совмещенной ЦМР, используя  $a = 10$  в качестве порога аккумуляции стока. Полученные потенциальные водотоки были расширены на одну ячейку в обоих случаях, чтобы компенсировать погрешность растеризации триангуляции, которое применяется на заключительном этапе основного блока конфляции. Затем исходные водотоки были растеризованы на ту же сетку, и было проанализировано перекрытие между ними и буферизованной дренажной сетью.

Статистика каппа зависит от доли ячеек водотоков, содержащихся в ячейках буферизованной дренажной сети. Таблица 3.10 содержит идентификатор каждого водотока, значения каппа-индекса согласия  $\kappa_0$  и его асимптотической стандартной ошибки (ASE)  $\hat{\sigma}_0$  для исходной ЦМР и, а также аналогичные статистики  $\kappa_c$  и  $\hat{\sigma}_c$  для совмещенной ЦМР. Полученные значения показывают, что наша методика систематически улучшает согласованность между водотоками и дренажной сетью, неявно представленной на поверхности ЦМР. Среднее значение каппа-индекса в результате конфляции увеличилось с 0.56 (умеренное) до 0.98 (почти идеальное). Таким образом, как визуальная, так и численная оценка результатов свидетельствует о том, что цель данной части диссертационного исследования достигнута. Полученная ЦМР может быть использована для создания географически достоверных карт, корректно совмещающих изображение рельефа и гидрографии местности.





**Рисунок 3.81** – ЦМР и водотоки для фрагментов Братиславы и Будапешта: (а) исходная ЦМР; (б) трансформированная ЦМР; (в) трансформированная и генерализованная ЦМР

**Таблица 3.10** – Статистика согласования между опорными линиями и дренажной сетью, извлеченной из ЦМР

ID	$\kappa_0$	$\hat{\sigma}_0$	$\kappa_c$	$\hat{\sigma}_c$
2	0.758	0.015	0.991	0.003
3	0.83	0.018	0.992	0.004
6	0.56	0.024	0.993	0.003
7	0.415	0.018	0.998	0.001
8	0.457	0.021	0.966	0.006
9	0.433	0.047	1	0
11	0.62	0.018	0.992	0.003
12	0.396	0.022	0.998	0.002
13	0.577	0.029	1	0
14	0.539	0.026	0.977	0.007
15	0.457	0.021	0.982	0.005
16	0.382	0.024	0.929	0.01
17	0.47	0.033	0.977	0.008
18	0.549	0.05	1	0
19	0.509	0.027	1	0
20	0.463	0.027	0.985	0.005
21	0.646	0.03	1	0
22	0.567	0.024	0.997	0.002
23	0.591	0.059	0.978	0.015
25	0.862	0.031	0.877	0.03
26	0.621	0.03	1	0
27	0.513	0.026	0.993	0.004
28	0.553	0.028	1	0
30	0.82	0.021	0.964	0.01
32	0.449	0.013	0.997	0.001
34	0.589	0.021	0.994	0.003
35	0.484	0.018	0.958	0.006
36	0.556	0.019	0.985	0.004

### 3.6.3 Обсуждение

В данном разделе представлена информация, необходимая для облегчения практического применения разработанной методики и учета ее ограничений. Оценка методики выполнена по четырем критериям: параметризация (как выбрать параметры), время обработки (сколько времени требуется для обработки данных), смещение и точность (насколько искажается полученная ЦМР относительно исходной), а также робастность (насколько эффективно метод работает на разных данных, в том числе в присутствии аномалий). Все выводы основаны на данных о рельефе и гидрографии, использованных в Разделе 3.6.2.

#### 3.6.3.1 Параметризация

**Радиус захвата**  $r$  ограничивает максимальное отклонение контрагента от водотока. Минимально возможное значение радиуса равно разрешению ЦМР  $R$ . В этом случае контрагенты будут следовать по ячейкам ЦМР точно под водотоками. Значения  $r$  меньше  $R$  воспрепятствуют выделению контрагентов, за исключением малореалистичного случая, когда водоток проходит точно по центрам ячеек ЦМР. Большие значения  $r$  гарантируют выделение контрагентов, но их достоверность будет под вопросом, если значение  $r$  близко к  $R$ .

Поскольку пространственная точность выделения контрагентов ограничена разрешением ЦМР, целесообразно задавать его кратным  $R$ :

$$r = k \times R, k \in \mathbb{N}^*. \quad (3.43)$$

В проведенном эксперименте использовался радиус захвата, равный 6000 метров, что соответствует  $k = 12$ . По умолчанию в разработанном программном обеспечении для конфляции установлено значение  $k = 10$ . Однако в каждом конкретном случае требуется его уточнение на основе визуальной оценки фактического несовпадения ЦМР и гидрографии. Чрезмерно большое значение  $k$  может привести к выявлению столь же не надежных контрагентов, что и при  $k = 1$ . Поскольку  $r$  определяет также размер области конфляции, большие значения коэффициента приведут к деформации существенной части ЦМР, что может представлять из себя нежелательный эффект.

*Минимальная аккумуляция стока*  $a$  определяет нижний порог площади водосбора контрагентов и, соответственно, их длину. Чем выше  $a$ , тем более значимые пути в дренажной сети будут идентифицированы как контрагенты. Поскольку они должны начинаться и заканчиваться в окрестности начальной и конечной точки соответствующих водотоков, то при задании большого значения  $a$  могут возникнуть ситуации, когда в окрестности не окажется ни одной дренажной ячейки, и в этом случае можно будет найти только контрагент наименьшей стоимости. Поэтому, если величина аккумуляции сто-

ка в соответствующем растре представляет собой количество дренируемых ячеек, то при выборе значения  $a$  в качестве эмпирического правила целесообразно использовать следующее неравенство:

$$a \leq k. \quad (3.44)$$

В нашем эксперименте использовалось значение  $a = 10$ , что удовлетворяет данному условию и позволяет получить густую дренажную сеть, подходящую для целей конфляции.

Штрафной вес  $w$  определяет, насколько строго контрагент минимальной стоимости будет следовать дренажной сети, определенной  $a$ . Чем меньше  $w$ , тем легче алгоритму «перепрыгивать» с одного потенциального водотока на другой в погоне за кратчайшим путем. Поскольку такое поведение не свойственно речным сетям, значение  $w$  должно быть достаточно большим, чтобы воспрепятствовать ему. Исходя из практического опыта, рекомендуется задавать  $w$  кратно 10:

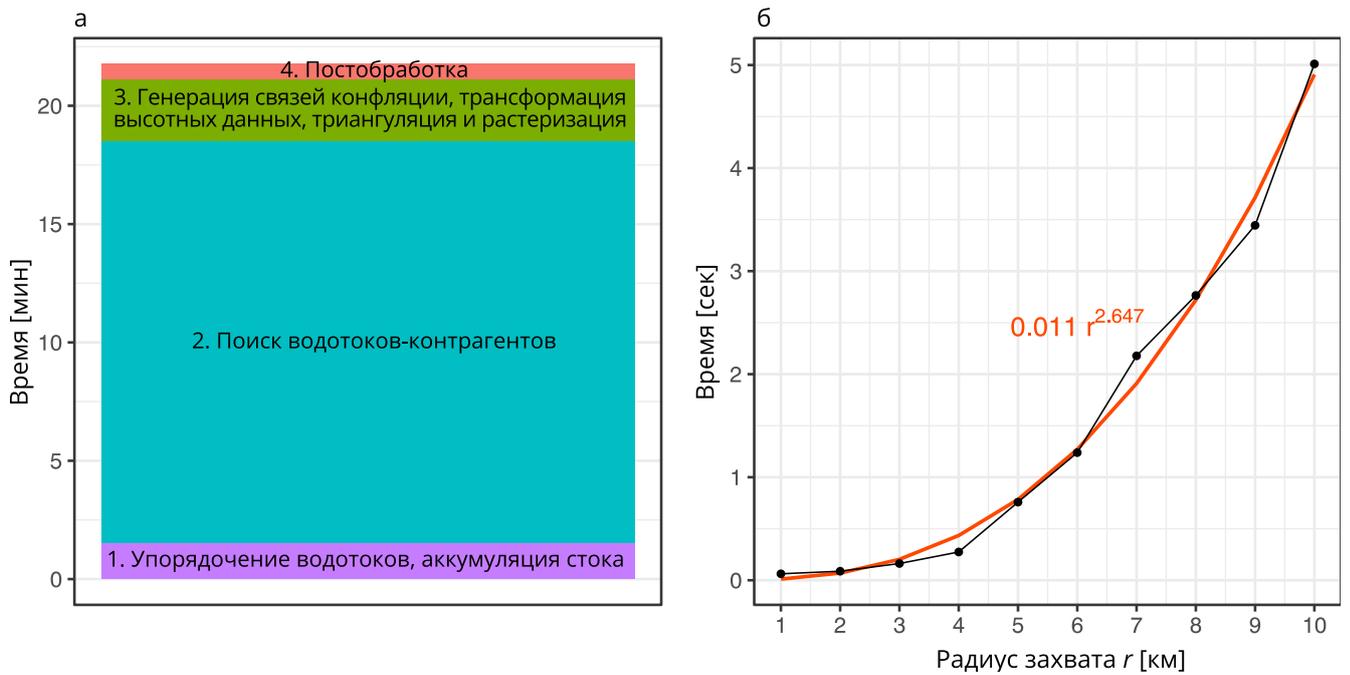
$$w = m \times 10, \quad m \in \mathbb{N}^*. \quad (3.45)$$

В нашем эксперименте использовалась величина  $w = 30$ , что соответствует  $m = 3$ . Это же значение установлено по умолчанию в разработанном программном обеспечении для конфляции. Меньшие значения часто приводят к ситуациям, когда контрагент следует точно по ячейкам под линией водотока или достаточно часто перескакивает между линиями дренажного графа, что выглядит географически неправдоподобно.

### 3.6.3.2 Время обработки

Общее время обработки тестовой ЦМР в соответствии с рабочим процессом, представленным на Рисунке 3.78, составило 21.5 минуты. Структура вычислительных затрат представлена на Рисунке 3.82. Процесс инициализируется упорядочиванием линий и расчетом раstra накопления аккумуляции стока, что заняло примерно 1.5 минуты. Следующий этап, выделение водотоков-контрагентов, является вычислительно наиболее затратным и занял 17 минут (79% общего времени обработки). Оставшиеся этапы конфляции до постобработки заняли 2.5 минуты, а постобработка была завершена за 0.5 минут.

Выполнение непосредственно трассировки контрагентов заняло около 9 минут, а оставшиеся 8 минут из 17 были заняты дополнительными действиями, включая оценку качества. В текущей реализации трассировка кандидатов в обычные контрагенты (которая выполняется в обязательном порядке для каждого водотока) осуществляется значительно быстрее, чем трассировка контрагента наименьшей стоимости (33 секунды и 8 минут 29 секунд в сумме соответственно). Стоит отметить, что при реализации трас-



**Рисунок 3.82** – Время обработки: (а) общее, (б) время трассировки контрагента для опорной линии с  $ID = 13$  в зависимости от радиуса захвата  $r$

сировка обычных контрагентов была запрограммирована на чистом языке Python, в то время как построение контрагентов минимальной стоимости состоит из вызовов скомпилированных функций алгебры карт `arcgis`, которые имеют некоторые накладные расходы на выполнение.

При заданном наборе водотоков и ЦМР продолжительность трассировки контрагентов показала зависимость по степенному закону от величины радиуса захвата  $r$ , что показано для водотока с  $ID = 13$  на Рисунке 3.82 (ось  $r$  для удобства размечена в километрах). При этом продолжительность трассировки контрагента минимальной стоимости не показала зависимости времени выполнения от радиуса захвата.

### 3.6.3.3 Смещение и точность

Процесс конфляции включает трансформацию местоположения высотных данных, в связи с чем возникает вопрос, насколько велико соответствующее смещение в горизонтальном и вертикальном измерениях. Для ответа на этот вопрос были выбраны данные, расположенные внутри области конфляции ( $\sim 29\%$  от всех точек), после чего рассчитаны фактические вектора смещения  $(dx, dy, dz)$  между их положениями до и после трансформации. Компоненты  $(dx, dy)$  определяют величину горизонтального смещения  $dxy = \sqrt{dx^2 + dy^2}$ . Третья компонента  $dz$  — это вертикальное смещение, которое представляет собой разницу между значениями  $z$  до и после трансформации в вершине векторного объекта, который несет в себе высотные данные. Т.е. по сути оценивается то, какую ошибку по высоте вносит восстановление поверхности после трансформации.

Таблица 3.11 и Рисунок 3.83 представляют статистику смещений в табличной и гра-

фической форме. Так как фактические смещения определяются связями конфляции, были рассчитаны статистики для  $(dx, dy)$  и  $dxy$  этих векторов тоже. Таблица 3.11 показывает, что длина вектора смещения в своей массе меньше размера ячейки ЦМР, причем по величине как среднего, так и медианного значения  $dxy$ , которые оба меньше 500 метров. При этом вектора фактических смещений в среднем короче, чем связи конфляции, что ожидаемо, поскольку реальная сила, приложенная к каждой точке, представляет собой суперпозицию нескольких связей конфляции, а величина силы уменьшается до нуля вблизи тождественных связей.

При этом наблюдается значительная асимметрия в смещении, которая отражается в компонентах векторов  $dx$  и  $dy$ . Если  $dx$  слабо отрицательна и близка к нулю, то  $dy$  имеет ярко выраженное положительное значение, что означает, что высотные данные в процессе конфляции систематически смещаются в положительном направлении оси  $Y$ . Этот факт наглядно иллюстрируют графики двумерной плотности вероятности для  $(dx, dy)$  на Рисунке 3.83а-б. Для оценки использовалось гауссовское ядро с шириной полосы пропускания  $\sigma = 750$  метров. Поверхность плотности для горизонтального смещения имеет максимум около нулевого значения, но при этом обе поверхности имеют длинный хвост в положительном направлении оси  $Y$ . Это наблюдение показывает систематический сдвиг между ЦМР и гидрографией, который может быть уменьшен путем предварительного глобального (аффинного) преобразования ЦМР.

Несмотря на то, что и для связей конфляции, и для векторов смещения по данным Таблицы 3.11 наибольшая длина близка к радиусу захвата (6000 м), ее интерквартильный размах ( $IQR = Q_3 - Q_1$ ) весьма компактен и близок к разрешению ЦМР. Эмпирическая функция распределения  $dxy$  (Рисунок 3.83в) показывает, что 66% всех высотных данных были смещены на расстояние, не превышающее разрешения ЦМР, а 95% всех высотных данных не были смещены дальше 1480 м, что примерно равно расстоянию, покрываемому тремя ячейками ЦМР.

Источником вертикального смещения в нашем случае является процедура, в ходе которой результирующий растр восстанавливается путем растеризации триангуляционной поверхности (если бы растеризация не требовалась, вертикальное смещение было бы равно нулю). Центры результирующих ячеек не совпадают с расположением вершин высотных данных, поэтому в процессе растеризации они интерполируются. Согласно Таблице 3.11 и среднее, и медиана  $dz$  очень близки к нулю, а интерквартильный размах, определяемый  $Q_1 = -1.52$  и  $Q_3 = 1.76$ , настолько узок, что может считаться незначительным для ЦМР с таким большим размером ячейки.

График плотности вероятности для  $dz$  на Рисунке 3.83г подтверждает, что большинство значений сосредоточено в нескольких метрах вокруг нулевого смещения. Эффективный диапазон вертикальных смещений, определяемый значениями не далее  $1,5IQR$  от  $Q_1$  и  $Q_3$ , составляет около  $(-6.5, 6.5)$  и обозначен «усами ящика» в нижней части Рисунка 3.83г. Эмпирическая функция распределения  $|dz|$  (Рисунок 3.83д) пока-

Таблица 3.11 – Статистика смещения

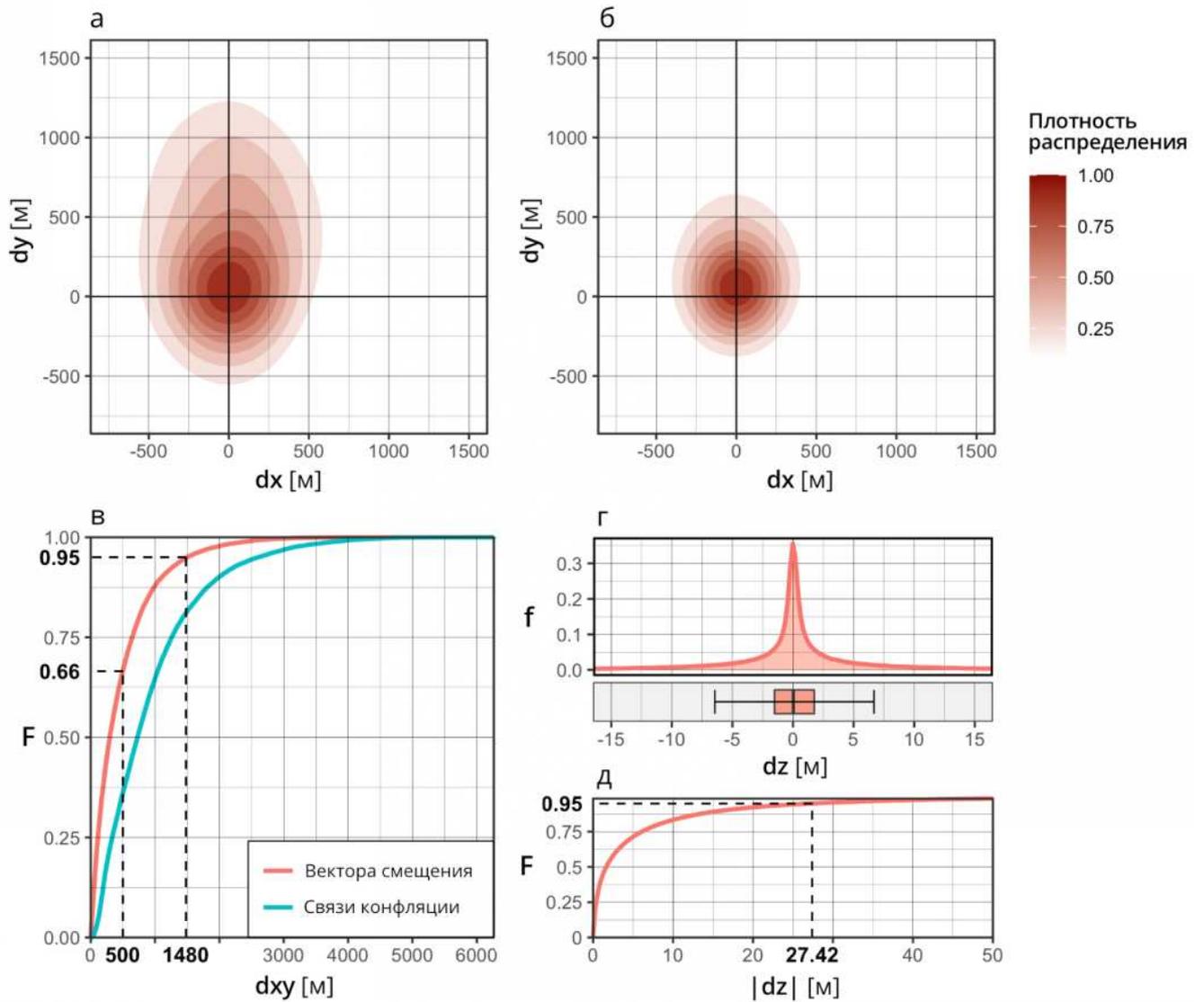
Переменная	Связи конфляции	Вектора смещения
dx [м]		
Среднее (СКО)	-49 (761)	-28 (397)
Медиана (Q1, Q3)	-16 (-302, 225)	-4 (-126, 92)
Размах	[-5404, 6210]	[-4461, 3747]
dy [м]		
Среднее (СКО)	301 (926)	179 (533)
Медиана (Q1, Q3)	256 (-111, 803)	98 (-22, 389)
Размах	[-4675, 4390]	[-4675, 4390]
dxу [м]		
Среднее (СКО)	936 (808)	467 (506)
Медиана (Q1, Q3)	726 (334, 1259)	302 (114, 643)
Размах	[4, 6264]	[0, 5137]
dz [м]		
Среднее (СКО)	–	0.07 (13.77)
Медиана (Q1, Q3)	–	0.03 (-1.52, 1.76)
Размах	–	[-466.49, 390.49]
Итого		
Общее число	22606	393622

зывает, что 95% всех высотных данных были смещены по вертикали не более чем на 27.42 м, что также вполне приемлемо для ЦМР низкого пространственного разрешения, покрывающей горную местность.

В заключение этого раздела следует отметить, что несмотря на то, что процесс конфляции может значительно исказить поверхность, фактические смещения оказались весьма умеренными, что потенциально может сохранить ЦМР пригодной не только для картографических, но и для аналитических приложений.

Аналитические приложения требуют от нас перехода от простой характеристики смещения к более полновесному понятию точности ЦМР. При этом не вполне ясно, как именно должно оцениваться изменение точности ЦМР, когда одни её части остаются неподвижными, а другие смещаются. Обеспечивает ли анализ векторов смещения надежную оценку изменения точности ЦМР – вопрос, который нуждается в дальнейшей проработке.

Разработанный метод является жадным в том смысле, что стремится к точному совмещению водотоков и их контрагентов, и поэтому допускает любое смещение в пределах заданного радиуса захвата. Для аналитических целей может быть желательно ограничить величину смещения меньшим значением. Этого можно добиться, как уменьшив радиус захвата, так и ограничив длину связей конфляции. В последнем случае все векто-



**Рисунок 3.83** – Распределение характеристик смещения при конфляции: а) плотность распределения компонент  $dx, dy$  связей конфляции; б) плотность распределения компонент  $dx, dy$  векторов смещения; в) функция распределения  $F$  длин  $dxy$  связей конфляции и векторов смещения; г) плотность распределения  $f$  вертикальных смещений  $dz$ ; функция распределения  $F$  модуля вертикального смещения  $|dz|$ .

ра связей, длина которых превышает заданный порог, могут быть принудительно укорочены. Адаптивное изменение размера области конфляции в зависимости от локального отклонения контрагента от водотока также может улучшить качество результата и минимизировать избыточные искажения.

Очевидно также, что трансформация ЦМР должна влиять на значения геоморфометрических и гидрологических характеристик в области конфляции. Анализ этих изменений выходит за рамки настоящего раздела диссертационного исследования, посвященного картографическому применению конфляции ЦМР. Однако такой анализ должен стать одним из основных направлений дальнейшего развития и оценки ограниченной предложенной методики.

#### 3.6.3.4 Робастность

Под робастностью можно понимать способность метода хорошо работать в различных ситуациях и противостоять сложностям, которые могут возникнуть из-за сочетания входных данных и выбранных параметров. Свойства входных данных, влияющие на эффективность метода, можно рассматривать с количественной и качественной точек зрения.

**Количественный** аспект входных данных связан с их размерностью, которая включает в себя пространственный домен, уровень детализации и объем данных.

Пространственный домен определяется множеством местоположений, в которых определены анализируемые объекты. Для гидрографии это множество равно геометрии линий, а для ЦМР — всем ячейкам, в которых есть данные (исключая ячейки NODATA). И то и другое можно более грубо задать охватом — предельными значениями координат. Конфляция может быть выполнена только области пересечения охватов гидрографии и ЦМР. При этом для обеспечения корректности процедуры необходимо обрезать оба набора данных прямоугольником области пересечения.

Уровень детализации гидрографических и высотных данных можно определить как количество информации, представленной в этих наборах данных на единицу площади. Он зависит от пространственного разрешения (плотности точек и размера ячейки), granularity линий и ЦМР, а также прочих особенностей их представления заданном пространственном разрешении. В Разделе 3.6.2 было продемонстрировано, что несоответствие между детальным рельефом и генерализованной гидрографией эффективно устраняется разработанной методикой. Однако необходимы дополнительные эксперименты для оценки качества результатов в обратной ситуации, когда гидрография более подробна.

Размер данных влияет на время их обработки. Очевидным решением для работы с большими наборами данных является разбиение их на части (тайлы) и обработка тайлов по отдельности. Текущая реализация предложенной методики не предусматривает

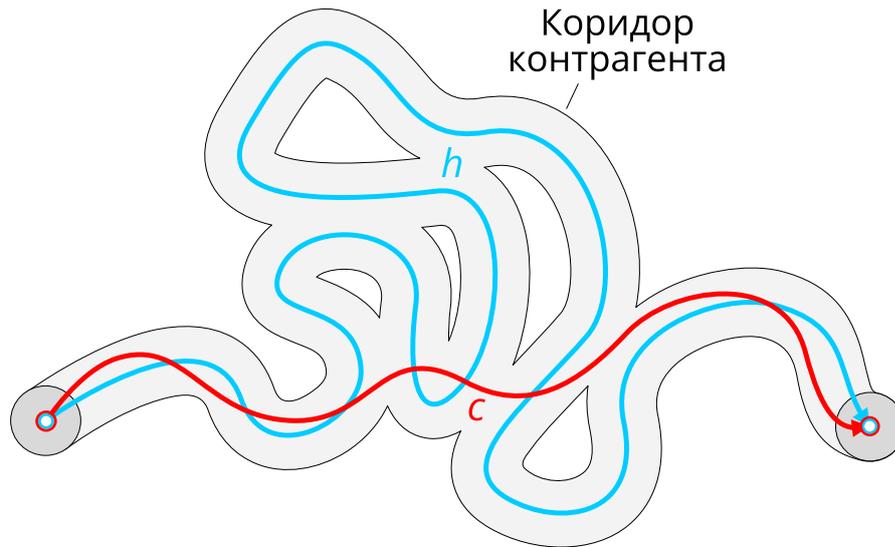
такого решения, этот вопрос будет одним из важнейших направлений будущих исследований. Ожидается, что проблемной частью реализации будет разрешение граничных эффектов, которые будут возникать из-за независимой конфляции тайлов.

**Качественный** аспект входных данных связан с их топологической корректностью и в основном существенен для гидрографической составляющей конфляции ЦМР. Разработанная методика не накладывает никаких ограничений на входные гидрографические данные. Если линии неправильно оцифрованы или не представляют собой топологически корректную сеть, методика все равно будет работать. Подобная робастность достигается за счет продуманной организации алгоритма. На этапе порядковой классификации топологические связи между гидрографическими линиями учитываются, но не обязаны быть не пустыми.

Если линия изолирована (не имеет общих точек с другими линиями), то она просто получает порядок 1 по классификации Хака. На этапе извлечения контрагентов сам процесс трассировки отделен от применения топологических правил 1 — 5, а значит, не зависит и от пространственных отношений входных линий. При этом контрагент всегда трассируется в направлении водотока, даже если он оцифрован в неправильном направлении. Это свойство гарантирует, что связи конфляции будут распределены в правильном порядке между начальными и конечными точками контрагента и водотока. Остальные этапы совмещения до постобработки вообще не затрагивают водотоки. Если же они тщательно подготовлены, то метод будет выполнять конфляцию в полном соответствии с их топологической структурой. Поэтому можно сказать, что предложенная методика способна принимать лучшие свойства входных данных и устойчива к ошибкам в них.

В то же время возможны сложные ситуации, в которых текущая реализация процесса конфляции будет неэффективной. Поскольку стандартная операция алгебры карт, используемая для извлечения контрагентов наименьшей стоимости, не может учитывать ни расстояние Хаусдорфа, ни расстояние Фреше, к ним можно применить только слабое ограничение, опирающееся на расчет направленного расстояния Хаусдорфа. На Рисунке 3.84 приведен пример водотока с близко расположенными меандрами, для которого полученный контрагент недостоверен. Потенциально для выделения более качественных контрагентов можно использовать более специфические методы, разработанные для конфляции векторных карт и способные построить кратчайший путь, ограниченный расстоянием Фреше [Gheibi, Maheshwari, Sack, 2016; Chambers *et al.*, 2020]. Однако, поскольку соответствующие алгоритмы часто основаны на динамическом программировании и рекурсивных процедурах, необходимо также оценить их вычислительную производительность в случае цифровых моделей рельефа, состоящих из тысяч и миллионов ячеек.

Следует также подчеркнуть, что если пространственная корректировка водотоков более предпочтительна из-за более высокой точности ЦМР, то она может быть легко достигнута с помощью разработанного метода. Для этого необходимо перевернуть связи



**Рисунок 3.84** – Ненадежный контрагент наименьшей стоимости на меандрирующем участке водотока конфляции и вместо данных о рельефе выполнить смещение объектов гидрографии.

### 3.6.4 Итоги

Цифровые модели рельефа часто используются в сочетании с гидрографическими данными для пространственного анализа и картографирования. Качество и точность результатов в значительной степени зависят от их пространственного соответствия. В то же время, растущее разнообразие источников пространственных данных, в том числе по *детальности*, делает несогласованность между разнородными данными широко распространенной проблемой, которая обычно решается путем конфляции. При этом до настоящего момента частный случай конфляции цифровых моделей рельефа и гидрографических данных не получил должного внимания. В настоящем разделе диссертации разработано комплексное решение этой проблемы, включающее методику, алгоритмы и программное обеспечение.

Основные выводы настоящего раздела могут быть обобщены следующим образом:

1. Процесс конфляции может быть основан на определении путей на поверхности ЦМР, наиболее похожих на эталонные гидрографические линии — водотоков-контрагентов. Для оценки сходства и ограничения кандидатов в контрагенты можно использовать несколько мер, в том числе направленное расстояние Хаусдорфа, расстояние Хаусдорфа и расстояние Фреше, но из практических соображений по умолчанию используется самое слабое ограничение — направленное расстояние Хаусдорфа.
2. Для сохранения топологических связей водотоков они должны быть упорядочены. Правильное упорядочение позволяет установить однозначную последовательность выделения контрагентов, при которой могут быть воспроизведены

отношения подчинения между водотоками, отражающие топологию сети. В работе показано удобство порядковой классификации Хака, с уточнениями для разветвленного русла.

3. Комбинация подходов, основанных на использовании линий тока и путей наименьшей стоимости, позволяет выделять контрагенты в разнообразных случаях, включая нестандартные, такие как разветвленное русло и неправильное направление оцифровки гидрографических линий.
4. Топологические отношения между упорядоченными водотоками могут быть воспроизведены в сети контрагентов с помощью ряда топологических правил. Эти правила обеспечивают корректное расположение узлов.
5. Выполнение конфляции применительно к векторным данным (точкам или структурным линиям) позволяет абстрагироваться от формата ЦМР и использовать стандартные алгоритмы конфляции. Еще одним преимуществом векторного подхода является легкая интеграция этапа конфляции в процесс генерализации ЦМР.

Разработанная методика позволяет эффективно согласовывать ЦМР с водотоками, что демонстрируется на многочисленных картографических примерах и подтверждается значительным улучшением статистики согласия между водотоками и дренажной сетью на поверхности ЦМР. По сравнению с существующими подходами к конфляции и прорезанию водотоков разработанный подход не создает новых форм рельефа на поверхности ЦМР, а распознает уже существующие и соответствующие гидрографическим линиям, что является значительным достижением.

Цифровые модели рельефа, трансформированные с помощью представленной в настоящем параграфе методики, могут быть использованы для создания высококачественных карт, объединяющих данные о рельефе и гидрографии различной детальности.

### 3.7 Взаимная координатная привязка цифровых геоизображений разной детальности<sup>9</sup>

**Координатная привязка** родственна конфляции, но носит более общий характер, поскольку связана с полноценной трансформацией данных при неизвестной системе координат. Она может осуществляться по опорным точкам, и в этом случае она требует задания их местоположения в системе координат с известными параметрами. Перспективным направлением компьютерного зрения, которое может быть применено в данной области, является *технология сопоставления изображений* (англ. *image matching*). Она яв-

<sup>9</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Мухаметшин, Самсонов, Лурье, 2024]

ляется подходом к поиску ключевых признаков на двух изображениях с целью нахождения уникальных точек, присутствующих на паре изображений [Brown, 1992]. Применительно к геоизображениям данный подход к настоящему моменту не получил должного развития. В связи с этим представляется целесообразным оценить его применимость при решении ранее обозначенной проблемы координатной привязки.

В исследовании [Мухаметшин, Самсонов, Лурье, 2024] была разработана модель координатной привязки геоизображений — карт, космических снимков и их сочетаний — с использованием технологии *LoFTR* (*Local Feature Matching with Transformers*), реализующей поиск и сопоставление опорных точек с использованием компьютерного зрения и опциональной кластеризации изображений [Sun et al., 2021]. Особенностью модели заключается в том, что она способна в том числе работать с изображениями разной детальности и тематики.

### 3.7.1 Методика

#### 3.7.1.1 Общие сведения про алгоритм LoFTR

В процессе анализа изображений алгоритм LoFTR [Sun et al., 2021] извлекает плотное поле возможных совпадений и позже уточняет хорошие совпадения до субпиксельной точности. Это означает, что каждая точка в паре изображений потенциально может быть сопоставлена с другой. В соответствии со схемой работы *LoFTR* выделяется 4 модуля, которые последовательно применяются в процессе работы (Рисунок 3.85). Так, в первом модуле исходными данными являются 2 изображения, из которых извлекаются карты признаков двух уровней с помощью процедуры свертки. С каждой новой итерацией производится уменьшение размерности карт признаков, что позволяет снизить вычислительные затраты до более низкого уровня в последующих модулях.

Во втором модуле полученные карты признаков преобразовываются в одномерные массивы, которым затем присваивается кодировка в соответствии с их местоположением на изображении. Далее в модуле *LoFTR* осуществляется перекрестная проверка значений для нахождения взаимосвязей между ними (как внутри одного изображения, так и между ними). После этого к двум изображениям применяется градиент в соответствии с полученной информацией о возможном взаимном расположении похожих точек (похожим областям на изображениях присваивается одинаковый цвет).

В третьем модуле точки предварительно сопоставляются на основе схожести их векторов признаков, а в четвертом — предварительные совпадения уточняются до субпиксельного уровня. Для этого вокруг каждой точки выделяется локальное окно, в пределах которого точка с одного изображения сравнивается со всеми точками с другого изображения в пределах данного окна. Между наиболее похожими точками устанавливается соответствие [Sun et al., 2021].

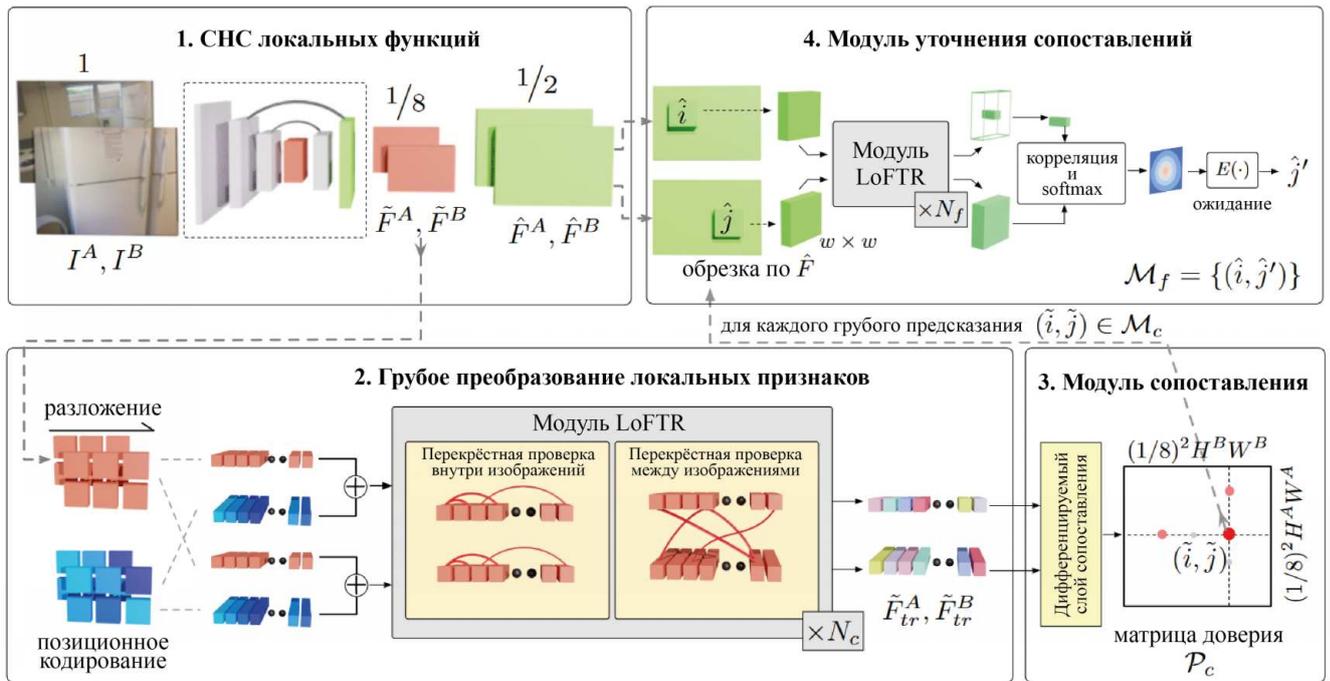


Рисунок 3.85 — Схема работы алгоритма LoFTR. Адаптировано с [Sun et al., 2021].

### 3.7.1.2 Координатная привязка

Описанный алгоритм LoFTR используется в разработанной модели координатной привязки на этапе обнаружения опорных точек. Модель можно разделить на несколько модулей, через которые последовательно проходят промежуточные данные (Рисунок 3.86). В соответствии с приведённой схемой, они могут быть описаны следующим образом:

1. Загрузка привязанной и непривязанной карт в виде массивов чисел.
2. Опциональный выбор проведения кластеризации (в случае отказа от нее, дальнейшая работа будет проводиться с исходными картами, в другом случае, необходимо указать единое число кластеров, на которые будут кластеризованы карты; данный этап необходим для предоставления возможности выделения более контрастных границ между разными классами объектов, приведенными на картах); наиболее оптимальное число кластеров определяется путем итеративного процесса, требующего анализа и экспертной оценки (проводится оценка, насколько корректно кластеры соответствуют реальным объектам на карте с учетом особенностей картографических данных — например, для карты землепользования может быть достаточно небольшого числа кластеров, соответствующих основным типам покрытия (лесная и травянистые растительности, водные объекты, городские территории)).
3. Преобразование карт в тензоры (разложение на три канала цветового пространства RGB и определение числа фрагментов, на которые будут поделены растры)

- подготовка к процедуре свертки.
- 4. Загрузка тензоров в LoFTR, в котором производится обработка изображений в соответствии с ранее описанными деталями.
- 5. Получение списка соответственных (опорных) точек и извлечение их пиксельных координат с обеих карт.
- 6. Преобразование пиксельных координат в географические (пиксельные координаты точек непривязанной карты соотносятся с географическими координатами тех же точек на привязанной карте).
- 7. Осуществление привязки непривязанной карты путем присвоения географических координат точек с указанием системы координат (СК) и проекции привязанной карты.
- 8. Трансформация результирующей привязанной карты с использованием одного из предварительно выбранных типов трансформации.
- 9. Оценка точности привязки путем расчета невязки, а также с помощью ручной и визуальной проверки корректности определения соответственных точек.

### 3.7.2 Реализация и апробация

Реализация модели была выполнена на языке программирования Python с использованием библиотек: *OpenCV*, *NumPy*, *Sys*, *Kornia*, *Torch*, *Matplotlib*, *GDAL*, *Os*, *Time*. Отдельные блоки кода, представленные в файле, проводят последовательные этапы обработки исходных данных.

Тестирование модели производилось на 10 парах геоизображений, которые представляли собой электронные, отсканированные и веб-карты, а также спутниковые снимки *Sentinel-2* и *Landsat-8*. Их источниками выступили: Национальный атлас РФ, региональные атласы, веб-карта «Яндекс.Карты» и сайт Геологической службы США. Комбинации пар формировались исходя из возможных ситуаций, в которых может понадобиться взаимная географическая привязка двух геоизображений.

В каждой паре с помощью модуля привязки QGIS вручную было привязано одно изображение, на основе которого затем автоматизировано будет привязано второе. Параметры точности привязки представлены в Таблице 3.12. Целевая проекция — равноугольная цилиндрическая проекция Меркатора, выбор которой был обусловлен тем, что она является наиболее компромиссной, т. к. используется на веб-картах и, следовательно, в некоторых парах геоизображений. Также она ориентирует изображения на север и при ограниченных размерах территорий (охват примеров локальный и региональный) не обладает большой вариацией искажений. Таким образом обеспечивалось более качественное совпадение привязанного и непривязанного изображений.

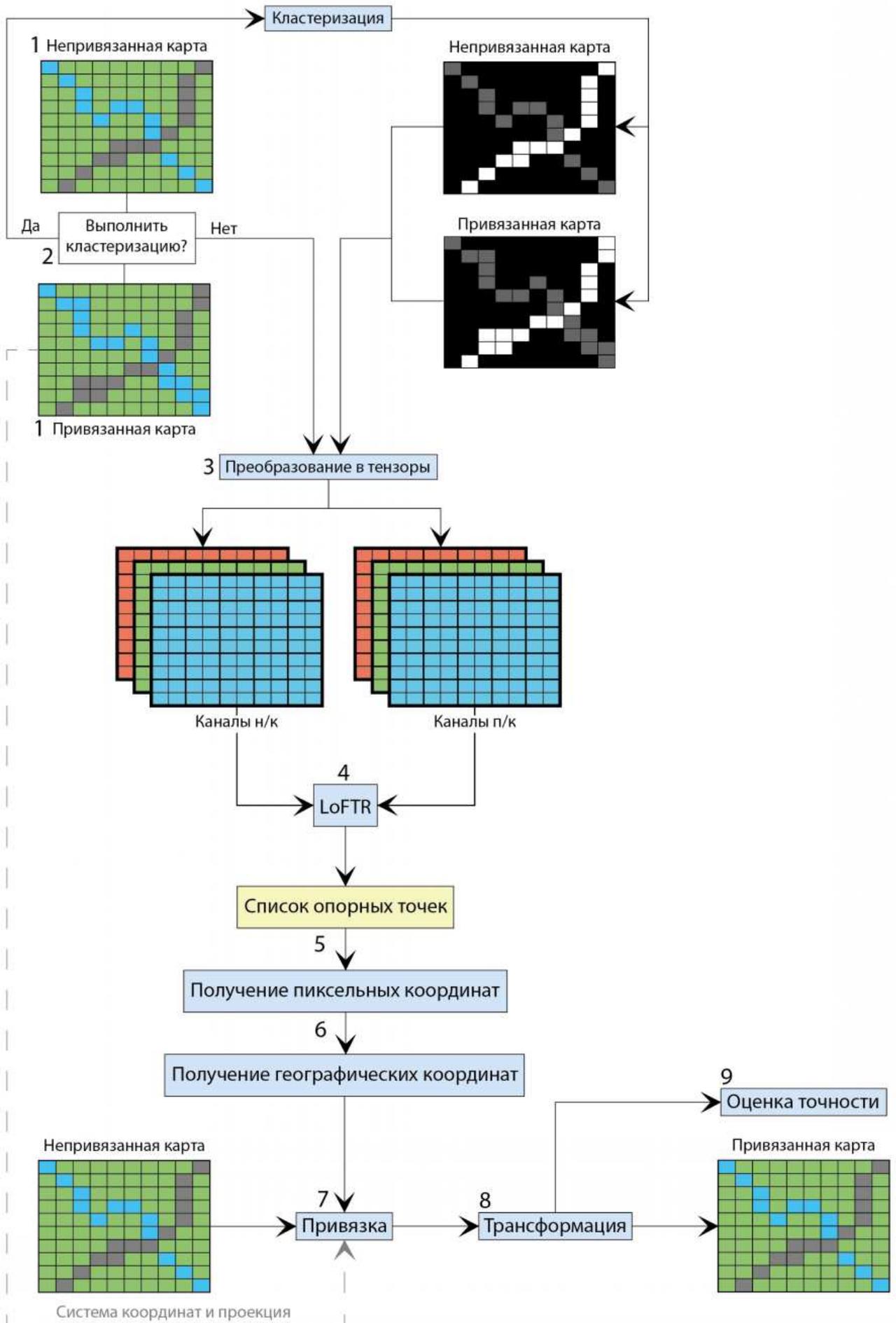


Рисунок 3.86 – Модель координатной привязки геоизображения

**Таблица 3.12** – Точность ручной привязки одного изображения из пары

№ пары	Название изображения, привязываемого вручную	Количество опорных точек	Тип трансформации	Общая невязка, пиксели
1	Европейская часть России (1:7 500 000)	10	Полиномиальная II степени	1
2	Фрагмент веб-карты «Яндекс.Карты» на территорию Иркутской области	5	Аффинная	0.5
3	Западная Сибирь (1:7 500 000)	29	Полиномиальная III степени	2.6
4	Овражность Алтайского края (1:2 000 000)	25	Полиномиальная III степени	2.3
5	Европейская часть России (Север) (1:2 500 000)	26	Полиномиальная III степени	1.3
6	Административная карта Вологодской области (1:2 500 000)	33	Тонкостенный сплайн	0
7	Спутниковый снимок Landsat-8 на центральную часть оз. Байкал	3	Аффинная	0
8	Фрагмент веб-карты «Яндекс.Карты» на центральную часть г. Санкт-Петербург	3	Аффинная	0
9	Спутниковый снимок Sentinel-2 на северный участок Германии	5	Аффинная	0.4
10	Западный спутниковый снимок Landsat-8 на территорию Греции у залива Термаикос	3	Аффинная	0

Тип трансформации определялся в зависимости от особенностей изображения: *аффинная* (отсутствие необходимости сильного преобразования изображения ввиду схожести пропорций контуров со вторым изображением в паре в данном масштабе), *полиномиальная II степени* (увеличение степени схожести контуров со вторым изображением в паре за счет минимизации влияния особенностей проекции), *полиномиальная III степени* и *естественный сплайн* применялись в более сложных случаях, когда требовалось сильное растяжение растра.

Далее пары привязанных и непривязанных изображений были загружены в разработанную модель, которая производила поиск и сопоставление общих точек, на основании чего затем осуществляла географическую привязку одного изображения на основании другого, преобразуя пиксельные координаты в спроецированные на местности (в данном случае — в проекции Меркатора). Среднее время выполнения скрипта для одной пары геоизображений зависит от объема входных растров, однако для представленных пар изображений составило 15,27 секунд для компьютера с 64-разрядной операционной системой *Windows* и процессором *Intel Core i5-10500H 2.50GHz (12 CPUs)*, *ОЗУ 8 ГБ*, видеокартой *NVIDIA GeForce GTX 1650 4 ГБ*.

### 3.7.3 Результаты и их анализ

Оценка качества выделения точек производилась тремя способами:

1. *Визуальная оценка положения эпиполярных (соединительных) линий.* Например, если изображения имеют одинаковую ориентировку, то линии должны идти приблизительно в одном направлении (параллельно) и не пересекаться друг с другом.
2. *Количественная оценка точности трансформации — невязки.* Результат работы модели — привязанное изображение и опорные точки, которые были автоматизировано отобраны для привязки. Последние загружались в модуль привязки QGIS, где устанавливалась величина невязки при используемых параметрах трансформации непривязанного изображения.
3. *Количественная оценка точности сопоставления опорных точек.* В процессе работы модели извлекались изображения, на которых были идентифицированы опорные точки и эпиполярные линии, соединяющие их между собой. Из опорных точек на двух геоизображениях составлялась случайная выборка из 10 точек, между которыми вручную были проведены эпиполярные линии. Полученный ручной результат сравнивался с автоматизированным. Таким образом оценивалась величина векторов смещения (далее — *BC*), характеризующая точность обнаружения точек алгоритмом LoFTR.

Таким образом, была выполнена автоматизированная привязка различных комби-

наций 10 пар геоизображений с помощью разработанной модели. Полученные числовые характеристики процедур привязки представлены в Таблице 3.13, анализ которой приведен далее, после вычисления зависимостей. Количество пар геоизображений было увеличено в связи с тем, что некоторые непривязанные растры были дополнительно преобразованы — повернуты или обрезаны до более локальных областей (пары геоизображений 3, 5, 8). *Обратим также внимание на то, что во многих случаях изображения различаются своими масштабами, а значит обладают разной детальностью.*

Между значениями невязки и ВС был рассчитан *ранговый коэффициент корреляции Спирмена*, который позволяет оценить взаимосвязь двух наборов чисел без проверки их на наличие нормального распределения. Значение равно 0.95 ( $p < 0.00001$  (результат значим при  $p < 0.05$ )), что говорит о наличии очень высокой связи между данными показателями. Чем менее точно точка определена моделью (большое значение ВС), тем большая невязка получается в результате трансформации.

На Рисунке 3.87а приведена диаграмма рассеяния, которая также демонстрирует, что рассматриваемые показатели имеют сильную связь — с возрастанием значений невязки возрастает значение ВС (достоверность аппроксимации 0.804).

Стоит отметить, что подавляющее большинство точек на изображениях было выделено с высокой степенью точности (практически идентично качеству ручного выделения). Полученные значения ВС включают в себя также значения точек, которые были определены моделью недостаточно корректно. Однако таких точек встречается небольшое число — в среднем 91.9% опорных точек определяются моделью верно (Таблица 3.13).

Следует зафиксировать, что коэффициент корреляции между процентом верно сопоставленных опорных точек и значениями невязки составляет 0.63 ( $p = 0.022$  (результат значим при  $p < 0.05$ )). Полученное значение говорит о наличии обратной связи между двумя параметрами. Так, с увеличением процента верно сопоставленных моделью точек наблюдается снижение величины невязки, что отражено на Рисунке 3.87б. Достоверность аппроксимации в данном случае достигает 0.656.

Таким образом, в работе модели прослеживается логика классической процедуры ручной привязки. Далее показатели невязки и ВС будут употребляться единым термином — *точность*. Некоторые пары геоизображений, описанные в Таблице 3.13, приведены далее (Рисунки 3.88 — 3.90, синие точки — ключевые, зеленые точки — отфильтрованные опорные, красные — исключенные из списка опорных).

**Таблица 3.13** – Характеристика процедур привязки различных комбинаций геоизображений (пары 1-5)

№	Пара геоизображений	Количество ключевых точек	Количество опорных точек ( $N_{ref}$ )	Количество опорных точек после ручной фильтрации ( $N_{ver}$ )	Тип трансформации	Невязка, пиксели	Средняя длина вектора смещения, пиксели	*Процент верно определённых моделью точек	Размер, пиксели (разрешение, метры)
1	«Европейская часть России» (1 : 7 500 000), «Европейская часть России (Север)» (1 : 2 500 000)	686	121	121	Полиномиальная II степени	3	3	100	914×621 (3656), 624×435(4630)
2	Фрагмент карты «Яндекс.Карты», «Физическая карта Юга области» (1 : 2 000 000)	599	91	88	Полиномиальная III степени	3.6	5	97	624×755(1699), 624×757(1249)
3	Фрагмент карты «Яндекс.Карты», повёрнутая «Физическая карта Юга области» (1 : 2 000 000)	355	35	32	Аффинный	4.1	6	91	624×755(1699), 624×757(1660)
4	«Западная Сибирь» (1 : 7 500 000), «Почвообразующие породы» (1 : 6 000 000)	367	118	113	Полиномиальная III степени	3.1	4	96	624×905(5296), 447×624(6301)
5	«Западная Сибирь» (1 : 7 500 000), фрагмент «Почвообразующие породы» (1 : 6 000 000)	153	18	16	Аффинный	6	5	89	624×905(5296), 244×254(9200)

\* после ручной фильтрации (отобранных для привязки)  $100 * N_{ver} / N_{ref}$

**Таблица 3.14** – Характеристика процедур привязки различных комбинаций геоизображений (пары 6-10)

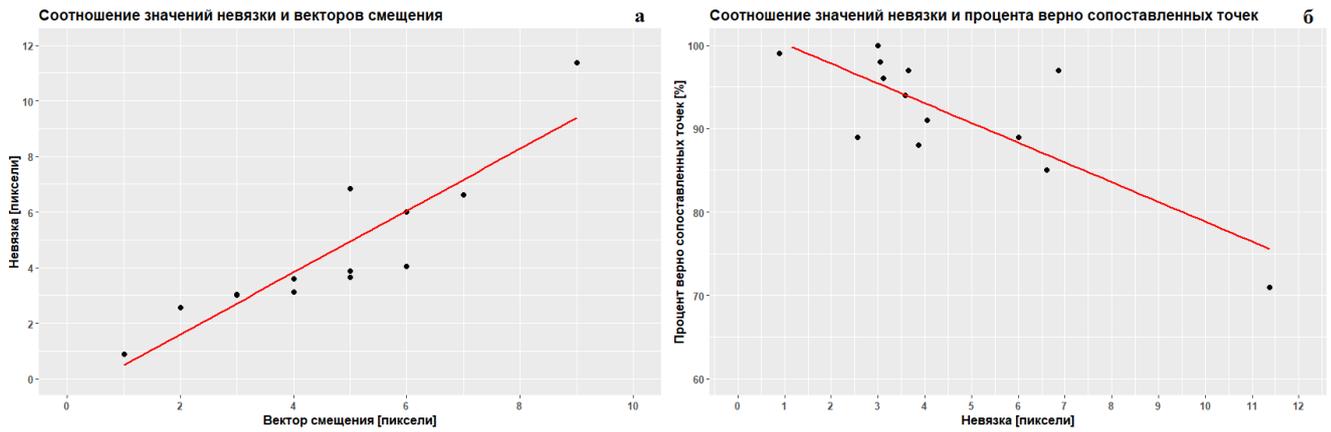
№	Пара геоизображений	Количество ключевых точек	Количество опорных точек ( $N_{ref}$ )	Количество опорных точек после ручной фильтрации ( $N_{ver}$ )	Тип трансформации	Невязка, пиксели	Средняя длина вектора смещения, пиксели	*Процент верно определённых моделью точек	Размер, пиксели (разрешение, метры)
6	«Овражность» (1 : 2 000 000), «Совхозы Алтайского края» (1 : 1 000 000)	427	51	50	Полиномиальная III степени	3	3	98	624×567(1265), 614×624(1198)
7	«Европейская часть России (Север)» (1 : 2 500 000), «Почвенная карта» (1 : 2 000 000)	326	150	128	Полиномиальная III степени	6.6	7	85	695×583(2428), 624×531(2268)
8	«Европейская часть России (Север)» (1 : 2 500 000), фрагмент «Почвенная карта» (1 : 2 000 000)	118	25	22	Аффинный	3.9	5	88	695×583(2428), 624×179(2163)
9	«Административная карта» (1 : 2 500 000), «Карта растительности» (1 : 2 500 000)	222	63	61	Аффинный	6.9	5	97	624×274(2404), 624×242(2473)
10	Спутниковый снимок Landsat-8 на центральную часть оз. Байкал, «Леса» (1 : 4 000 000)	168	17	12	Аффинный	11.4	9	71	551×610(671), 624×629(393)

\* после ручной фильтрации (отобранных для привязки)  $100 * N_{ver} / N_{ref}$

**Таблица 3.15** – Характеристика процедур привязки различных комбинаций геоизображений (пары 11-13)

№	Пара геоизображений	Количество ключевых точек	Количество опорных точек ( $N_{ref}$ )	Количество опорных точек после ручной фильтрации ( $N_{ver}$ )	Тип трансформации	Невязка, пиксели	Средняя длина вектора смещения, пиксели	*Процент верно определённых моделью точек	Размер, пиксели (разрешение, метры)
11	Фрагмент карты «Яндекс.Карты», «План столичного города Санкт-Петербурга»	207	33	31	Аффинный	3.6	4	94	624×509 (16), 624×531 (16)
12	Спутниковый снимок Sentinel-2, спутниковый снимок Landsat-8	508	172	171	Аффинный	0.9	1	99	624×606(311), 624×427(392)
13	Западный спутниковый снимок Landsat-8, восточный спутниковый снимок Landsat-8	580	157	140	Аффинный	2.6	2	89	710×720(440), 710×720(435)

\* после ручной фильтрации (отобранных для привязки)  $100 * N_{ver} / N_{ref}$



**Рисунок 3.87** — а) Диаграмма рассеяния между значениями невязки и ВС; б) Диаграмма рассеяния между значениями невязки и процентами точек, отобранных для привязки (красная линия – линия тренда)

Лучшие значения точности наблюдаются при привязке *спутниковых снимков: различной детальности с большим перекрытием и одинаковой детальности с небольшим перекрытием* — невязка 0.9 и 2.6 пикселя и ВС 1 и 2 пикселя для пар геоизображений 12, 13 соответственно (Таблица 3.13). Это означает, что местоположения соответственных точек на двух спутниковых снимках можно определить более точно, чем на двух различных картах или между спутниковым снимком и картой, т. к. при поиске точек на изображениях сравнивается фактическое состояние одного и того же объекта — земной поверхности. В случае сравнения карт, каждая из которых имеет отличное друг от друга содержание, используемые условные знаки и уровень генерализации, достоверность определения корректного местоположения соответственных точек снижается.

Высокая точность наблюдается при привязке *карт с небольшими отличиями в содержании* — общегеографических карт: между собой, с веб-картами, с тематическими картами. Примерами таких пар геоизображений являются: 1 (невязка — 3, ВС — 3), 2 (невязка — 3.6, ВС — 5), 4 (невязка — 3.1, ВС — 4), 7 (невязка — 6.6, ВС — 7), 11 (невязка — 3.6, ВС — 4) (Таблица 3.13). Это объясняется тем, что тематические карты по-своему представляют рисунок объектов картографируемого явления. Он может быть сегментирован на полигоны различного размера, формы, цвета. К тому же на фоне тематического содержания может теряться общегеографическое, которое в таких случаях будет выступать основным при поиске ключевых точек.

Менее высокие значения точности получаются при привязке *карт с разным тематическим содержанием*, но при условии наличия объектов общегеографического содержания, которые могут послужить источниками качественных отличительных признаков (яркие границы, водные объекты, береговая линия и т. д.). Примерами таких случаев являются процедуры привязки пар карт 6 (невязка — 3, ВС — 3), 9 (невязка — 6.9, ВС — 5) (Таблица 3.13). Однако всё же при использовании *разных тематик карт точность заметно снижается*. Особенно в случае присутствия на разных картах полигонов различного

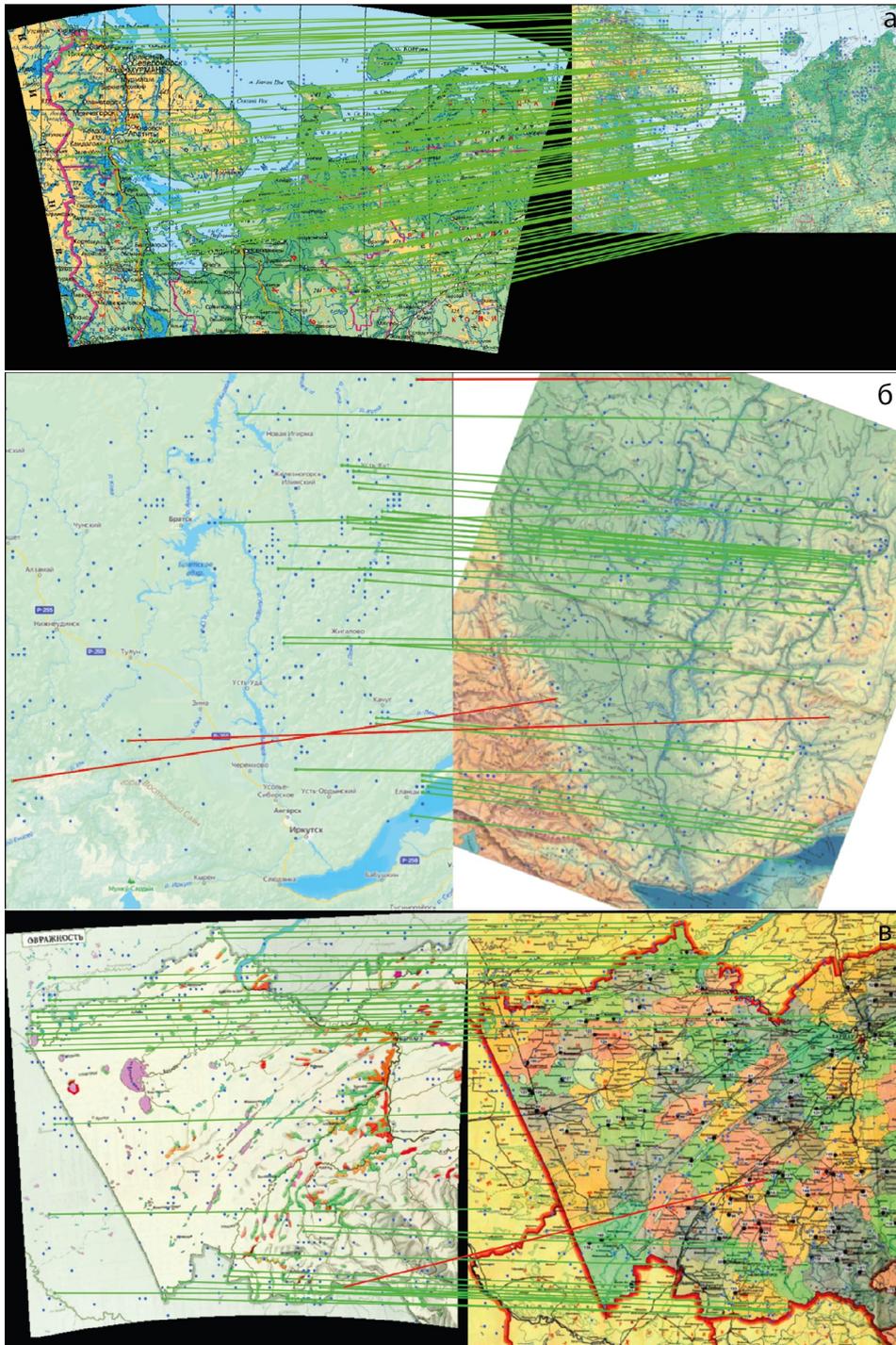
размера. В такой ситуации модель не может установить однозначное соответствие между общегеографическим содержанием двух карт, т. к. оно «теряется» в излишне детализированном рисунке одной из карт.

Наиболее низкое значение точности было получено при привязке *спутникового снимка и отсканированной карты* — пара геоизображений 10 (невязка — 11.4, ВС — 9) (Таблица 3.13). Это объясняется тем, что они имеют различные формы представления информации о Земной поверхности. В первом случае это её фактическое изображение, во втором — общегеографическая основа, интерпретированная в виде условных знаков и дополненная тематическим содержанием. В таком случае целесообразно воспользоваться вспомогательным средством — кластеризацией, которая выполнит перевод двух изображений к единой форме представления. Так, до применения кластеризации не было обнаружено ни одной верной ключевой точки. После сегментации изображений на 6 классов верно было идентифицировано 10 из 14 точек. Таким образом получилось значительно повысить точность определения опорных точек и минимизировать результирующие искажения.

Также стоит сказать, что точность падает при привязке *повернутой карты* или *фрагмента с меньшей площадью перекрытия* с привязанным геоизображением. В первом случае (пара 3 (невязка — 4.1, ВС — 6)) (Таблица 3.13) модель менее эффективно определяет соответственные точки при рассмотрении изображений с разной ориентировкой. Данный тезис также подтверждает первая пара карт, где в центральной части было выделено большое число точек, а в периферийных — нет, что стало возможным благодаря разной проекции карт. В непривязанной карте с конической проекцией по бокам изображение меняет ориентировку, поворачиваясь в соответствии с линиями параллелей. В первом же изображении применена проекция Меркатора, которая располагает все элементы содержания по прямой линии параллели. Вторая ситуация была рассмотрена на примерах пар карт 5 (невязка — 6, ВС — 5) и 8 (невязка — 3.9, ВС — 5) (Таблица 3.13). Так, если в первой паре карт точность упала, то во второй возросла (в сравнении с полными фрагментами карт). Вероятнее всего это может быть связано с привязкой территории с меньшей площадью, что привело к общему уменьшению искажений одного изображения относительно другого.

Сравним полученный результат с ранее рассмотренными работами [Cléri, Pierrot-Deseilligny, Vallet, 2014] и [Burt et al., 2020]. Первый метод является содержательным и основан на выделении структурных линий на аналоговых снимках и их привязке по ним. Тестирование производилось по 4 примерам, средняя невязка составила 9.03 пикселя. Второй метод осуществлял автоматизированную привязку карт по считанным координатам сетки с учетом специфики используемых проекций. Использование углов карты и пересечений сетки позволило достичь высокой точности в 1–4 пикселя, что существенно меньше полученного результата.

Так, в авторской модели средняя величина невязки составила 4.5 пикселя (что сопо-



**Рисунок 3.88** – Сопоставление пар геоизображений: а) №1; б) №3; в) №6 (Таблица 3.13)

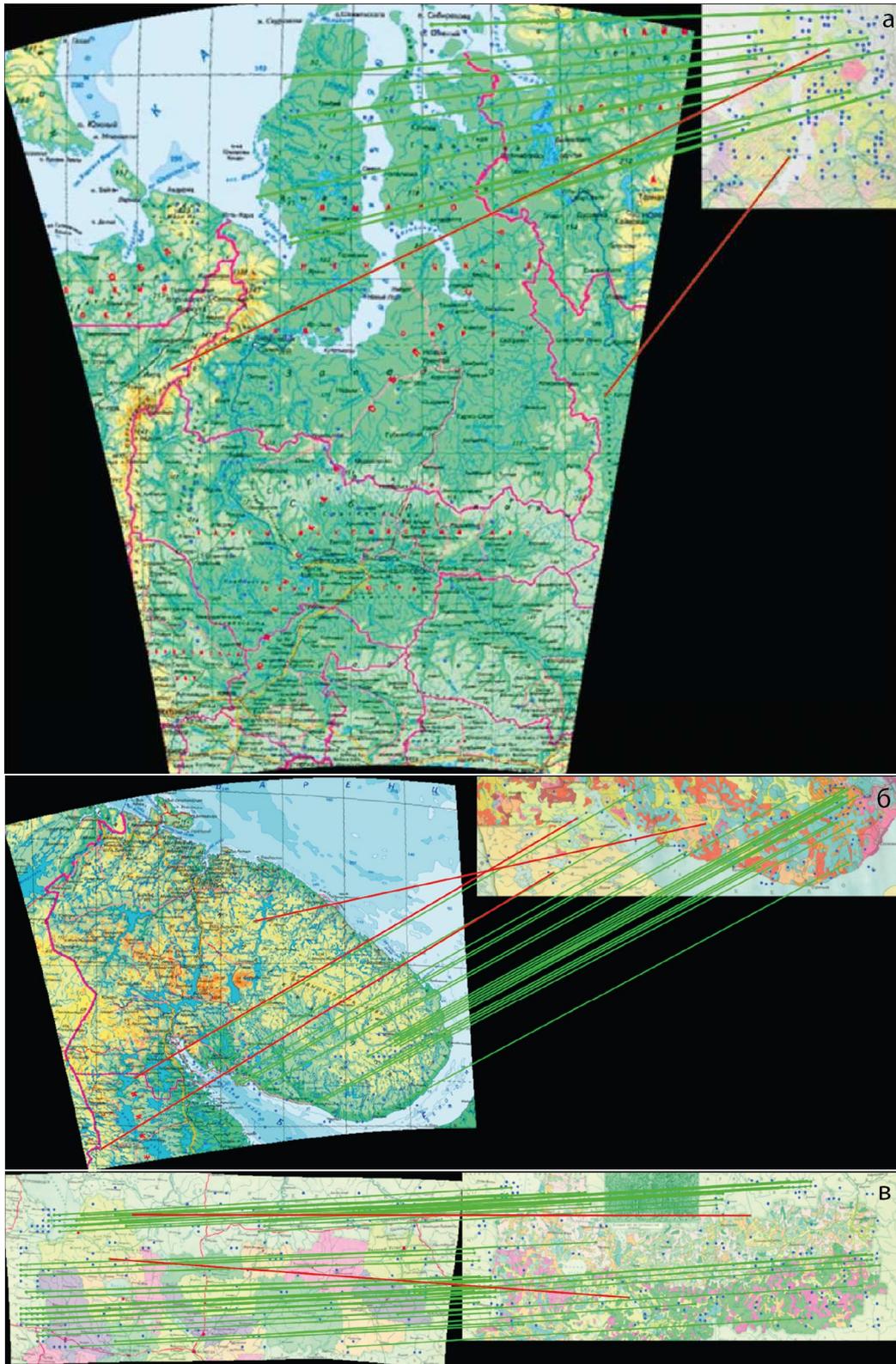


Рисунок 3.89 – Сопоставление пар геоизображений: а) №5; б) №8; в) №9 (Таблица 3.13)

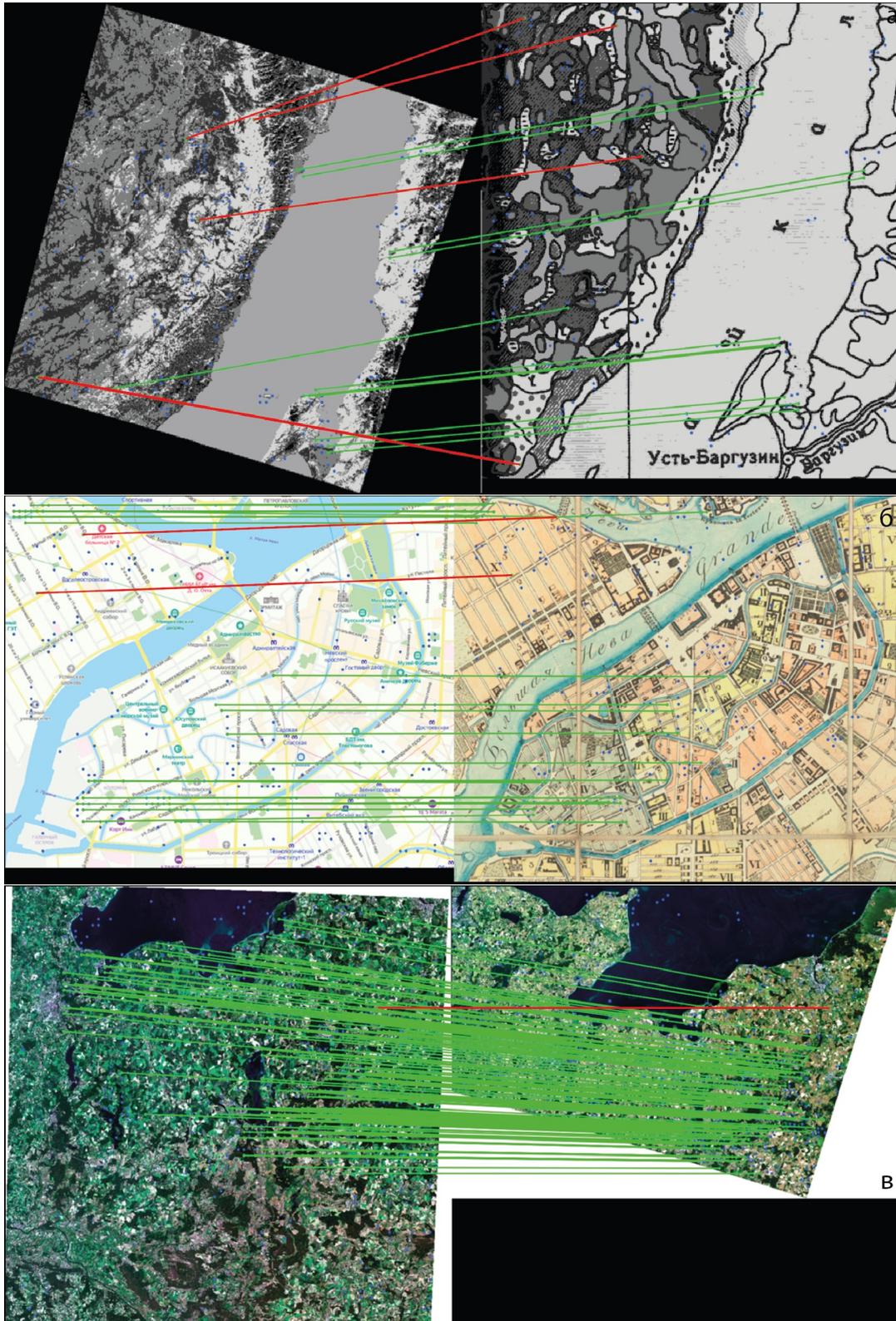


Рисунок 3.90 – Сопоставление пар геоизображений: а) №10 (6 кластеров); б) №11; в) №12 (Таблица 3.13)

ставимо с показателями точности современных подходов автоматизированной привязки) при среднем значении вектора смещения в 4.6 пикселя и достоверности выделения опорных точек в 91,9%. Было выявлено наличие сильной прямой связи между значениями невязки и векторов смещения. Ранговый коэффициент корреляции Спирмена составил 0.95 при  $p < 0.00001$ . Между значениями невязки и процентами отобранных опорных точек была найдена обратная связь, выраженная в виде того же коэффициента корреляции и равная  $-0.63$  при  $p = 0.022$ .

### 3.7.4 Итоги

В настоящем блоке диссертационного исследования разработана модель координатной привязки картографических и снимковых геоизображений с использованием компьютерного зрения. Экспериментальные исследования показали, что лучшие результаты точности получились при привязке пар спутниковых снимков, далее — привязке пар общегеографических карт, затем — общегеографических и тематических, после — пар тематических, худшие — при привязке карты к спутниковому снимку. Главный фактор точности — степень схожести содержания. Чем содержание двух карт имеет большее сходство (даже в общегеографических элементах при сравнении пар тематических карт), тем легче модели находить соответственные точки на двух геоизображениях.

Полученный результат также включает в себя показатели точности сложных примеров с привязками повернутой отсканированной карты к веб-карте, неполных фрагментов тематических карт к общегеографическим, а также привязку отсканированной карты к спутниковому снимку Landsat-8 при помощи кластеризации изображений. Показано, что процедура кластеризации способна повысить качество привязки в случае наличия пар разнородных изображений, причем как по масштабу (детальности), так и по тематике.

Таким образом, передовые технологии компьютерного зрения, а в частности сопоставления изображений, могут быть эффективно применены в картографии и геоинформатике для решения практических задач и минимизации труда специалиста при координатной привязке геоизображений.

## 3.8 Результаты и выводы

По результатам исследований, представленных в настоящей главе был достигнут значительный прогресс в понимании и автоматизации различных аспектов обработки пространственных данных, связанных с их детальностью. Исследованы понятия плотности и гранулярности пространственных данных, разработаны методы управления де-

тальностью картографического изображения линий и цифровых моделей рельефа, получены новые методы конфляции и координатной привязки пространственных данных и геоизображений различной детальности.

Основные выводы по Главе 3 заключаются в следующем:

1. Плотность пространственных данных зависит от характеристик территории и определяется множеством факторов, таких как соотношение типов наземного покрова (land cover), сложности его рисунка, уровня хозяйственной освоенности и других факторов. При этом внутри уровней детализации топографических баз пространственных данных наблюдается соответствующая вариабельность, которая может быть смоделирована с привлечением внешних данных о территории.
2. Анализ цифровых карт разных масштабов от 1 : 200 000 до 1 : 8 000 000 показывает, что ряд геометрических характеристик пространственных объектов, таких как количество точек, количество и суммарная площадь изгибов линий; средняя площадь и периметр полигона меняются монотонным образом (возрастают). Аналогичное поведение показывает характеристика гранулярности цифровых моделей рельефа. Это позволяет использовать данные характеристики детальности пространственных данных в качестве ограничивающих метрик при выполнении геометрического упрощения линий и генерализации цифровых моделей рельефа..
3. Контроль детальности линий при геометрическом упрощении осуществляется путем использования метрик, проявляющих монотонное поведение, при этом возможно достижение сопоставимого изменения детальности разными алгоритмами путем построения регрессионной модели между их параметрами. Это позволяет применять разные алгоритмы к линиям разной морфологии и при этом достигать однородной детальности результатов.
4. Детальность цифровых моделей рельефа целесообразно оценивать по соотношению их гранулярности и пространственного разрешения. Если первая характеристика отражает средний размер форм рельефа, представленных в модели, то вторая характеристика задает нижний порог для этого размера и одновременно определяет предельную графическую точность изображения.
5. При визуализации цифровых моделей рельефа способами, требующими преобразования модели данных, такими как горизонтали и штрихи, управление детальностью изображения реализуется путем вставки дополнительных элементов (таких как дополнительные горизонтали), а также вариации различных параметров, таких как высота сечения рельефа, шаг расстановки штрихов, предварительное сглаживание ЦМР и т.д.
6. Согласование данных разной детальности возможно путем конфляции (разработанной для сочетания цифровых моделей рельефа и гидрографических дан-

ных), а также методами компьютерного зрения (исследованными применительно к взаимной координатной привязке карт разной детальности, представленных в виде цифровых изображений).

Обобщение этих выводов позволяет сформулировать третье защищаемое положение:

### **Защищаемое положение №3**

*Детальность* пространственных данных в геометрической части определяется характеристиками цифровых моделей географических объектов, которые монотонно изменяются в процессе их генерализации. Направленное воздействие на эти характеристики с помощью операторов генерализации позволяет достигать требуемой детальности картографического изображения, а моделирование этих характеристик даёт количественные оценки детальности и возможность её управляемого изменения. Согласование данных разной детальности осуществляется с помощью конфляции (совмещения) их соответствующих элементов, а их взаимная координатная привязка — с использованием методов компьютерного зрения.

## 4 Мультимасштабное картографирование

В главе рассмотрен принцип *мультимасштабности*, постулирующий необходимость интерактивной генерализации пространственных данных в соответствии с масштабом их визуализации в цифровой среде. В отличие от бумажных карт, масштаб изображения на цифровых устройствах представляет собой переменную величину, что открыло возможность создания нового типа карт — *мультимасштабных*, обеспечивающих переменную детальность картографического изображения. Для определения оптимального соответствия между масштабом визуализации, детальностью используемых для этого пространственных данных и способами их изображения необходимы методологические основания.

### 4.1 Основные понятия и принципы мультимасштабного картографирования<sup>1</sup>

Впервые основания мультимасштабного картографирования разработаны автором и изложены в монографии [Самсонов, 2011], где в том числе введены следующие понятия:

**Мультимасштабное картографирование (МК)** — это создание и использование электронных карт переменной детализации, обеспечивающих представление объектов во множестве масштабов. Мультимасштабные карты отличаются введением таких элементов математической основы как масштабный ряд и масштабный диапазон.

**Масштабный ряд** — упорядоченная последовательность масштабов карты, на которых производится ее составление. Используется для контроля детализации и оформления карты, а также отображения ключевых уровней иерархии отображаемого явления.

**Масштабный диапазон** — числовой отрезок, границы которого определяются крайними элементами (минимальным и максимальным) масштабного ряда. В пределах масштабного диапазона осуществляется генерализация изображения. За его пределами — простое уменьшение и увеличение масштаба без изменения детализации.

<sup>1</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Лурье, Самсонов, 2010; Самсонов, Подольский, Юрова, 2013]



**Рисунок 4.1** – Масштабный диапазон и масштабный ряд карты [Самсонов, 2011]

Организация данных для мультимасштабного картографирования может быть осуществлена множеством способов. При большом объеме данных эта задача решается средствами организации базы данных. База данных (БД) — ключевая компонента любого геоинформационного исследования. Согласно теории баз данных, в процессе их проектирования выделяют три уровня: концептуальный, логический и физический. Концептуальный уровень проектирования предусматривает [Лурье, 2016]:

- определение и описание исследуемых объектов или явлений;
- установление способа представления географических объектов в базе данных;
- выбор базовых типов пространственных объектов (точки, линии, ареалы, ячейки раstra);
- решение вопроса о способе представления размерности и взаимосвязей реального мира в БД;
- определение необходимых источников данных, требований к их качеству.

В свою очередь содержание базы данных определяется сутью явления, характером его пространственного распространения и задачами, для которых создается БД. К отличительной особенности логической структуры БД для обеспечения МК относится представление данных с различной детализацией. Отсюда возникает понятие мультимасштабной базы данных.

**Мультимасштабная база пространственных данных (МБД)** — база пространственных данных, содержащая множество уровней детализации. Используется для информационного обеспечения МК, позволяет выполнить ресурсоемкие операции генерализации заранее.

В стандартных геореляционных базах данных ГИС использование множественных представлений неэкономно, однако избежать этого пока не удастся, так как соответствующие программные методы перехода между ними еще недостаточно разработаны [Лурье, 2016]. Построение МБД позволяет обойти ограничения существующих алгоритмов генерализации, среди которых — низкая производительность, не позволяющая использовать их в режиме реального времени (что также упирается в производительность аппаратного обеспечения), а также невысокая достоверность результатов, которая требует

их ручной коррекции. Переключение между уровнями детализации БД позволяет имитировать интерактивную генерализацию, которая теоретически может осуществляться на основе одного детального набора данных.

Число уровней детализации БД определяется масштабным диапазоном и масштабным рядом карты. Использование для каждого масштаба ряда своего уровня детализации БД может привести к нескольким негативным эффектам: существенному увеличению объема хранимых данных, дублированию данных при идентичности представления объектов на разных уровнях, необходимости каскадного обновления всех уровней при актуализации хотя бы одного из них [Самсонов, 2009а]. Поэтому обычно число уровней минимизируется, а для генерации промежуточных уровней разрабатываются методы генерализации «на лету», которые пока представляют собой простейшие негеометрические операции: отбор, исключение, обобщение характеристик (на уровне атрибутов). Масштабный диапазон, в пределах которого удается достичь приемлемого качества генерализации на основе одного и того же уровня детализации БД, называют диапазоном применимости уровня [Cecconi, Weibel, Barrault, 2002]. За пределами диапазонов применимости используются соседние уровни детализации БД.

При формировании мультимасштабных карт методы генерализации пространственных данных могут использоваться на двух разных уровнях: на *уровне базы данных* (предварительная подготовка) и на *уровне карты* (генерализация «на лету») [Самсонов, 2009а]. На первый уровень выносятся ресурсоемкие операции, которые не могут быть выполнены в режиме реального времени и/или требуют ручной верификации результатов. Выполняется поиск и *интеграция* источников данных, наилучшим образом соответствующих представленным уровням детализации. Неохваченные данными уровни заполняются путём генерализации заполненных [Самсонов, 2009б; Лурье, Самсонов, 2010].

В исследовании [Самсонов, 2011] показано, что количество уровней БД для МК рельефа должно соответствовать принципу двукратного уменьшения масштаба: если между уровнями масштабного ряда разница в 2 раза и более, то для каждого из них должен быть подготовлен свой уровень детализации БД. Для слоев другой тематики число уровней может быть уменьшено путем использования методов генерализации «на лету» с учетом диапазонов применимости. Следует учитывать, что состав слоев между уровнями детализации может изменяться. Например, в низкодетальных уровнях, предназначенных для мелких масштабов, могут быть исключены слои, содержащие информацию о конкретных объектах на земной поверхности, и заменены на синтетические слои, показывающие ареалы распространения явления. Специфика мультимасштабных карт заключается в оперировании большими объемами данных, поэтому необходимо следить за тем, чтобы используемые форматы хранения слоев в БД поддерживали методы оптимизации пространственного поиска: построение пространственных индексов; разбиение данных на блоки и т. д.

**Основные задачи** мультимасштабного картографирования [Самсонов, 2012]:

1. Моделирование пространственно-временных многоуровневых структур в организации географических явлений, объектов, процессов и отношений;
2. Синтез информации о географической среде путем применения картографической генерализации как научного метода исследования;
3. Совершенствование принципов оформления карт с учетом визуального восприятия масштабных переходов;
4. Интеграция карт и данных различной детализации в единые мультимасштабные модели и хранилища данных.

Систематическое решение этих задач требует декомпозиции механизма построения мультимасштабных карт на три **компоненты**, которые исследованы в настоящей работе на примерах карт различной тематики:

1. *Уровни (масштабные уровни, уровни детализации)*, которые определяются сочетанием пространственных данных определенной детальности и метода их визуализации в заданном масштабе.
2. *Преобразования*, которые определяют логику изменений происходящих между масштабными уровнями при смене масштаба визуализации.
3. *Анимации*, которые визуализируют преобразования в виде плавных переходов.

В основе проектирования мультимасштабных карт лежит опыт традиционного картографирования в разных масштабах. В работах [Подольский, Самсонов, 2012; Юрова, Самсонов, 2012; Самсонов, Подольский, Юрова, 2013] на основе анализ многочисленных карт и научной литературы выделены следующие свойства **преобразований способов изображений** при изменении масштаба картографирования:

1. Логичность и последовательность изменения графических переменных; смешение цветов, увеличение и уменьшение прозрачности, толщины линий должно происходить по возможности плавно, без резких скачков;
2. Циклическое заимствование графических переменных при переходе на более высокий уровень единиц картографирования; в простых способах изображения этому предшествует постепенное сближение символов, которое не позволяет сделать замену резкой.
3. Переход от аналитических способов к комплексным и синтетическим; значки и ареалы заменяются картодиаграммами, типологическими значками, что соответствует переходу на суперординатный уровень таксономической и партномической иерархий.
4. Постепенное уменьшение информационной нагрузки; при переходе на более высокий уровень таксономической классификации сохраняются основные графиче-

ские переменные и удаляются второстепенные.

5. Постепенное уменьшение графической нагрузки; в частности, оно выражается в уменьшении размеров значков, толщины линий, кегля шрифтов и т.д.

Так, например, принцип *последовательности* при масштабировании картодиаграмм обеспечивает постепенное, естественное уменьшение их среднего размера при масштабировании (Рисунок 4.2).

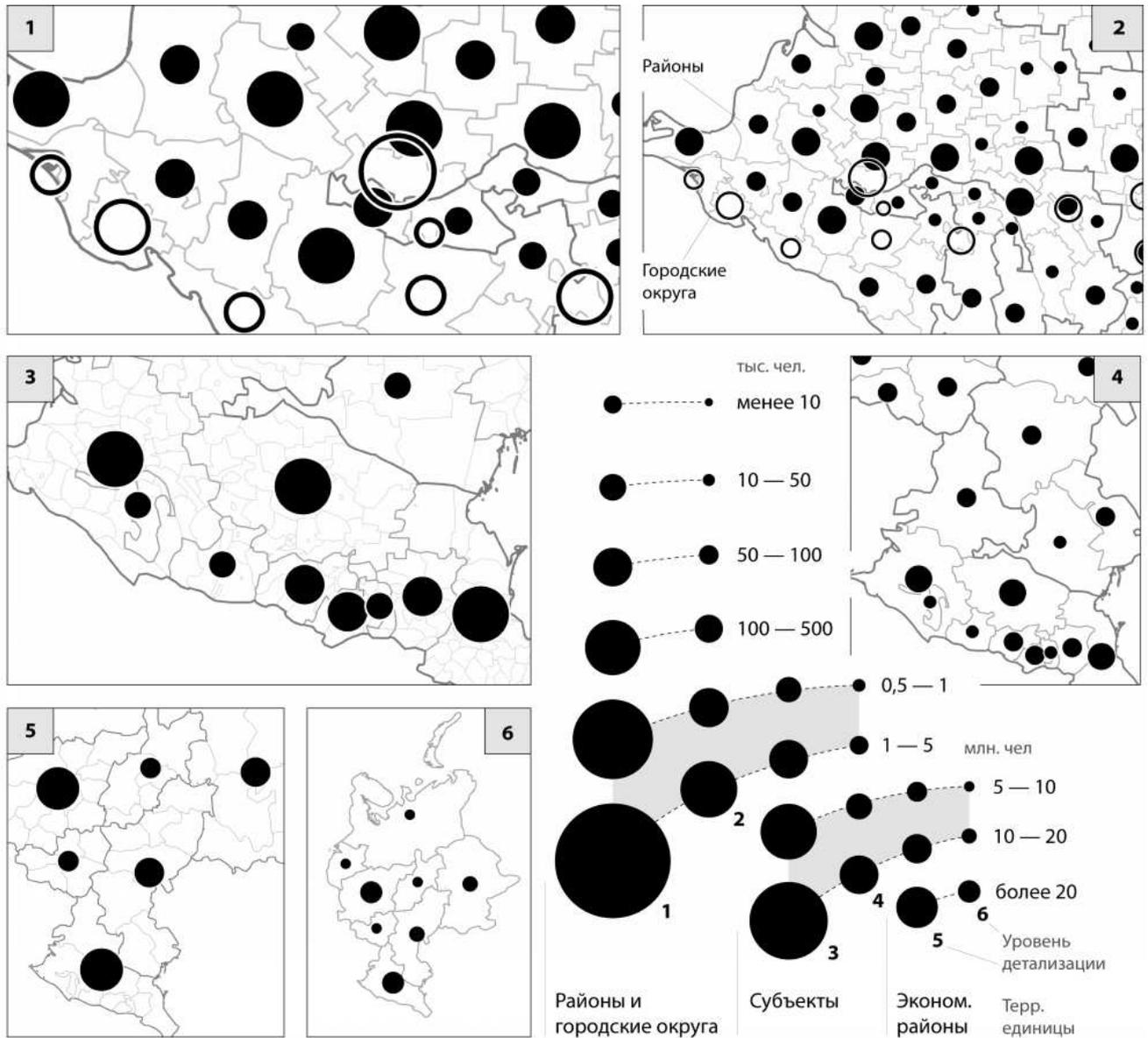


Рисунок 4.2 – Преобразование картодиаграмм численности населения на мультимасштабной карте [Самсонов, Подольский, Юрова, 2013]

## 4.2 Вычисление параметров визуализации цифровых моделей дорожной сети и структуры населенных пунктов на мультимасштабных картах<sup>2</sup>

Важным аспектом изображения на **мультимасштабных общегеографических картах** является необходимость согласованного отображения множества слоев, образующих пространственные структуры [Touya, Potié, Mackaness, 2023]. Предпринятые ранее способы решения проблемы преобразования способов изображения путем введения многоуровневой символики [Choi, Hwang, 2009] решают лишь проблему ее организации, но не отвечают на вопросы, при каких масштабах и как должно происходить изменение символики.

Для исследования этой проблемы был разработан способ вычисления масштабных уровней и преобразований при визуализации цифровых моделей дорожной сети и структуры населенных пунктов на мультимасштабных картах [Самсонов, Подольский, 2014]. В топографическом картографировании структура населенных пунктов показывается блоками (в крупных масштабах — кварталами, далее — блоками кварталов) и только некоторые важные проезды выделяются цветом автомобильных дорог [Верещака, 2002]. Таким образом, уличная сеть распознается при чтении карты по пробелам между блоками. На электронных картах распространен противоположный подход, когда структура населенного пункта распознается по рисунку уличной сети, а функциональное назначение блоков (селитебные, промышленные, парковые территории и т. д.) определяется подстилающим слоем землепользования. При этом уличная и транспортная сети за пределами населенных пунктов составляют единое целое (например, проспект Энгельса в Санкт-Петербурге переходит в Новоприозерское шоссе без смены условного знака). С одной стороны, это наделяет оформление определенной гибкостью, поскольку иерархичность транспортной (уличной) сети уже никак не связана с геометрией объектов застройки. С другой стороны, характеры уличной сети в городах и транспортной сети за их пределами могут существенно различаться, и их объединение под маской одного символа даже при условии сопоставимой значимости не всегда будет географически достоверным.

Предлагаемая методика и алгоритм автоматизации отображения дорожной сети и структуры населенных пунктов на основе их цифровых моделей опирается на ряд начальных условий и допущений:

- исходное оформление карты в самом крупном масштабе проработано пользователем; задача алгоритма — транслировать это оформление в более мелкие масштабы, применяя подходящие операторы;

<sup>2</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Самсонов, Подольский, 2014]

- генерализация дорожной сети при достижении критических масштабов реализуется путем исключения целых классов дорог;
- для отображения дорожной сети используются простые линии без обводок; цвет линий для разных классов остается постоянным, меняется только их толщина;
- алгоритм предназначен для работы в крупных и средних масштабах картографирования, что позволяет отложить вопросы, связанные с отображением и генерализацией систем объектов.

#### 4.2.1 Преобразование толщины линий и вычисление критических масштабов

В процессе движения от крупных масштабов к мелким при отсутствии генерализации объектов возрастает графическая нагрузка. Наступает момент, когда необходимо исключать часть объектов для обеспечения читаемости карты. Масштаб, в котором возникает такая необходимость, называется критическим. Критический масштаб может быть определен как для одного объекта, так и в обобщенном виде для группы (слоя) объектов или уровня детализации на основе их среднестатистических размеров.

Одним из основных критериев для исключения или объединения линейных объектов является графическая нагрузка. За единицу измерения графической нагрузки можно принять отношение площади  $S_r$ , занимаемой знаками линий, к площади карты  $S$ . Поскольку графическая нагрузка тесно связана не только с конфигурацией объектов, но и с символикой, необходим алгоритм преобразования символов линий при исключении классов, который будет с этой характеристикой взаимоувязан. На величину нагрузки влияет прежде всего толщина линий (постоянство цвета классов и смена толщины). Была предложена следующая стратегия преобразования: при переходе на новый уровень детализации исключаются дороги низшего класса, толщина символов смещается на одну ступень вверх, а их цвет остается прежним (Рисунок 4.3).

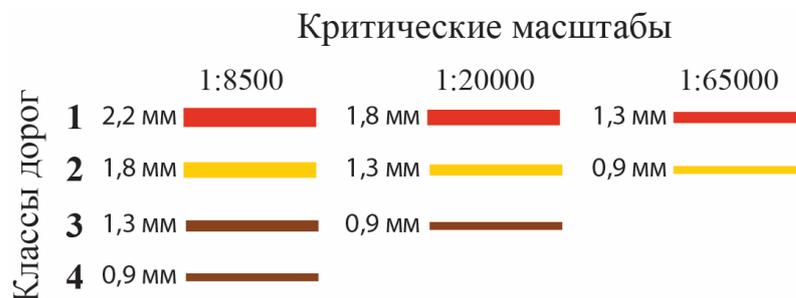


Рисунок 4.3 – Изменение толщины линий дорожной сети

Точка перехода на следующий уровень детализации определяется критическим масштабом, в котором при изначально заданном оформлении графическая нагрузка, создаваемая слоем дорог, достигнет предельно допустимой величины  $p_{\max}$ . Для вычисления фактической нагрузки необходимо знать площадь символов на карте.

Математически точно эта операция может быть проделана путем построения вокруг линий буферных зон с радиусом  $r_i = w_i/2$ , где  $w_i$  – толщина линии для  $i$ -го класса дорог, объединения полученных буферов, подсчета площади результирующей фигуры и взятия отношения этой площади к площади картографируемой территории.

В настоящем исследовании использовалась методика приближенной оценки, основанная на вычислении густоты линий каждого класса. Густота  $d_{ij}$  (км/км<sup>2</sup>) для каждого класса дорог  $C_i$  в ячейке  $j$  оценивалась по ячейкам регулярной сети как отношение суммы длин линий  $L_i$  к площади ячейки  $S$ :  $d_{ij} = L_{ij}/S$ . Для снижения влияния эффектов, связанных с неравномерностью распределения дорог, оценка густоты должна осуществляться по регулярной сетке (Рисунок 4.4). Густота оценивается только по тем ячейкам, в которых присутствуют объекты улично-дорожной сети. Искомой интегральной характеристикой является средняя густота дорожной сети каждого класса  $d_i = \sum_{j=1}^n d_{ij}/n$ , где  $n$  – число ячеек.



Рисунок 4.4 – Вычисление густоты дорожной сети по регулярной сетке (км/км<sup>2</sup>)

Зная толщину линии  $w_i$  каждого класса дорог, можно найти знаменатель масштаба  $M$ , в котором среднее отношение площади знаков дорожной сети к площади ячейки достигнет заданного значения  $p_{max}$  (обратим внимание на то, что произведение густоты и толщины линий в знаменателе – безразмерная величина):

$$M = \frac{p_{max}}{\sum_{i=1}^k w_i d_i}, \quad (4.1)$$

где  $k$  – число классов дорог.

Данная методика позволяет вычислить критический масштаб при фиксированном размере ячейки. Однако с уменьшением масштаба и прореживанием дорожной сети возникает необходимость увеличивать размер ячейки. В противном случае суммарная площадь знаков дорог будет стремиться к суммарной площади всех непустых ячеек, а расчетная графическая нагрузка – к 100%. Слишком крупный размер ячейки напротив, приведет к учету обширных территорий без транспортной сети и лишит смысла использо-

вание регулярной сетки. Эту проблему можно решить следующим образом:

1. После получения первого критического масштаба  $M_1$  производится переход на следующий структурный уровень дорожной сети и смещение толщин линий.
2. Вычисляется второй критический масштаб  $M_2$  с теми же параметрами сетки, далее — масштабный коэффициент  $m = M_1/M_2$ .
3. Рассчитывается новый размер ячейки сетки:  $C_2 = mC_1$ , где  $C_1$  и  $C_2$  — исходный и расчетный размеры ячейки.
4. Второй критический масштаб  $M_2$  вычисляется заново с новым размером ячейки.

Описанные шаги формируют циклический алгоритм, рассчитывающий критические масштабы и толщины линий дорог для каждого структурного уровня.

Остается не ясным, каков же должен быть изначальный шаг сетки для первого уровня детализации. Для его выбора можно выполнить анализ ближайшего соседства, определив среднее расстояние от каждой линии до ближайшей. Размер сетки можно далее вычислить как функцию от полученного расстояния. Дальнейшие исследования должны определить оптимальный путь вычисления.

#### 4.2.2 Согласование слоя типов землепользования с дорожной сетью

Фрагментарная согласованность границ полигонов типов землепользования и линий улиц ухудшает читаемость карты. Например, при удалении улицы появляется возможность объединить полигоны, которые с ней граничат. Помимо этого, физическое разбиение слоя землепользования линиями дорог позволяет оценивать площадь получающихся территориальных единиц и корректировать параметры оформления и генерализации в соответствии с ней (Рисунок 4.5).

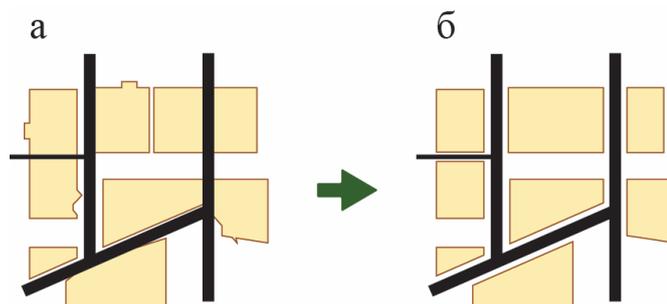


Рисунок 4.5 — Слой землепользования и дорожной сети до (а) и после (б) согласования

Для согласования детализации слоев дорожной сети и структуры населенного пункта был разработан линейный алгоритм. Он выполняется для каждого критического масштаба и соответствующих ему набора классов дорог и их толщин и состоит из нескольких шагов:

1. Полигоны упрощаются в соответствии с масштабом (пользователь задает величи-

ну допуска, видимую на мониторе в миллиметрах, что позволяет использовать ее во всех масштабах).

2. Для исключения мелких и осколочных полигонов проводится агрегирование; барьерами для агрегирования выступают дороги.
3. Вокруг слоя дорог строятся буферные зоны, ширина которых соответствует ширине символа дороги соответствующего класса.
4. В слое полигонов функциональных типов территории вырезаются дырки по построенным буферным зонам.

Поскольку данные OpenStreetMap имеют внутреннее согласование между слоями для наиболее крупных масштабов, первый уровень детализации дорожной сети соответствует исходному полигональному слою, а генерализованный с помощью дорожной сети первого уровня детализации — ее второму уровню. Такой сдвиг позволяет избежать резких скачков детализации, так как исключенный класс дорог отражен в структуре полигонального слоя до перехода на следующий уровень детализации

### 4.2.3 Техническая реализация

Описанные алгоритмы для вычисления критических масштабов и толщин линий, а также согласования дорожной сети и полигональных слоев типов землепользования были реализованы в виде скрипта на языке программирования Python с использованием библиотеки ArcPy. Скрипт обеспечивает работу инструмента в среде ArcGIS. Исходными данными и параметрами, необходимыми для работы инструмента, являются:

- размер (в метрах) ячейки регулярной сетки, по которой будет рассчитываться плотность дорожной сети;
- максимально допустимая величина графической нагрузки; среднее для всех ненулевых ячеек отношение общей площади знаков дорог к площади ячейки;
- экстенд, участок местности, в пределах которого производятся расчеты; таблица соответствия классов дорог и исходных толщин линий;
- количество главных классов; главные классы не исключаются, т. е. когда количество классов становится равным указанному количеству главных классов, процесс генерализации останавливается;
- минимально допустимая видимая на мониторе площадь полигона (части полигона) в квадратных миллиметрах; выступает в качестве параметра при генерализации полигонального слоя;
- полигональный слой объектов гидрографии; используется в качестве дополнительного барьера при агрегировании полигонов;
- линейный слой объектов гидрографии; используется в качестве дополнительно-

го барьера при агрегировании полигонов.

При работе инструмента выполняется следующая последовательность действий, представленная на Рисунке 4.6.

1. Создается регулярная сетка с заданным пользователем размером ячейки.
2. С помощью инструмента Intersect (пересечение) слой дорожной сети разрезается границами ячеек сетки.
3. Подсчитывается суммарная длина дорог каждого класса внутри каждой ячейки.
4. По Формуле 4.1 вычисляется масштаб, в котором при заданных толщинах линий среднее отношение суммарной площади знаков дорог внутри ячейки к ее площади достигнет критического значения.
5. С использованием полученного масштаба и заданного пользователем допуска полигональный слой структуры населенного пункта упрощается и агрегируется (в качестве барьеров выступают слои дорожной сети и гидрографии).
6. Вокруг слоя дорожной сети строятся буферные зоны (дистанция буферизации — толщина линии дороги в масштабе карты).
7. Генерализованный полигональный слой обрезается построенными буферными зонами.
8. Исключается последний класс дорог, значения толщин линий смещаются таким образом, чтобы исключить самую толстую.
9. Повторяются п. 3 и п. 4.
10. Рассчитывается новый размер ячейки сетки.
11. Возвращаются к п. 1.

Инструмент работает до тех пор, пока количество классов дорог не уменьшится до количества главных классов.

Результатами работы алгоритма являются:

- таблица, в которой для каждого структурного уровня дорожной сети указан критический масштаб и толщины линий, соответствующие каждому классу;
- набор генерализованных полигональных слоев структуры населенного пункта.

#### 4.2.4 Результаты

Для тестирования инструмента использовались данные OSM на территорию Санкт-Петербурга. Поскольку алгоритмы оперируют значениями площадей и длин объектов, была выбрана проекция UTM<sup>3</sup> (36 зона), обеспечивающая минимальные искажения пло-

<sup>3</sup>Здесь и в прочих разделах диссертации преимущественно используются преобразования к системам координат, которые не предполагают трансформации датума. В частности, поскольку данные OSM хранятся в географических координатах относительно датума WGS 84, переход к UTM позволяет остаться на том же датуме, в

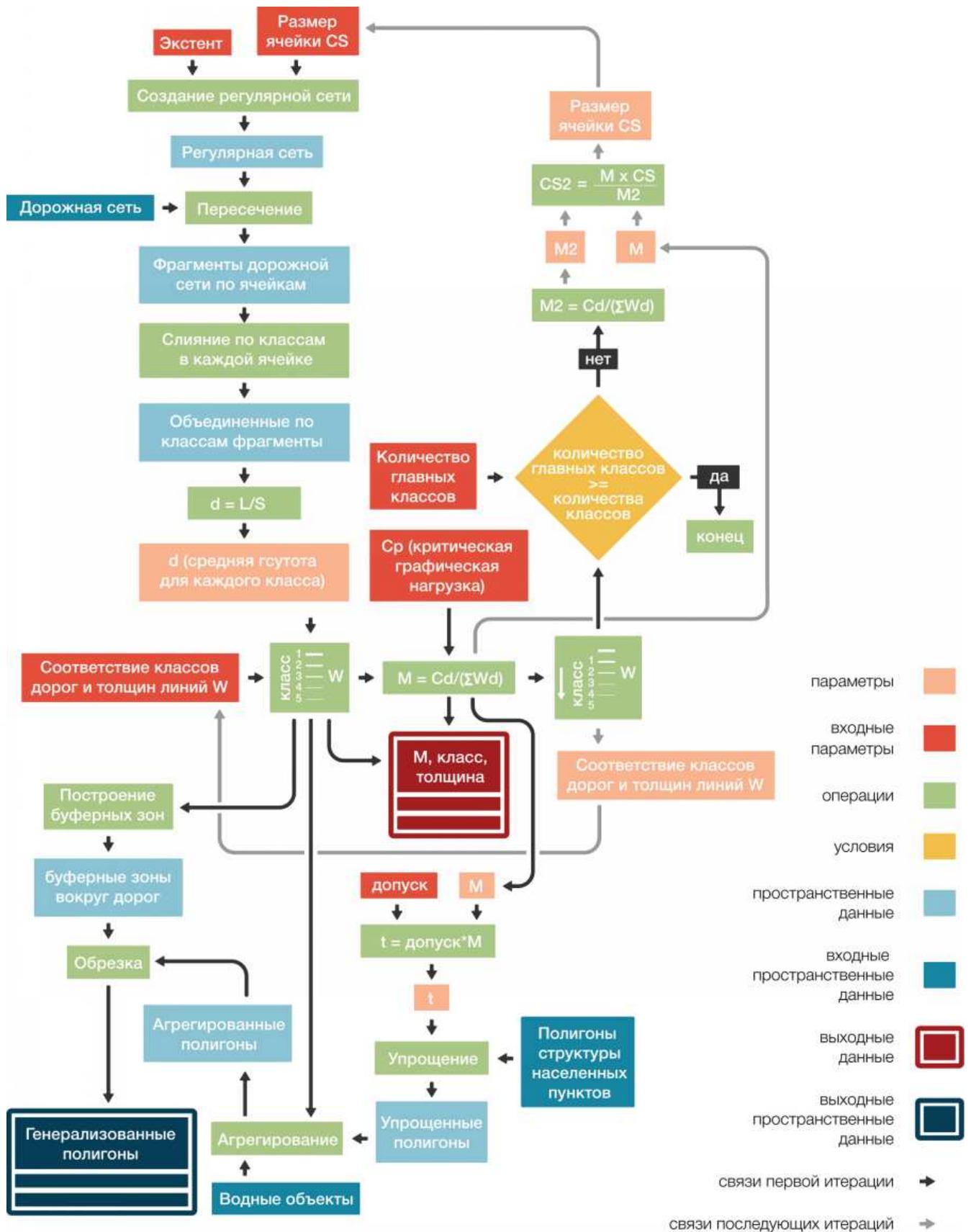


Рисунок 4.6 – Схема работы инструмента *ComputeScales*

Таблица 4.1 – Результат работы инструмента *ComputeScales*

Класс дорог (CLASS)	Критический масштаб (SCALE)	Толщина линии, (WIDTH)
Уровень детализации 1		
1	1 : 5 938	2.2
2	1 : 5 938	1.8
3	1 : 5 938	1.3
4	1 : 5 938	0.9
5	1 : 5 938	0.6
Уровень детализации 2		
1	1 : 19 372	1.8
2	1 : 19 372	1.3
3	1 : 193 72	0.9
4	1 : 193 72	0.6
Уровень детализации 3		
1	1 : 54 952	1.3
2	1 : 54 952	0.9
3	1 : 54 952	0.6
Уровень детализации 4		
1	1 : 167 601	0.9
2	1 : 167 601	0.6

щадей и длин. Данные OSM были разбиты на следующие слои:

- дорожная сеть (выделено пять классов дорог);
- типы землепользования (разделения по типам не проводилось, все типы участвовали в формировании структуры населенного пункта равносильно);
- строения (отображаются полностью на первом структурном уровне, с 50% прозрачности — на втором, с 75% — на третьем);
- растительность;
- полигональная гидрография.

Пяти классам дорог были назначены толщины от 2.2 до 0.6 мм. Исходный размер ячейки варьировался от 200 м до 2 км. Опытным путем был определен оптимальный размер — 500 м. По результатам работы алгоритма была составлена Таблица 4.1, отражающая вычисленные критические масштабы и толщины линий, а также набор генерализованных слоев землепользования для уровней детализации 2–4.

Примеры оформления мультимасштабной карты, полученные с помощью разрабо-

---

то время как переход к проекции Гаусса-Крюгера предполагал бы трансформацию к ГСК-2011. Теоретически, можно было бы также использовать проекцию Гаусса-Крюгера на эллипсоиде WGS84, однако такая практика мало распространена.

танного инструмента, представлены на Рисунке 4.7. Цветовое решение карты прорабатывалось вручную. Данные иллюстрации демонстрируют срезы непрерывного процесса преобразования на всем масштабном диапазоне. Наиболее удачно, на наш взгляд, работает методика вычисления критических масштабов с учетом оформления. В то же время, результаты не лишены недостатков. В ряде случаев объединение ранее разделенных полигонов не произошло, что связано с большим расстоянием между ними. Также очевидно, что генерализация путем исключения классов приводит к избыточному удалению дорог и утрате полноты рисунка транспортной сети. В дальнейшем методика будет модифицирована с использованием маршрутной генерализации [Samsonov, Krivosheina, 2012], которая должна будет позволить сохранять все дороги, важные на каждом уровне детализации. Однако это потребует соответствующей модификации метода изменения толщины линий.

#### 4.2.5 Итоги

В главе предложена методика и набор алгоритмов, позволяющих частично автоматизировать визуализацию цифровых моделей дорожной сети и структуры населенных пунктов на мультимасштабных общегеографических картах. Методика основана на допущении, что структура населенного пункта может определяться дорожной сетью. Основной упор сделан на вычисление критических масштабов отображения дорожной сети с учетом допустимой графической нагрузки и при заданной толщине линий. Для определения масштабов предложена и апробирована схема изменения толщины линий, при которой достижение критического масштаба сопровождается удалением самого низкого класса дорог, а сами толщины смещаются на одну ступень вниз. Также в методику включена процедура согласования полигональной структуры населенных пунктов с дорожной сетью.

В полученных результатах выявлены недостатки, которые необходимо разрешить в будущих исследованиях. В частности, удаление полного класса дорог не является оптимальным, поскольку приводит к скачкообразному изменению подробности изображения дорожной сети.

Поставленные в работе допущения будут в дальнейшем сняты для разработки более комплексной автоматизации составления мультимасштабных карт, включающей полноценную генерализацию дорожной сети и автоматический выбор большинства параметров оформления. Представленные результаты демонстрируют, что для дальнейшего развития автоматизации мультимасштабного картографирования необходимо не только развитие автоматизации геометрической и графической генерализации, но и их взаимопроникновение.



**Рисунок 4.7** – Изображения, полученные по данным OSM с использованием разработанной методики. Масштабы: (а) – 1:5 000 (уровень детализации 1); (б) – 1:10 000 (уровень детализации 2); (в) – 1:25 000 (уровень детализации 3); (г) – 1:50 000 (уровень детализации 3); (д) – 1:100 000 (уровень детализации 4); (е) – 1:200 000 (уровень детализации 4). ((с) Участники OpenStreetMap – картографические данные

## 4.3 Мультимасштабное гипсометрическое картографирование на основе цифровых моделей рельефа<sup>4</sup>

На картах физико-географической тематики часто используется способ изолиний, который позволяет визуализировать распределение географических полей. Примерами могут служить гипсометрические карты и карты характеристик ветрового волнения в морях. Мультимасштабные варианты таких карт предполагают генерализацию растровых и триангуляционных покрытий, лежащих в основе таких карт, а также выбор оптимальной шкалы сечения для каждого масштаба картографирования. Таким образом обеспечивается логическая последовательность масштабных уровней карты и их преобразований.

Методика МК в приложении к рельефу требует комплексирования методов моделирования рельефа и его генерализации при построении мультимасштабных баз данных, а также автоматизации способов изображения рельефа при визуализации ЦМР. В рамках данной части исследований была создана база пространственных данных и мультимасштабная гипсометрическая карта, охватывающая Европейскую территорию России в масштабах 1 : 25 000 — 1 : 50 000 000 (с частичным покрытием в крупных масштабах) [Лурье, Самсонов, 2010; Samsonov, 2011; Самсонов, 2011]. При составлении карты использован метод масштабных групп с выделением 11 масштабных уровней, соответствующих уровням детализации созданной базы данных.

### 4.3.1 Объекты и территориальный охват картографирования

Для реализации принципов мультимасштабного картографирования применительно к гипсометрическому изображению рельефа был выбран большой и геоморфологически разнообразный регион, захватывающий такие значимые географические объекты как Большой Кавказ, Восточно-Европейская равнина, Уральские горы. Задача проекта апробации — продемонстрировать возможность плавного интерактивного перехода от макро- к мезоуровню картографирования и изучения рельефа в пределах отдельных участков выбранного региона с использованием разработанной методики.

### 4.3.2 Средства реализации и программное обеспечение

В качестве программной платформы для реализации проекта апробации использовалась настольная ГИС ArcGIS Desktop 10. Выбор этой платформы обусловлен возможностью реализации всех стадий предложенной методики и наличием лицензионного про-

<sup>4</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Лурье, Самсонов, 2010; Samsonov, 2011]

граммного обеспечения.

### 4.3.3 Математическая основа

С учетом общего охвата карты в качестве обзорного был выбран масштаб 1 : 50 000 000. В этом масштабе на экране компьютера обеспечивается полный охват картографируемой территории, а также частичный охват соседних географических областей (Фенноскандия, Польская равнина, Карпаты, Малая Азия, Западно-Сибирская низменность).

Поскольку при изучении мезоформ рельефа необходимо обеспечить масштаб 1 : 200 000 и крупнее, нижняя граница масштабного диапазона расширена до этой величины, а по возможности — до масштаба 1 : 25 000. Масштабный ряд карты построен исходя из предложенного принципа 2 — 2.5-кратного увеличения масштаба.

Крупно- и среднемасштабное картографирование всей территории потребовало бы поиска, хранения и обработки гигантских объемов данных. Поэтому в пределах каждой формы были выделены и внесены в схему разграфки ограниченные участки. Схема разграфки масштабных уровней представлена на Рисунке 4.8. Далее был произведен поиск наиболее подходящих по детализации и охвату источников данных. Сплошное покрытие всей территории обеспечивается вплоть до масштаба 1 : 2 500 000. Начиная с масштаба 1 : 1 000 000 покрытие характеризуется уменьшением площади и фрагментацией. Масштабный ряд и используемые источники представлены в Таблице 4.2.

Для представления данных в БД и последующего картографирования было решено использовать равноугольную цилиндрическую проекцию Меркатора с центральным меридианом 10° в.д., главной параллелью 50° с.ш. и эллипсоид WGS-84. Выбор математической и геодезической основы обусловлен требованиями, сформулированными в параграфе 2.5.2, а также возможностью свободного расширения охвата карты для покрытия всей территории России и далее мира. При этом не потребуется изменение параметров проекции.

### 4.3.4 Исходные данные

В проекте были использованы разнородные источники данных, что потребовало их интеграции в базе данных и предварительной оценки точности и достоверности.

*Цифровые топографические карты (ЦТК)* 1 : 25 000, 1 : 50 000, 1 : 100 000, 1 : 200 000, 1 : 500 000 и 1 : 1 000 000 составлены на основе бумажных топокарт в соответствии с требованиями к качеству [ГОСТ Р 51608-2024, 2024] и уже имеют необходимую точность и достоверность. Карты были изначально представлены в виде бесшовного покрытия без математической основы, при этом координаты хранились в геодезических широтах и долготах относительно эллипсоида Красовского. Все данные были спроецированы в вы-

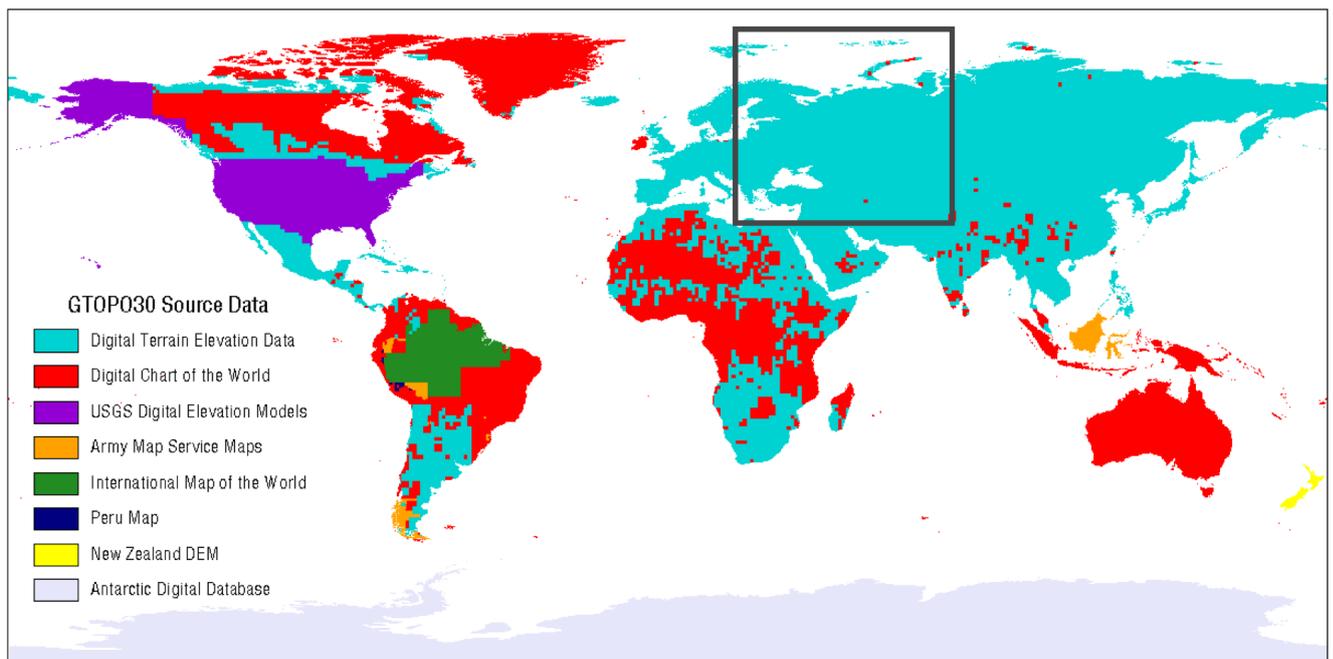


**Рисунок 4.8** – Разграфка масштабных уровней карты

бранную математическую и геодезическую основу и приведены в соответствие с предложенной логической структурой БД. Далее для построения ЦМР были использованы горизонтали и гидрография карт масштабов 1 : 25 000 — 1 : 200 000.

Горизонталы с ЦТК масштабов 1 : 500 000 и 1 : 1 000 000 в построении ЦМР не использовались. ЦМР для этих масштабов было решено получить путем генерализации более детальных моделей 1:200 000 и ASTER GDEM с привлечением данных о гидрографии, что преследовало две цели: получение более информативных и достоверных моделей, отражающих характер структурных линий и сопоставление производных горизонталей с эталонными образцами цифровых и бумажных топокарт.

*Цифровая модель рельефа суши GTOPO30*, имеющая разрешение 30'' или около 1 км по меридиану, использовалась в мелких масштабах картографирования. Ранее уже упоминалось о том, что эта модель имеет низкую географическая достоверность отдельных областей [Gamache, 2004]. Однако в первую очередь это касается тех участков, которые созданы на основе карт (в частности, DCW), где среднеквадратическая погрешность высот может составлять 100 и более метров. Картографируемая территория (Рисунок 4.9) попадает в область высококачественных данных DTED, которые, согласно документации, обеспечивают погрешность высот 30 м [USGS, 1996], что близко к эталонной точности изображения рельефа на карте 1 : 1 000 000. При этом модель недостаточно информативна для масштаба 1 : 1 000 000, а ее разрешение недостаточно подробно для четкого воспроизведения характера структурных линий способом отмывки. Поэтому в масштабе 1 : 1 000 000 использовались генерализованные модели ASTER GDEM и 1 : 200 000, а GTOPO30 применялась в масштабах 1 : 2 500 000 и мельче.



**Рисунок 4.9** — Источники данных, использованные при построении модели GTOPO30. Прямоугольником выделена область покрытия карты

*Цифровая модель рельефа суши ASTER GDEM* с разрешением 30 м была использована для картографирования Кавказа в масштабах 1 : 100 000 — 1 : 1 000 000. Согласно официальной документации среднеквадратическая ошибка модели по высоте в 95% случаев не превышает 20 метров [METI, NASA, 2009]. Для пустынных и горных слабо залесенных территорий точность, как правило, выше и может составлять порядка 5 — 10 метров по высоте, что соответствует эталонной точности изображения рельефа на карте масштаба 1 : 100 000. Сравнение с моделью SRTM подтвердило хорошее качество модели на выбранный участок. На территорию Среднерусской возвышенности модель GDEM имеет наоборот чрезвычайно низкое качество, изобилует артефактами, поэтому покрытие этой области высотными данными в средних масштабах осуществлялось на основе данных ЦТК 1 : 200 000.

Все картографические источники использовались для извлечения данных не только о рельефе и гидрографии, но и об остальных элементах содержания, включая геодезические пункты, поселения, транспортную и социально-экономическую инфраструктуру, растительность и грунты. Полученный состав слоев базы данных обеспечивает возможность мультимасштабного общегеографического картографирования территории в целом и рельефа в частности.

После проведения интеграции исходных данных цифровые модели в БД покрывали масштабы 1 : 25 000 — 1 : 200 000. Для обеспечения остальных масштабов и территорий картографирования была произведена генерализация ЦМР.

#### 4.3.5 Генерализация ЦМР

Тестирование представленного ранее алгоритма структурно-морфологической генерализации ЦМР (Раздел 2.5) на небольших участках в пределах каждой территории позволили методом проб и ошибок выработать оптимальные параметры генерализации.

В процессе тестирования было установлено, что длина тальвега и пороговое значение аккумуляции тока зависят и от степени уменьшения и от исходного разрешения модели (масштаба картографирования) и от того масштаба, которому сопоставлено исходное разрешение. При одной и той же степени генерализации (разрешение 10 и 50 метров, 100 и 500) параметры могут быть отличными, в чем, по-видимому, проявляется отличие генерализации рельефа в крупных, средних и мелких масштабах.

В средних масштабах картографирования прослеживается большая чувствительность параметров к морфологическим особенностям рельефа. В частности, площадь междуречий в равнинно-эрозионном и плоскогорном рельефе, как правило, больше, чем в альпийском. Это означает, что в среднем тальвеги у своего истока должны иметь большее накопление тока в равнинно-эрозионном рельефе и меньшее — в высокогорном альпийском. Если не учесть это правило, потеряется характер расчленения. В мелких масштабах эта разница уже не так заметна, поскольку междуречья сильно

расчлененного рельефа подвергаются слиянию из-за удаления форм, в результате чего по отношению к разрешению ЦМР их ширина увеличивается.

Выбираемые параметры также чувствительны к детализации исходной ЦМР. Модели, полученные путем генерализации ЦМР высокого разрешения либо на основе данных ДЗЗ, обычно более детализированы, чем те, что построены на основе горизонталей при аналогичном разрешении.

Отличия в морфологии и источниках данных, в частности, определили использование разных параметров генерализации ЦМР на Среднерусскую возвышенность (исходная модель построена на основе карты 1 : 200 000) и на Приэльбрусье (модель GDEM получена на основе данных ДЗЗ).

Особый случай представляет собой генерализация модели, разрешение которой уже близко к целевому. В частности, с этой проблемой пришлось столкнуться при получении ЦМР для масштаба 1 : 2 500 000. Модель GTOPO имеет разрешение 1 км, что для 1 : 1 000 000 слишком грубо, но уже приемлемо для 1 : 2 500 000. В этом случае допуски по аккумуляции тока и длине были ослаблены, иначе полученная модель была бы недостаточно детализированной.

При значительной генерализации рельефа, когда масштаб уменьшался в 5 и более раз, оказалось возможным не использовать полный набор исходных высот, а произвести предварительную передискретизацию модели, чтобы привести ее к более грубому разрешению. Передискретизация позволяет значительно сократить время генерализации ЦМР на большие территории, поскольку количество обрабатываемых узлов уменьшается в геометрической прогрессии. Этот факт особенно важен для алгоритма генерализации тальвегов, время работы которого пропорционально  $O(n^2)$ , где  $n$  — число ячеек ЦМР.

При уменьшении разрешения в 2 раза не наблюдается ощутимого смещения структурных линий с точки зрения целевого масштаба, когда генерализация происходит в 5 и более раз. В то же время число ячеек ЦМР уменьшается в 4 раза. При этом экспериментально было установлено, что даже при разнице разрешений передискретизированной и целевой ЦМР в 1.5 раза алгоритмом обеспечивается необходимое качество результата. Используемые параметры генерализации представлены в Таблице 4.2.

Масштабы 1 : 1 000 000 и крупнее изначально были обеспечены векторными данными топографических карт. Для заполнения уровней детализации, соответствующих более мелким масштабам, данные ЦТК были генерализованы путем отбора объектов, упрощения очертаний и обобщения характеристик.

**Таблица 4.2 – Параметры генерализации ЦМР**

Масштаб	Aster GDEM				ЦТК 1:200 000				GTOPO30			
	$R_0$	$R$	$A$	$L$	$R_0$	$R$	$A$	$L$	$R_0$	$R$	$A$	$L$
1:50 000 000	–	–	–	–	–	–	–	–	2000	10000	300	80
1:20 000 000	–	–	–	–	–	–	–	–	2000	5000	150	40
1:10 000 000	–	–	–	–	–	–	–	–	2000	2500	75	20
1:5 000 000	–	–	–	–	–	–	–	–	1000	1500	75	25
1:2 500 000	–	–	–	–	–	–	–	–	1000	1000	20	10
1:1 000 000	100	250	150	45	200	250	125	20	–	–	–	–
1:500 000	100	150	50	20	100	150	125	20	–	–	–	–
1:200 000	30	50	50	30	–	–	–	–	–	–	–	–

$R_0$  – исходное разрешение, м. Курсивом выделена передискретизация

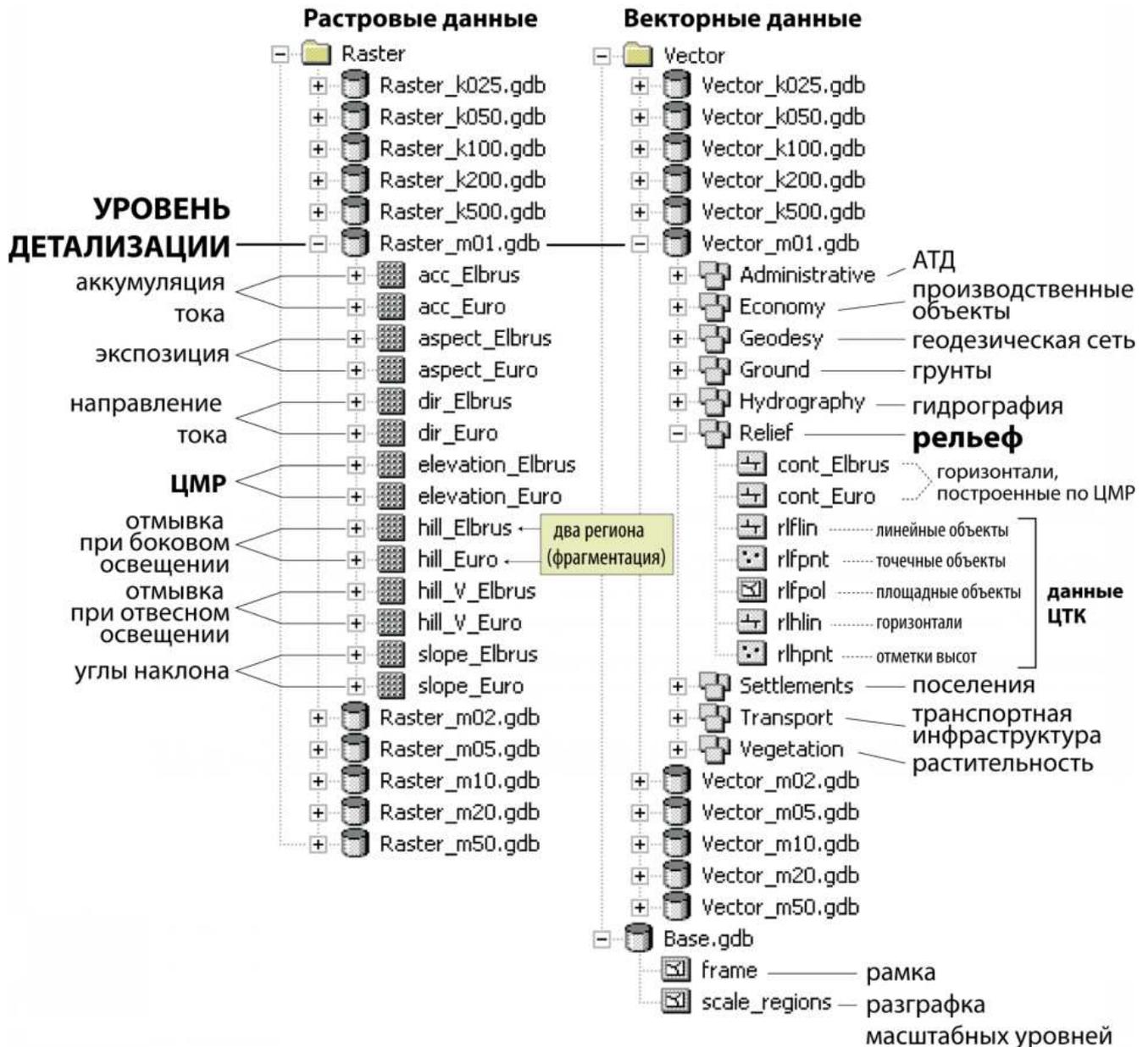
$R$  – целевое разрешение, м

$A$  – минимальная аккумуляция тока

$L$  – минимальная длина тальвега (число ячеек ЦМР)

### 4.3.6 Структура базы данных

После проведения интеграции и генерализации данных была получена база, обладающая искомой структурой и наполнением (Рисунок 4.10).



**Рисунок 4.10** – Структура и состав базы данных (на примере уровня детализации для масштаба 1 : 1 000 000 с фрагментацией покрытия на два региона)

Все данные поделены на уровни детализации, для удобства работы векторные и растровые данные были разнесены по отдельным группам, а растровые данные в случае фрагментации покрытия (масштабы 1:1 000 000 и крупнее) также разделены соответствующим образом на отдельные слои. Это позволило использовать ресурсы БД более экономно. Так, в частности, для масштаба 1:200 000 ЦМР на Кавказ имеет рекомендованное (параграф 2.6.2) разрешение 50 м, в то время как для Среднерусской возвышенности оказалось достаточным разрешение 100 м в силу менее расчлененного рельефа. Векторные

данные внутри уровней детализации сгруппированы по тематике.

### 4.3.7 Составление карты

Составление карты производилось согласно предложенной последовательности. При этом процесс значительно упростился благодаря тому, что структура подготовленной базы данных во многом повторяет структуру слоев карты.

#### 4.3.7.1 Разработка структуры и порядка слоев

При составлении карты был использован принцип группировки слоев: сначала по масштабу, далее — по тематике. Разработанная структура представлена на Рисунке 4.11.

В приведенном примере раскрыта масштабная группа 1 : 5 000 000, внутри нее раскрыты слои вспомогательных горизонталей, рек, озер и водохранилищ, а также населенных пунктов. Список слоев одновременно выполняет и функцию легенды (что является общей особенностью электронных карт), поэтому отдельное составление легенды не производилось. Состав слоев масштабных групп в целом повторяет данный, но постепенно видоизменяется от масштаба к масштабу. Так, например, в масштабах 1 : 1 000 000 и крупнее появляется возможность отображения данных о рельефе ЦТК (площадных, линейных, точечных форм рельефа), а в масштабах 1 : 10 000 000 и мельче не предусмотрено отображение дорог, поскольку они мешают восприятию рельефа.

Каждой масштабной группе был назначен свой диапазон применимости согласно Таблице 4.2. Можно обратить внимание (на Рисунке 4.11), что все масштабные группы кроме 1 : 2 500 000 отмечены серым флажком, в то время как текущая группа выделена черным. Серый цвет означает, что группа (или слой) находится вне диапазона применимости и не может быть отображена в текущем масштабе. Если группе или слою назначен диапазон, переключение групп и изменение цвета галочки при смене масштаба происходит автоматически. Поскольку данная карта насчитывает порядка 100 слоев, это значительно облегчает работу с ней. Для того чтобы ознакомиться с условными обозначениями для текущего вида карты, достаточно найти активную масштабную группу.

#### 4.3.7.2 Шкалы послыной и градиентной окраски

Для достижения наибольшего пластического эффекта и информативности изображение рельефа было составлено комбинацией способов горизонталей, цветовой окраски (послойной и градиентной), а также отмывки при боковом и отвесном освещении.

Путем эксперимента было установлено, что в масштабах 1:2 500 000 и мельче оптимально использование послыной окраски. В более крупных масштабах разделение на цветовые слои уже не целесообразно, поскольку для достижения пластического эффекта

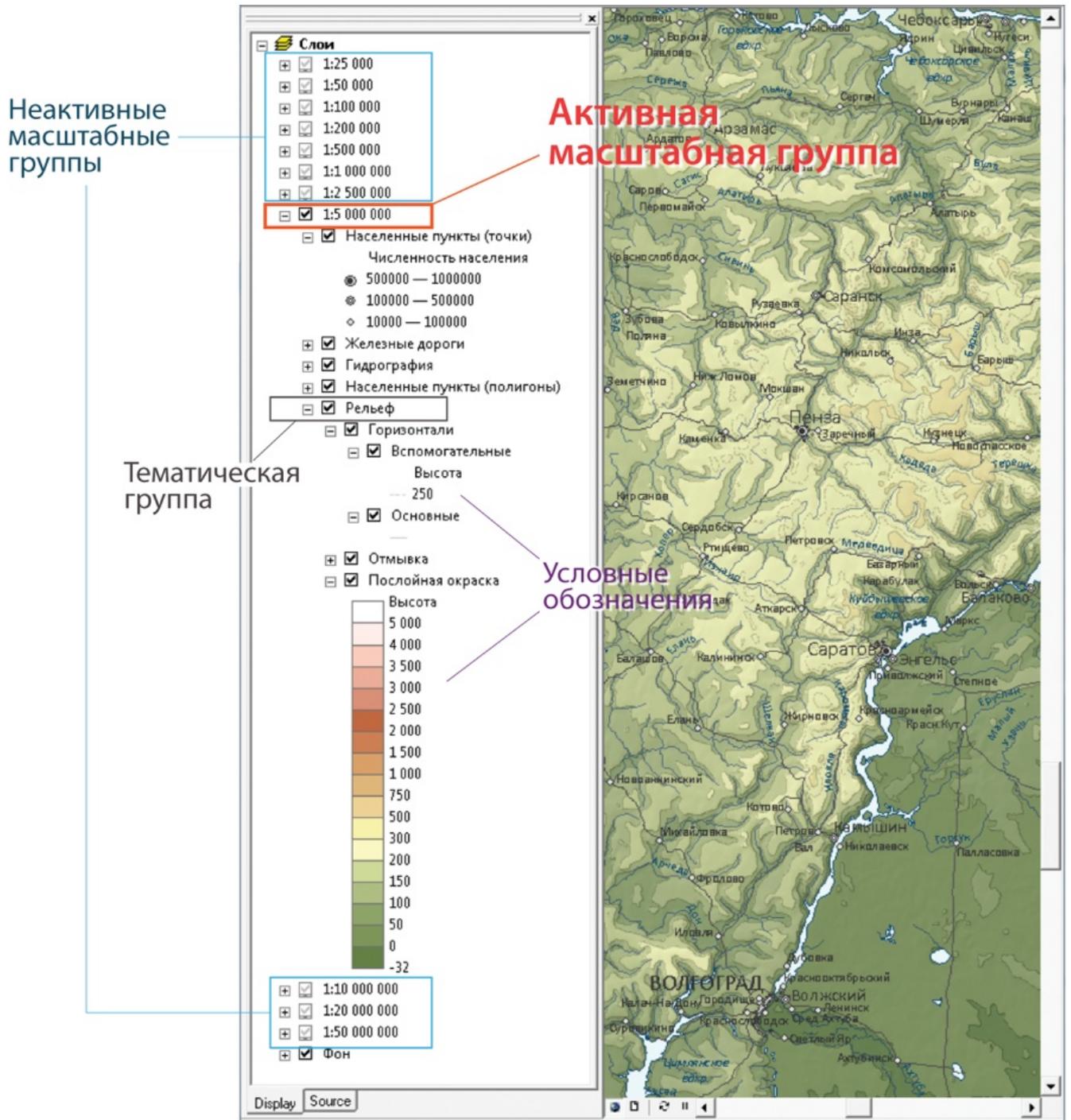


Рисунок 4.11 – Структура слоев мультимасштабной карты (на примере масштаба 1 : 2 500 000)

понадобилось бы большое количество слоев (20 и более), что осложняет их сопоставление с легендой. Соответственно, в масштабах 1:2 500 000 и мельче источником информации о высоте является цвет, и горизонталы даны без подписей. В более крупных масштабах используется градиентная цветовая окраска, а высоты на карте определяются по горизонталям, основные из которых подписаны.

Шкалы сечения послойной окраски была подобраны путем анализа углов наклона по ЦМР в каждом масштабе с учетом преимущественного заложения не менее 0,5 мм, а также на основе визуальной оценки ЦМР, представленной в различных шкалах сечения.

При разработке цветовых шкал был реализован принцип преемственности цветового ряда. На выбор цветов оказывали влияние следующие факторы, имеющие место при укрупнении масштаба:

1. Увеличение площади цветовых пятен.
2. Разделение слоев на промежуточные градации.
3. Изменение соотношения соседних интервалов (в случае если текущий интервал остался без изменения, а соседний подвергся дроблению).

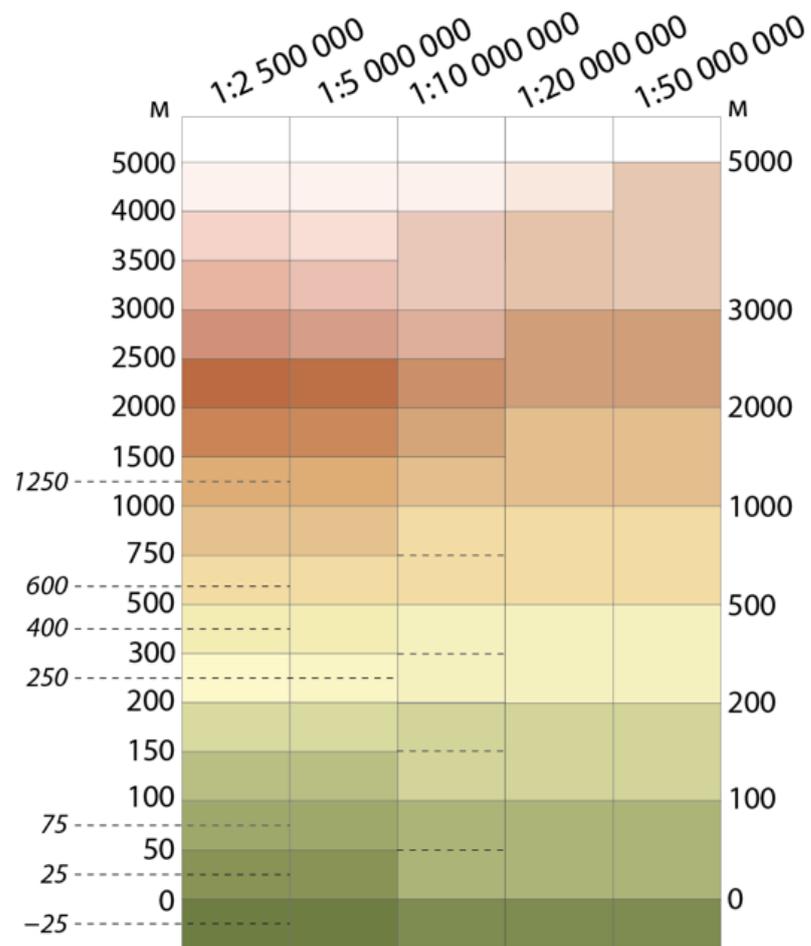
Разработанные шкалы (Рисунок 4.12) являются примером комбинированного цветового ряда, в котором до высот 2000 — 3000 метров используется возрастание насыщенности и теплоты тона, а выше — увеличение светлоты. В шкалах избегаются излишне насыщенные и яркие цвета, от которых быстро утомляется зрение при работе с электронными картами. При этом шкалы ориентированы на комбинацию с отмывкой, что также определяет их осветляющийся характер в верхней части спектра.

На основе диаграммы можно легко установить, например, что постепенная цветовысотная дифференциация горных территорий происходит уже в масштабах 1 : 20 000 000 и 1 : 10 000 000, в то время как дифференциация равнин в диапазоне 0 — 300 м становится возможной только в масштабе 1 : 5 000 000. В частности, это позволяет наглядно и пластично охарактеризовать рельеф Восточно-Европейской равнины.

В масштабах 1 : 1 000 000 и крупнее была использована градиентная окраска рельефа, при этом были взяты базовые цвета со шкалы 1 : 2 500 000. Для территории Среднерусской возвышенности использовались цвета диапазона 0 — 500 м, а для Приэльбрусья — полная шкала цветов.

#### **4.3.7.3 Шкалы сечения горизонталей в масштабах 1 : 2 500 000 и мельче**

Эти масштабы карты являются гипсометрическими. При разработке шкал для них, было стремление достичь максимальной информативности в изображении закономерностей распределения высот по территории. В этих масштабах шкала сечения горизонталей повторяет шкалу сечения послойной окраски, однако там, где это необходимо для



**Рисунок 4.12** – Диаграмма шкал послонной окраски для мультимасштабной карты рельефа. Пунктиром показаны вспомогательные уровни горизонталей

отражения морфологии рельефа (и при этом позволяет масштаб), введены вспомогательные горизонталы, которые не заполняют все пространство карты, а обрываются при узких заложениях либо при их нерепрезентативности. Проведение вспомогательных горизонталей потребовало экспертной оценки и ручного редактирования. Для этого была разработана следующая последовательность действий:

1. Построение горизонталей в шкале сечения, соответствующей шкале послойной окраски. Выделение областей, недостаточно информативно представленных в данной шкале. Определение абсолютных высот и превышений в этих областях путем анализа ЦМР, установление необходимых уровней вспомогательных горизонталей.
2. Построение вспомогательных горизонталей по ЦМР.
3. Ручное редактирование горизонталей путем их «разрезания». Сегменты, попавшие в области с узкими заложениями, помечаются в специальном атрибутивном поле как «не отображаемые». Таким образом, линии физически не удаляются и по-прежнему хранятся в базе данных, но отображаются только необходимые их участки.
4. Визуализация вспомогательных горизонталей особым условным обозначением (штриховой линией).

В масштабе  $1 : 50\,000\,000$  используется шкала  $0 - 100 - 200 - 500 - 1000 - 2000 - 3000 - 5000$  м. При переходе к масштабу  $1 : 20\,000\,000$  эти уровни дополняются горизонталью  $4000$  м.

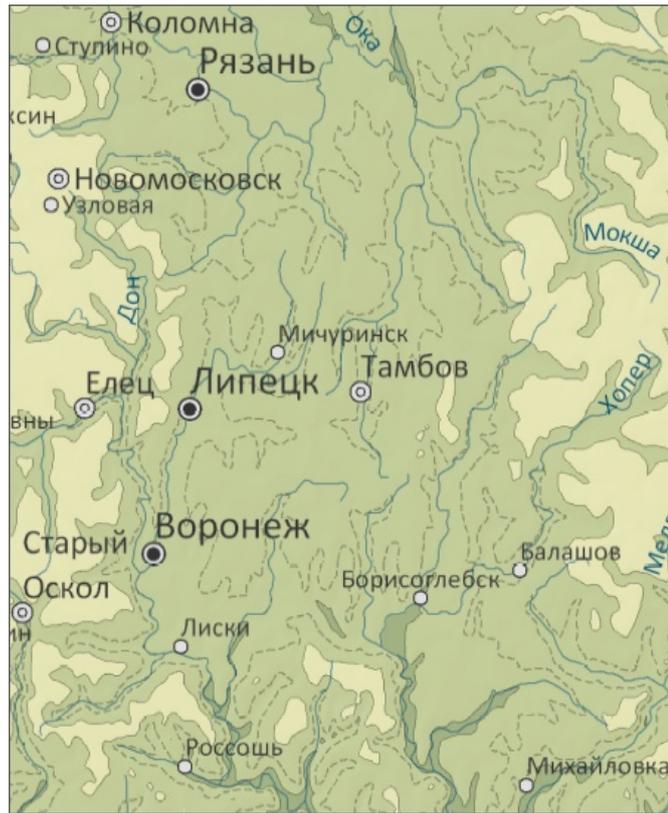
В масштабе  $1 : 10\,000\,000$  используется шкала  $0 - 100 - 200 - 500 - 1000 - 1500 - 2000 - 2500 - 3000 - 4000 - 5000$  м. В этом масштабе введены вспомогательные горизонталы  $50, 150, 300$  и  $750$  м.

Горизонталь  $50$  м позволила отразить рельеф Северодвинской, Печорской, Кубано-Приазовской и Причерноморской низменностей, а также северной части полуострова Крым.

Горизонталь  $150$  м существенно усилила информативность изображения рельефа, будучи близкой к характерной для Восточно-Европейской равнины высоте  $170$  м. В частности, она проявила рельеф Окско-Донской равнины, который в основной шкале сечения полностью отсутствовал (Рисунок 4.13). Благодаря этому уровню гораздо лучше отрисован рельеф Украины между Донецким кряжем и Подольской возвышенностью, а также степные районы к югу от Среднерусской возвышенности.

Горизонталь  $300$  м проявила ранее не выражающийся рельеф Подольской возвышенности, Среднего Урала и Мугоджар, позволила лучше отрисовать рельеф Южного Урала, западных предгорий Северного Урала, а также предгорий Кавказа и Карпат.

Горизонталь  $750$  м повысила информативность изображения Южного Урала, подчеркнула продольный характер расположения хребтов.



**Рисунок 4.13** — Изображение Окско-Донской равнины вспомогательной горизонталью 150 м в масштабе 1 : 10 000 000

В масштабах 1 : 5 000 000 и 1 : 2 500 000 используется полная шкала 0 — 50 — 100 — 150 — 200 — 300 — 500 — 750 — 1000 — 1500 — 2000 — 2500 — 3000 — 3500 — 4000 — 4500 — 5000 м. В масштабе 1 : 5 000 000 введена дополнительная горизонталь 250 м, которая проявила рисунок междуречий Среднерусской и Приволжской возвышенностей, а также предгорий Урала. Эта же горизонталь сохраняется и в масштабе 1 : 2 500 000 (Рисунок 4.14), будучи дополненной следующими уровнями: —25, 25, 75, 250, 400, 600 и 1250 м.

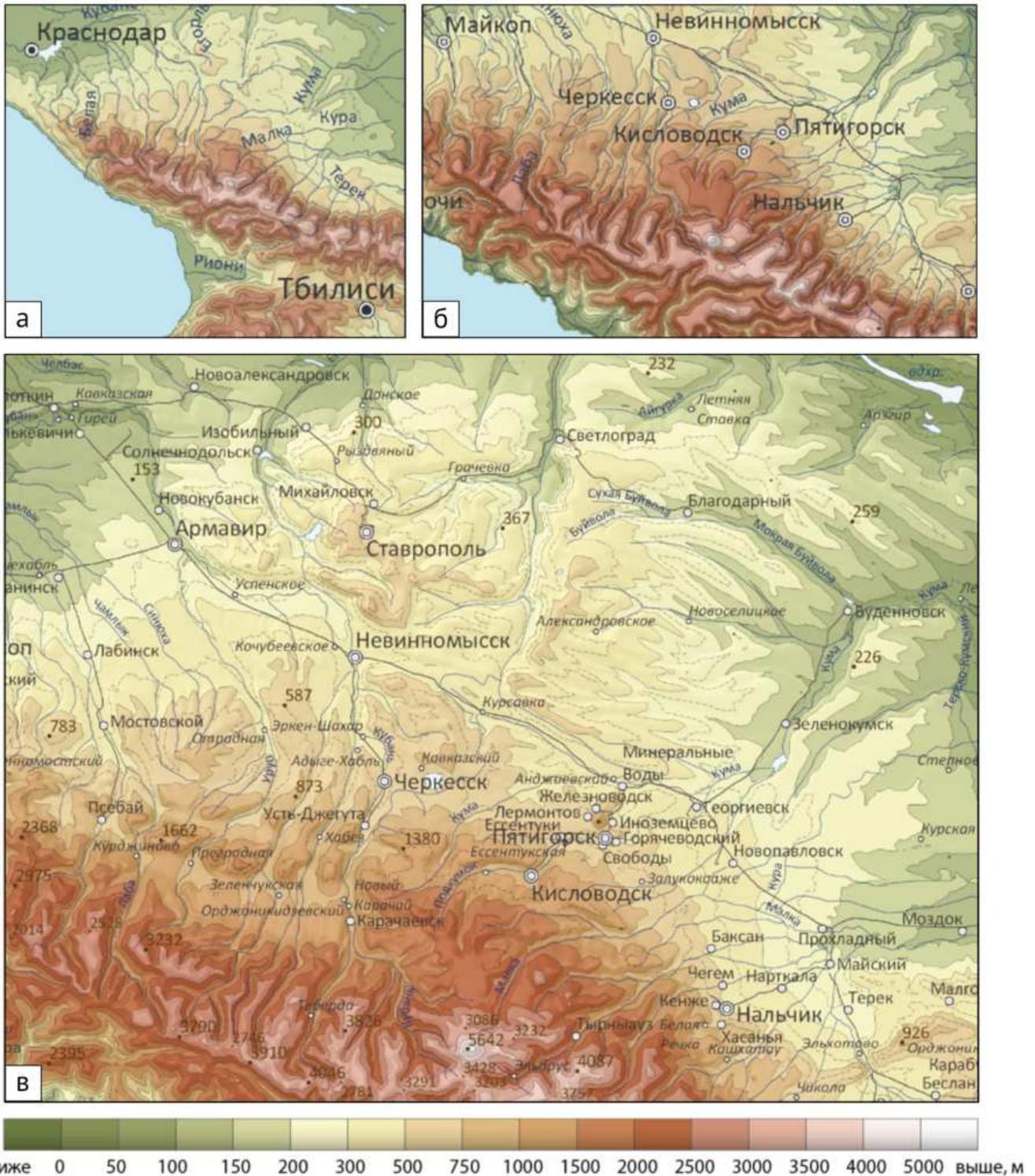
Горизонтали —25 и 25 м позволили отрисовать рельеф Прикаспийской низменности.

Горизонтали 25 и 75 метров дополнили 50-ю горизонталь на Печорской, Кубано-Приазовской и Причерноморской низменностях.

Горизонталь 400 м оконтуривала наиболее высокие участки междуречий Бугульминско-Белебеевской возвышенности, а также позволила значительно подробнее охарактеризовать рельеф Среднего Урала, предгорий Южного Урала и Кавказа (Рисунок 4.15).

Горизонтали 600 и 1250 м выделили не выражающиеся в сечении формы рельефа Южного Урала и Северного Кавказа.





**Рисунок 4.15** – Изображение северных предгорий Большого Кавказа в масштабах (а) 1 : 10 000 000; (б) 1 : 5 000 000; (в) 1 : 2 500 000 с использованием вспомогательных горизонталей 250, 400, 600 и 1250 м

ЦМР, построенной на основе ЦТК 1 : 200 000. ЦМР оказалась выположенной между горизонталями, исходное сечение которых равно 20 м.

В данной группе масштабов горизонтали подписаны, будучи основным источником информации о высоте. Градиентная цветовая окраска выполняет в основном чисто пластическую роль. В то же время, приблизительная оценка высот по ней также возможна. Населенные пункты в этих масштабах специально даны в приглушенных тонах, а горизонтали идут поверх, что позволяет акцентировать внимание на изображении рельефа

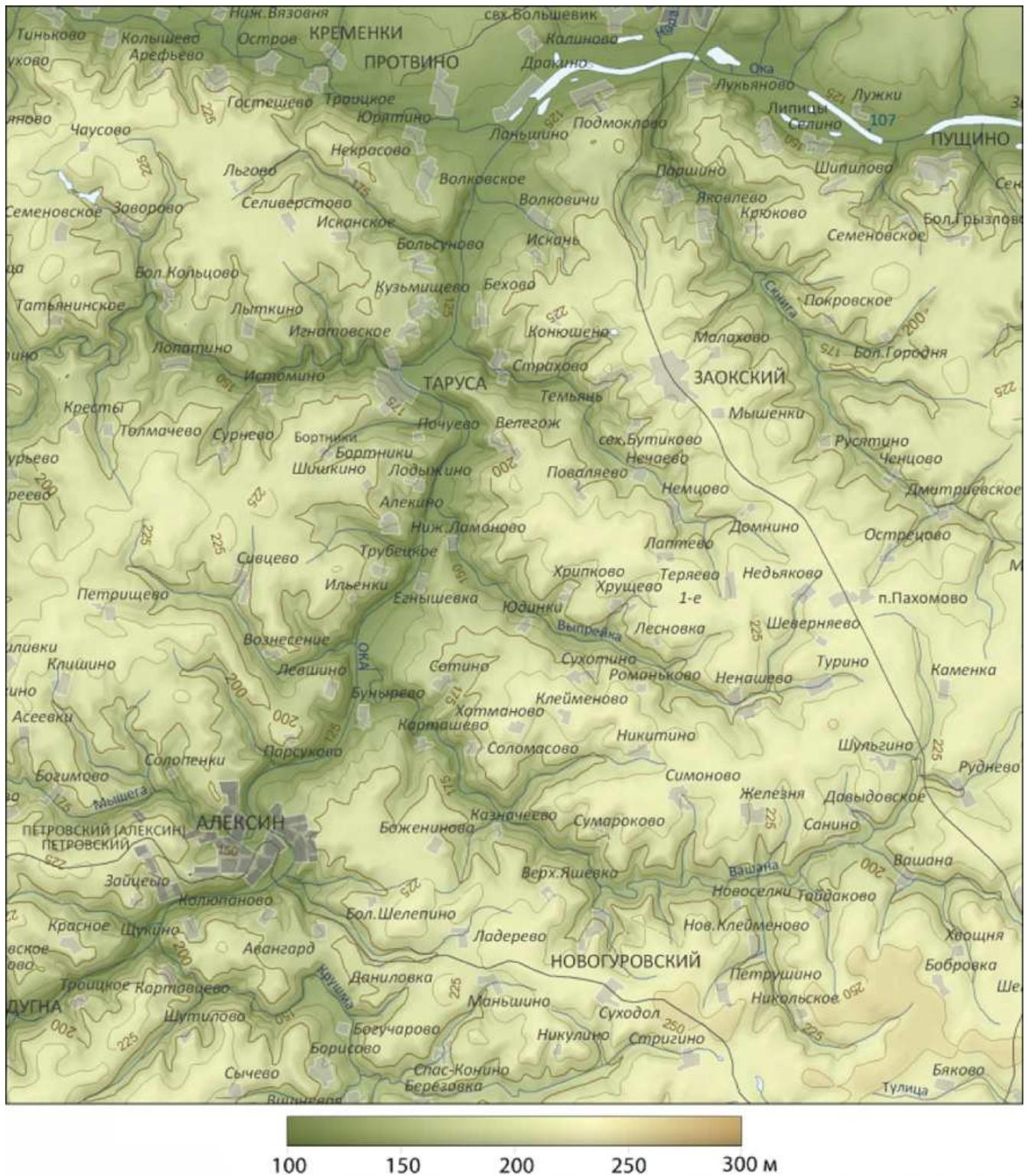
Была также произведена оптимизация геометрического представления горизонталей в БД, поскольку они содержали избыточное количество точек. Это обусловлено как высоким разрешением самих ЦМР, так и тем, что алгоритм построения горизонталей в ArcGIS использует триангуляцию ячеек. Для решения этой проблемы горизонтали были обработаны с помощью алгоритма Дугласа-Пейкера (в ArcGIS этот инструмент носит название Point Remove).

Путем экспериментов было установлено, что оптимальный результат достигается при ширине допуска отклонения 0.1 мм в масштабах 1 : 200 000 и крупнее и 0.05 мм в более мелких масштабах. При этом визуальное отличие от исходных линий не наблюдается, но число узлов горизонталей было уменьшено на порядок, а скорость отрисовки и надписывания возросла в несколько раз.

#### 4.3.8 Итоги

В настоящем разделе диссертации было показано, как принцип мультимасштабности реализуется при гипсометрическом картографировании. Созданная карта содержит 11 уровней детализации, каждый из которых может быть использован для изучения рельефа с определенным уровнем обобщения в зависимости от охвата и размеров форм рельефа. Все масштабы отличаются преемственностью условных обозначений и цветовых шкал, что отвечает принципам составления мультимасштабных карт и обеспечивает сопоставление изображений в разных масштабах. Диапазон 1 : 2 500 000 — 1 : 50 000 000 является собственно гипсометрическим и предназначен для изучения и анализа общих закономерностей строения рельефа на макроуровне. При этом масштабы 1 : 2 500 000 и 1 : 5 000 000 отличаются высокой степенью детализации, что обеспечивается подробной шкалой цветовой окраски и введением дополнительных горизонталей.

Масштабы 1 : 1 000 000 и 1 : 500 000 обеспечивают плавный переход к детальному исследованию рельефа Кавказа и южной половины Восточно-Европейской равнины в масштабах 1 : 25 000 — 1 : 200 000. Наличие готовых моделей углов наклона и экспозиции позволяет в интерактивном режиме получать эти характеристики в точках интереса. Разрешение и точность моделей в этих масштабах обеспечивают проведение морфометрического анализа рельефа. Модели направления и аккумуляции тока позволяют в



**Рисунок 4.16** – Изображение равнинно-эрозионного рельефа в масштабе 1 : 500 000. Сечение горизонталей 25 м

интерактивном режиме получать для точки интереса (или группы точек) водосборный бассейн.

Карта может быть использована также для привязки различных природных границ (тектоники, растительности, почв, ландшафтов и т.д.) к формам рельефа определенного порядка и установления связи между ними. Для этого необходимо добавить изучаемые границы в список слоев карты и расположить их поверх масштабных групп. При изменении масштаба карты будет происходить переключение масштабных групп, сопровождающееся генерализацией изображения рельефа, при этом детализация изучаемого слоя будет оставаться неизменной. Это позволит определить масштаб (и соответствующий иерархический уровень форм рельефа), в котором наблюдается наиболее близкое совпадение контуров изображений. Пространственная корреляция между ними может быть получена и автоматизированно путем оверлея соответствующих слоев в базе данных.

## 4.4 Мультимасштабная визуализация цифровых моделей характеристик ветрового волнения<sup>5</sup>

### 4.4.1 Введение

На современном этапе развития науки огромные массивы океанологических и метеорологических данных получают по результатам моделирования. При этом специфика процессов в океане заключается в существенном возрастании их изменчивости вблизи берегов, что стимулирует использование моделей на нерегулярных триангуляционных сетках, в которых плотность узлов возрастает в узких проливах и на малых глубинах [Cavaleri *et al.*, 2007]. Результаты подобного моделирования представляют сложность с точки зрения мультимасштабного картографирования, поскольку требуют генерализации не только в вертикальном измерении (при уменьшении масштаба), но также и в горизонтальном (с учетом значительного градиента детальности, обусловленного нерегулярностью сетки). Исследование возможностей отображения таких пространственных данных на мультимасштабных картах было выполнено в два этапа. На первом этапе принципы, аналогичные созданной гипсометрической карте, были применены при создании *Веб-атласа ресурсов волновой и ветровой энергии морей России* [Мысленков *et al.*, 2018; Мысленков *et al.*, 2020; Myslenkov *et al.*, 2023]. В частности, для каждой характеристики было спроектировано три масштабных уровня, кратно отличающихся интервалом изолиний, что позволило реализовать принцип последовательного преобразования изображения. На втором этапе был впервые предложен и реализован оператор утрирования высокоизменчивых прибрежных участков [Карташов, Самсонов, Мысленков, 2024],

<sup>5</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Myslenkov *et al.*, 2023]

который основан на пересчете значений в узлах с учетом расстояния от берега, градиента расстояния от берега, градиента непосредственно самого поля, а также расстояний до соседних узлов триангуляции.

#### 4.4.2 Исходные данные

Моделирование характеристик волн проводилось *С. А. Мысленковым* с использованием модели WAVEWATCH [2019] на нескольких нерегулярных сетках для прибрежных морей России (Рисунок 4.17). Пространственное разрешение составляло 10–20 км в центральных открытых районах морей и 300–1400 м в прибрежных зонах. Использовано 7 расчетных доменов, характеристика которых приведена в Таблице 4.4. Нерегулярная триангуляционная сетка для Баренцева и Карского морей включала Североатлантический регион, а для дальневосточных морей — северную часть Тихого океана. Сгущение триангуляции вблизи берегов позволило получить высокое пространственное разрешение не только в прибрежных зонах, но и в узких проливах. При использовании подобной сетки волновая энергия может свободно проходить из Тихого океана через проливы Курильской гряды и Алеутских островов. Пример такой конфигурации представлен на Рисунке 4.18. Поля ветрового волнения, как выход модели, были получены с дискретностью раз в 3 часа с 1979 по 2019 год (всего 41 год) для Баренцева, Карского, Берингова, Охотского и Японского/Восточного морей и с 1979 по 2020 год (всего 42 года) для Черного и Каспийского морей. По результатам моделирования *С. А. Мысленковым* получены поля среднемноголетних (за весь период) значений высоты значительных волн, длины, периода волн и потока волновой энергии. Подробная статистика по точности моделирования приведена в [Myslenkov et al., 2023]

#### 4.4.3 Геоинформационное моделирование

Результаты моделирования волн и ветра должны быть представлены в виде моделей пространственных данных, которые могут быть обработаны программным обеспечением ГИС и представлены в картографическом виде. Такая трансформация требует выбора оптимальных моделей данных и программирования технологического перехода между исходными и пространственными данными. Основным требованием к такому переходу является минимизация ошибки представления — то есть, ГИС-представление должно максимально точно отражать данные моделирования. Второе требование — возможность визуализации полученной модели данных в картографическом виде. Для представления геополей традиционно используется метод изолиний с послойной окраской [Берлянт, 1986]. Соответственно, выбор подходящей модели пространственных данных должен принимать в расчет необходимость извлечения изолиний. Третье и последнее требование — возможность выполнять пространственные запросы по параметрам волн

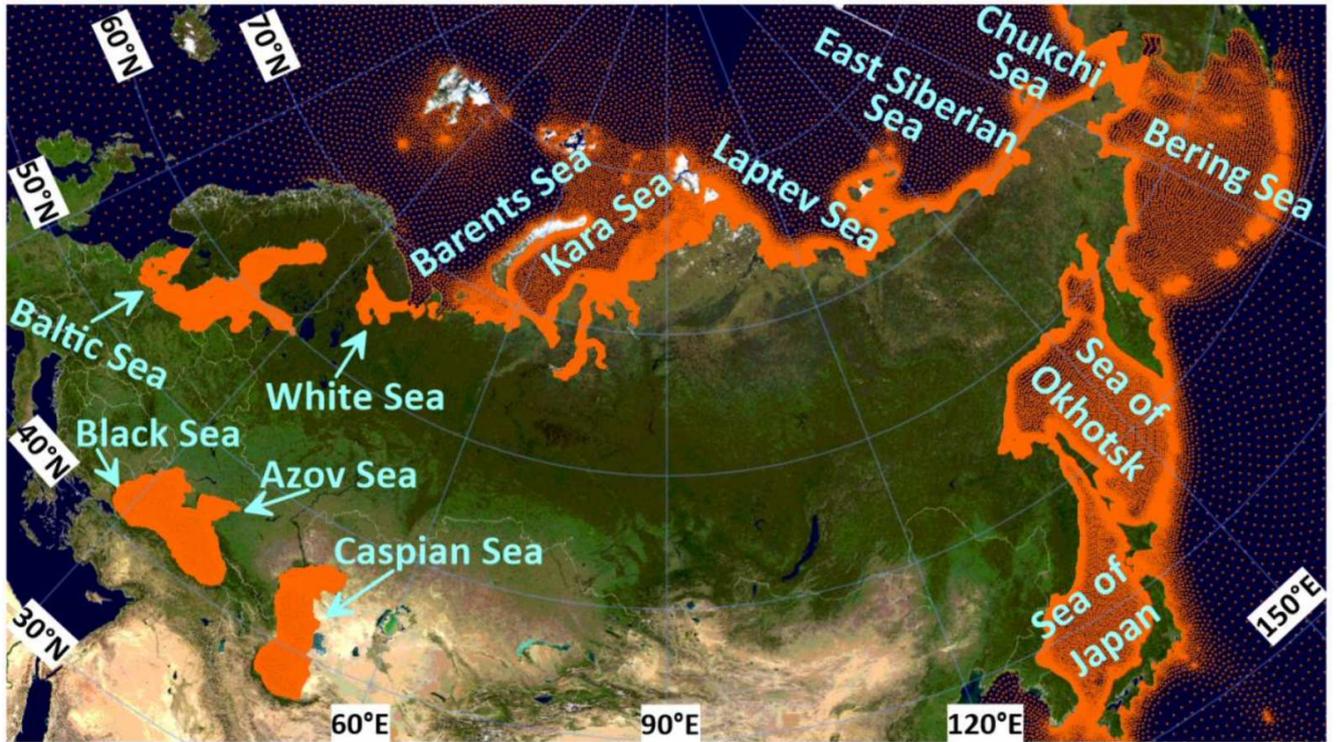
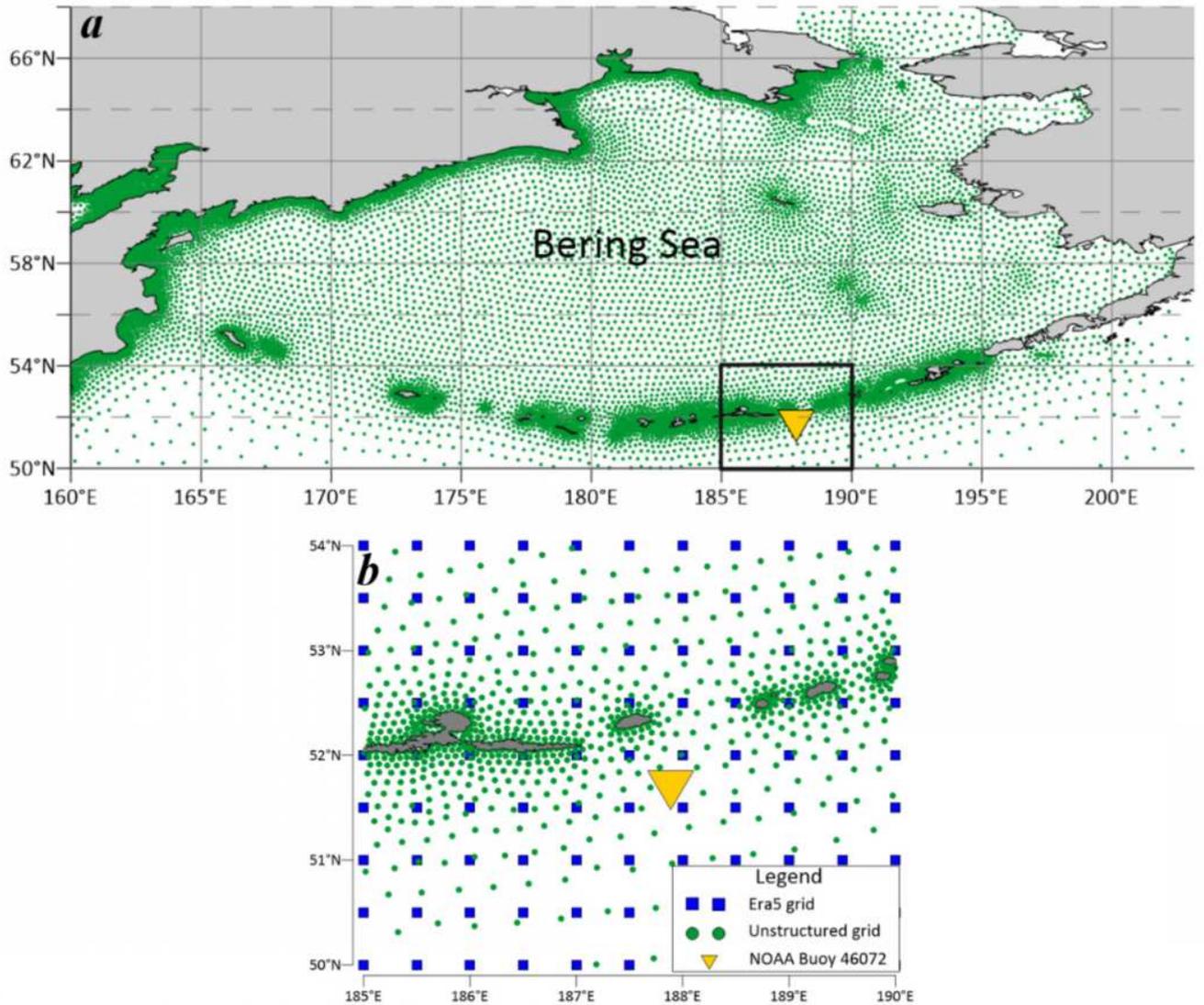


Рисунок 4.17 – Расчетные узлы триангуляционных сеток, использованных при моделировании [Myslenkov *et al.*, 2023]

Таблица 4.4 – Вычислительные домены

Домен	Включенные моря	Число узлов	Пространственное разрешение, км		Особенности
			Прибрежная зона	Открытое море	
1	Карское, Баренцево	37729	0.7	10-20	Западный сектор Российской Арктики + Северная Атлантика
2	Лаптевых, Восточно-Сибирское, Чукотское	36176	0.8	10-15	Восточный сектор Российской Арктики до 87° с.ш. + море Бофорта
3	Беринга, Охотское, Японское	69333	1.4	15-25	Северная часть Тихого океана
4	Черное, Азовское	59690	0.3	5-10	–
5	Каспийское	17529	0.8	10	–
6	Балтийское	34985	0.4	10	–
7	Белое	26587	0.4	10	Баренцево море включено



**Рисунок 4.18** – Конфигурация модели WAVEWATCH для Берингова моря: (а) общий охват; (б) фрагмент Алеутской гряды [Myslenkov *et al.*, 2023]

и ветра, то есть пользователь должен иметь возможность получить значение в любой заданной точке. Соответственно, результирующая модель данных должна относиться к классу покрытий в понимании ISO/OGC [Baumann, Hirschorn, Masó, 2017].

Модель WAVEWATCH, использованная для расчетов характеристик волнения, оперирует на нерегулярной триангуляционной сетке. В то же время, визуализация триангуляционных структур данных не поддерживаются современными стандартами веб-сервисов OGC(WMS, WFS), а также программными библиотеками визуализации данных (такими как OpenLayers, Leaflet, Mapbox), которые необходимы для картографической визуализации в веб-среде. Следствием данного технологического ограничения является необходимость преобразования триангуляционных поверхностей (TIN) в вид, подходящий для картографического представления. В рамках исследования было исследовано два подхода: растровый [Мысленков *et al.*, 2018] и векторный [Мысленков *et al.*, 2020].

#### 4.4.3.1 Растровый подход

Данный подход заключается в преобразовании триангуляционной модели в растровую. Для этого был использован метод иерархических базисных сплайнов (ИБС) [Lee, Wolberg, Shin, 1997]. В данном методе каждый участок поверхности размером  $4 \times 4$  ячейки представляет собой бикубическую поверхность. Высоты в ячейках получаются, исходя из условия прохождения поверхности через исходные точки, а также минимизации суммы квадратов высот (поскольку существует бесконечное число различных поверхностей на сетке  $4 \times 4$ , которые проходят через заданную точку). При иерархическом подходе поверхность строится в несколько итераций с последовательным переходом на более детальное разрешение сетки. При этом на каждой последующей итерации аппроксимации подвергаются остатки между исходными значениями в точках и теми, которые получаются по бикубической поверхности. Достоинством метода иерархических базисных сплайнов является то, что поверхность получается сразу для всех узлов, нет необходимости решать систему линейных уравнений для каждого узла сетки. Метод является локальным: исходные точки, удаленные от текущего узла цифровой модели рельефа (ЦМР) далее чем на 2 узла, не оказывают на нее влияния. В результате этого метод ИБС является быстрым и эффективным в вычислительном плане. Помимо этого, мультимасштабность метода позволяет эффективно использовать его при интерполяции данных, распределенных неравномерно по пространству, а также получать генерализованные варианты поверхности путем интерполяции с малым числом итераций.

Для оценки необходимого разрешения выходного растра в процессе интерполяции необходимо вычислить для массива исходных точек среднее расстояние до ближайшего соседа (Nearest Neighbor Distance, NND), которое дает оценку пространственной частоты дискретизации. Согласно теореме Котельникова, потерь данных можно избежать,

если частота дискретизации (сигнала, данных) будет вдвое больше максимальной пространственной частоты (частоты Найквиста). Интерпретируя это утверждение в терминах растрового анализа, можно сказать, что разрешение растра должно быть, по крайней мере, в 2 раза выше, чем минимальное расстояние до ближайшего соседа в анализируемом множестве точек. Произведенные расчеты показали, что для представленных на акваторию Черного моря данных достаточно разрешения 200 метров.

Наконец, важным, но малозаметным для пользователя аспектом построения растровой модели является вопрос выбора области распределения. Для океанологических показателей специфична обрезка этой области контуром береговой линии. При этом в разных масштабах используются представления береговой линии соответствующей детализации. Это означает, что интерполируемая поверхность должна иметь область значений, перекрывающую территорию суши таким образом, чтобы при ее маскировании береговой линией различной детализации не появлялись пустые ячейки. Исходя из этого, полученные растровые поверхности имеют размер  $9000 \times 5400$  ячеек, а их версии, маскированные контуром береговой линии, имеют размер  $8000 \times 4400$  ячеек.

При использовании растровой модели послойная окраска получается путем окраски ячеек растра, а изолинии извлекаются на лету. При этом для увеличения скорости извлечения для каждого отображаемого на картах показателя рассчитаны 8 уровней растровой пирамиды в целях оптимизации скорости и качества отображения в мелких масштабах картографирования. Растровая пирамида представляет собой множество растровых поверхностей, полученных путем последовательного увеличения размера ячейки исходного растра (передискретизация) с соответствующим перевычислением отнесенных к каждой ячейке величин. Каждый уровень пирамиды получен передискретизацией исходного растра с удвоением исходного разрешения методом билинейной интерполяции. Результирующие узлы сетки совпадают с центрами ячеек, образованных четырьмя соседними узлами исходной сетки. Таким образом, в зависимости от масштабного уровня карты появляется возможность извлечения изолиний из соответствующего по детализации уровня растровой пирамиды.

Выполнение же третьего условия, касаемого получения данных в любой произвольной точке обеспечивается автоматически тем, что растровая модель относится к покрытиям [Baumann, Hirschorn, Masó, 2017].

#### 4.4.3.2 Векторный подход

Растр, полученный на акваторию Черного моря имеет размер 48, 6 миллионов узлов. Однако модель для Черного и Азовского морей, согласно Таблице 4.4, содержит всего 59690 узлов. Это означает, что растровая структура данных крайне неэффективна для представления результатов в ГИС, в особенности на высоких уровнях детализации, где не используются растровые пирамиды. Поскольку требуемое картографическое пред-

ставление волновых характеристик является изолинейным, было предложено компромиссное решение, основанное на сочетании двух моделей пространственных данных: триангуляционной (TIN) и векторной.

Сначала триангулированная поверхность была реконструирована из координат сетки моделирования. Поскольку моделирование проводилось независимо для каждого моря, возникла проблема интеграции данных для всех соседних морей. Для получения бесшовного покрытия ГИС-данными каждой характеристики волн и ветра в обработку был введен промежуточный этап. Сначала точки моделирования для всех морей были объединены в один слой TIN. Затем этот слой редактировался для удаления возможных краевых эффектов в областях, где области моделирования совпадают. В результате получилась структура данных TIN, идентичная той, что используется в модели WW3.

Изолинии по триангуляционной модели были выделены с шагом значений, характерным для каждой характеристики волнения:

1. Максимальная высота значительных волн, сечение 1 м.
2. Максимальная высота волн, сечение 2 м.
3. Средняя высота волн, сечение 0.1 м.
4. Средняя энергия волн, сечение 0.5 — 2 кВт/м.
5. Средняя длина волн, сечение 2 — 5 м.
6. Средний период волн, сечение 0.5 с.
7. Обеспеченность волновой энергией более 1 кВт/м волнового фронта, сечение 5%.

При использовании линейной интерполяции поверхность внутри каждого треугольника представляется как часть плоскости, проходящей через его точки. Это приводит к тому, что изолинии представляют собой отрезки внутри каждого треугольника и могут быть извлечены очень быстро.

#### 4.4.4 Картографическое представление

##### 4.4.4.1 Базовые картографические стили

Для представления энергетических характеристик в картографическом виде было выполнено проектирование картографических стилей отображения. Цветовые шкалы униполярных и биполярных показателей для скалярных величин были прописаны в формате SLD (StyledLayerDescriptor) на основе системы цветовых палитр Color Brewer [Harrower, Brewer, 2003].

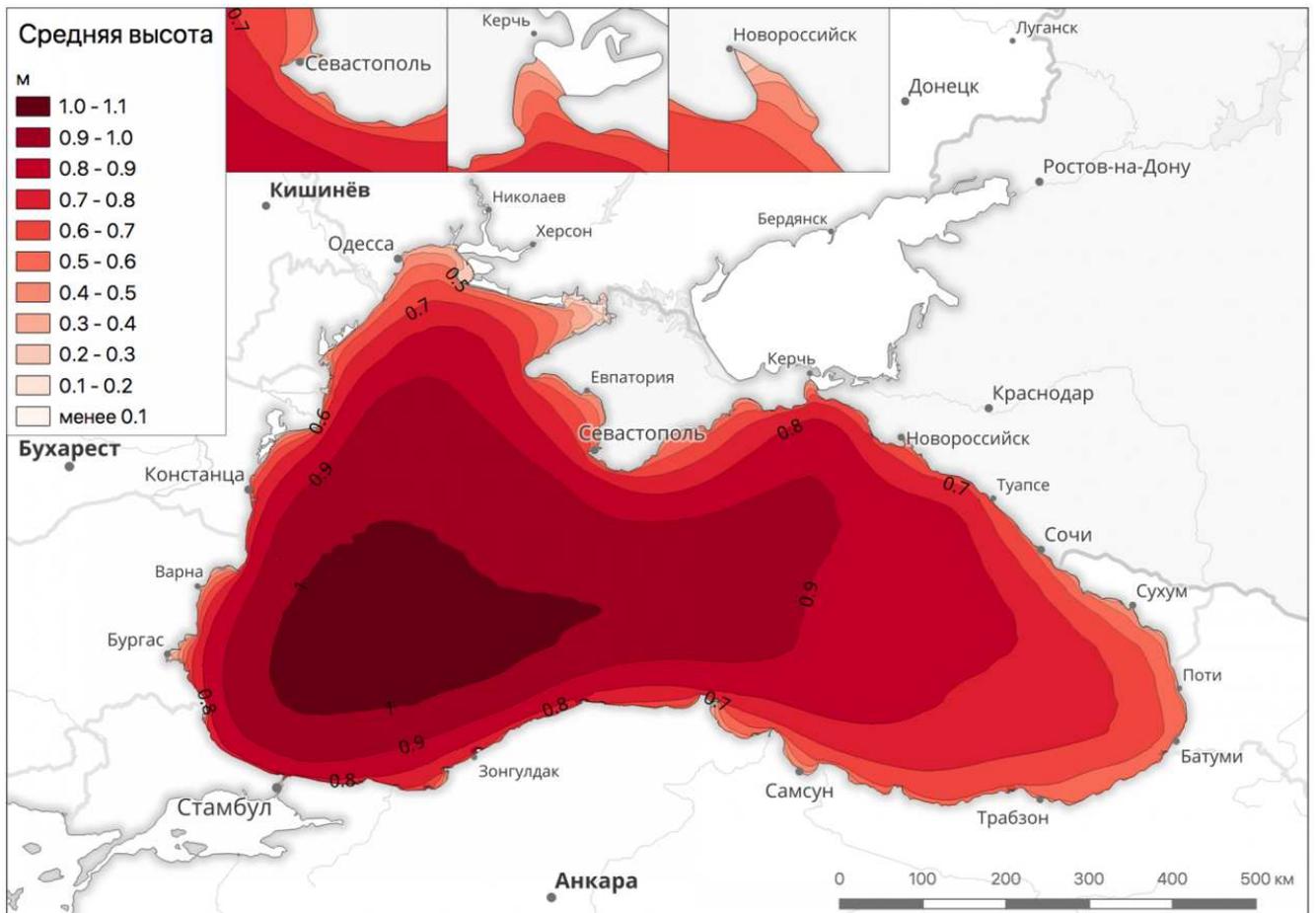
- Униполярные показатели имеют один характерный вектор изменения величины (от минимума к максимуму). Для их картографирования используются последовательные шкалы изменения светлоты и насыщенности цвета при возмож-

ном изменении цветового тона. Для базовых физических характеристик, таких как высота, длина и период волны, скорость ветра, предложено использовать однотонные последовательные шкалы (зеленого, синего, красного, фиолетового, оранжевого тона). Для производных характеристик, отражающих непосредственно величину энергии, а также ее обеспеченность, предложено использовать более сложные многоцветные шкалы. Данное разделение способствует наглядному разделению показателей при их визуальном анализе по картам.

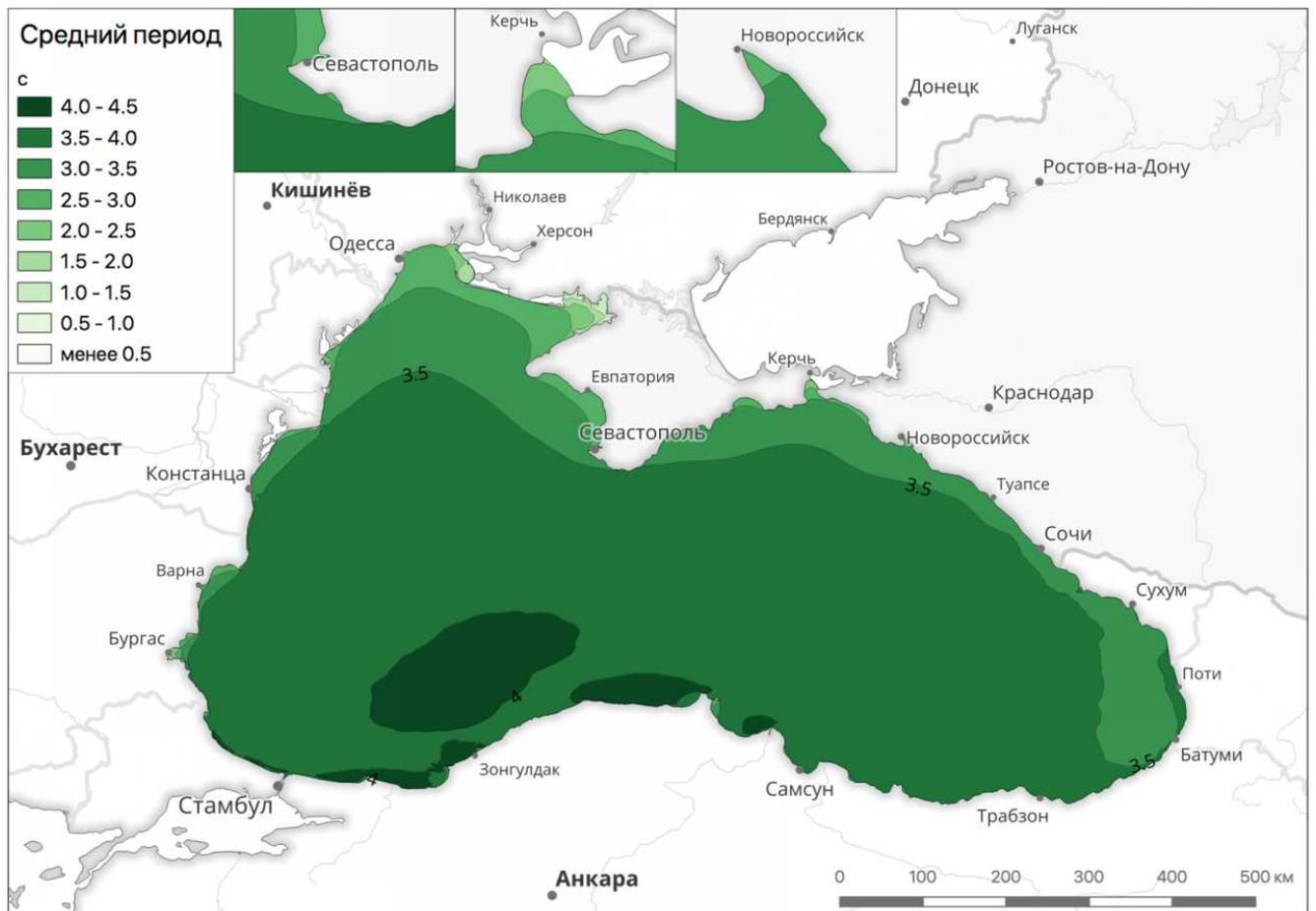
- Биполярные показатели характеризуются наличием нейтрального значения (обычно нулевого), относительно которого существует два характерных вектора изменения величины показателя (обычно в положительную и отрицательную сторону). К таким показателям можно отнести величину долгопериодных (многолетних) изменений анализируемой величины. Для картографирования биполярных показателей традиционно используются расходящиеся цветовые шкалы изменения светлоты и насыщенности цвета. При наличии значительного количества градаций в таких шкалах возможно добавление нового цветового тона в каждый из градиентов.

На основе данных о характеристиках волнения, волновой и ветровой энергии с использованием выбранных шкал были созданы прототипы картографических изображений для акватории Черного моря. Примеры визуализации данных представлены на Рисунках 4.19 — 4.23. Базовые физические характеристики (высота, длина, период волны) отображены с использованием последовательных однотонных шкал. Карты средней величины энергии и обеспеченности составлены с использованием более сложных многоцветных цветовых градиентов.

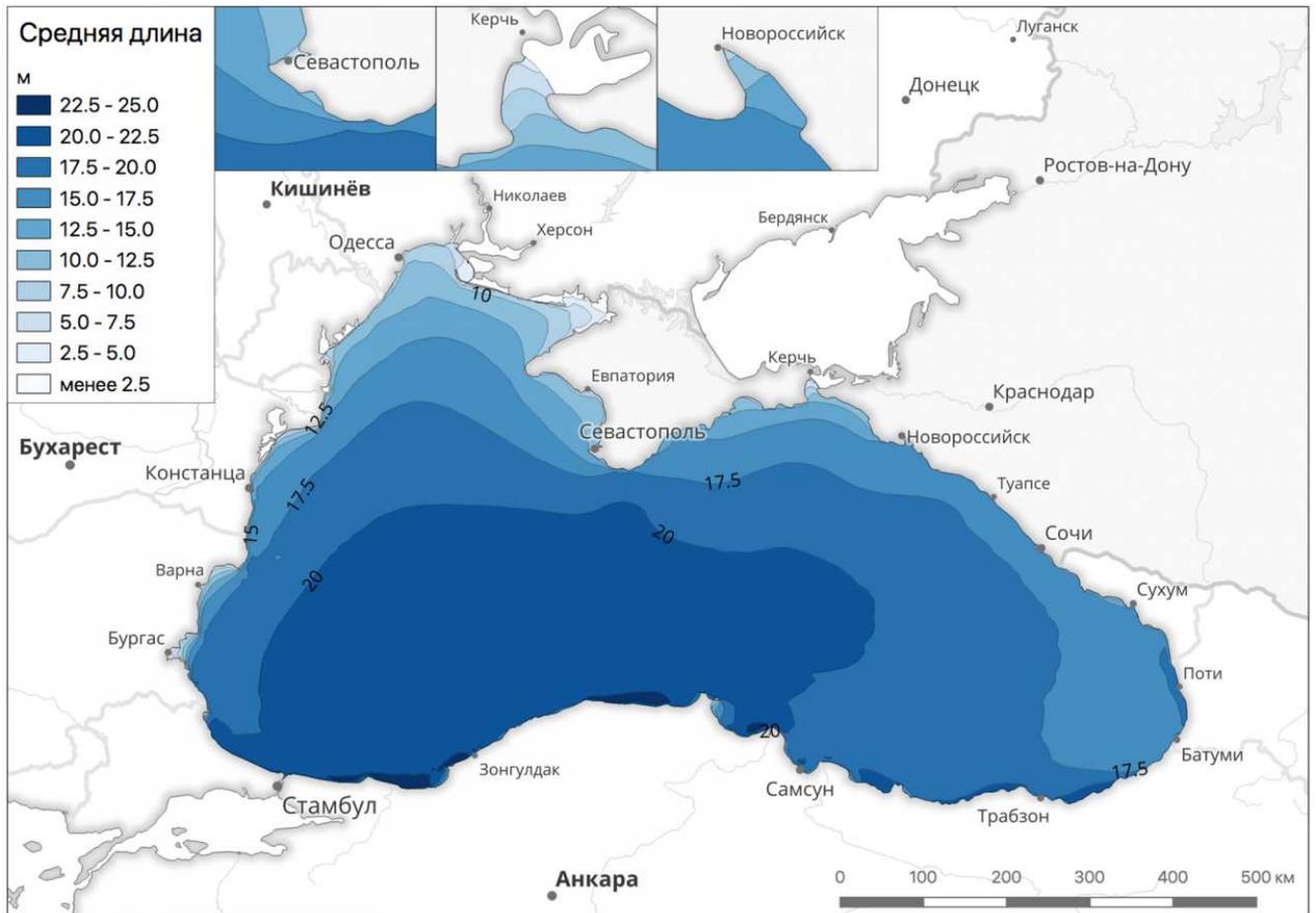
Среднемноголетняя высота значительных волн меняется от 0.5 — 1.0 м, максимальные значения наблюдаются в западной части Черного моря. В прибрежной зоне Российского сектора максимальные значения (0.7 м) наблюдаются у Южного берега Крыма и в районе м. Утриш (Рисунок 4.19). Средняя длина волн меняется от 10 до 22 м, максимальные значения наблюдаются в центральной и южной части моря. Средний период меняется от 2.7 с в северо-западной части моря до 4 с в центральной и южной части моря. Среднемноголетнее распределение волновой энергии в Черном море приведено на Рисунке 4.22 (поток волновой энергии через 1 м волнового фронта). В западной части Черного моря наблюдаются значения 3 — 4 кВт/м, в восточной части приблизительно в два раза меньше. Среднегодовая обеспеченность энергией более 1 кВт/м за весь период представлена на Рисунке 4.23. Видно, что только в 50% случаев в западной части моря поток энергии превышает 1 кВт/м. В восточной части обеспеченность около 20 — 40%. Эти карты позволяют оценить волновой энергетический потенциал конкретного участка побережья, однако нуждаются в генерализации при отображении в мультимасштабной среде.



**Рисунок 4.19** – Распределение среднегогодового значения высоты значительных волн



**Рисунок 4.20** – Распределение среднегогодечного значения высоты значительных волн



**Рисунок 4.21** – Распределение среднегогодового значения высоты значительных волн

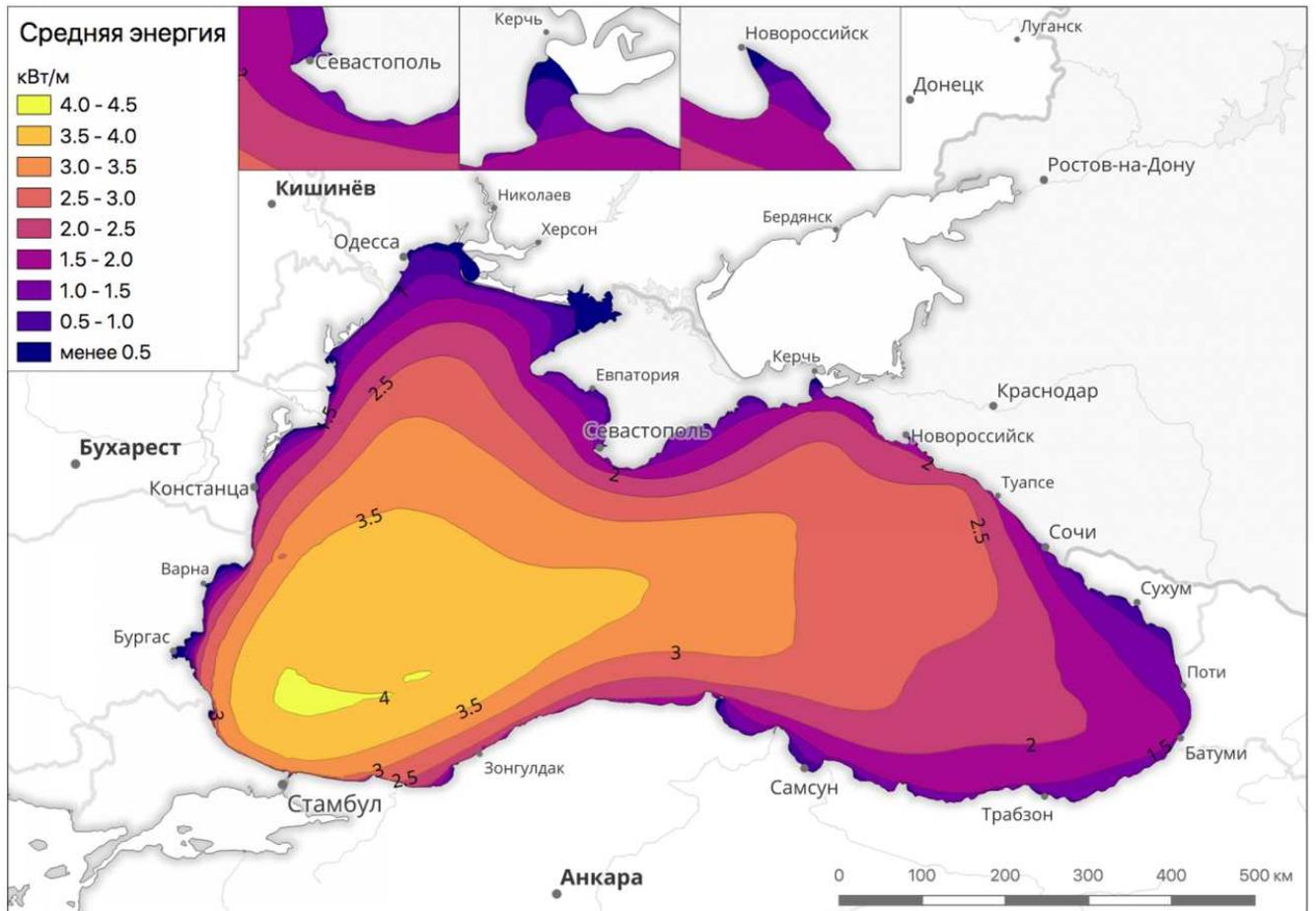
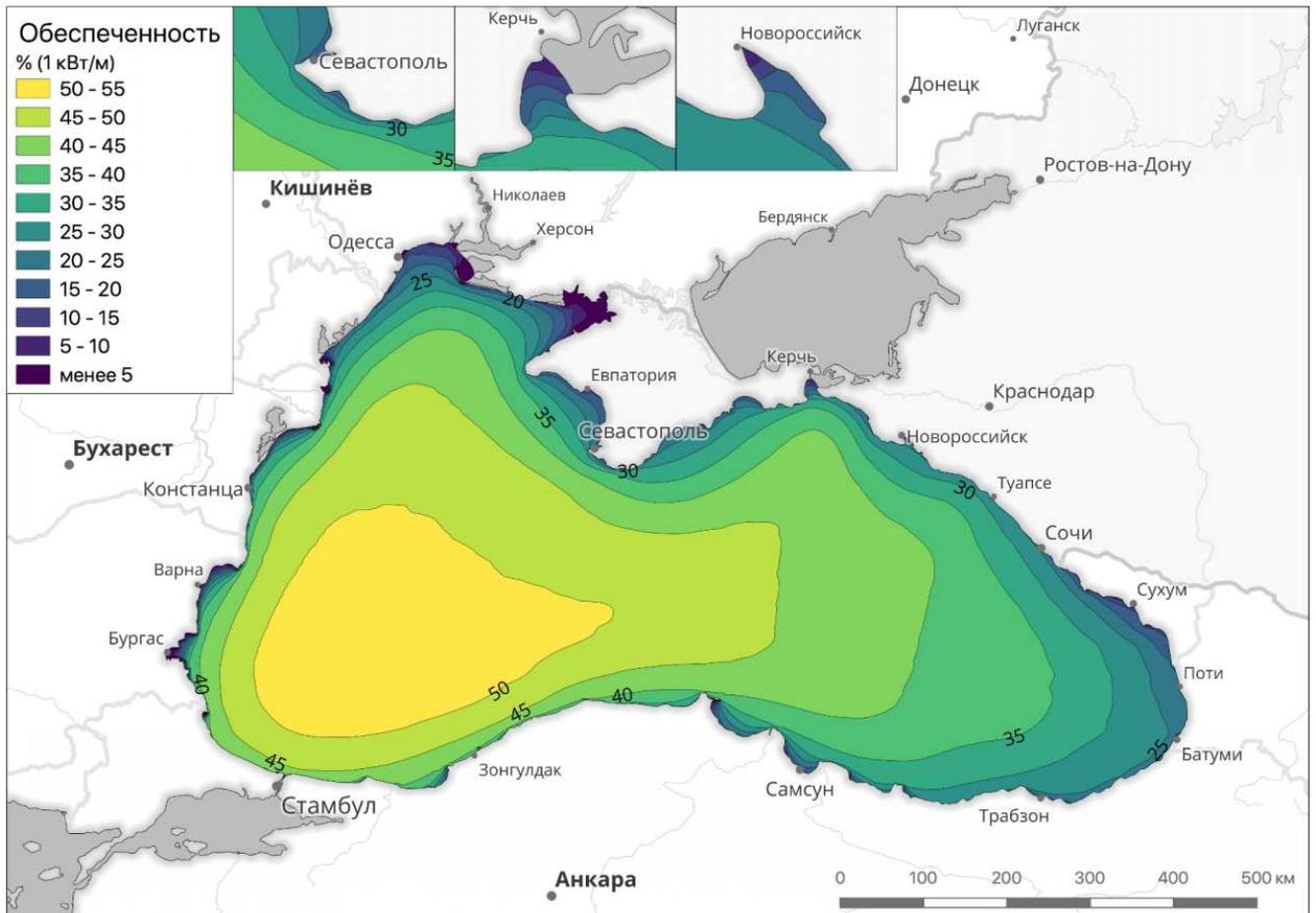


Рисунок 4.22 – Распределение среднегогодового значения высоты значительных волн



**Рисунок 4.23** – Распределение среднемноголетней обеспеченности волновой энергией

#### 4.4.4.2 Мультимасштабное представление

Для реализации принципа мультимасштабности применительно к отображению данных о ветровом волнении было использовано изменение сечения, а также сглаживание и утрирование подстилающей триангуляционной модели.

##### Изменение сечения

При картографировании геополей характеристик ветрового волнения использован принцип двукратного увеличения сечения. Данный принцип означает, что при достижении порогового масштабного уровня сечение укрупняется в два раза, при этом соответствующим образом адаптируется цветовая шкала, используемая для визуализации. Пример визуализации максимальной значимой высоты волны с использованием этого принципа представлен для Дальневосточного домена на Рисунке 4.24. Первый уровень детализации отображается на масштабных уровнях карты 2 — 3, второй — на масштабных уровнях карты 4 — 5, третий — на масштабных уровнях 6 и более крупных.

Необходимость отображения данных по всем морям в единой шкале сечения изолиний, даже при условии адаптации сечения к масштабному уровню, вызывает трудности, поскольку масштабы характеристик могут значительно отличаться. Так, например, средняя энергия волн в западной части Баренцева моря может превышать 30 кВт/м, в то время как для Черного моря она не превышает 4.5 кВт/м. Использование малого интервала, подходящего для Черного моря, приведет к перегрузке изображения на акватории Баренцева моря. В свою очередь применение крупного интервала, подходящего для Баренцева моря, на акватории Черного моря приведет к пустому и неинформативному изображению. Для разрешения ситуации был использован следующий графический прием: основную информационную нагрузку на карте играет цветовая шкала, которая соответствует текущему интервалу изолиний. А на акваториях морей с малой амплитудой характеристик в каждом интервале проводятся дополнительные изолинии на половине или четверти основного сечения. При этом цвет в пределах основного (крупного) интервала сохраняется постоянным.

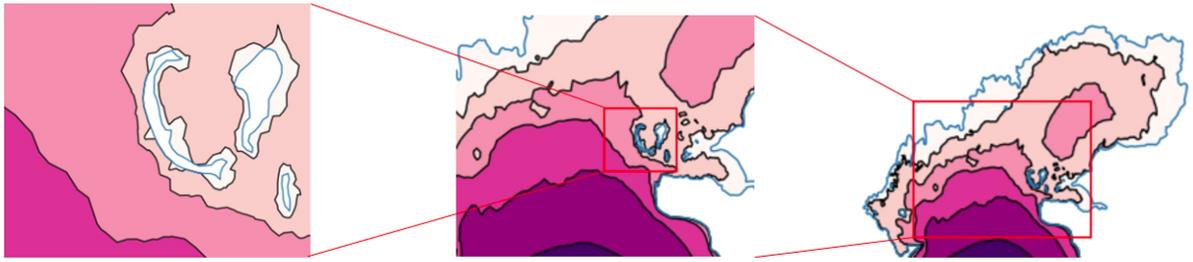
##### Сглаживание и утрирование

При отображении получаемых изолиний в мелких масштабах, возникает проблема избыточной точности контуров изолиний, а также невозможность разглядеть происходящие близ берегов изменения Рисунок 4.25.

Разработанная для решения этой проблемы методика предполагает последовательное применение к триангуляционной модели операторов сглаживания и утрирования, которые соответствуют классическим принципам генерализации.

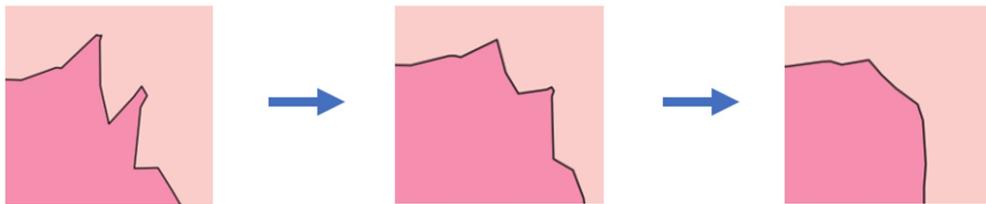
Сглаживание представляет собой упрощение геометрии контуров получаемых на триангуляционной модели изолиний за счет переопределения значения характе-





**Рисунок 4.25** – Проблема избыточной детальности изолиний при отображении прибрежных районов

ристики в каждом узле сетки как среднее значение в рассматриваемой и соседних с ней точках. Такая операция приводит к удалению локальных неоднородностей поля и, таким образом, к упрощению рисунка изолиний Рисунок 4.26.



**Рисунок 4.26** – Пример упрощения контура изолинии с помощью оператора сглаживания

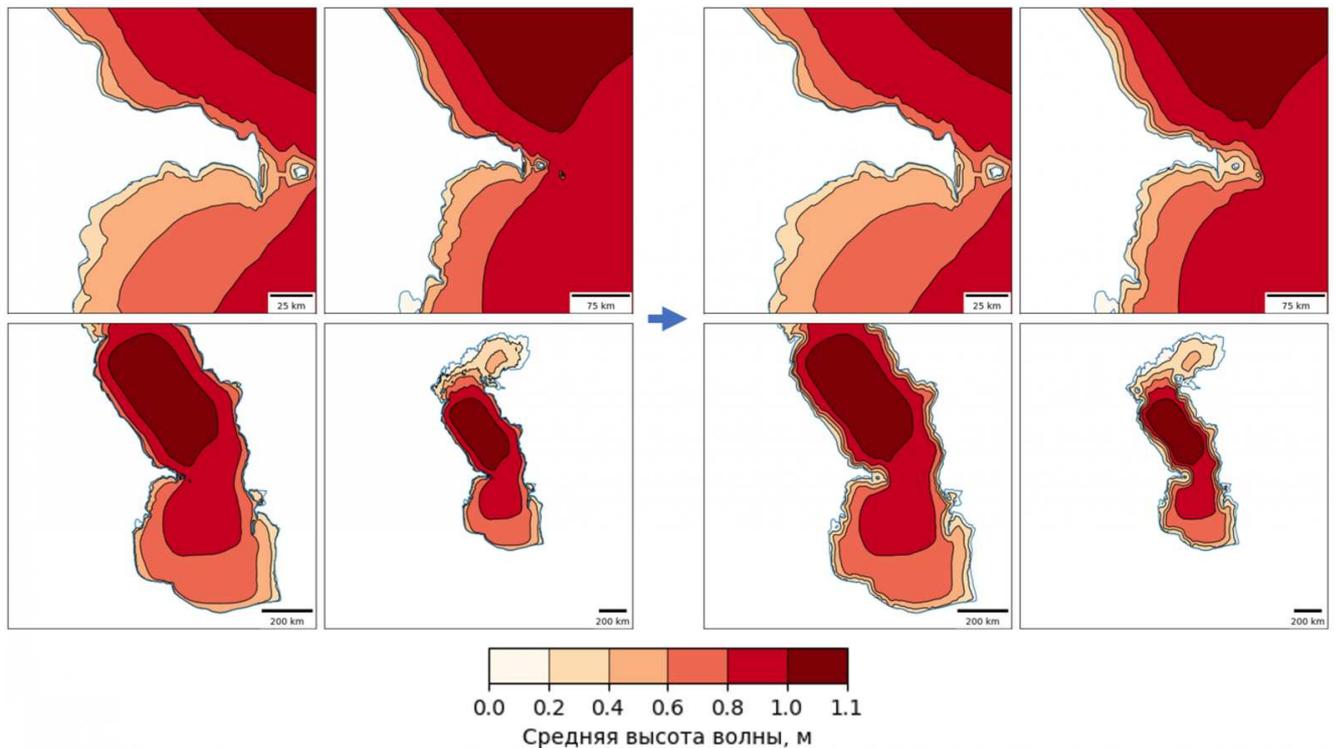
Оператор утрирования [Карташов, Самсонов, Мысленков, 2024] решает задачу невидимости локальных прибрежных изменений в мелких масштабах и увеличивает расстояние между горизонталями в местах наибольшей изменчивости поля характеристик Рисунок 4.27. Данный оператор, как и оператор сглаживания, не изменяет геометрию модели, а лишь пересчитывает нужным образом значения характеристик во всех её точках. Величина изменения характеристики в этом операторе прямо пропорциональна локальному градиенту поля, который рассчитывается по нормали к сфере, построенной на вершине и её соседях [Hu *et al.*, 2022] и обратно пропорциональна расстоянию до берега. Влияние каждого соседнего узла прямо пропорционально расстоянию до него и обратно пропорционально косинусу угла между градиентом расстояния от берега и вектором, соединяющим соседний узел с текущим. Использование градиента поля как катализатора утрирования позволяет не затрагивать и не искажать спокойные участки поля.



**Рисунок 4.27** – Пример преувеличения площади участков поля с низкими значениями характеристик для обеспечения их видимости с помощью оператора утрирования

В рамках задачи генерализации характеристик ветрового волнения были опробованы разные методы сглаживания, учитывающие соседей разного порядка и расстояния

от изменяемой точки до соседей. И, исходя из метода экспертной оценки результатов сглаживания, сделан вывод о наибольшей эффективности взвешенного усреднения с учетом всех соседей до 3 порядка с коэффициентами равными обратному расстоянию до изменяемой вершины. Результаты применения разработанной методики, сочетающей оба оператора Рисунок 4.28 показывают возможность ее применения для проведения автоматизированной генерализации триангуляционных моделей характеристик ветрового волнения в целях мультимасштабного картографирования.



**Рисунок 4.28** — Результаты автоматизированной генерализации триангуляционной модели средней высоты волны Каспийского моря для разных масштабов (слева негенерализованные изображения, справа — генерализованные с применением сглаживания и утрирования)

#### 4.4.5 Итоги

В настоящем разделе диссертации было показано как принцип мультимасштабности реализуется при картографировании характеристик ветрового волнения. Для этого предложена методика, которая опирается на восстановление цифровой модели (растровой или триангуляционной) и дальнейшее извлечение из нее изолиний, соответствующих масштабному уровню визуализации. Проблема генерализации данных при этом решается либо путем построения пирамид (для растрового подхода), либо путем сглаживания (для триангуляционного подхода). При этом показана избыточность растровой модели для представления данных моделирования на нерегулярной сетке. Введенный оператор утрирования позволяет выявить особенности прибрежного поля в мелких масштабах картографирования.

## 4.5 Построение цифровых моделей полимагистралей при мультимасштабном картографировании транспортных сетей<sup>6</sup>

В области **социально-экономического картографирования** большое внимание уделяется каркасу системы расселения, который образован населенными пунктами и транспортными сетями, а также системе административно-территориального деления, которая является основой статистического учета. Эти системы обладают иерархическими свойствами, образуя объекты разных пространственных масштабов. В частности, *Полимагистралю* представляют собой «единение параллельно идущих линий различных видов общего и специального транспорта в территориально-сближенные и однонаправленные пучки» [Воробьев, 1973; Полян, 1988; Полян, 2014].

В рамках настоящего исследования впервые разработана методика и технология автоматизированного выделения полимагистралей [Самсонов, Прохорова, 2019]. Полимагистралю являются полезной высокоуровневой абстракцией, позволяющей идентифицировать линейные кластеры путей сообщения. Для их выделения и использования необходима не только соответствующая геометрическая процедура, но также определение их места в масштабном ряду картографирования. Решение этой проблемы в рамках мультимасштабного картографирования требует наличия спецификации уровней детализации: определения включаемых классов объектов, подробности их представления, оптимальных масштабов визуализации.

### 4.5.1 Методика

Задача мультимасштабного картографирования транспортных сетей — отражение полимасштабности в организации транспортной сети путем последовательного обобщения и визуализации ее структуры, а также введения собирательных абстракций, таких как полимагистралю и транспортные коридоры, которые дают представление о сопряженности путей сообщения в обзорных масштабах рассмотрения. В основе реализации мультимасштабного подхода к картографированию транспортной сети лежит разработка спецификации уровней детализации.

Основное внимание при выделении уровней картографирования уделяется составу, пространственным и семантическим характеристикам отображаемых путей сообщения, принципам их отбора и генерализации. Целесообразно выделить четыре таких уровня, при этом разграничение присущих им масштабных диапазонов достаточно условно: они могут корректироваться в зависимости от географических особенностей территории.

<sup>6</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Самсонов, Прохорова, 2019]

*Первый* уровень — это, как правило, исходный картографический материал с минимальной степенью отбора: сеть железных дорог общего пользования, автомагистрали, автомобильные дороги с усовершенствованным покрытием, грунтовые автомобильные дороги, водные пути сообщения, трубопроводы. Для высокого уровня сложности структуры транспортной сети на картах 1-го уровня находят отражение первые три вида транспортных путей. Масштабный диапазон — 1 : 1 000 000–1 : 2 500 000.

*Второй* уровень сохраняет автомагистрали и шоссе, железные дороги, трубопроводы и водные пути (генерализация сети путей сообщений по двум направлениям – отбор и упрощение очертаний) — все магистральные пути сообщения. При сложной структуре транспортной сети – если не удастся при проведении отбора обозначить конкретные количественные критерии, то генерализация выполняется исходя из принципа сохранения структуры ведущих международных и внутригосударственных транспортных путей широтного и меридионального направлений. При низком уровне сложности сохраняются автомагистрали и скоростные дороги с покрытием, соединяющие транспортные узлы; допустимы отрезки и сегменты дорог к населенным пунктам, удаленным от основных дорог до 30 км. Основным критерием для отбора второстепенных дорог является ее значимость (соединяет ли населенные пункты и транспортные узлы, является ли единственной дорогой в малонаселенном регионе). Масштабные диапазоны — 1 : 2 000 000 — 1 : 4 000 000; 1 : 4 000 000 — 1 : 8 000 000; 1 : 8 000 000 — 1 : 15 000 000.

*Третий* уровень характеризуется значительной степенью отбора и геометрического упрощения, при его создании выделяются, прежде всего основные магистрали — ведущие линии в коммуникационной сети, характеризующиеся наибольшей пропускной способностью. Кроме отдельных видов путей сообщения сеть дорог на этом уровне представлена также полимагистралями. Масштабные диапазоны — 1 : 4 000 000 — 1 : 8 000 000; 1 : 8 000 000 — 1 : 15 000 000; 1 : 15 000 000 — 1 : 100 000 000. С помощью выделения магистралей и полимагистралей, которые отражают важнейшие экономические связи, ярко проявляется линейная концентрация территориальной структуры хозяйства.

*Четвертый* уровень предполагает изображение в качестве отдельных путей сообщений лишь автодорог самого высокого класса, полимагистралей, а также транспортных коридоров. Масштабные диапазоны картографирования — 1 : 15 000 000 — 1 : 100 000 000, 1 : 100 000 000 — 1 : 200 000 000.

Проведение генерализации при смене уровня детализации требует правильного учета факторов, влияющих на конфигурацию транспортной сети: природных — рельефа, конфигурации (очертаний) береговой линии, материка и речной сети, экономико-географических — плотности населения, экономической эффективности отдельных путей, размещения транспортных узлов, основных направлений транспортных потоков, территории, обслуживаемой транспортной сетью, политико-географических — начертания границ.

Учет взаимосвязи между различными элементами содержания карты, например,

между населенными пунктами и путями сообщения (устранение второстепенных населенных пунктов) влечет исключение местных путей, связывающих эти пункты с магистральными линиями. В процессе отбора населенных пунктов — изображение или не изображение определяется в зависимости от их удаленности от показываемых дорог. При этом сеть населенных пунктов обобщается в соответствии с дорожной сетью. Непременным условием является необходимость сохранить густоту дорог [Прохорова, 2010]. При составлении карт масштаба 1 : 4 000 000 и мельче первоочередной задачей становится сохранение и визуальное подчеркивание иерархической структуры дорожной сети, то есть на смену однообразному графическому оформлению дорог приходит система условных знаков с четкой дифференциацией по иерархическим ступеням [Самсонов, Кривошеина, 2015b].

Одним из важных приемов генерализации является введение на третьем, иногда втором уровне детализации понятия полимагистрали. Критерий сближенности при этом будет зависеть от масштаба визуализации. Так, например, в масштабах 1 : 4 000 000 — 1 : 8 000 000 оптимально использовать расстояние в пределах 10 — 20 км. При составлении карты «Транспорт и его влияние на природную среду» [и его влияние на природную среду Транспорт, 1998] масштаба 1 : 60 000 000 объединение разных магистралей в полимагистраль на отдельных участках происходило при 70-километровом интервале.

Важнейшие условия при выделении полимагистралей — сонаправленность линий разных видов транспорта, а также возможность соединения полимагистралей с собственно дорогами и выбора в каждом конкретном случае наиболее подходящего вида и структуры условного знака, позволяющего различить состав полимагистралей. Различные структурные типы полимагистралей (железнодорожно-речные, автомобильно-речные, комплексные и т. д.) выделяются исходя из специфики их функционирования. Основной компонентный тип — железнодорожно-автомобильный, который рассмотрен в настоящей статье.

Автоматизированный поиск элементов транспортной сети, участвующих в формировании полимагистрали, может быть осуществлен путем нахождения участков линий, которые при заданном расстоянии поиска  $d$  отклоняются друг от друга на расстояние, не превышающее  $d$ . Сформулируем данный принцип для случая двух линий.

Пусть дана линия  $A$ . Введем систему линейных координат, такую, что координатой любой точки  $p \in A$  будет являться криволинейное расстояние вдоль  $A$  от начальной точки  $p_0 \in A$  до  $p$ . Обозначим через  $A_{pq}$  участок линии  $A$  между точками  $p$  и  $q$ . Аналогичным образом введем систему линейных координат и участок для линии  $B$  и точек  $v$  и  $w$ , принадлежащих ей. Тогда при заданном расстоянии  $d$  будем считать формирующими полимагистраль участки  $A_{pq}$  и  $B_{vw}$ , если расстояние от любой точки  $a \in A_{pq}$  до ближайшей к ней точки  $b \in B_{vw}$  не превышает  $d$ , и наоборот. Иначе говоря, это такие пары участков линий  $A$  и  $B$ , расстояние Хаусдорфа между которыми  $d_H(A_{pq}, B_{vw}) \leq d$ .

Расширим понятие полимагистрали на случай  $N > 2$  линий. Пусть даны линии

$A_i, i = 1 \dots N$ , образующие множество  $\Omega = \cup A_i$ . Тогда  $\Omega$  является полимагистралью шириной  $d$ , если выполнено условие  $d_H(A_i, A_j) \leq d, \forall i, j \in \{1, \dots, N\}$ . То есть все линии упакованы в коридор шириной  $d$ . Представим теперь, что множество  $\Omega$  состоит из трех параллельных линий, отстоящих друг от друга на расстояние 15 км. Будет ли их совокупность являться полимагистралью при  $d = 20$  км? Пары соседних линий — да, однако суммарная ширина найденного коридора составляет 30 км, что превышает заданный порог. Для разрешения таких ситуаций можно «ослабить» определение полимагистрали следующим образом. Будем считать множество линий  $\Omega$  полимагистралью при заданном расстоянии  $d$ , если для любого  $i \in \{1, \dots, N\}$  существует такое  $j \neq i, j \in \{1, \dots, N\}$ , что  $d_H(A_i, A_j) \leq d$ . При данном определении полимагистрали будет достаточно, чтобы Хаусдорфово расстояние до ближайшей линии в пучке полимагистрали не превышало  $d$ . Для обозначения отличия полимагистралей, полученных исходя из двух данных определений, объекты первого типа (упакованные в коридор заданной ширины) будем называть *компактными полимагистралями*.

Целесообразно выдвинуть следующие требования к технологии (алгоритму) выделения полимагистралей:

1. Поиск полимагистралей осуществляется в соответствии с формальным определением.
2. Алгоритм поиска полимагистралей должен работать для произвольного количества линейных объектов и их типов.
3. Результатом поиска должно быть множество линейных объектов, в котором найденные полимагистрали представлены в виде одной осевой линии.
4. Объекты, не образующие полимагистрали, сохраняют свою атрибутивную информацию (вид транспорта).

При такой постановке задачи нахождение осевых линий полимагистрали может быть осуществлено различными путями: посредством слияния входящих сегментов или путем выделения площади, занимаемой полимагистралью [Edelsbrunner, Kirkpatrick, Seidel, 1983] и построения ее срединной оси [Mekhedov, Mestetskiy, 2010]. Дополнительно можно добавить условие параллельности линий, однако на практике строгая параллельность выполняется только в крупных масштабах, и на обзорных уровнях детализации это условие только усложнит решение задачи.

#### 4.5.2 Результаты и обсуждение

На основе предложенных требований была разработана технология построения компактных полимагистралей, состоящих из железных и автомобильных дорог (*бимагистралей*). Технологическая схема алгоритма представлена на Рисунке 4.29.

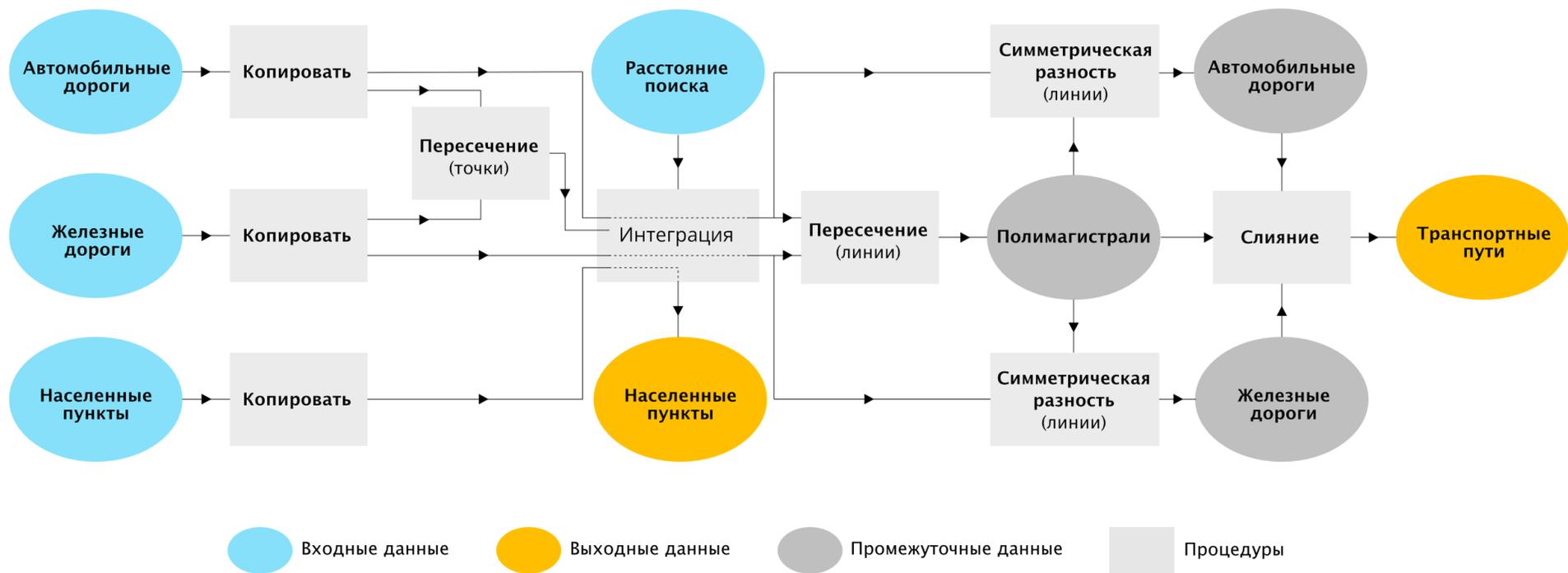


Рисунок 4.29 – Технологическая схема (алгоритм) выделения полимагистралей

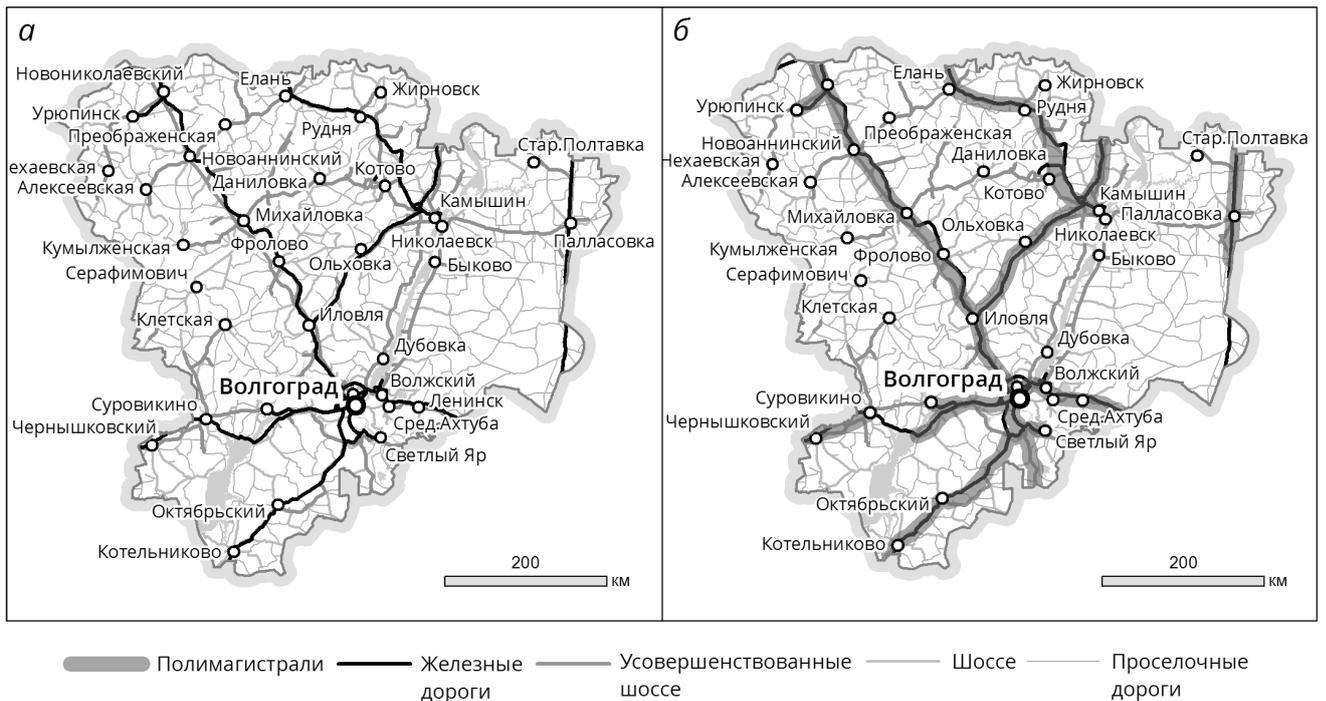
Ключевой компонентой, обеспечивающей нахождение полимагистралей, является операция *Интеграции* (Integrate), которая совмещает вершины и линейные сегменты близко расположенных объектов (точек, линий, границ полигонов), если они находятся на расстоянии, не превышающем  $d$ . Количество интегрируемых объектов при этом не ограничено. Также в процессе интеграции совмещению подвергаются вершины самого объекта — происходит стягивание ребер, узлы которых расположены на расстоянии, меньшем  $d$ . Эта особенность позволяет реализовать одновременно со слиянием соседних линий их геометрическое упрощение. Обратим внимание, что в процедуре участвуют также точки населенных пунктов и узлы пересечения автомобильных и железных дорог, выделяемые функцией *Пересечения* (Intersect). Это позволяет, по возможности, сохранить корректность топологических отношений в системе «населенные пункты — транспортная сеть».

Таким образом, использование интеграции позволяет полностью удовлетворить требования 1, 3 и частично 2. Расширение возможностей работы алгоритма для включения произвольного количества типов объектов возможно путем введения дополнительных параметров. Остальные компоненты технологии реализуют чтение, подготовку и сохранение данных, а также обеспечение выполнения требования 4. В частности, после интеграции данных выполняются операции оверлея данных *Пересечения* (Intersect) и *Симметрической разности* (Symmetrical Difference), позволяющие выделить полимагистралю (совпадающие участки автомобильных и железных дорог), а также автомобильные и железные дороги (участки дорог, не совпадающие с полимагистралями). Результирующие линии объединяются с помощью операции *Объединения* (Merge) в один набор данных с сохранением атрибутивной информации линий, что удобно для последующего управления результатами и их отображения. Еще одним выходным параметром инструмента являются точки населенных пунктов, которые могут подвергаться смещению в процессе интеграции.

Реализация технологии осуществлена путем разработки скрипта на языке программирования Python и его интеграции в виде инструмента для ГИС-пакета *ArcGIS Desktop 10.5*. Технология была апробирована на примере поиска полимагистралей по данным общегеографических карт России масштабов 1 : 1 000 000 (предоставлена компанией Дата+) и 1 : 2 500 000 (ВСЕГЕИ). В качестве тестовой области были выбраны Волгоградская область (1 : 1 000 000) и Центральный федеральный округ (1 : 2 500 000).

Результаты экспертного (ручного) и автоматизированного выделения полимагистралей на территорию Волгоградской области приведены на Рисунке 4.30 и Рисунке 4.31 соответственно.

В процедуре участвовали все населенные пункты (на рисунке показаны не все в целях обеспечения читаемости), а также дороги с исходного картографического материала. При анализе полученных изображений можно наблюдать рост количества полимагистралей с увеличением параметра  $d$ . Наиболее адекватным и приближенным к эксперт-



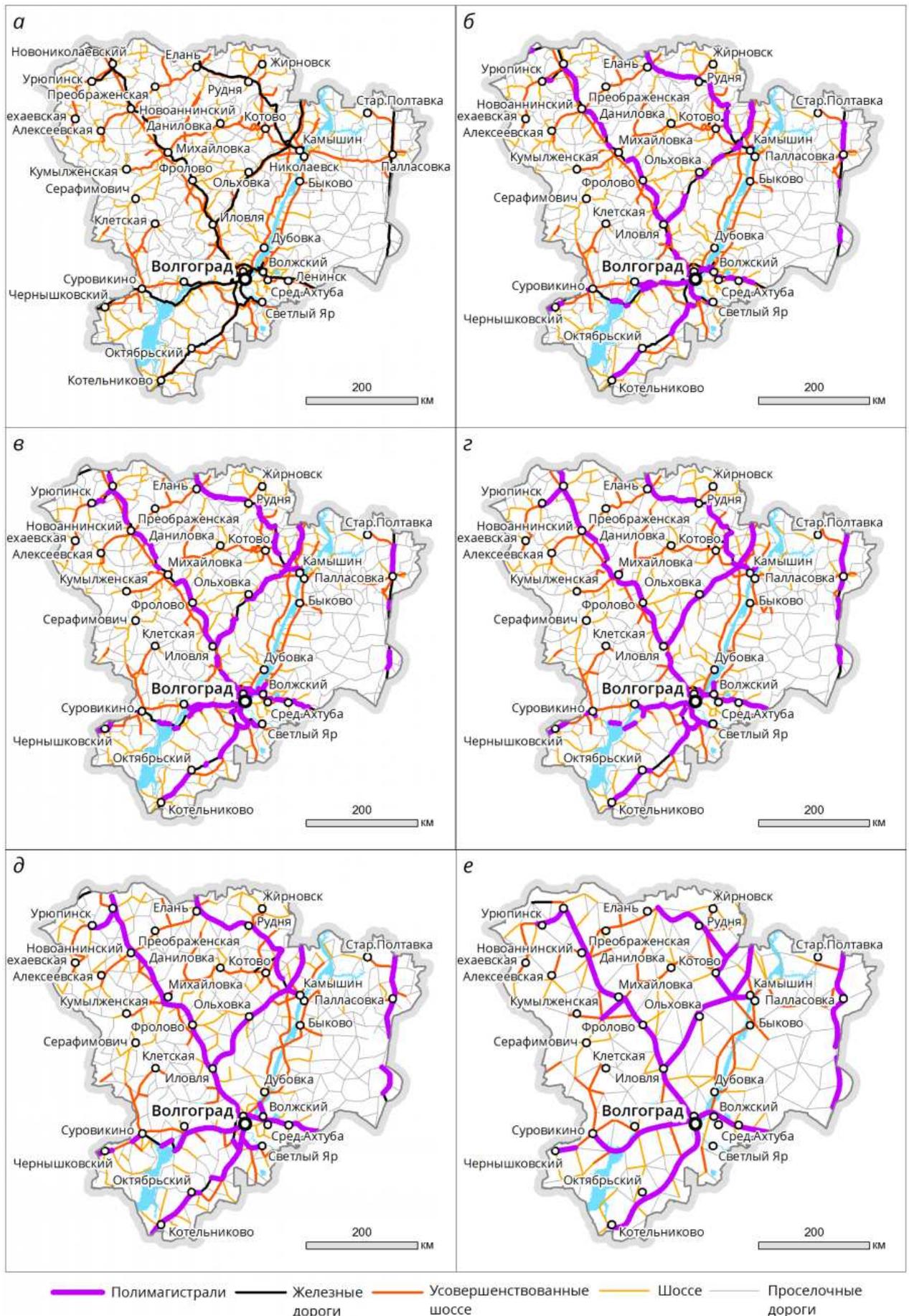
**Рисунок 4.30** – Полимагистралы, автомобильные и железные дороги Волгоградской области. (а) исходные данные (масштаб 1 : 1 000 000); (б) результат ручного выделения полимагистралей

ному выделению полимагистралей представляется результат, полученный при расстоянии поиска 4 км. При дальнейшем увеличении расстояния (6 км) начинаются значительные сдвиги линий, приводящие к неестественной изломанности полимагистралей, а при больших расстояниях (12 км) транспортная сеть приобретает черты абстрактного графа.

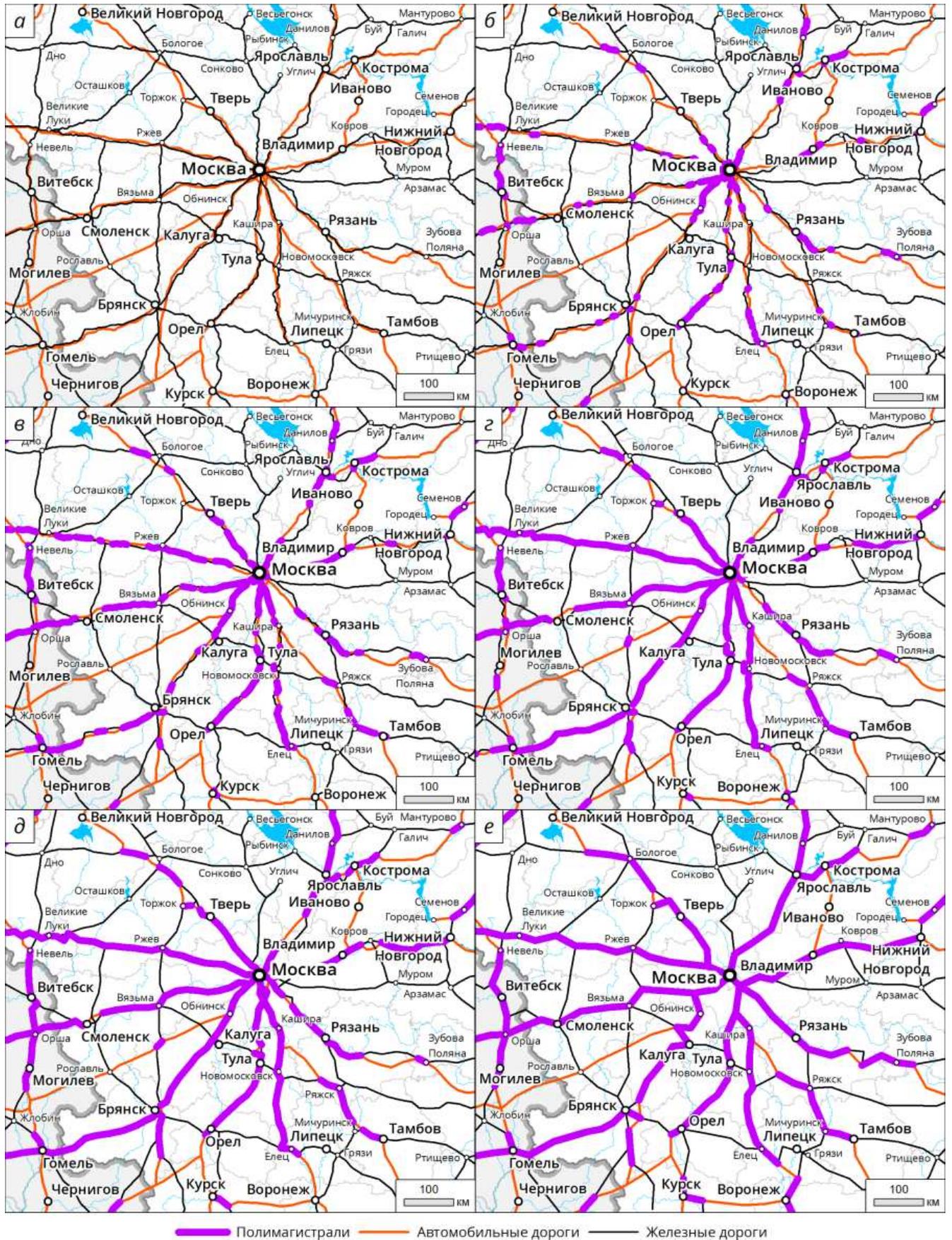
На Рисунке 4.32 представлены результаты поиска полимагистралей по данным карты масштаба 1 : 2 500 000. В этом эксперименте были использованы все элементы транспортной сети с исходного картографического материала, но из населенных пунктов были сохранены преимущественно транспортные узлы и тупиковые поселения.

На этом примере наблюдается схожий рост полимагистралей при увеличении расстояния поиска. Первое оптимальное значение достигается при расстоянии поиска 2 км, когда заполняются многие пробелы на основных направлениях, но линии при этом не приобретают угловатый характер. При расстоянии 4 км полимагистраль образуется на Рижском, Смоленском и Брянском направлениях, 6 км – Воронежском. При расстоянии поиска 12 км почти все радиальные транспортные артерии преобразуются в полимагистралы на длительном протяжении, их геометрические очертания приобретают абстрактные формы.

Таким образом, можно сделать вывод, что с использованием предложенной технологии можно получить несколько уровней детализации представления транспортной сети с последовательным введением полимагистралей при увеличении расстояния поиска. Предложенная технология может быть использована в рамках разработанной специфици-



**Рисунок 4.31** – Полимагистралы, автомобильные и железные дороги Волгоградской области: (а) исходные данные с топографической карты масштаба 1 : 1 000 000; (б)–(е) полимагистралы, полученные при расстоянии поиска 1, 2, 4, 6, 12 км



**Рисунок 4.32** – Полимагистралы, автомобильные и железные дороги центра Европейской части России: а) исходные данные с топографической карты масштаба 1 : 2 500 000; б–е) полимагистралы, полученные при расстоянии поиска 1, 2, 4, 6, 12 км

кации уровней детализации мультимасштабной карты транспортной сети. В то же время, заметим, что использование стандартной для ГИС процедуры интеграции при выделении полимагистралей обладает своими особенностями. Так, например, при увеличении расстояния слияние линий может происходить непоследовательно, то есть, одна и та же линия может присоединяться к разным соседним линиям при изменении расстояния поиска, что является недостатком конкретной реализации, используемой функции интеграции. Соответственно, одним из направлений дальнейшего совершенствования предложенной технологии должна быть разработка более робастного алгоритма выделения осевой линии полимагистралей.

### 4.5.3 Итоги

По итогам выполненных работ получены следующие результаты:

1. Показана логическая связь между понятием полимагистралей как способа представления пространственно сопряженных участков транспортных путей разного типа и процессом генерализации картографического изображения на мультимасштабных картах.
2. Разработана четырехуровневая спецификация уровней детализации для мультимасштабных карт транспортных сетей, в которой определено место полимагистралей на третьем и четвертом уровне детализации.
3. Дано формализованное определение полимагистралей, необходимое для автоматизации процесса их выделения методами пространственного анализа. На основе него предложена технология выделения полимагистралей, основанная на пространственной интеграции линейных объектов и операций оверлея (пространственного наложения). Произведена алгоритмизация и реализация технологии на языке программирования Python.
4. На примерах данных цифровых карт масштабов 1:1 000 000 и 1:2 500 000 апробировано выделение полимагистралей путем объединения сопряженных путей автомобильных и железных дорог, показано сходство в экспертном и автоматизированном способе выделения полимагистралей.

Дальнейшее развитие предложенного подхода видится в разработке нового, более совершенного алгоритма выделения полимагистралей; разработке методики мультимасштабного картографирования транспортных сетей, основанной в том числе на практическом опыте создания мультимасштабных тематических карт транспорта. Многоуровневое картографическое представление с выделением полимагистралей и параметризацией их характеристик предоставляет возможность анализировать полимасштабную структуру транспортной сети страны, представить ее основные векторы напряжения и коридоры в наглядной графической форме. Результаты могут найти

применение как в среде исследователей, специализирующихся в географии транспорта, так и при решении производственных транспортных задач.

## 4.6 Визуализация пространственных данных способами картограмм и картодиаграмм на мультимасштабных картах<sup>7</sup>

Социально-экономические карты часто основываются на наборах статистических данных, которые положены на сетку административно-территориального деления (АТД), образующую партономическую иерархию [Tversky, 1989], или привязаны к объектам точечной локализации. Для отображения пространственных данных, привязанных к сетке АТД, чаще всего используются картодиаграммы (для абсолютных показателей) или картограммы (для относительных показателей), для точечных данных большую популярность приобрело кластерное отображение с группировкой точек по пространственной близости, которое по сути представляет собой традиционный точечный способ.

Эффекты, связанные с построением и визуальной интерпретации *картограмм и картодиаграмм* во множестве масштабов являются малоизученными [Beconyté et al., 2022]. Данная проблема была исследована в работах [Юрова, Самсонов, 2012; Самсонов, Юрова, 2014; Samsonov, 2015], результатом которых явилась методика, позволяющая определять масштабы смены иерархических уровней административно-территориального деления (АТД), а также выполнять преобразования шкал цветов и размеров, способствующих оптимальной интерпретации социально-экономических карт, созданных способами картограмм и картодиаграмм. В дополнение к этому исследованию была разработана методика анимационных переходов между масштабными уровнями при использовании картограмм и картодиаграмм [Самсонов, 2024].

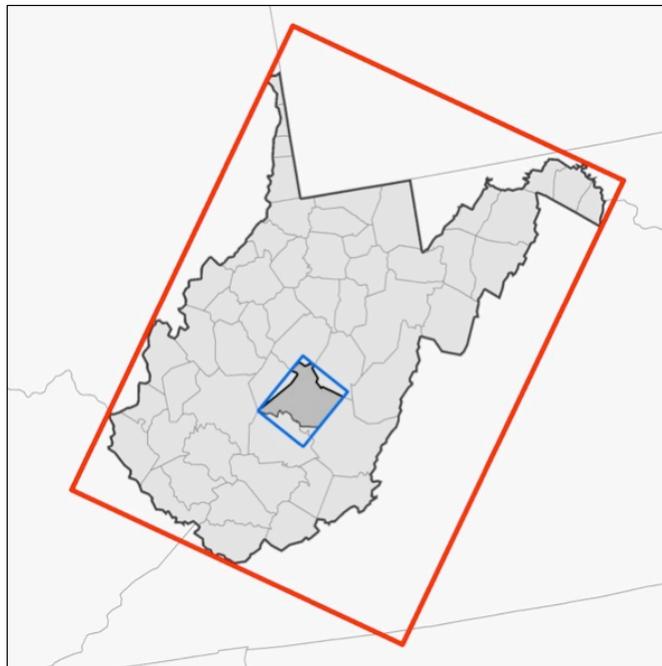
### 4.6.1 Определение критических масштабов

Под критическим масштабом понимается минимально допустимый масштаб, в котором объект может быть визуализирован в силу его размеров, подробности геометрического представления, графической нагрузки карты и т. д. Обобщенный критический масштаб может быть также определен для слоя или уровня детализации целиком с учетом среднестатистических характеристик объектов. При достижении критического масштаба необходимо переключение карты на следующий уровень детализации данных. Несмотря на существование эффективных структур данных, таких как *tGAP* [van Oosterom, Meijers, 2014], позволяющих непрерывно менять детализацию изображения,

<sup>7</sup>При подготовке данного раздела диссертации использованы следующие публикации, выполненные автором лично или в соавторстве, в которых, согласно Положению о присуждении ученых степеней в МГУ, отражены основные результаты, положения и выводы исследования: [Самсонов, Юрова, 2014]

в случае административно-территориального деления такой подход неприемлем, так как объединение и исключение единиц картографирования недопустимо. Это означает, что необходима смена территориального уровня картографирования.

Основным фактором критического масштаба является размер объектов. Для оценки размеров может быть использована техника минимального по площади ограничивающего прямоугольника *MBR* (*minimum bounding rectangle*) [Самсонов, Подольский, Юрова, 2013]. В этой технике ширина среднестатистического прямоугольника сравнивается с минимально допустимым линейным размером объекта (Рисунок 4.33). При этом рекомендуется учитывать искажения используемой проекции. Для вычисления знаменателя критического масштаба используется соотношение  $M = W/W_{min}$ , где  $M$  — знаменатель критического масштаба;  $W$  — средняя ширина MBR;  $W_{min}$  — минимально допустимый размер «среднестатистического» объекта.



**Рисунок 4.33** — Построение минимальных ограничивающих прямоугольников для определения критических масштабов отображения

#### 4.6.2 Выбор и преобразование цветовых шкал для картограмм

При создании цветовых шкал картограмм важную роль играют три аспекта — определение числа классов, границ классов и выбор цветов. Коснемся первых двух моментов лишь кратко, сосредоточившись более на цвете. Рекомендуемое число градаций в шкале для их хорошей различимости, как правило, не должно превышать семи [Brewer, 2015]. Это ограничение справедливо для каждого масштабного уровня. Однако с изменением иерархического уровня на более высокий происходит сужение диапазона значений относительного статистического показателя. Кроме того, каждый последующий

более высокий уровень иерархии содержит меньше территориальных образований, чем предыдущий. Поэтому при уменьшении масштаба число градаций следует сокращать пропорционально уменьшению размаха вариации и числа единиц. Промежуточные границы классов выбираются в зависимости от особенностей распределения показателя на каждом уровне. Заметим, что сравнимость уровней будет существенно выше, если сделать границы классов постоянными (что, однако, удается не всегда в силу существенных отличий в распределениях).

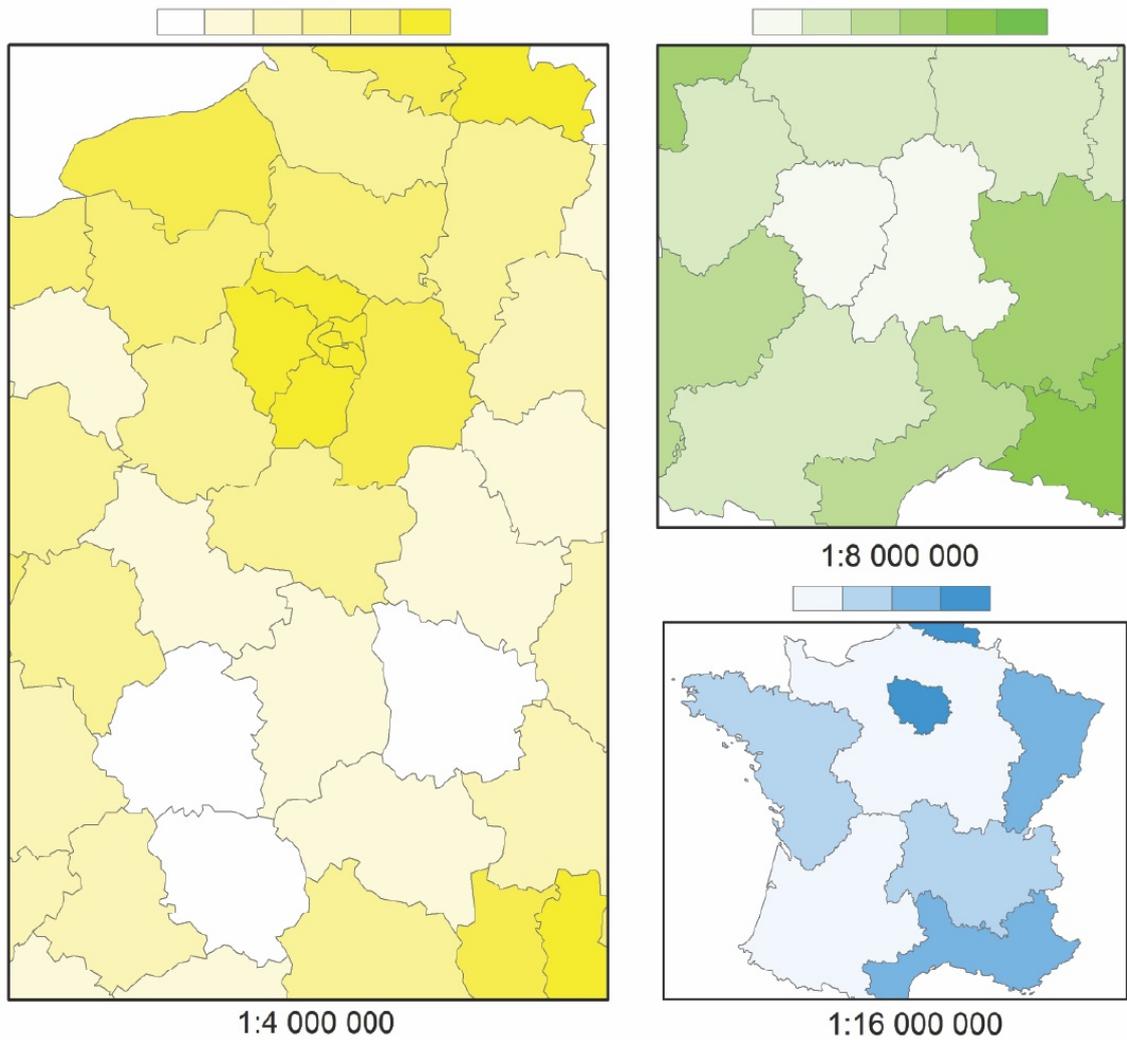
Множество факторов влияют на выбор цветов при составлении и оформлении карт: восприятие цвета, соотношение цветового решения с отображаемыми данными, их структурой, предпочтение пользователя цветовому оформлению. Главным фактором является то, что цветовое решение карты не должно мешать правильному восприятию информации.

В картографии для построения шкал используются три цветовые схемы: *последовательная* (sequential), *расходящаяся* (diverging) и *качественная* (qualitative) [Brewer, 2015]. Для отображения показателей, характеризующихся постепенным увеличением или уменьшением значений при отсутствии «нейтрального» значения в распределении, наилучшим образом подходят последовательные шкалы. Рассмотрим возможности их применения на мультимасштабных картах.

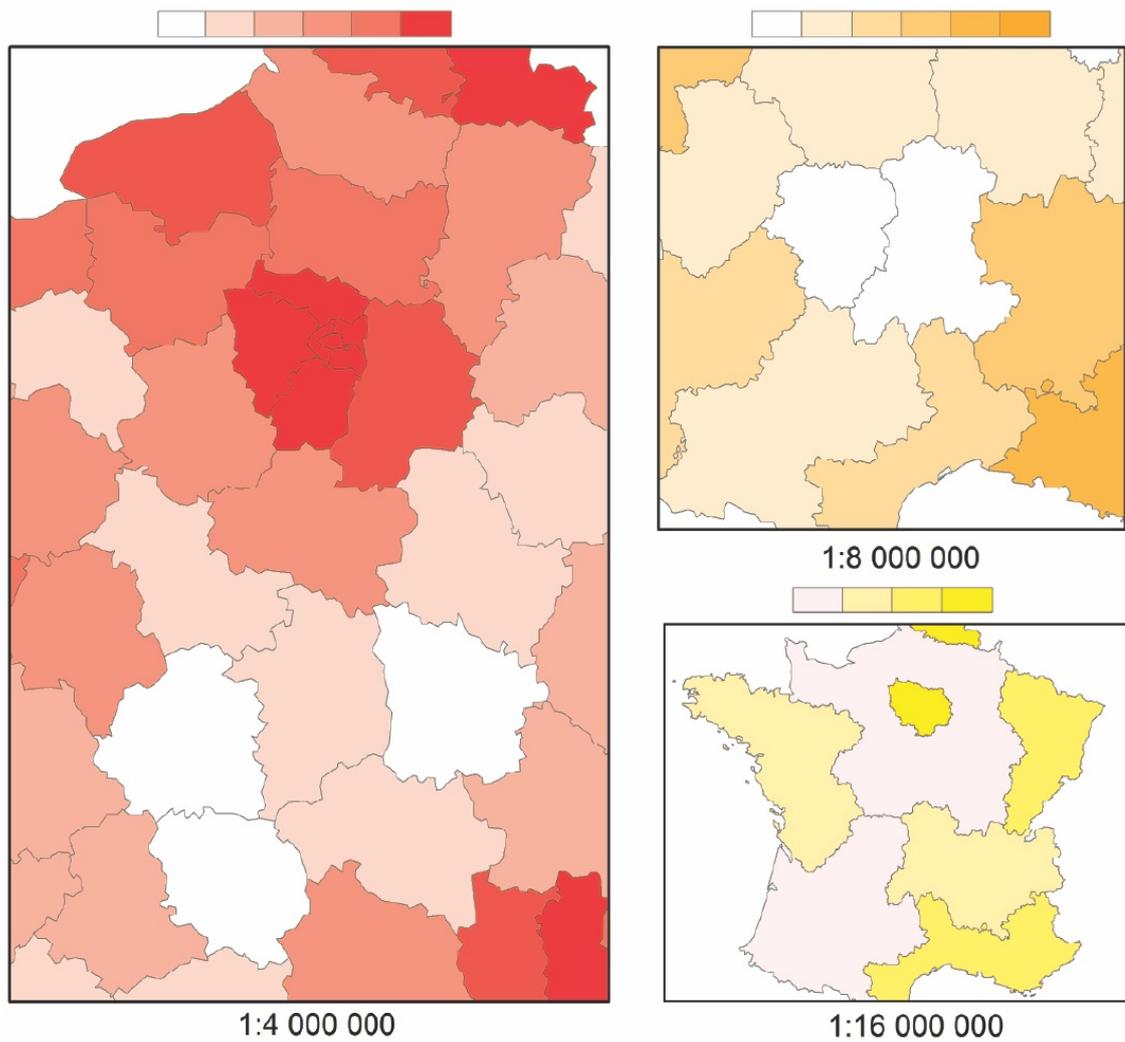
Экспериментально было опробовано несколько способов построения цветовых шкал для картограмм. В первом подходе (Рисунок 4.34) используются принципиально разные цветовые тона для разных уровней детализации. Такой подход подчеркивает факт переключения на другой иерархический уровень. В то же время он мешает сравнению значений на разных уровнях, что является одним из основных методов использования мультимасштабных карт. Кроме того, резкая смена цвета создает ощущение, что на разных уровнях картографируются разные характеристики

Схожая ситуация характерна при использовании родственных цветовых тонов (желтый — оранжевый — красный) на разных картономических уровнях (Рисунок 4.35). В данном случае цветовой переход выглядит более логично, однако сравнение распределения значений показателя по-прежнему затруднительно.

На основании экспериментов был сделан вывод, для шкал разных уровней логичнее всего использовать один цветовой тон в случае последовательно меняющегося показателя. Однако применение одного и того же цветового градиента (даже при разном числе классов) не позволит отразить изменение граничных значений и размаха вариации показателя между уровнями АТД. Поэтому была предложена следующая оптимизация. Произвести анализ исходных данных по всем уровням детализации и выделить уровень, который характеризуется наибольшим размахом вариации (для относительных показателей это наиболее детальный уровень), установить для него цветовую шкалу. Тогда изменение контрастности шкалы (т. е. сужение диапазона светлоты и/или насыщенности) при переходе на следующий уровень АТД будет соответствовать изменению размаха ва-



**Рисунок 4.34** – Фрагменты карты плотности населения Франции в различных масштабах. Для каждого уровня детализации использована шкала с разным цветовым тоном



**Рисунок 4.35** – Фрагменты карты плотности населения Франции в различных масштабах. Для каждого уровня детализации использована отдельная шкала. В качестве основных цветов при построении шкал применены родственные цветовые тона

риации. Этот способ позволяет создать систему взаимосвязанных между собой шкал, которые привлекают внимание пользователя к изменению диапазона значений в данных.

Для автоматизации данного подхода предлагается следующая методика. В качестве исходных величин выступают:

- $v_i$  — массив значений картографируемого показателя на уровне  $i = 1, 2, 3, \dots, n$ , где  $n$  — число уровней детализации;
- $V = \cup_{i=1}^n v_i$  — полный массив значений, включающий данные по всем уровням детализации;
- $S^{min}, S^{max}$  — значения насыщенности цветового тона в диапазоне от 1 до 100, соответствующие минимальному и максимальному значению показателя в массиве  $V$ . Значения  $S^{min}$  и  $S^{max}$  выбираются автоматически.

Искомыми величинами являются  $s_i^{min}, s_i^{max}$  — значения насыщенности цветового тона, соответствующие минимальному и максимальному значению показателя в массиве  $v_i$ .

Расчеты производятся следующим образом. Вычисляют коэффициенты  $k_i$ :

$$k_i^{min} = \frac{\min(v_i)}{\min(V)}, k_i^{max} = \frac{\max(v_i)}{\max(V)}; \quad (4.2)$$

а также минимальное и максимальное значения  $s_i$  для выбранного уровня:

$$s_i^{min} = S^{min} (k_i^{min})^r, s_i^{max} = S^{min} (k_i^{min})^r. \quad (4.3)$$

Параметр  $r$  отвечает за степень чувствительности цветовой шкалы к изменению размаха вариации значений. Пример изображения при  $r = 0, 1$  представлен на Рисунке 4.36.

Аналогичный подход может быть использован и в случае симметрично распределенных относительно «нейтрального» значения показателей, например рождаемости. Для них должны использоваться расходящиеся шкалы. При этом изменение разброса значений между уровнями детализации также будет влиять на контрастность цветовой шкалы.

### 4.6.3 Выбор и преобразование шкал размеров картодиаграмм

Изменение размеров картодиаграмм на мультимасштабных социально-экономических картах будет рассмотрено на примере круговых диаграмм как наиболее часто используемых в картографии, и построение которых хорошо поддается автоматизации. Рекомендации касательно числа классов остаются аналогичными и для способа картодиаграмм. Выбор границ классов, а также реклассификация показателя в случае

ФРАНЦИЯ. ПЛОТНОСТЬ НАСЕЛЕНИЯ

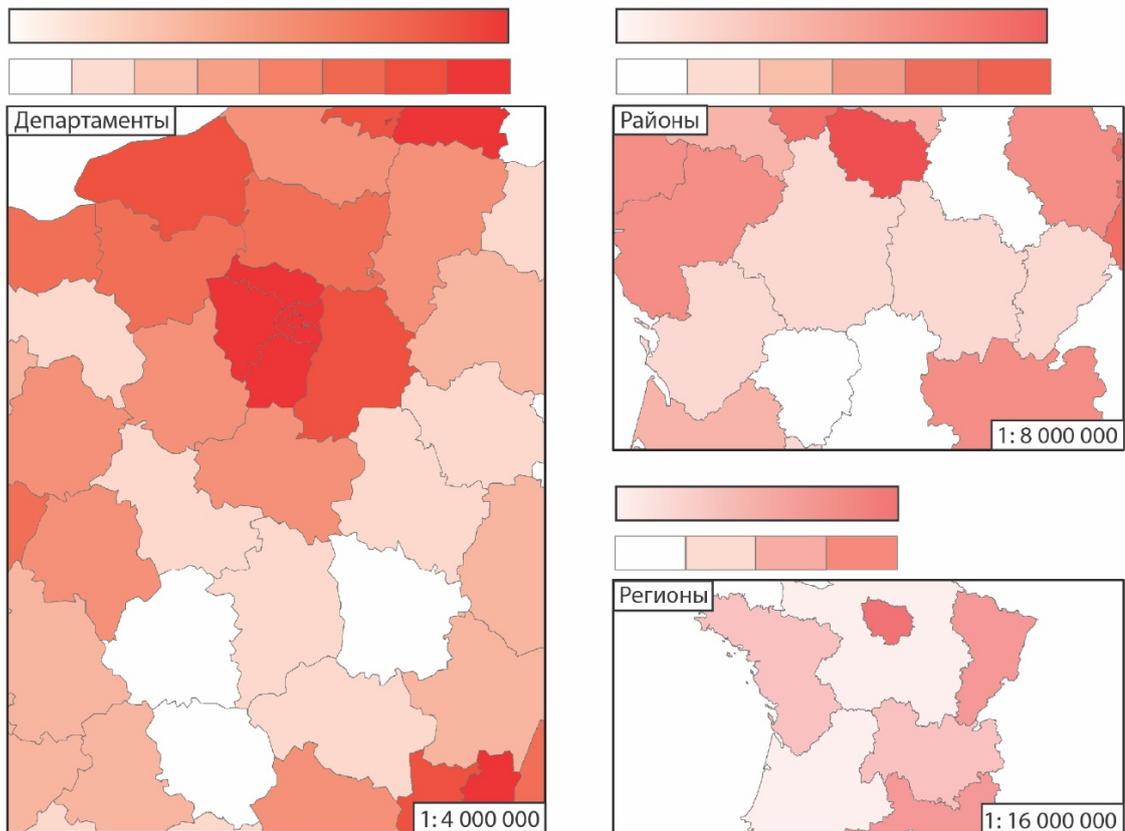


Рисунок 4.36 – Использование одноцветных адаптивных шкал

секторных диаграмм являются предметом отдельного исследования.

В данном исследовании фокус был сосредоточен на условных шкалах как наиболее универсальных и простых в построении. В случае использовании таких шкал диаметр кружка увеличивается между классами не пропорционально значению показателя, а в фиксированное число раз  $k$ , лежащее обычно в диапазоне от 1, 2 до 1, 6 [Прохорова, 2010].

Использование мультимасштабности требует установки правил изменения шкалы размеров диаграмм в зависимости от нескольких факторов: соотношения масштабов визуализации, смены единиц картографирования и их средней площади, изменения размаха вариации показателя (в случае смены уровня детализации), графической нагрузки (долей площади диаграмм, занимаемой на карте).

Представим, что пользователь меняет масштаб карты. Поскольку уровни АТД обладают критическими масштабами отображения, в зависимости от величины масштабного перехода может произойти либо сохранение текущего уровня, либо его смена.

В первом случае необходимо определить, будут ли размеры диаграмм меняться пропорционально уменьшению изображения или же будут оставаться постоянными. Увеличение и уменьшение размеров позволяют создать иллюзию приближения и отдаления от территории, что повышает коммуникативные качества карты. Во втором

случае возникает сложность, связанная с резким увеличением (уменьшением) территориальной единицы и возможно требующая пропорционального масштабирования диаграмм.

Отметим также, что критические масштабы отображения картодиаграмм могут не совпадать с критическими масштабами отображения самих территориальных единиц. Например, после переключения картодиаграмм с районной сетки на областную, границы районов могут продолжать показываться в некотором диапазоне уменьшения масштаба.

С учетом этой совокупности факторов было предложено четыре стратегии изменения шкал картодиаграмм. Во всех стратегиях необходимо предварительно установить минимальный  $D_{min}$  и максимальный  $D_{max}$  допустимый размер диаграммы наиболее крупной градации в шкале (вне зависимости от уровня). Относительно него будут автоматически получены диаметры остальных градаций. Для упрощения расчетов коэффициент  $k$ , характеризующий степень увеличения значков внутри шкалы, задан и равен везде 1,3. Стратегиям были присвоены двузначные коды, которые соответствуют возможности изменения размера (первая цифра) и компенсации иерархического перехода (вторая цифра). 0 означает «нет», 1 означает «да».

*Стратегия 0–0.* При изменении масштаба и уровня детализации диаметр наиболее крупной градации остается неизменным. Меняется только число градаций в соответствии с размахом вариации и числом единиц, и, соответственно, диаметр диаграммы для наименьшей градации. Переключение на следующий уровень детализации происходит, когда средний размер территориальной единицы становится равным  $D_{min}$ .

*Стратегия 0–1.* При изменении масштаба и постоянном уровне детализации диаметр наиболее крупной градации остается неизменным. Для масштабных переходов, сопровождающихся сменой уровня детализации, вводится коэффициент увеличения  $r$ , который является функцией среднего отношения размеров территориальных единиц и значений показателя между соседними уровнями АТД (см. далее). Он позволяет отразить увеличение значений показателя при агрегации данных и одновременно визуально компенсировать рост территориальной единицы. Переключение на следующий уровень детализации при уменьшении масштаба происходит, когда средний размер территориальной единицы становится равным  $rD_{min}$ .

*Стратегия 1–0.* В отличие от первых двух стратегий визуализацию проще программировать от мелких масштабов к крупным. Для каждого уровня детализации заранее определяются масштабы, в которых средний размер территориальной единицы будет равен  $D_{min}$ . При увеличении масштаба для первого уровня детализации диаметр диаграмм увеличивается с  $D_{min}$  до  $D_{max}$  как функция отношения масштабов вплоть до достижения критического масштаба следующего, более детального уровня, после чего происходит переключение на следующий уровень и размер максимальной градации снова становится равным  $D_{min}$ . Процесс повторяется. Для всех уровней детализации (кро-

ме самого детального) увеличение диаграмм происходит таким образом, чтобы размер максимальной градации достиг  $D_{max}$  в момент переключения на следующий уровень детализации. А для самого детального уровня увеличение происходит прямо пропорционально масштабу, и по достижению размера  $D_{max}$  диаметр остается постоянным.

*Стратегия 1–1.* Визуализация программируется также от мелких масштабов к крупным. При переключении уровня детализации размер становится равным  $D_{min} = D_{min}/r$  и далее увеличивается вплоть до  $D_{max} = D_{max}/r$  (переменным  $D_{min}$  и  $D_{max}$  переприсваиваются новые значения). После перехода на максимальный уровень детализации размер диаграмм аналогично стратегии 1–0 увеличивается вплоть до  $D_{max}$  и далее останавливается. Эта стратегия представляет собой симбиоз стратегий 0–1 и 1–0. Пример уровней мультимасштабной карты, выполненной в данной технике, показан на Рисунке 4.37.

Одним из основных вопросов в стратегиях 0–1 и 1–1 является вычисление коэффициента  $r$ . В работе [Самсонов, Подольский, Юрова, 2013] для переходов район — субъект и субъект — экономический район этот коэффициент имел значение 1.5 и 2 соответственно. Дальнейшие исследования позволят вывести обоснованную методику его расчета. Обратим внимание также на то, что в данных стратегиях вычислялись критические масштабы для визуализации картодиаграмм, а не самих территориальных единиц. Цены по их размерам могут быть разными.

Техника минимального ограничивающего прямоугольника позволяет также, не придерживаясь какой-либо из стратегий, «вписывать» шкалу размеров картодиаграмм в среднестатистические размеры территориальных единиц для визуализации в любом заданном масштабе таким образом, чтобы диаграммы не вылезали за пределы территориальных единиц. Для этого достаточно сопоставить максимальную по размеру градацию и среднюю ширину территориальной единицы на экране в заданном масштабе. Пример решения этой задачи представлен на Рисунке 4.38.

При использовании стратегий 1–0 и 1–1 может быть эффективно использована также *прозрачность знаков* для снижения графической нагрузки при увеличении диаграмм. Для этого в процессе увеличения масштаба прозрачность увеличивается с 0% до  $t\%$ , где  $t$  — допустимый уровень прозрачности знака. Данный прием, а также все стратегии визуализации картодиаграмм, были реализованы в приложении *Multimapper* на языке программирования *Java*. Пример визуализации с использованием стратегии 1–0 и прозрачности показан на Рисунке 4.39.

#### 4.6.4 Отображение внутренней неоднородности

При анализе внутренней неоднородности территориальных единиц можно показать, насколько вложенные в них единицы различаются по значению исследуемого показателя. Например, как области различаются внутри экономических районов и феде-

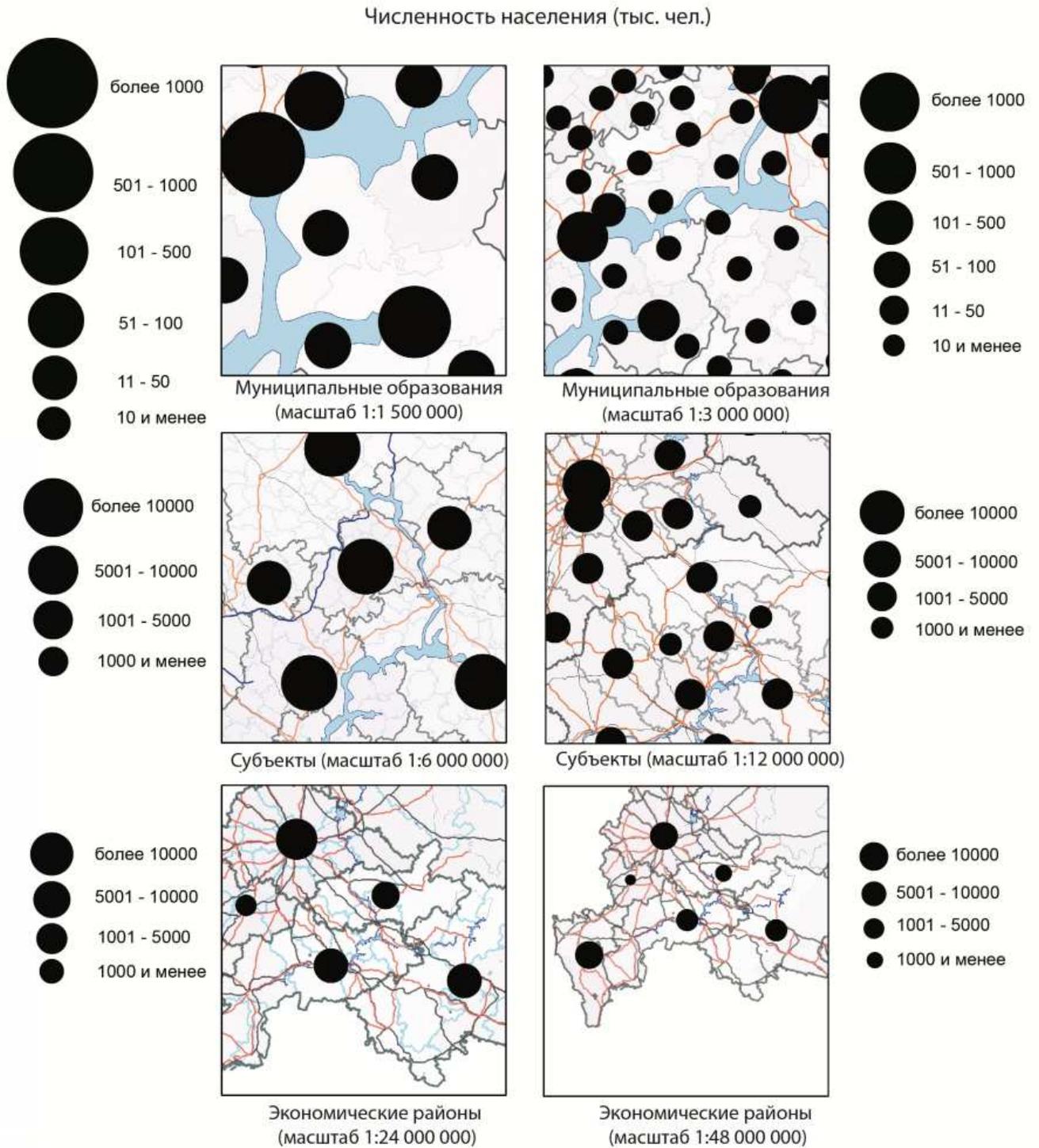


Рисунок 4.37 – Преобразование картодиаграмм (стратегия 1–1)

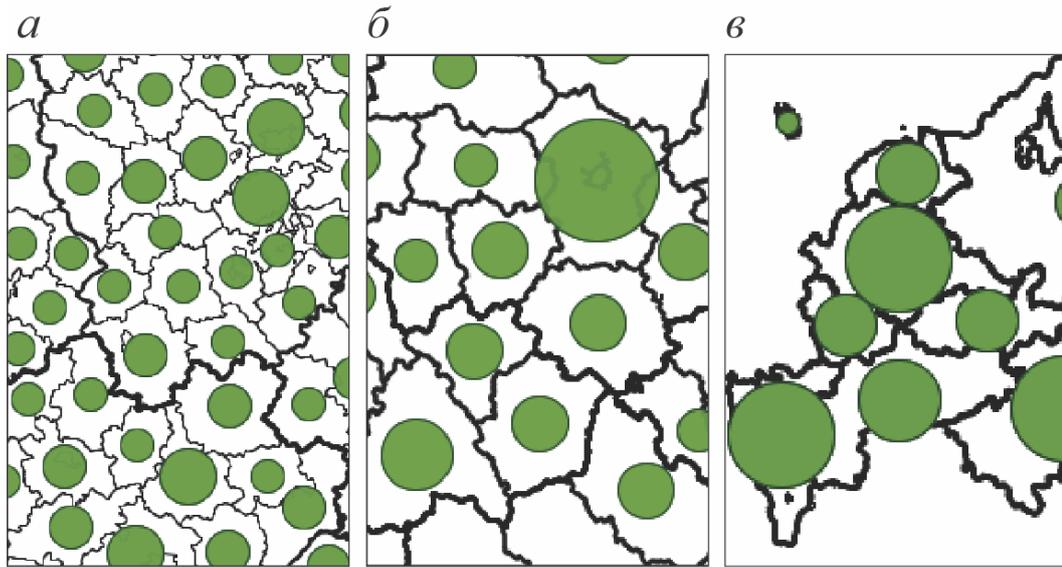


Рисунок 4.38 – Вписывание диаграмм в размеры территориальных единиц с помощью техники MBR

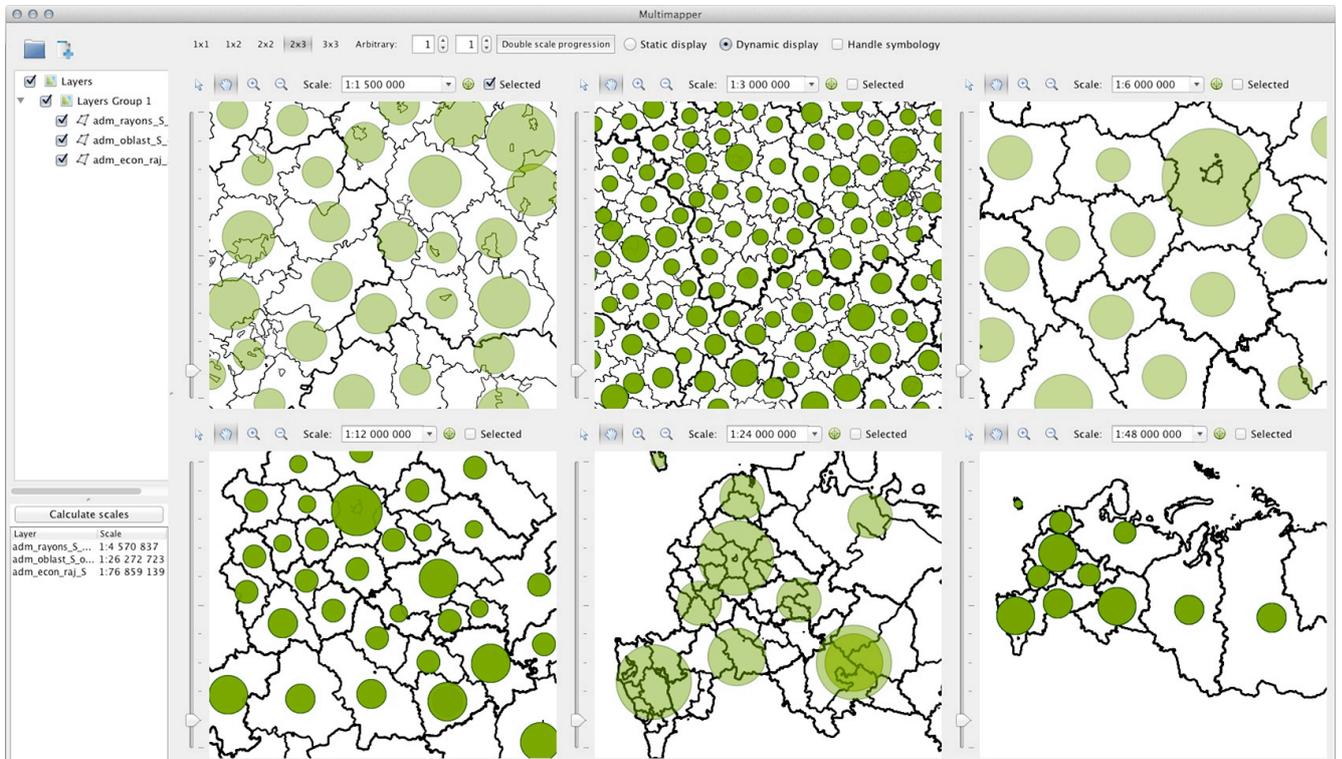


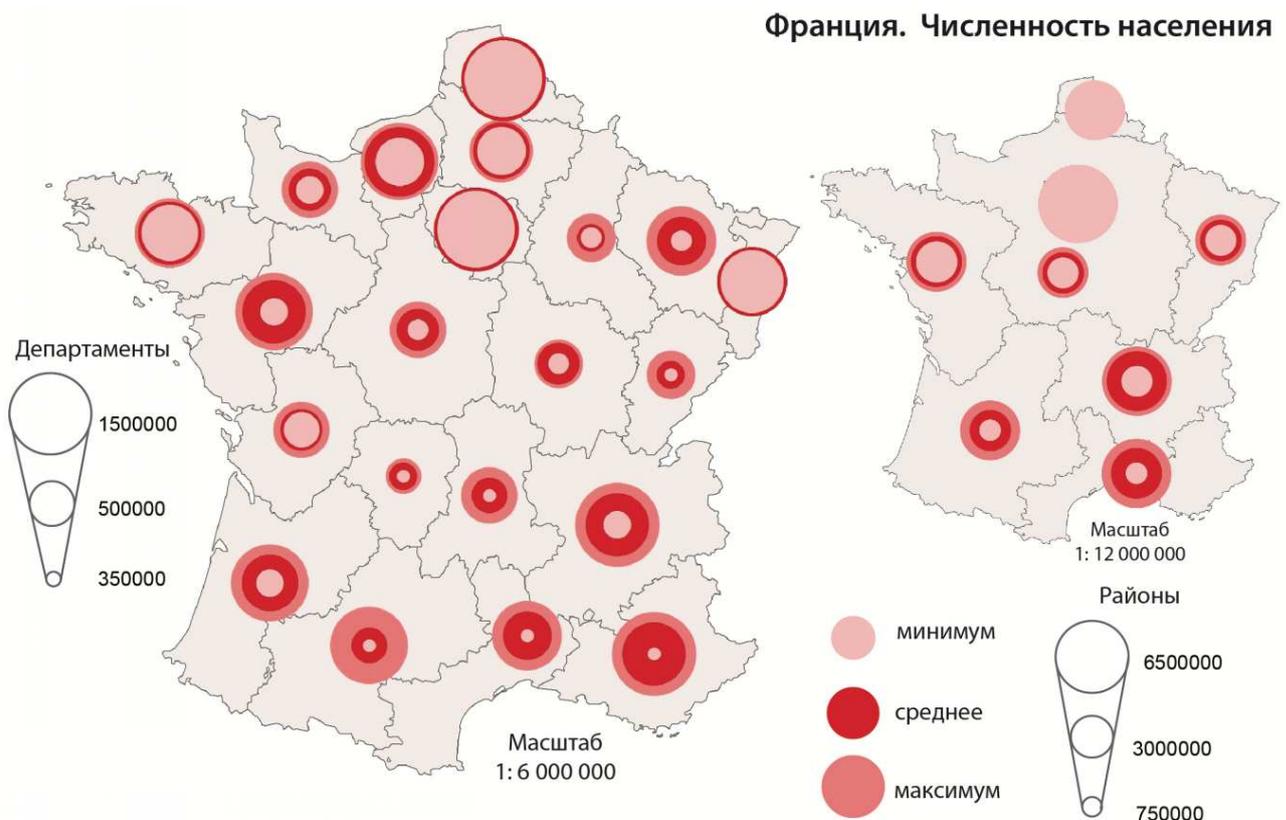
Рисунок 4.39 – Реализации прозрачности картодиаграмм в приложении \*Multimapper\*

ральных округов. Таким образом, можно выстроить дерево внутренней неоднородности иерархической сетки АТД. В качестве показателя может участвовать как некая тематическая характеристика, так и собственно размер самих единиц.

Для визуализации неоднородности предлагаются *вложенные диаграммы*. В одном значке совмещаются среднее, максимальное и минимальное значения показателя по субординатным (подчиненным) единицам. Алгоритм построения карт с использованием этого метода отображения данных включает в себя несколько этапов обработки данных:

1. Вычисление среднего, минимального и максимального значений показателя.
2. Классификация показателя и построение шкалы значков.
3. Выбор цветового оформления.

Для графических образов картодиаграмм используются оттенки одного цветового тона с изменением светлоты и насыщенности тона. Для среднего значения использована диаграмма самого насыщенного цвета, что позволяет акцентировать основное внимание на ней. Диаграмма для максимального значения окрашена светлее и менее насыщено, наконец, наиболее нейтральный оттенок имеет минимальное значение (Рисунок 4.40).



**Рисунок 4.40** – Вложенные картодиаграммы для отображения неоднородности

Метод достаточно выразителен при значительном размахе вариации, позволяет оценить асимметрию распределения относительно среднего значения, однако неэффективен в случае близости минимального и максимального значений (см. северные депар-

таменты на Рисунке 4.40). Для исключения влияния выбросов вместо среднего значения может быть использована медиана, а минимум и максимум заменены на, соответственно, нижнюю и верхнюю квартиль распределения показателя.

#### 4.6.5 Анимации

Исследования в области визуализации данных обеспечили возможность программирования анимационных переходов между графическими представлениями информации, которые способствуют пониманию того, как трансформируются данные при смене метода визуализации [Heer, Robertson, 2007; Ruchikachorn, Mueller, 2015; Kim, Heer, 2021]. Недавняя работа [Lee *et al.*, 2023] показала, что такие плавные переходы возможны между пространственными (значки, картограммы) и непространственными (диаграмма рассеяния, столбчатая диаграмма) представлениями. При этом в картографии область применения анимаций традиционно ограничена показом пространственно-временной динамики явлений [DiBiase *et al.*, 1992; Берлянт, Ушакова, 2000]. Применение же анимаций для визуализации переходов между способами картографического изображения остается недостаточно изученным.

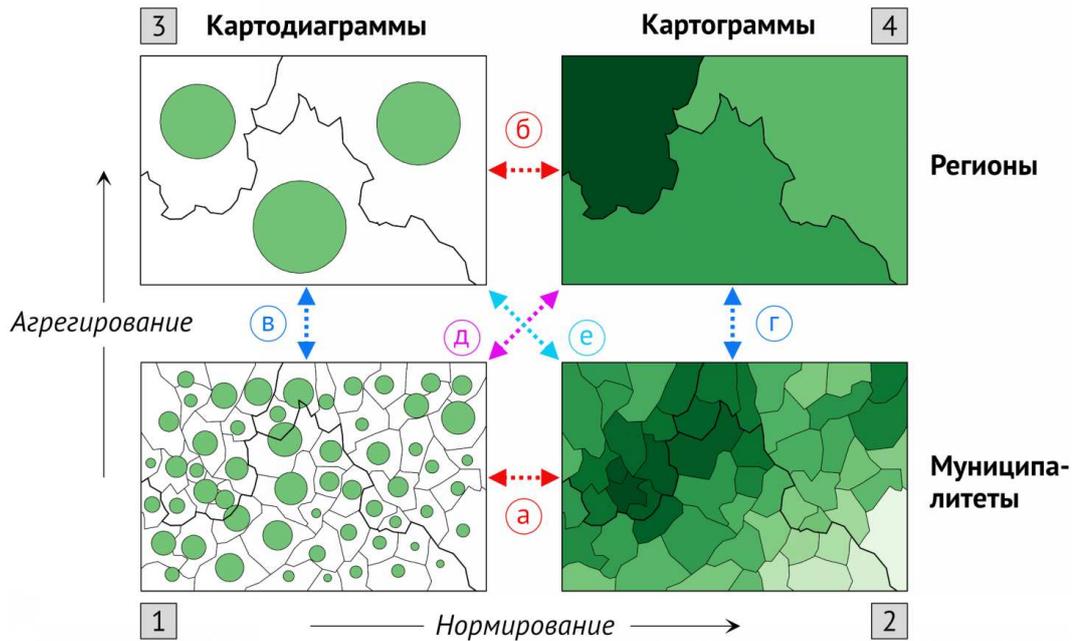
Данная проблема была исследована в работе [Самсонов, 2024], в который впервые получен метод графического перехода между картограммами и картодиаграммами на разных масштабных уровнях, который обеспечил наглядную визуализацию происходящих при этом преобразований данных — стандартизации (деления на площадь или относенный к ней показатель) и агрегирования (вычисления сводных статистик по группам территориальных единиц).

Для достижения поставленной цели была разработана система переходов, представленная на Рисунке 4.41:

1. Переходы *a*, *б* — связаны со сменой способа изображения и предполагают переключение реальной (полигон) и концептуальной (точка) геометрий территориальной единицы, а также графических переменных цвет и размер; применяется [де]нормирование данных.
2. Переходы *в*, *г* — связаны с переходом между соседними иерархическими уровнями территориальных единиц; способ изображения остается прежним, но происходит переключение между геометриями иерархических уровней; применяется [дез]агрегирование данных.
3. Переходы *д*, *е* — смешанные, комбинируют свойства переходов *a—г*.

В процессе анимации между состояниями происходит интерполяция координат границ и цвета заливки/обводки соответствующих элементов.

В процессе переходов происходит декомпозиция способов изображения на графические элементы, включающие геометрическую границу, цвет заливки и цвет обводки.



**Рисунок 4.41** — Исследуемая система картографических представлений (1-4) и переходов между ними (a-e)

В случае картограмм эти компоненты касаются площадных единиц картографирования, в случае картодиаграмм — символов, используемых для визуализации (в нашем случае это кружок). В процессе анимации между состояниями происходит интерполяция координат границ и цвета заливки/обводки соответствующих элементов. В метод заложена возможность нелинейного изменения скорости интерполяции за счет степенного параметра, который позволяет сфокусировать внимание пользователя карты на начальном либо конечном состоянии. Общая форма преобразования имеет в данном случае вид:

$$s_t = (1 - t^\lambda)s_0 + t^\lambda s_1, \quad (4.4)$$

где  $s_0, s_1$  — начальное и конечное состояние графического элемента,  $t \in [0, 1]$  — параметр и  $\lambda$  — показатель степени, по умолчанию равный 1. При  $\lambda > 1$  анимация будет более быстрой в области конечного состояния ( $t = 1$ ), а при  $\lambda < 1$  — в области начального состояния ( $t = 0$ ).

Для нахождения оптимального соответствия между точками на границе полигонов и точками на границах кружков (окружностях) используется метод, разработанный в [Veltman, 2019], а для разбиения кружков регионов на элементы, пропорциональные объему показателя в муниципалитетах и имеющие схожее относительное положение, применяется метод древовидной диаграммы Вороного [Balzer, Deussen, 2005] с пространственными ограничениями.

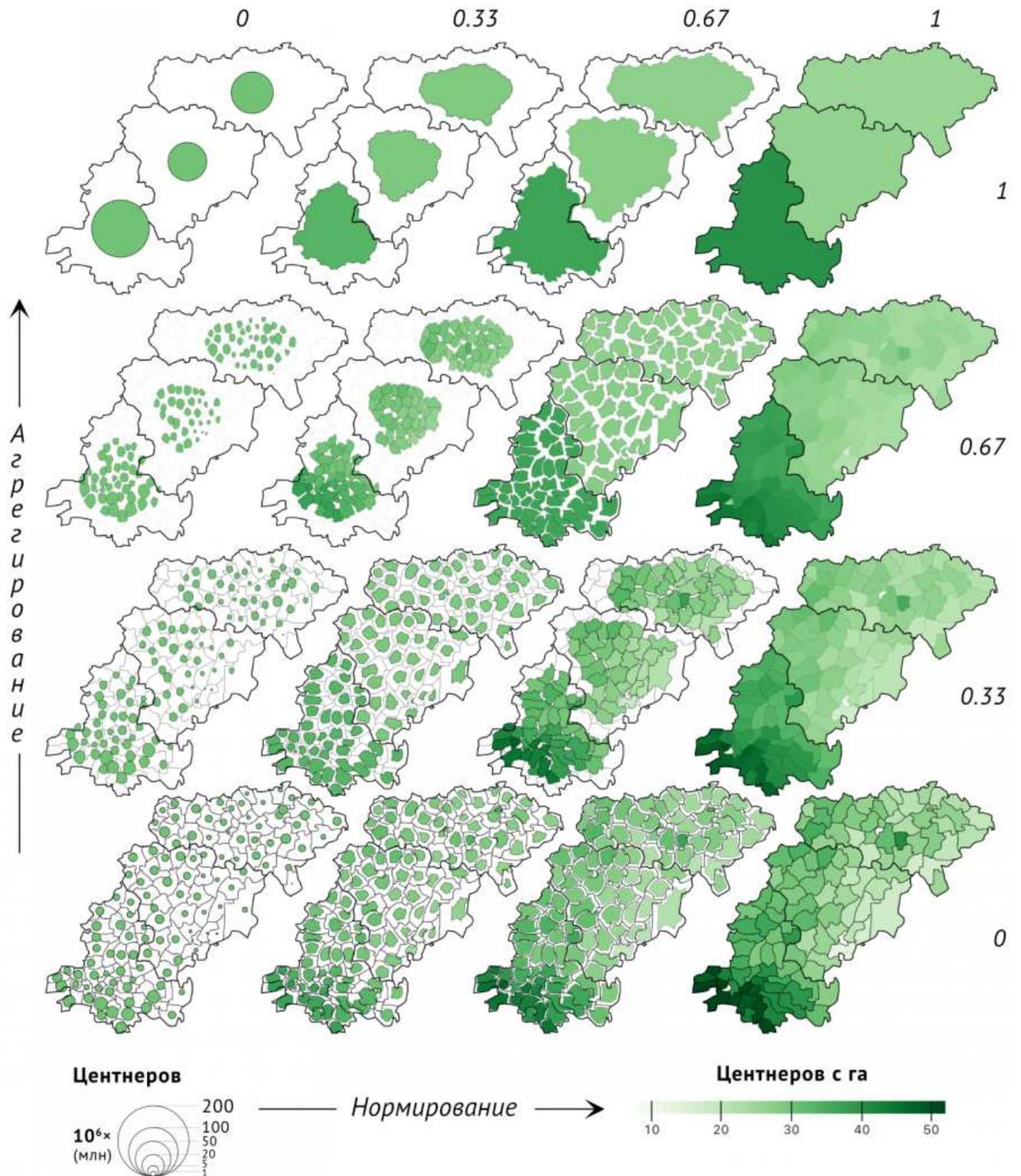
Разработанный метод был реализован на языке программирования *JavaScript* в интерактивной среде *Observable* [Bostock, 2025] с использованием программной библиотеки *flubber* [Veltman, 2019] для интерполяции геометрий. Апробация произведена на при-

мере данных по урожайности зерновых культур в Саратовской, Волгоградской и Ростовской областях за 2017 год [Росстат, 2024]. Представления полученных анимаций представлены на Рисунке 4.42. Апробация произведена на примере данных по урожайности зерновых культур в Саратовской, Волгоградской и Ростовской областях за 2017 год. Снимки состояний полученных анимаций представлены на Рисунке 4.42.

Проанализируем результаты с точки зрения поставленной цели — наглядного отображения преобразований данных при переходе между способами изображения и уровнями территориальных единиц. На Рисунке 4.42 видим, что стандартизация на обоих масштабных уровнях выражается в растяжении диаграммы до площади соответствующей территориальной единицы, что наглядно отражает процесс нормирования на площадь. Агрегирование нормированных значений на картограммах приводит к усреднению цветов по регионам, что отражает процесс осреднения сгруппированных данных. В то же время агрегирование абсолютных значений визуализируется более выразительно — в форме объединения диаграмм муниципалитетов в диаграммы регионов. Смешанные преобразования, изображенные по диагоналям на Рисунке 4.42, демонстрируют возможности одновременного отображения нормирования и агрегирования значений. При этом следует отметить, что каждое смешанное преобразование может быть представлено как комбинация прямых преобразований (например,  $e = a + b$ ).

#### 4.6.6 Итоги

Анализ современного состояния социально-экономического веб-картографирования показывает, что предлагаемые методы не учитывают специфику масштабных переходов и особенности распределения показателя на разных уровнях сетки АТД. В главе предложены приемы использования картограмм и картодиаграмм, позволяющие учесть данные особенности. Методика корректировки насыщенности в цветовых шкалах картограмм позволяет показать изменение размаха вариации показателя при смене уровня картографирования. Предложенные стратегии поведения шкал картодиаграмм при постоянном или меняющемся уровне картографирования требуют всестороннего тестирования и открывают ряд новых вопросов в применении этих способов на мультимасштабных картах. В обеих методиках активно используется метод минимальных ограничивающих прямоугольников для определения критических масштабов отображения территориальных единиц. Показано, что данный метод может также быть использован для вписывания диаграмм внутрь единиц АТД. Для характеристики неоднородности распределения показателя по вложенным единицам АТД каждого уровня предложен способ вложенных картодиаграмм, который наглядным образом показывает соотношение минимума, максимума и среднего значения. Проведенный анализ позволяет утверждать, что поставленная цель наглядного отображения преобразований данных методом анимаций была достигнута. Полученные результаты имеют потен-



**Рисунок 4.42** – Анимации между представлениями данных об урожайности зерновых в виде картограмм и картодиаграмм (в Саратовской, Волгоградской и Ростовской областях России за 2017 г.). Значения параметра интерполяции подписаны по столбцам и строкам.

циал дальнейшего развития и применения, в том числе в технологиях формирования мультимасштабных тематических карт.

## 4.7 Результаты и выводы

Подводя итог краткому изложению содержания четвертой главы диссертации, можем заключить, что в ней впервые не только закреплены авторские положения теории мультимасштабного картографирования, но также показано на разнообразных примерах, как эти положения реализуются в задачах общегеографического, физико-географического и социально-экономического картографирования.

Основные выводы по Главе 4 заключаются в следующем:

1. Создание мультимасштабных карт опирается на построение баз пространственных данных со множеством уровней детализации, а также проектирование масштабных уровней карты, преобразований между ними и анимаций этих преобразований. Преобразования являются интерактивным откликом карты на изменение ее масштаба пользователем, а их поведение должно подчиняться правилам логичности и последовательности изменений.
2. При визуализации цифровых моделей дорожной сети и структуры населенных пунктов на мультимасштабных картах необходим последовательный отбор объектов дорог в соответствии с предельными значениями густоты, а также их согласование с квартальной структурой для каждого масштабного уровня.
3. Мультимасштабное картографирование на основе цифровых моделей геополей (рельефа, параметров ветрового волнения) предполагает использование нескольких уровней детализации поверхностей, а также последовательного изменения сечения и цветовых шкал при использовании способа изолиний. Введение дополнительных горизонталей, а также утрирование цифровой модели в участках высокой вариабельности позволяет локально повысить детальность и читаемость изображения.
4. Обзорные уровни мультимасштабных карт транспортных сетей могут опираться на введение полимагистралей, объединяющих сонаправленные пути сообщения. Алгоритм построения полимагистралей должен опираться на поиск линий, проходящих в пределах заданного коридора, при этом его ширина определяется целевым масштабом картографирования.
5. Использование картограмм и картодиаграмм на мультимасштабных картах требует вычисления предельных масштабов визуализации единиц картографирования, а также последовательного изменения шкал размеров и цветов, при котором обеспечивается компенсация резкого увеличения абсолютных значений в

результате суммирования, а также снижения разброса относительных значений в результате осреднения. Эти преобразования данных могут быть явно визуализированы посредством анимаций, что повышает наглядность мультимасштабных карт, на которых применяются данные способы изображения.

Обобщение этих выводов позволяет сформулировать четвертое защищаемое положение:

#### **Защищаемое положение №4**

*Мультимасштабность* картографического изображения обеспечивается интерактивным откликом процессов генерализации пространственных данных и преобразования графических средств и способов изображения на изменение масштаба карты. Анимации переходов между масштабными уровнями карты позволяют явным образом визуализировать геометрические и атрибутивные преобразования пространственных данных и повысить информативность мультимасштабных карт.

## Заключение

В результате выполненного исследования был пройден новый этап в развитии методологии генерализации пространственных данных для целей картографирования. Основные выводы заключаются в следующем:

1. Разработанные методы генерализации пунктов системы расселения и дорожной сети позволяют сохранить пространственный рисунок, плотность и связность соответствующих элементов при выполнении их отбора, а также выделить основные полимагистралы путем совмещения векторных линий, соответствующих сопоставленным путям сообщения.
2. Для сохранения характера формы и соседства линейных объектов необходимо выполнять сегментацию линий на участки различного морфологического типа, а также объединять их в сетевую структуру (граф), учитывающую взаимные пересечения. При этом для генерализации естественных, схематичных и ортогональных сегментов следует использовать алгоритмы Ли-Оупеншоу, Дугласа-Пейкера и ортогональный соответственно, а для сохранения топологии — представление линий в виде графа.
3. Географически правдоподобная мелкомасштабная генерализация цифровых моделей рельефа (ЦМР) достигается путем восстановления обобщенной поверхности по тальвегам и водоразделам с последующим расширением долин и междуречий для усиления их визуального распознавания, при необходимости согласования результирующих отметок высот с заданным множеством генерализованных объектов гидрографии.
4. Генерализация пространственных данных урбанизированных территорий позволяет выполнять построение регулярно-сеточных баз данных городской среды путем агрегирования разнородной пространственной информации для численных моделей прогноза погоды, климатических и эколого-геохимических моделей.
5. Детальность пространственных данных в геометрической части может быть оценена по совокупности характеристик объектов и их множеств, монотонно реагирующих на последовательную генерализацию с учетом количественных характеристик сложности ландшафта; при этом достижение сопоставимой детальности разными алгоритмами может быть осуществлено путем предложенного метода регрессионного моделирования зависимости между их параметрами.

6. Детальность цифровых моделей рельефа для целей картографирования определяется совокупностью их пространственного разрешения и гранулярности. Относительная гранулярность ЦМР для мелкомасштабных карт должна составлять 5—6 мм. При визуализации ЦМР на детальность изображения влияет как детальность ЦМР, так и параметры визуализации. В частности, при использовании горизонталей и штрихов детальность варьируется путем генерации дополнительных горизонталей с учетом автоматической оценки заложения и центральности, а также путем регулировки степени сглаживания, интервала и шага расстановки штрихов.
7. Автоматическое совмещение (конфляция) цифровой модели рельефа (ЦМР) и линейных объектов гидрографии разной детальности может быть достигнуто путем сдвига узлов ЦМР в направлении, локально определяемом суперпозицией векторов, заканчивающихся на гидрографических линиях и начинающихся на водотоках-контрагентах — путях на поверхности ЦМР, наиболее похожих по пространственному положению на объекты гидрографии и выделяемых автоматическим путем. Координатная привязка геоизображений произвольной детальности может быть осуществлена с использованием компьютерного зрения.
8. Построение мультимасштабных карт предполагает выполнение генерализации пространственных данных, а также преобразований графических средств и способов изображения, раскрывающих многоуровневые (в том числе, иерархические) аспекты картографируемого явления при переходе между масштабами. При этом преобразования могут носить не только вертикальный, но и горизонтальный характер. В частности, оператор утрирования разработанный для отображения высокоизменчивых прибрежных участков характеристик ветрового волнения, позволяет повысить наглядность карт путем смещения изолиний от берега.
9. Разработанные методы отображения структуры населенных пунктов, гипсометрического картографирования и картографирования характеристик ветрового волнения, построения полимагистралей, вычисления параметров картограмм и картодиаграмм позволяют автоматизировать составление мультимасштабных карт соответствующего содержания. При этом анимационные переходы между способами картографического изображения позволяют явным образом визуализировать происходящие при этом преобразования пространственных данных, а их применение на мультимасштабных картах обеспечивает возможность плавного перехода между масштабными уровнями.

## Публикации по теме диссертации

Статьи в рецензируемых научных изданиях, индексируемых в базе ядра Российского индекса научного цитирования «eLibrary Science Index»:

1. Лурье И. К., **Самсонов Т. Е.** Структура и содержание базы пространственных данных для мультимасштабного картографирования // *Геодезия и картография*. – 2010. – № 11. – С. 17-23. EDN: [SOAXEZ / 0,81 п.л.](#) (IF SJR = 0.225, Q3). Вклад соискателя 80%: Лурье И. К. принадлежит постановка задачи; все результаты получены соискателем.
2. **Самсонов Т. Е.**, Подольский А. С., Юрова Н. Д. Преобразования способов изображения и их соотношение с иерархией единиц картографирования // *Вестник Московского университета. Сер.5 География*. – 2013. – № 4. – С. 15-23. EDN: [RDWYWV / 1,04 п.л.](#) (IF SJR = 0.258, Q3). Вклад соискателя 50%: Подольский А. С. и Юрова Н. Д. выполнили подбор источников и анализ карт; постановка задачи, систематизация и обобщение результатов выполнены соискателем.
3. **Самсонов Т. Е.**, Подольский А. С. Автоматизация вычисления параметров отображения дорожной сети и структуры населенных пунктов на мультимасштабных картах // *Геодезия и картография*. – 2014. – № 9. – С. 21-28. DOI: [10.22389/0016-7126-2014-891-9-21-28](#) / 0,5 п.л. (IF SJR = 0.225, Q3). Вклад соискателя 50%: Подольский А. С. выполнил техническую реализацию методики, обработку данных и первичный анализ результатов; постановка задачи, разработка методики и обобщение результатов выполнены соискателем.
4. **Самсонов Т. Е.**, Юрова Н. Д. Применение картограмм и картодиаграмм на мультимасштабных социально-экономических картах // *Геодезия и картография*. – 2014. – № 11. – С. 30-38. DOI: [10.22389/0016-7126-2014-893-11-30-38](#) / 0,56 п.л. (IF SJR = 0.225, Q3). Вклад соискателя 50%: Юрова Н. Д. выполнила техническую реализацию методик, обработку данных и первичный анализ результатов; постановка задачи, разработка методик и обобщение результатов выполнены соискателем.
5. **Самсонов Т. Е.** Якимова О. П., Алексеев В. В., Богаевская В. Г., Горохов А. А., Князев В. Н., Преображенская М. М., Ухалов А. Ю., Эдельсбруннер Х. Алгоритм геометрического упрощения множества линий путем стягивания ребер графа с сохранением топологии // *Геодезия и картография*. – 2014. – № 3. – С. 29-36. DOI: [10.22389/0016-7126-2014-885-3-29-36](#) / 0,5 п.л. (IF SJR = 0.225, Q3). Вклад соискателя 50%: Якимова О. П. и другие соавторы предложили использовать стягивание ребер для геометрического упрощения, реализовали его и выполнили обработку данных; постановка задачи, подготовка данных, разработка методики оценки и анализ результатов выполнены соискателем.

6. **Самсонов Т. Е.**, Кривошеина А. М. Автоматизация отбора населенных пунктов с учетом пространственной неравномерности их распределения для целей мелко-масштабного картографирования // *Известия вузов. Геодезия и аэрофотосъемка*. – 2015а. – № 1. – С. 74-82. EDN: [RTPVYG](#) / 0,56 п.л. (IF РИНЦ = 0.234, K1). Вклад соискателя 50%: Кривошеина А. М. выполнила техническую реализацию методики, обработку данных и первичный анализ результатов; постановка задачи, разработка методики и обобщение результатов выполнены соискателем.
7. **Самсонов Т. Е.**, Кривошеина А. М. Генерализация дорожной сети в мелких масштабах картографирования с сохранением связности населенных пунктов // *Известия вузов. Геодезия и аэрофотосъемка*. – 2015b. – № 2. – С. 23-30. EDN: [TTAFZL](#) / 0,5 п.л. (IF РИНЦ = 0.234, K1). Вклад соискателя 50%: Кривошеина А. М. выполнила техническую реализацию методики, обработку данных и первичный анализ результатов; постановка задачи, разработка методики и обобщение результатов выполнены соискателем.
8. **Самсонов Т. Е.**, Тригуб К. С. Картографирование локальных климатических зон Москвы по космическим снимкам // *Геодезия и картография*. – 2018. – Т. 79. – № 6. – С. 14-25. DOI: [10.22389/0016-7126-2018-936-6-14-25](#) / 0,63 п.л. (IF SJR = 0.225, Q3). Вклад соискателя 50%: Тригуб К. С. выполнила техническую реализацию методики, обработку данных и первичный анализ результатов; постановка задачи, разработка методики и обобщение результатов выполнены соискателем.
9. **Самсонов Т. Е.**, Прохорова Е. А. Генерализация транспортных сетей при мультимасштабном картографировании // *Известия Российской академии наук. Серия географическая*. – 2019. – № 1. – С. 107-117. DOI: [10.31857/S2587-556620191107-117](#) / 0,69 п.л. (IF SJR = 0.217, Q2). Вклад соискателя 75%: Прохорова Е. А. предложила масштабные диапазоны картографирования транспортных сетей и определила место полимагистралей в них; постановка задачи автоматизированного выделения полимагистралей, разработка и техническая реализация методики ее решения, обработка данных, анализ результатов выполнены соискателем.
10. Мухаметшин А. Р., **Самсонов Т. Е.**, Лурье И. К. Координатная привязка геоизображений с использованием компьютерного зрения // *Геодезия и картография*. – 2024. – № 5. – С. 24-36. DOI: [10.22389/0016-7126-2024-1007-5-24-36](#) / 0,81 п.л. (IF SJR = 0.225, Q3). Вклад соискателя 30%: Мухаметшин А. Р. участвовал в разработке методики, выполнил ее реализацию, обработку данных и первичный анализ результатов; И.К.Лурье участвовала в первичном анализе результатов; соискатель поставил задачу, участвовал в разработке методики и выполнил обобщение результатов.
11. **Samsonov T.** Morphometric Mapping of Topography by Flowline Hachures // *The Cartographic Journal*. – 2014. – Vol. 51. – № 1. – pp. 63-74. DOI: [10.1179/1743277413Y.00000](#)

/ 0,75 н.л. (IF JIF = 1.000, Q3). Вклад соискателя 100%.

12. **Samsonov T.** Review of Abstracting Geographic Information In a Data Rich World: Methodologies and Applications of Map Generalisation // *Cartographic Perspectives*. – 2015. – no. 80. – pp. 54–57. DOI: [10.14714/CP80.1308](https://doi.org/10.14714/CP80.1308) / 0,25 н.л. (IF SJR = 0.133, Q4). Вклад соискателя 100%.
13. **Samsonov T.**, Konstantinov P., Varentsov M. Object-oriented approach to urban canyon analysis and its applications in meteorological modeling // *Urban Climate*. – 2015. – Vol. 13. – pp. 122-139. DOI: [10.1016/j.uclim.2015.07.007](https://doi.org/10.1016/j.uclim.2015.07.007) / 1,13 н.л. (IF JIF = 6.000, Q1). Вклад соискателя 50%: Константинову П. И. и Варенцову М. И. принадлежит постановка задачи, обобщение результатов и их апробация в климатическом моделировании; разработка методики, ее техническая реализация, обработка данных и первичный анализ результатов выполнены соискателем.
14. **Samsonov T.**, Yakimova O. Shape-Adaptive Geometric Simplification of Heterogeneous Line Datasets // *International Journal of Geographical Information Science*. – 2017. – Т. 31. – № 8. – pp. 1485-1520. DOI: [10.1080/13658816.2017.1306864](https://doi.org/10.1080/13658816.2017.1306864) / 2,25 н.л. (IF JIF = 4.300, Q1). Вклад соискателя 75%: Якимова О. П. выполнила разработку вычислительных алгоритмов и техническую реализацию методики; постановка задачи, разработка методики, обработка данных, анализ и обобщение результатов выполнены соискателем.
15. **Samsonov T.**, Koshel S., Walther D., Jenny B. Automated placement of supplementary contour lines // *International Journal of Geographical Information Science*. – 2019. – Vol. 33. – № 10. – pp. 2072-2093. DOI: [10.1080/13658816.2019.1610965](https://doi.org/10.1080/13658816.2019.1610965) / 1,38 н.л. (IF JIF = 4.300, Q1). Вклад соискателя 50%: Кошель С. М. и Йенни Б. принимали участие в разработке методики; Вальтер Д. О. выполнил подбор примеров и первичный анализ результатов, а также основную работу по технической реализации методики и обработке данных; соискателю принадлежит постановка задачи, участие в разработке и технической реализации методики, участие в обработке данных, а также обобщение результатов.
16. **Samsonov T. E.** Automated Conflation of Digital Elevation Model with Reference Hydrographic Lines // *ISPRS International Journal of Geo-Information*. – 2020. – Vol. 9. – № 5 (334). – pp. 1-40. DOI: [10.3390/ijgi9050334](https://doi.org/10.3390/ijgi9050334) / 2,56 н.л. (IF JIF = 2.800, Q2). Вклад соискателя 100%.
17. **Samsonov T. E.**, Varentsov M. I. Computation of City-descriptive Parameters for High-resolution Numerical Weather Prediction in Moscow Megacity in the Framework of the COSMO Model // *Russian Meteorology and Hydrology*. – 2020. – Vol. 45. – № 7. – pp. 515-521. DOI: [10.3103/s1068373920070079](https://doi.org/10.3103/s1068373920070079) / 0,44 н.л. (IF JIF = 1.400, Q4). Вклад соискателя 50%: Варенцов М. И. участвовал в постановке задачи, разработке методики агрегирования, выполнил ее апробацию; соискатель участвовал в постановке зада-

чи и разработке методики агрегирования, а также выполнил подготовку исходных данных. [Русскоязычная версия: **Самсонов Т. Е.**, Варенцов М. И. Определение параметров городской среды Московского мегаполиса для детализированного численного прогноза погоды в рамках модели COSMO // *Метеорология и гидрология*. — 2020. — № 7. — С. 112–119. EDN: [IRDBCD](#) / 0,5 п.л. (IF РИНЦ = 1.415, K1). Вклад соискателя 50%.]

18. **Samsonov T.**, Yakimova O. Regression modeling of reduction in spatial accuracy and detail for multiple geometric line simplification procedures // *International Journal of Cartography*. — 2020. — Vol. 6. — № 1. — pp. 47-70. DOI: [10.1080/23729333.2019.1615745](#) / 1,5 п.л. (IF JIF = 0.400, Q4). Вклад соискателя 75%: Якимова О. П. выполнила разработку вычислительных алгоритмов и техническую реализацию методики; постановка задачи, разработка методики, обработка данных, анализ и обобщение результатов выполнены соискателем.
19. Varentsov M., **Samsonov T.**, Demuzere M. Impact of Urban Canopy Parameters on a Megacity's Modelled Thermal Environment // *Atmosphere*. — 2020. — Vol. 11. — № 12 (1349). — pp. 1-31. DOI: [10.3390/atmos11121349](#) / 2,0 п.л. (IF JIF = 2.500, Q3). Вклад соискателя 20%: Варенцову М. И и Демузере М. принадлежит постановка задачи, разработка методики, моделирование и анализ результатов; соискателем выполнены обработка и анализ данных, необходимых для моделирования.
20. **Samsonov T.** Granularity of Digital Elevation Model and Optimal Level of Detail in Small-Scale Cartographic Relief Presentation // *Remote Sensing*. — 2022. — Vol. 14. — № 5 (1270). — pp. 1-40. DOI: [10.3390/rs14051270](#) / 2,56 п.л. (IF JIF = 4.200, Q1). Вклад соискателя 100%.
21. Kosheleva N. E., Vlasov D. V., Timofeev I. V., **Samsonov T. E.**, Kasimov N. S. Benzo[a]pyrene in moscow road dust: pollution levels and health risks // *Environmental Geochemistry and Health*. — 2023. — Vol. 45 (1287). — pp. 1669-1694. DOI: [10.1007/s10653-022-01287-9](#) / 1,63 п.л. (IF JIF = 3.200, Q2). Вклад соискателя 20%: Кошелевой Н. Е. и другим соавторам принадлежит постановка задачи, разработка методики, сбор полевых данных, моделирование и анализ результатов; соискателем выполнена подготовка данных, необходимых для моделирования.
22. Myslenkov S., **Samsonov T.**, Shurygina A., Kiseleva S., Arkhipkin V. Wind waves web atlas of the Russian seas // *Water*. — 2023. — Vol. 15, no. 11 (2036). — pp. 1-18. DOI: [10.3390/w15112036](#) / 1,19 п.л.(IF JIF = 3.000, Q2). Вклад соискателя 20%: Мысленков С. А., Киселева С. В. и Архипкин В. С. выполнили работы, связанные с моделированием характеристик ветрового волнения и разработкой необходимых для этого методик; Шурыгина А. А. выполнила основной объем работ по технической реализации атласа; разработка методик визуализации, картографических стилей оформления, архитектуры атласа принадлежат соискателю.

23. **Samsonov T.**, Yakimova O., Potemkin D., Guseva O. Spatial variation of feature density in multiscale topographic data // *GEOGRAPHY. ENVIRONMENT. SUSTAINABILITY*. – 2023. – Vol. 16. – № 1. – pp. 86-102. DOI: [10.24057/2071-9388-2022-127](https://doi.org/10.24057/2071-9388-2022-127) / 1,06 п.л. (IF SJR = 0.349, Q2). Вклад соискателя 60%: Якимова О.П. и Потемкин Д.А. выполнили разработку вычислительных алгоритмов и техническую реализацию методики; Гусева О.А. подготовила исходные данные; постановка задачи, разработка методики, обработка данных, анализ и обобщение результатов выполнены соискателем.

## Монографии

24. **Самсонов Т. Е.** Мультимасштабное картографирование рельефа: общегеографические и гипсометрические карты. — *LAP Lambert Academic Publishing Saarbrucken*, 2011. — 208 с.

## Прочие издания

25. **Самсонов Т. Е.** Мультимасштабное картографирование — новое направление картографии // *Современная географическая картография* / под ред. И.К.Лурье и В.И.Кравцовой. — Дата+ Москва: 2012. — С. 21–35.
26. **Самсонов Т. Е.** Картографические методы визуализации и генерализации цифровых моделей рельефа // *Геоморфологи: Современные методы и технологии цифрового моделирования рельефа в науках о Земле.* / под ред. А.В.Кошкарёва — Вып. 6. — Медиа-ПРЕСС Москва: 2016. — С. 9–18.
27. **Самсонов Т. Е.** Современная геоинформатика и ее роль в развитии картографии и взаимодействии географических наук // *Картография в цифровую эпоху* / под ред. В. М. Котлякова, И. К. Лурье, А. А. Медведева, Т. Е. Самсонова. — Вопросы географии, Т. 144 — Издательский дом Кодекс, Москва: 2017. — С. 52–72.
28. **Самсонов Т. Е.**, Ривин Г. С., Варенцов М.И. Моделирование метеорологических условий городской атмосферы // *Климат Москвы в условиях глобального потепления* / под ред. Кислова А.В. — Издательство МГУ. — Москва, 2017. — с. 33-89.
29. Мысленков С.А., **Самсонов Т.Е.**, Киселева С.В., Архипкин В.С., Шестакова А.А., Умнов П.М. Разработка веб-атласа доступной энергии волн и ветра прибрежной зоны морей России: информационное и картографическое обеспечение // *Международный научный журнал Альтернативная энергетика и экология*. — 2018. — № 16-18. — С. 39–54. DOI: [10.15518/isjaee.2018.16-18.039-054](https://doi.org/10.15518/isjaee.2018.16-18.039-054)
30. Мысленков С. А., **Самсонов Т. Е.**, Шурыгина А. А., Киселева С. В., Архипкин В. С., Шестакова А. А., Суркова, Г. В., Сильвестрова К. П., Умнов П. М. Разработка веб-атласа доступной энергии волн и ветра прибрежной зоны морей

России: веб-интерфейс и аналитические функции // *Международный научный журнал Альтернативная энергетика и экология*. — 2020. — № 7-18 (330-341). — С. 44–59. DOI: [10.15518/isjaee.2020.07-18.044-059](https://doi.org/10.15518/isjaee.2020.07-18.044-059)

31. **Samsonov T.** Multiscale Hypsometric Mapping // *Lecture Notes in Geoinformation and Cartography* / ed. by A. Ruas. — Vol. 1. — Springer: Berlin, Heidelberg — 2011. — pp. 497–520.
32. **Samsonov T., Jenny B.** Small-scale and Multi-scale Relief Mapping // *Mountain Cartography: 16 Years ICA Commission on Mountain Cartography*. — Vol. 21 of Wiener Schriften zur Geographie und Kartographie. — University of Vienna Vienna, Austria: 2015. — pp. 77–86.
33. **Samsonov T.** Spatial Analysis and Modelling // *The Routledge Handbook of Geospatial Technologies and Society*. — Abingdon, Oxfordshire, UK, England: Routledge (Taylor and Francis), 2023. — pp. 263–290.

### Свидетельства о регистрации прав на программы для ЭВМ

1. Преображенская М.М., Алексеев В.В., **Самсонов Т.Е.** Программа для вычисления топологических характеристик географических карт. *Свидетельство о регистрации № 2016616516 от 15 июня 2016 г.*
2. Якимова О.П., **Самсонов Т.Е.** MapSimplify 2.0 — научно-исследовательская программа для картографической генерализации линейных объектов смешанного характера. *Свидетельство о регистрации № 2017610341 от 10 января 2017 г.*
3. Якимова О.П., **Самсонов Т.Е.** Научно-исследовательская программа согласования параметров алгоритмов геометрического упрощения для получения сходных по детализации результатов. *Свидетельство о регистрации № 2021611919 от 8 февраля 2021 г.*
4. **Самсонов Т.Е.**, Варенцов М.И. Программный комплекс подготовки данных о геометрии городской застройки и подстилающей поверхности для детализированного численного прогноза погоды и метеорологического моделирования. *Свидетельство о регистрации № 2024687849 от 21 ноября 2024 г.*

### Материалы конференций

1. Кривошеина А. М., **Самсонов Т. Е.** Сопряженная генерализация населенных пунктов и транспортной сети в мелких масштабах картографирования // *ИнтерКарто. ИнтерГИС*. — 2012. — Т. 18. — С. 419–425.
2. Куканова Е. А., Константинов П. И., **Самсонов Т. Е.** Создание универсальной модели пространственного анализа для вычисления характеристик урбанизирован-

- ной подстилающей поверхности с использованием ГИС-технологий // *Материалы Международной научно-практической конференции Современные проблемы географии, экологии и природопользования, г. Волгоград, 25-26 апреля 2012 г.* — Волгоград, 2012. — С. 143–147.
3. Подольский А. С., **Самсонов Т. Е.** Принципы оформления мультимасштабных общегеографических карт // *ИнтерКарто. ИнтерГИС.* — 2012. — Т. 18. — С. 71–81.
  4. Юрова Н. Д., **Самсонов Т. Е.** Мультимасштабное социально-экономическое картографирование на примере производства строительных материалов в России // *ИнтерКарто. ИнтерГИС.* — 2012. — Т. 18. — С. 469–476.
  5. **Самсонов Т. Е.**, Мысленков С. А., Торопов П. А. Генерализация карт градиентного поля при картографировании ветра и морских течений в разных масштабах // *Геоинформационное картографирование в регионах России: материалы V Всероссийской научно-практической конференции.* — Цифровая полиграфия Воронеж: 2013. — С. 122–125.
  6. Халиуллина А. Р., **Самсонов Т. Е.** Мультимасштабная гипсометрическая карта России // *Геоморфология и картография: Материалы XXXIII Пленума Геоморфологической комиссии РАН.* — Изд-во Саратовского ун-та Саратов: 2013. — С. 106–110.
  7. **Самсонов Т. Е.** Новые цифровые модели рельефа NYPISO для мелкомасштабного картографирования // *Материалы V Международной конференции молодых ученых и специалистов памяти академика А. П. Карпинского (28 февраля – 3 марта 2017 г., ВСЕГЕИ, Санкт-Петербург).* — Изд-во ВСЕГЕИ Санкт-Петербург: 2017. — С. 687–689.
  8. Якимова О.П., **Самсонов Т.Е.**, Гусева О.А., Потёмкин Д.А. Исследование метрик визуальной сложности изображения на примере карт различного масштаба и детализации // *Геодезия, картография, геоинформатика и кадастры. Наука и образование. Сборник III Всероссийской научно-практической конференции.* — Издательство РГПУ им. А.И. Герцена СПб: 2019. — С. 376–383.
  9. Мысленков С.А., **Самсонов Т.Е.**, Шурыгина А.А., Киселева С.В., Сильвестрова К.П., Архипкин В.С. Особенности разработки веб-атласа доступной ветровой и волновой энергии морей России // *Возобновляемые источники энергии: Материалы Всероссийской научной конференции и XII молодежной школы с международным участием : сборник / Отв. ред. С.В. Киселева, Ю.Ю. Рафикова / Под ред. Ю. Ю. Рафикова.* — Наука, Москва: 2020. — С. 347–351.
  10. **Самсонов Т.Е.** Автоматизированное совмещение цифровой модели рельефа и векторных данных о гидрографической сети // *Геоинформационное картографирование в регионах России: материалы XI Всероссийской научно-практической конференции (Воронеж, 23-24 ноября 2020 г.)* — Воронеж: Цифровая полиграфия,

2020. — С. 294–301

11. Самсонов Т. Е., **Якимова О. П.** Анализ эффективности ограничивающих метрик алгоритмов геометрического упрощения на основе экспертной оценки детализации линий // *ИнтерКарто. ИнтерГИС.* — 2021. — Т. 27, № 2. — С. 253–267.
12. Якимова, О. П., **Самсонов, Т. Е.**, Потёмкин, Д. А., Усманова, Э. А. Инструменты для оценки детализации карт для геоинформационной системы QGIS // *ИнтерКарто. ИнтерГИС.* — 2021. — Т. 27, № 2. — С. 268–279.
13. Карташов Г.А., **Самсонов Т.Е.**, Мысленков С.А. Генерализация триангуляционных моделей характеристик ветрового волнения // *Геодезия, картография, геоинформатика и кадастры. Инновации в науке, образовании и производстве: Материалы V Всероссийской науч.-практ. конф. 2024 г., Санкт-Петербург* / науч. ред. И. Е. Сидорина. — Издательско-полиграфическая ассоциация вузов — СПб: 2024. — С. 522–527.
14. **Самсонов Т.Е.** Анимационные переходы между картографическими способами изображения: на примере картограмм и картодиаграмм // *Цифровая география : Материалы II Всероссийской научно-практической конференции с международным участием, Пермь, 25-28 сентября 2024 г.* — Пермь: Пермский государственный национальный исследовательский университет, 2024. — С. 140-143.
15. **Самсонов Т.Е.**, Ужегов М.В., Энтин А.Л. Оценка точности воспроизведения основных направлений стока в растровых моделях низкого пространственного разрешения // *Геодезия, картография, геоинформатика и кадастры. Инновации в науке, образовании и производстве: Материалы V Всероссийской науч.-практ. конф. 2024 г., Санкт-Петербург* / науч. ред. И. Е Сидорина. — Санкт-Петербург: Издательско-полиграфическая ассоциация высших учебных заведений, 2024. — С. 537–545.
16. **Samsonov T.**, Krivosheina A. Joint generalization of city points and road network for small-scale mapping // *Proceedings of Seventh International Conference on Geographic Information Science GIScience, September 18-21, 2012.* — Columbus, Ohio: 2012. — P. 1–7.
17. Raposo P., **Samsonov T.** Towards general theory of raster data generalization // *Proceedings of 17th ICA Workshop on Generalisation and Multiple Representation, Vienna, Austria, 23th September, 2014.* — 2014. — P. 1–10.
18. **Samsonov T. E.**, Konstantinov P. I. OpenStreetMap data assessment for extraction of urban land cover and geometry parameters required by urban climate modeling // *Extended Abstract Proceedings of the GIScience 2014, September 23-26, Vienna, Austria.* — Vol. 40 of GeoInfo Series. — Hochschülerschaft, TU Vienna Vienna, Austria: 2014. — P. 395–399.

19. **Samsonov T. E.**, Yakimova O. P. Geometric simplification of administrative borders with mixture of irregular and orthogonal segments // Proceedings of 19th ICA Workshop on Generalisation and Multiple Representation, 14 June 2016. — Helsinki, Finland: 2016. — P. 1–12.
20. **Samsonov T.**, Trigub K. Towards computation of urban local climate zones (lcz) from openstreetmap data // Proceedings of the 14th International Conference on GeoComputation, 4th-7th September 2017. — Leeds, UK: 2017. — P. 1–9.

## Список литературы

1. **Баранский Н. Н.** *Экономическая география*. — 2-е изд. — Москва: Географиздат, 1960. — 367 с.
2. **Берлянт А. М.** *Образ пространства: карта и информация*. — Москва: Мысль, 1986. — 237 с.
3. **Берлянт А. М.** *Картографический метод исследования*. — 2-е изд. — Москва: Издательство МГУ, 1988. — 251 с.
4. **Берлянт А. М.** *Теория геоизображений*. — Москва: ГЕОС, 2006. — 262 с.
5. **Берлянт А. М.** *Картография*. — 4-е изд. — Москва: КДУ, 2014. — 448 с.
6. **Берлянт А. М., Мусин О. Р., Собчук Т. В.** *Картографическая генерализация и теория фракталов*. — Москва: Издательство МГУ, 1998. — 136 с.
7. **Берлянт А. М., Ушакова Л. А.** *Картографические анимации*. — Москва: Научный мир, 2000. — 108 с.
8. **Бородин А. В.** К вопросу об отборе населенных пунктов на общегеографических картах. // *Сборник научно-технических и производственных статей. Вып. 18*. — Москва: Геодезиздат, 1948.
9. **Бугаевский Л. М., Подольская Е. С.** Теоретические аспекты картографической генерализации населенных пунктов при составлении обзорно-топографических карт // *Известия высших учебных заведений. Геодезия и аэрофотосъемка*. — 2004. — № 2. — С. 67–79.
10. **Бугаевский Л. М., Подольская Е. С.** Картографическая генерализация дорожной сети при составлении обзорно-топографических карт // *Известия высших учебных заведений. Геодезия и аэрофотосъемка*. — 2005. — № 1. — С. 84–97.
11. **Будыко М. И.** *Тепловой баланс земной поверхности*. — Ленинград: Гидрометеоздат, 1956. — 255 с.

12. **Варенцов М. И., Самсонов Т. Е., Кислов А. В., Константинов П. И.** Воспроизведение острова тепла Московской агломерации в рамках региональной климатической модели COSMO-CLM // *Вестник Московского университета. Серия 5: География.* — 2017. — № 6. — С. 25–37.
13. **Верещака Т. В.** *Топографические карты: научные основы содержания.* — Москва: МАИК Наука/Интерпериодика, 2002. — 319 с.
14. **Воробьев А. А.** Политранспортные магистрали // *Вестник Московского университета. Серия 5. География.* — 1973. — № 3. — С. 88–91.
15. **Гарцман Б. И., Шекман Е. А., Ли К. Т.** Порядковая классификация речных водоразделов на основе обработки цифровых моделей рельефа // *География и природные ресурсы.* — 2016. — № 4.
16. **Геоинформатика.** *Толковый словарь основных терминов* / под ред. А. М. Берлянта, А. В. Кошкарёва. — Москва: ГИС-Ассоциация, 1999. — 204 с.
17. **Говоров М. О., Хорев А. Г.** Об интегрировании различных обобщенных картографических данных в объектно-ориентированном мультимасштабном ГИС (на примере геоморфологических карт) // *Труды международной конференции INTERCARTO ГИС для экологического развития.* — Новосибирск: СГТА, 1997. — С. 272–282.
18. **Гольдман Л. М.** *Альбом образцов изображения рельефа на топографических картах.* — Москва: Главное управление геодезии и картографии при совете министров СССР, 1968. — 59 с.
19. **ГОСТ Р 10032-2007.** *Эталонная модель управления данными.* — Москва: Стандартинформ, 2009.
20. **ГОСТ Р 51608-2024.** *Карты цифровые топографические. Требования к качеству.* — Москва: Российский институт стандартизации, 2024.
21. **ГОСТ Р 52438-2005.** *Географические информационные системы. Термины и определения.* — Москва: Стандартинформ, 2005.
22. **ГОСТ Р 57657-2017.** *Пространственные данные. Спецификация информационного продукта.* — Москва: Стандартинформ, 2018.
23. **ГОСТ Р 70316-2022.** *Пространственные данные. Пространственная привязка по географическим идентификаторам.* — Москва: Российский институт стандартизации, 2022.

24. **ГОСТ Р 70846.4-2023.** *Национальная система пространственных данных. Система классификации и кодирования. Разработка и применение систем классификации и кодирования пространственных данных. Общие требования.* — Москва: Российский институт стандартизации, 2023.
25. **Евтеев О. А.** *Проектирование и составление социально-экономических карт.* — Москва: Издательство МГУ, 1999. — 224 с.
26. **Заруцкая И. П.** *Методы составления рельефа на гипсометрических картах.* — Москва: Издательство геодезической литературы, 1958. — 216 с.
27. **Заруцкая И. П., Сваткова Т. Г.** *Проектирование и составление карт. Общегеографические карты.* — Москва: Издательство МГУ, 1982. — 238 с.
28. **Иванов А. Г., Булыгина О. А.** Автоматизация процессов выбора способа изображения картографируемых объектов и явлений // *Геодезия и картография.* — 2012. — № 10. — С. 27–32.
29. **Иванов В. В.** О программировании отбора населенных пунктов на топографических картах // *Геодезия и картография.* — 1964. — № 2. — С. 52–63.
30. **Карташов Г. А., Самсонов Т. Е., Мысленков С. А.** Генерализация триангуляционных моделей характеристик ветрового волнения // *Материалы V Всероссийской научно-практической конференции "Геодезия, картография, геоинформатика и кадастры. Инновации в науке, образовании и производстве".* — Санкт-Петербург: Издательско-полиграфическая ассоциация высших учебных заведений, 2024. — С. 522–527.
31. **Кравцова В. И.** *Генерализация аэрокосмического изображения: континуальные и дискретные снимки.* : Изд-во Моск. ун-та М, 2000. — 256 с.
32. **Лурье И. К.** *Геоинформационное картографирование. Методы геоинформатики и цифровой обработки космических снимков.* — 3-е изд. — Москва: Книжный дом Университет, 2016. — 424 с.
33. **Лурье И. К., Самсонов Т. Е.** Структура и содержание базы пространственных данных для мультимасштабного картографирования // *Геодезия и картография.* — 2010. — № 11. — С. 17–23.
34. **Лурье И. К., Самсонов Т. Е.** Развитие геоинформационного образования на географическом факультете МГУ им. М. В. Ломоносова // *ArcReview.* — 2011. — № 3. — С. 12–13.
35. **Михайлов В. Н., Добролюбов С. А.** *Гидрология : учебник для вузов.* — Москва, Берлин: Директ-Медиа, 2017. — 752 с.

36. **Мухаметшин Т. Е., Самсонов Т. Е., Лурье И. К.** Координатная привязка геоизображений с использованием компьютерного зрения // *Геодезия и картография*. — 2024. — Т. 1007, № 5. — С. 24–36.
37. **Мысленков С. А., Самсонов Т. Е., Киселева С. В., Архипкин В. С., Шестакова А. А., Умнов П. М.** Разработка веб-атласа доступной энергии волн и ветра прибрежной зоны морей России: информационное и картографическое обеспечение // *Международный научный журнал Альтернативная энергетика и экология*. — 2018. — № 16–18. — С. 39–54.
38. **Мысленков С. А., Самсонов Т. Е., Шурыгина А. А., Киселева С. В., Архипкин В. С., Шестакова А. А., Суркова Г. В., Сильвестрова К. П., Умнов П. М.** Разработка веб-атласа доступной энергии волн и ветра прибрежной зоны морей России: веб-интерфейс и аналитические функции // *Международный научный журнал Альтернативная энергетика и экология*. — 2020. — № 7–18. — С. 44–59.
39. **Подольский А. С., Самсонов Т. Е.** Принципы оформления мультимасштабных общегеографических карт // *ИнтерКарто-ИнтерГИС — 18: Устойчивое развитие территорий: теория ГИС и практический опыт. Материалы международной конференции, 26 июня – 4 июля 2012 г / под ред. С. П. Евдокимова*. — Смоленск: СмолГУ, 2012. — С. 71–81.
40. **Полян П. М.** *Методика выделения и анализа опорного каркаса населения*. — Москва: ИГ АН, 1988. — 220 с.
41. **Полян П. М.** *Территориальные структуры – урбанизация – расселение: теоретические подходы и методы изучения*. — Москва: Новый хронограф, 2014. — 783 с.
42. **Прохорова Е. А.** *Социально-экономические карты*. — Москва: КДУ, 2010. — 424 с.
43. **Пузаченко Ю. Г.** Приложение теории фракталов к изучению структуры ландшафта // *Известия РАН. Серия географическая*. — 1997. — № 2. — С. 24–40.
44. **Пузаченко Ю. Г., Онуфреня И. А., Алещенко Г. М.** Количественные методы классификации форм рельефа // *Известия РАН. Серия географическая*. — 2002. — № 6. — С. 17–25.
45. **Россия.** *Общегеографическая карта масштаба 1:8 000 000*. — Омск: Омская картографическая фабрика, 2012.
46. **Росстат.** *База данных показателей муниципальных образований* [Электронный ресурс]. URL: <https://rosstat.gov.ru/storage/mediabank/munst.htm> (дата обращения: 16.02.2025).

47. **Руководство.** *По картографическим и картоиздательским работам. Часть 1. Составление и подготовка к изданию топографических карт масштабов 1:25 000, 1:50 000, 1:100 000.* — Москва: РИОВТС, 1978. — 131 с.
48. **Руководство.** *По картографическим и картоиздательским работам. Часть 2. Составление и подготовка к изданию топографических карт масштабов 1:200 000, 1:500 000.* — Москва: РИОВТС, 1980. — 167 с.
49. **Руководство.** *По картографическим и картоиздательским работам. Часть 3. Составление и подготовка к изданию топографических карт масштаба 1:1 000 000.* — Москва: РИОВТС, 1985. — 148 с.
50. **Салищев К. А.** *Основы картоведения: общая часть.* — Москва: Редбюро ГУГК при СНК СССР, 1939. — 308 с.
51. **Салищев К. А.** *Картография.* — Москва: Высшая школа, 1966. — 214 с.
52. **Салищев К. А.** Генерализация в ее истории и современном развитии // *Итоги науки и техники. Сер. Картография. т. 5.* — Москва: ВИНТИ, 1972. — С. 6–23.
53. **Салищев К. А.** *Картоведение.* — 1-е изд. — Москва: Издательство МГУ, 1976. — 438 с.
54. **Салищев К. А.** *Картоведение.* — 3-е изд. — Москва: Издательство МГУ, 1990. — 400 с.
55. **Салищев К. А., Сухов В. И., Филиппов Ю. В.** *Составление и редактирование карт. Часть 1.* — Москва: Геозедиздат, 1947. — 192 с.
56. **Самсонов Т. Е.** *Мультимасштабные базы геоданных // Материалы конференции Геоинформационное картографирование в регионах России, 2-4 декабря 2009 года.* — Воронеж: ВГУ, 2009б. — С. 1–6.
57. **Самсонов Т. Е.** *Мультимасштабные базы геоданных для электронных карт // Пространственные данные.* — 2009а. — № 4. — С. 46–51.
58. **Самсонов Т. Е.** *Мультимасштабное картографирование рельефа: общегеографические и гипсометрические карты.* — Saarbrücken: LAP Lambert Academic Publishing, 2011. — 208 с.
59. **Самсонов Т. Е.** *Мультимасштабное картографирование — новое направление картографии // Современная географическая картография / под ред. И. К. Лурье, В. И. Кравцовой.* — Москва: Дата+, 2012. — С. 21–35.

60. **Самсонов Т. Е.** Анимационные переходы между картографическими способами изображения: на примере картограмм и картодиаграмм // *Материалы II Всероссийской научно-практической конференции с международным участием «Цифровая география», 25-28 сентября 2024 г.* — Пермь: Пермский государственный национальный исследовательский университет, 2024. — С. 141–144.
61. **Самсонов Т. Е., Кривошеина А. М.** Генерализация дорожной сети в мелких масштабах картографирования с сохранением связности населенных пунктов // *Известия высших учебных заведений. Геодезия и аэрофотосъемка.* — 2015b. — № 2. — С. 23–30.
62. **Самсонов Т. Е., Кривошеина А. М.** Автоматизация отбора населенных пунктов с учетом пространственной неравномерности их распределения для целей мелко-масштабного картографирования // *Известия высших учебных заведений. Геодезия и аэрофотосъемка.* — 2015a. — № 1. — С. 74–82.
63. **Самсонов Т. Е., Подольский А. С.** Автоматизация вычисления параметров отображения дорожной сети и структуры населенных пунктов на мультимасштабных картах // *Геодезия и картография.* — 2014. — № 9. — С. 21–28.
64. **Самсонов Т. Е., Подольский А. С., Юрова Н. Д.** Преобразования способов изображения и их соотношение с иерархией единиц картографирования // *Вестник Московского университета. Сер.5 География.* — 2013. — № 4. — С. 15–23.
65. **Самсонов Т. Е., Прохорова Е. А.** Генерализация транспортных сетей при мультимасштабном картографировании // *Известия Российской академии наук. Серия географическая.* — 2019. — № 1. — С. 107–117.
66. **Самсонов Т. Е., Тригуб К. С.** Картографирование локальных климатических зон Москвы по космическим снимкам // *Геодезия и картография.* — 2018. — Т. 79, № 6. — С. 14–25.
67. **Самсонов Т. Е., Юрова Н. Д.** Применение картограмм и картодиаграмм на мультимасштабных социально-экономических картах // *Геодезия и картография.* — 2014. — № 11. — С. 30–38.
68. **Самсонов Т. Е., Якимова О. П.** Анализ эффективности ограничивающих метрик алгоритмов геометрического упрощения на основе экспертной оценки детализации линий // *Интеркарто. ИнтерГИС.* — 2021. — Т. 27, № 2. — С. 253–267.
69. **Самсонов Т. Е., Якимова О. П., Алексеев В. В., Богаевская В. Г., Горохов А. А., Князев В. Н., Преображенская М. М., Ухалов А. Ю., Эдельсбруннер Х.** Алгоритм геометрического упрощения множества линий путем стягивания ребер графа с сохранением топологии // *Геодезия и картография.* — 2014. — № 3. — С. 29–36.

70. **Симонов Ю. Г.** *Морфометрический анализ рельефа.* — Москва-Смоленск: Изд-во Смоленского гос. ун-та, 1998. — 272 с.
71. **Скворцов А. В.** *Триангуляция Делоне и ее применение.* — Томск: Изд-во Томск. ун-та, 2002. — 128 с.
72. **Сухов В. И.** *Отбор и классификация путей сообщения.* — Москва: Геодезиздат, 1945. — 105 с.
73. **Сухов В. И.** *Составление и редактирование общегеографических карт.* — Москва: Геодезиздат, 1957. — 280 с.
74. **Тархов С. А.** *Эволюционная морфология транспортных сетей.* — Смоленск: Универсум, 2005. — 386 с.
75. **Транспорт и его влияние на природную среду.** Карта масштаба 1:60 000 000 // *Атлас "Природа и ресурсы Земли". Часть II.* — Вена—Москва: Ед Хёлцел, 1998. — С. 166.
76. **Трейвиш А. И.** Принцип полимасштабности в географии и страноведении // *Меняющаяся география зарубежного мира. Вопросы экономической и политической географии зарубежных стран.* / под ред. И. С. Ивановой, И. М. Кузиной, А. С. Фетисова. — Москва—Смоленск: Ойкумена, 2007. — С. 50–65.
77. **Флоринский И. В.** *Генерализация в картографии: краткий обзор проблемы.* — Пушкино: Пушкин. науч. центр, Ин-т почвоведения и фотосинтеза, 1991. — 54 с.
78. **Флоринский И. В.** Анализ планетарного рельефа Марса, Венеры и Луны по данным миссий Mars Global Surveyor, Magellan и Clementine // *Исследование Земли из космоса.* — 2009. — Т. 5. — С. 32–48.
79. **Хорошев А. В.** *Полимасштабная организация географического ландшафта.* — Москва: Товарищество научных изданий КМК, 2016. — 417 с.
80. **Юрова Н. Д., Самсонов Т. Е.** Мультимасштабное социально-экономическое картографирование на примере производства строительных материалов в России // *ИнтерКарто-ИнтерГИС — 18: Устойчивое развитие территорий: теория ГИС и практический опыт. Материалы международной конференции.* — Смоленск: СмолГУ, 2012. — С. 469–476.
81. **Якимова О. П., Эдельсбруннер Х., Алексеев В. В., Богаевская В. Г., Горохов А. А., Преображенская М. М., Князев В. Н.** *GenMap – научно-исследовательская программа для картографической генерализации.* — Свидетельство о регистрации: №2013618052 от 29.08.2013, 2013.

82. **Якимова О., Самсонов Т., Потемкин Д., Усманова Э.** Инструменты для оценки детализации карт для геоинформационной системы QGIS // *Интеркарто. Интер-ГИС*. — 2021. — Т. 27, № 2. — С. 260–279.
83. **Abler R. F.** The National Science Foundation National Center for Geographic Information and Analysis // *International journal of geographical information systems*. — 1987. — Vol. 1, No. 4. — P. 303–326.
84. **Adams B., Li L., Raubal M., Goodchild M. F.** A General Framework for Conflation // *Proceedings of GIScience 2010*. — Zurich, Switzerland: University of Zurich, 2010. — P. 1–5.
85. **Ai T.** The drainage network extraction from contour lines for contour line generalization // *ISPRS Journal of Photogrammetry and Remote Sensing*. — 2007. — Vol. 62, No. 2. — P. 93–103.
86. **Ai T., Ke S., Yang M., Li J.** Envelope generation and simplification of polylines using Delaunay triangulation // *International Journal of Geographical Information Science*. — 2017. — Vol. 31, No. 2. — P. 297–319.
87. **Ai T., Li J.** A DEM generalization by minor valley branch detection and grid filling // *ISPRS Journal of Photogrammetry and Remote Sensing*. — 2010. — Vol. 65, No. 2. — P. 198–207.
88. **Ai T., Liu Y. A.** A method of point cluster simplification with spatial distribution properties preserved // *Acta Geodaetica et Cartographica Sinica*. — 2002. — Vol. 31, No. 2. — P. 175–181.
89. **Alt H., Godau M.** Computing the Fréchet Distance Between Two Polygonal Curves // *International Journal of Computational Geometry & Applications*. — 1995. — Vol. 05. — P. 75–91.
90. **Arinaga R. A., Cheung K. F.** Atlas of global wave energy from 10 years of reanalysis and hindcast data // *Renewable Energy*. — 2012. — Vol. 39, No. 1. — P. 49–64.
91. **Arlati A.** Navigating urban futures: Exploring NbS upscaling discourses, practices, and relations in reimagining human-nature relationships // *Futures*. — 2024. — Vol. 161, No. 103403. — P. 1–16.
92. **Arrell K., Carver S.** Surface Roughness Scaling Trends // *Proceedings of Geomorphometry 2009. Zurich, Switzerland, 31 August – 2 September, 2009*. — Zurich, Switzerland: The International Society for Geomorphometry, 2009. — P. 120–123.

93. **Arundel S. T., Thiem P. T., Constance E. W.** Automated extraction of hydrographically corrected contours for the conterminous United States: The US Geological Survey US Topo product // *Cartography and Geographic Information Science*. — 2018. — Vol. 45, No. 1. — P. 31–55.
94. **Assunção R. M., Neves M. C., Câmara G., Da Costa Freitas C.** Efficient regionalization techniques for socio-economic geographical units using minimum spanning trees // *International Journal of Geographical Information Science*. — 2006. — Vol. 20, No. 7. — P. 797–811.
95. **Baklanov A., Grimmond C. S. B., Carlson D., Terblanche D., Tang X., Bouchet V., Lee B., Langendijk G., Kolli R. K., Hovsepyan A.** From urban meteorology, climate and environment research to integrated city services // *Urban Climate*. — 2018. — Vol. 23. — P. 330–341.
96. **Balzer M., Deussen O.** Voronoi Treemaps // *IEEE Symposium on Information Visualization, 2005. INFOVIS 2005*. — Minneapolis, MN, USA: IEEE, 2005. — P. 49–56.
97. **Band L. E.** Topographic Partition of Watersheds with Digital Elevation Models // *Water Resources Research*. — 1986. — Vol. 20, No. 1. — P. 15–24.
98. **Baumann P., Hirschorn E., Masó J.** *OGC Coverage Implementation Schema*. — OGC: Open Geospatial Consortium, 2017. — 77 p.
99. **Bayer T.** Automated Building Simplification Using a Recursive Approach // *Advances in Cartography and GIScience. Volume 1*. — Berlin, Heidelberg: Springer, 2009. — P. 121–146.
100. **Beard K.** Constraints on Rule Formation // *Map Generalisation: Making Rules for Knowledge Representation*. — UK: Longman Scientific & Technical, 1991. — P. 121–135.
101. **Bechtel B., Alexander P., Böhner J., Ching J., Conrad O., Feddema J., Mills G., See L., Stewart I.** Mapping Local Climate Zones for a Worldwide Database of the Form and Function of Cities // *ISPRS International Journal of Geo-Information*. — 2015. — Vol. 4, No. 1. — P. 199–219.
102. **Becker-Reshef I., Vermote E., Lindeman M., Justice C.** A generalized regression-based model for forecasting winter wheat yields in Kansas and Ukraine using MODIS data // *Remote Sensing of Environment*. — 2010. — Vol. 114, No. 6. — P. 1312–1323.
103. **Beconytė G., Balčiūnas A., Šturaitė A., Viliuvienė R.** Where Maps Lie: Visualization of Perceptual Fallacy in Choropleth Maps at Different Levels of Aggregation // *ISPRS International Journal of Geo-Information*. — 2022. — Vol. 11. — P. 1–13.

104. **Bedard Y., Bernier E.** Supporting multiple representations with spatial databases views management and the concept of VUEL // *Proceedings of ISPRS/ICA Joint Workshop on Multi-Scale Representations of Spatial Data*. — Ottawa: Canada, 2002. — P. 1–14.
105. **Benz S. A., Weibel R.** Road network selection for medium scales using an extended stroke-mesh combination algorithm // *Cartography and Geographic Information Science*. — 2014. — Vol. 41, No. 4. — P. 323–339.
106. **Bereuter P., Weibel R.** Real-time generalization of point data in mobile and web mapping using quadtrees // *Cartography and Geographic Information Science*. — 2013. — Vol. 40, No. 4. — P. 271–281.
107. **Bereuter P., Weibel R.** Variable-scale maps in real-time generalisation using a quadtree data structure and space deforming algorithms // *International Journal of Cartography*. — 2017. — Vol. 3, No. 1. — P. 134–147.
108. **Berg M. de, Kreveld M. van, Schirra S.** Topologically Correct Subdivision Simplification Using the Bandwidth Criterion // *Cartography and Geographic Information Science*. — 1998. — Vol. 25, No. 4. — P. 243–257.
109. **Bergman C., Oksanen J.** Conflation of OpenStreetMap and Mobile Sports Tracking Data for Automatic Bicycle Routing // *Transactions in GIS*. — 2016. — Vol. 20, No. 6. — P. 848–868.
110. **Bernardin T., Cowgill E., Kreylos O., Bowles C., Gold P., Hamann B., Kellogg L.** Crusta: A new virtual globe for real-time visualization of sub-meter digital topography at planetary scales // *Computers and Geosciences*. — 2011. — Vol. 37, No. 1. — P. 75–85.
111. **Bertin J.** *Semiology of Graphics: Diagrams, Networks, Maps*. — Redlands, CA: ESRI Press, 2010. — 456 p.
112. **Biljecki F., Ledoux H., Stoter J., Zhao J.** Formalisation of the Level of Detail in 3D City Modelling // *Computers, Environment and Urban Systems*. — 2014. — Vol. 48. — P. 1–15.
113. **Bingölbali B., Jafali H., Akpınar A., Bekiroğlu S.** Wave energy potential and variability for the south west coasts of the Black Sea: The WEB-based wave energy atlas // *Renewable Energy*. — 2020. — Vol. 154. — P. 136–150.
114. **Bolton T., Zanna L.** Applications of Deep Learning to Ocean Data Inference and Subgrid Parameterization // *Journal of Advances in Modeling Earth Systems*. — 2019. — Vol. 11, No. 1. — P. 376–399.
115. **Bonneau G. P.** Multiresolution analysis on irregular surface meshes // *IEEE Transactions on Visualisation and Computer Graphics*. — 1998. — Vol. 4, No. 4. — P. 365–378.

116. **Borges K. A. V., Laender A. H. F., Davis C. A.** Spatial data integrity constraints in object oriented geographic data modeling // *Proceedings of the seventh ACM international symposium on Advances in geographic information systems - GIS '99*. — Kansas City, Missouri, United States: ACM Press, 1999. — P. 1–6.
117. **Bostock M.** *Observable* [Электронный ресурс]. URL: <https://observablehq.com/product> (дата обращения: 10.02.2025).
118. **Bouvier M., Durrieu S., Fournier R. A., Renaud J.-P.** Generalizing predictive models of forest inventory attributes using an area-based approach with airborne LiDAR data // *Remote Sensing of Environment*. — 2015. — Vol. 156. — P. 322–334.
119. **Brassel K. E., Weibel R.** A review and conceptual framework of automated map generalization // *International Journal of Geographical Information Systems*. — 1988. — Vol. 2, No. 3. — P. 229–244.
120. **Brewer C. A.** *Designing Better Maps: A Guide for GIS Users*. — 2 ed. — Redlands, California: Esri Press, 2015. — 250 p.
121. **Brewer C. A., Akella M. K.** Multi-resolution Multi-scale Topographic Map Design: Toward a New Look for The National Map // *Proceedings of AutoCarto 2008*. — Shepherdstown, West Virginia: USA, 2008. — P. 1–12.
122. **Brewer C. A., Buttenfield B. P.** Framing Guidelines for Multi-Scale Map Design Using Databases at Multiple Resolutions // *Cartography and Geographic Information Science*. — 2007. — Vol. 34, No. 1. — P. 3–15.
123. **Brewer C. A., Buttenfield B. P.** Mastering map scale: Balancing workloads using display and geometry change in multi-scale mapping // *GeoInformatica*. — 2009. — Vol. 14, No. 2. — P. 221–239.
124. **Brewer C. A., Marlow K. A.** Color representation of aspect and slope simultaneously // *Proceedings of Eleventh International Symposium on Computer-Assisted Cartography (Auto-Carto-11)*. — Minneapolis, Minnesota: ASPRS, ACSM, 1993. — P. 328–337.
125. **Brewer C. A., Stanislawski L. V., Buttenfield B. P., Sparks K. A., McGilloway J., Howard M. A.** Automated thinning of road networks and road labels for multiscale design of The National Map of the United States // *Cartography and Geographic Information Science*. — 2013. — Vol. 40, No. 4. — P. 259–270.
126. **Brown L. G.** A survey of image registration techniques // *ACM Computing Surveys*. — 1992. — Vol. 24, No. 4. — P. 325–376.

127. **Brown W. G., Cosh M. H., Dong J., Ochsner T. E.** Upscaling soil moisture from point scale to field scale: Toward a general model // *Vadose Zone Journal*. — 2023. — Vol. 22 (2), No. e20244. — P. 1–14.
128. **Buchhorn M., Lesiv M., Tsendbazar N.-E., Herold M., Bertels L., Smets B.** Copernicus Global Land Cover Layers—Collection 2 // *Remote Sensing*. — 2020. — Vol. 12, No. 6. — P. 1044.
129. **Buchin K., Meulemans W., Renssen A. V., Speckmann B.** Area-Preserving Simplification and Schematization of Polygonal Subdivisions // *ACM Transactions on Spatial Algorithms and Systems*. — 2016. — Vol. 2, No. 1. — P. 1–36.
130. **Buchin K., Meulemans W., Speckmann B.** A New Method for Subdivision Simplification with Applications to Urban-area Generalization // *Proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. — New York, NY, USA: ACM, 2011. — P. 261–270.
131. **Buchin K., Sousa M. C., Döllner J., Samavati F., Walther M.** Illustrating Terrains using Direction of Slope and Lighting // *Proc. 4th ICA Mountain Cartography Workshop*. — Vall de Núria, Catalonia, Spain: ICC (Institut Cartogràfic de Catalunya), 2004. — P. 1–11.
132. **Burghardt D., Duchene C., Mackaness W.** *Abstracting Geographic Information in a Data Rich World: Methodologies and Applications of Map Generalisation* / ed. by D. Burghardt, C. Duchene, W. Mackaness. — Berlin, Heidelberg: Springer, 2014. — 450 p.
133. **Burt J. E., White J., Allord G., Then K. M., Zhu A.-X.** Automated and semi-automated map georeferencing // *Cartography and Geographic Information Science*. — 2020. — Vol. 47, No. 1. — P. 46–66.
134. **Buttenfield B. P.** Automating the Identification of Cartographic Lines // *Cartography and Geographic Information Science*. — 1987. — Vol. 14, No. 1. — P. 7–20.
135. **Buttenfield B. P.** A rule for describing line feature geometry // *Map Generalization: Making Rules for Knowledge Representation* / ed. by B. P. Buttenfield, R. B. McMaster. — London: Longman Scientific and Technical, 1991. — P. 150–171.
136. **Buttenfield B. P.** *Research Initiative 3: Multiple Representations. Closing Report*. — Buffalo: National Center for Geographic Information and Analysis, 1993.
137. **Buttenfield B. P., McMaster R. B.** *Map Generalization: Making rules for knowledge representation*. — New York: John Wiley & Sons, Inc, 1991. — 246 p.

138. **Buttenfield B. P., Stanislawski L. V., Brewer C. A.** Adapting Generalization Tools to Physiographic Diversity for the United States National Hydrography Dataset // *Cartography and Geographic Information Science*. — 2013. — Vol. 38, No. 3. — P. 289–301.
139. **Buys J., Messerschmidt H. J., Botha J. F.** Including known discontinuities directly into a triangular irregular mesh for automatic contouring purposes // *Computers and Geosciences*. — 1991. — Vol. 17, No. 7. — P. 875–881.
140. **Cámara M., López F.** Mathematical Morphology Applied to Raster Generalization of Urban City Block Maps // *Cartographica: The International Journal for Geographic Information and Geovisualization*. — 2000. — Vol. 37, No. 1. — P. 33–48.
141. **Carmichael T., Ailleres L.** Method and analysis for the upscaling of structural data // *Journal of Structural Geology*. — 2016. — Vol. 83. — P. 121–133.
142. **Cartographic generalization. Topographic maps.** — Switzerland: Swiss Society of Cartography, 1987. — 61 p.
143. **Cavaleri L., Alves J.-H. G. M., Arduin F., Babanin A., Banner M., Belibassakis K., Benoit M., Donelan M., Groeneweg J., Herbers T. H. C., Hwang P., Janssen P. A. E. M., Janssen T., Lavrenov I. V., Magne R., Monbaliu J., Onorato M., Polnikov V., Resio D., Rogers W. E., Sheremet A., McKee Smith J., Tolman H. L., Van Vledder G., Wolf J., Young I.** Wave modelling – The state of the art // *Progress in Oceanography*. — 2007. — Vol. 75, No. 4. — P. 603–674.
144. **Cecconi A., Weibel R., Barrault M.** Improving Automated Generalisation for On-Demand Web Mapping by Multiscale Databases // *Advances in Spatial Data Handling* / ed. by D. E. Richardson, P. van Oosterom. — Berlin, Heidelberg: Springer, 2002. — P. 515–531.
145. **Cetinkaya B., Aslan S., Sengun Y. S., Cobankaya O. N., Ilgin D. E.** Contour Simplification with Defined Spatial Accuracy // *Workshop of the ICA Commission on Map Generalisation and Multiple Representation*. — Portland: USA, 2006. — P. 1–7.
146. **Chambers E., Fasy B. T., Wang Y., Wenk C.** Map-Matching Using Shortest Paths // *ACM Transactions on Spatial Algorithms and Systems*. — 2020. — Vol. 6, No. 1. — P. 1–17.
147. **Chavent M., Kuentz-Simonet V., Labenne A., Saracco J.** ClustGeo: An R package for hierarchical clustering with spatial constraints // *Computational Statistics*. — 2018. — Vol. 33, No. 4. — P. 1799–1822.
148. **Chen C.-C., Knoblock C. A., Shahabi C.** Automatically Conflating Road Vector Data with Orthoimagery // *GeoInformatica*. — 2006. — Vol. 10, No. 4. — P. 495–530.

149. **Chen He X., Yung N. H. C.** Corner detector based on global and local curvature properties // *Optical Engineering*. — 2008. — Vol. 47. — P. 1–12.
150. **Chen J., Hu Y., Li Z., Zhao R., Meng L.** Selective omission of road features based on mesh density for automatic map generalization // *International Journal of Geographical Information Science*. — 2009. — Vol. 23, No. 8. — P. 1013–1032.
151. **Chen Y., Wilson J. P., Zhu Q., Zhou Q.** Comparison of drainage-constrained methods for DEM generalization // *Computers & Geosciences*. — 2012. — Vol. 48. — P. 41–49.
152. **Chen Z.** Contour Generalization by a 3-Dimensional Spatial Low-Pass Filtering // *Proceedings of GIS'87, Second International Conference on Geographic Information Systems*. — San Francisco, USA: USA, 1987. — P. 375–386.
153. **Chen Z., Guevara J. A.** Systematic selection of very important points (VIP) from digital elevation models for constructing triangular irregular networks // *Proceedings of the International Symposium on Computer-Assisted Cartography Auto-Carto VIII*. — Baltimore, Maryland: USA, 1987. — P. 475–484.
154. **Chen Z., Pu H., Wang B., Jiang G.-M.** Fusion of Hyperspectral and Multispectral Images: A Novel Framework Based on Generalization of Pan-Sharpening Methods // *IEEE Geoscience and Remote Sensing Letters*. — 2014. — Vol. 11, No. 8. — P. 1418–1422.
155. **Chen Z., Tobler W.** Quadtree representation of digital terrain // *Proceedings AutoCarto*. — London: England, 1986. — P. 475–484.
156. **Cheng G.** *Hierarchy Representation of Virtual Terrain Environment and Research into the Real Time Shading Technology*: Ph.D. thesis (in Chinese). — Zhengzhou, China: Zhengzhou Institute for Mapping and Surveying, 2000. — 133 p.
157. **Cheng X., Liu Z., Zhang Q.** MSLF: Multi-scale legibility function to estimate the legible scale of individual line features // *Cartography and Geographic Information Science*. — 2021. — Vol. 48, No. 2. — P. 151–168.
158. **Cheng X., Wu H., Ai T., Yang M.** Detail Resolution: A New Model to Describe Level of Detail Information of Vector Line Data // *Spatial Data Handling in Big Data Era. Advances in Geographic Information Science*. / ed. by C. Zhou, F. Su, F. Harvey, J. Xu. — Singapore: Springer Singapore, 2017. — P. 167–177.
159. **Cheng Y., Zhao L., Chakraborty T., Oleson K., Demuzere M., Liu X., Che Y., Liao W., Zhou Y., Li X.** U-Surf: A Global 1 km spatially continuous urban surface property dataset for kilometer-scale urban-resolving Earth system modeling // *Earth System Science Data Discussions [preprint]*. — 2024.

160. **Chimani M., Van Dijk T. C., Haunert J.-H.** How to eat a graph: Computing selection sequences for the continuous generalization of road networks // *Proceedings of the 22nd ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*. — Dallas Texas: ACM, 2014. — P. 243–252.
161. **Ching J., Mills G., See L., Alexander P., Bechtel B., Feddema J., Oleson K. L., Stewart I., Neophytou M., Chen F., Wang X., Hanna A.** Wudapt : Facilitating Advanced Urban Canopy Modeling for Weather , Climate and Air Quality Applications // *American Meteorological Society Symposium on Urban Environment, 2-7 February 2014*. — Atlanta, GA: USA, 2014. — P. 1–7.
162. **Choi J., Hwang C.-S.** Multi-scale Rendering with Geometry Collapse and a Symbol Knowledge Base // *The Cartographic Journal*. — 2009. — Vol. 46, No. 2. — P. 155–163.
163. **Christophe S., Mermet S., Laurent M., Touya G.** Neural map style transfer exploration with GANs // *International Journal of Cartography*. — 2022. — Vol. 8, No. 1. — P. 18–36.
164. **Chuvieco E.** *Fundamentals of satellite remote sensing: An environmental approach*. — Third ed. — Boca Raton London New York: CRC Press, 2023. — 415 p.
165. **Clarke K. C.** Scale-Based Simulation of Topographic Relief // *Cartography and Geographic Information Science*. — 1988. — Vol. 15, No. 2. — P. 173–181.
166. **Cléri I., Pierrot-Deseilligny M., Vallet B.** Automatic Georeferencing of a Heritage of old analog aerial Photographs // *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. — 2014. — Vol. II-3. — P. 33–40.
167. **Clifton-Brown J., Hastings A., Mos M., McCalmont J. P., Ashman C., Awty-Carroll D., Cerazy J., Chiang Y., Cosentino S., Cracroft-Eley W., Scurlock J., Donnison I. S., Glover C., Gołąb I., Greef J. M., Gwyn J., Harding G., Hayes C., Helios W., Hsu T., Huang L. S., Jeżowski S., Kim D., Kiesel A., Kotecki A., Krzyzak J., Lewandowski I., Lim S. H., Liu J., Loosely M., Meyer H., MurphyBokern D., Nelson W., Pogrzeba M., Robinson G., Robson P., Rogers C., Scalici G., Schuele H., Shafiei R., Shevchuk O., Schwarz K., Squance M., Swaller T., Thornton J., Truckses T., Botnari V., Vizir I., Wagner M., Warren R., Webster R., Yamada T., Youell S., Xi Q., Zong J., Flavell R.** Progress in upscaling *Miscanthus* biomass production for the European bio-economy with seed-based hybrids // *GCB Bioenergy*. — 2017. — Vol. 9, No. 1. — P. 6–17.
168. **Cobb M. A., Chung M. J., Foley III H., Petry F. E., Shaw K. B., Miller H. V.** A Rule-based Approach for the Conflation of Attributed Vector Data // *Geoinformatica*. — 1998. — Vol. 2, No. 1. — P. 7–35.
169. **Cohen J.** A Coefficient of Agreement for Nominal Scales // *Educational and Psychological Measurement*. — 1960. — Vol. 20, No. 1. — P. 37–46.

170. **Collier P., Forrest D., Pearson A.** The Representation of Topographic Information on Maps: The Depiction of Relief // *The Cartographic Journal*. — 2003. — Vol. 40, No. 1. — P. 17–26.
171. **Creutzig F., Bai X., Khosla R., Viguie V., Yamagata Y.** Systematizing and upscaling urban climate change mitigation // *Environmental Research Letters*. — 2020. — Vol. 15 (10), No. 100202. — P. 1–7.
172. **Creutzig F., Lohrey S., Bai X., Baklanov A., Dawson R., Dhakal S., Lamb W. F., McPhearson T., Minx J., Munoz E., Walsh B.** Upscaling urban data science for global climate solutions // *Global Sustainability*. — 2019. — Vol. 2 (e2). — P. 1–25.
173. **Crow W. T., Berg A. A., Cosh M. H., Loew A., Mohanty B. P., Panciera R., De Rosnay P., Ryu D., Walker J. P.** Upscaling sparse ground-based soil moisture observations for the validation of coarse-resolution satellite soil moisture products // *Reviews of Geophysics*. — 2012. — Vol. 50 (2). — P. 1–20.
174. **Dahinden T., Hurni L.** Development and Quality Assessment of Analytical Rock Drawings // *Proceedings of ICC 2007*. — Moscow: Russia, 2007. — P. 1–15.
175. **Dalang O.** *VectorBender* [Электронный ресурс]. URL: <https://github.com/olivierdalang/VectorBender> (дата обращения: 15.02.2025).
176. **Damen J., Kreveld M. van, Spaan B.** High Quality Building Generalization by Extending the Morphological Operators // *12th ICA Workshop on Generalisation and Multiple Representation, 20-21 June 2008*. — Montpellier, France: LIRMM, 2008. — P. 1–12.
177. **Danielson J. J., Gesch D. B.** *Global multi-resolution terrain elevation data 2010 (GMTED2010)*. : USGS, 2011.
178. **Daubechies I.** *Ten lectures on wavelets*. — Philadelphia, PA: Society for Industrial and Applied Mathematics, 1992. — 357 p.
179. **Davy P., Crave A.** Upscaling local-scale transport processes in large-scale relief dynamics // *Physics and Chemistry of the Earth, Part A: Solid Earth and Geodesy*. — 2000. — Vol. 25, No. 6–7. — P. 533–541.
180. **DCW.** *Digital Chart of the World (DCW)*. — Fairfax, VA: Defense Mapping Agency, 1992.
181. **De Floriani L.** A pyramidal data structure for triangle-based surface description // *IEEE Computer Graphics and Applications*. — 1989. — Vol. 9, No. 2. — P. 67–78.
182. **De Floriani L., Falcidieno B., Nagy G., Pienovi C.** A hierarchical structure for surface approximation // *Computers & Graphics*. — 1984. — Vol. 8, No. 2. — P. 183–193.

183. **De Floriani L., Puppo E.** A hierarchical triangle-based model for terrain description // *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*. — Berlin, Heidelberg: Springer, 1992. — P. 236–251.
184. **Dey T. K., Edelsbrunner H., Guha S., Nekhayev D. V.** Topology preserving edge contraction // *Publications de l'Institut Mathématique*. — 1999. — Vol. 66. — P. 23–45.
185. **DiBiase D., MacEachren A. M., Krygier J. B., Reeves C.** Animation and the Role of Map Design in Scientific Visualization // *Cartography and Geographic Information Systems*. — 1992. — Vol. 19, No. 4. — P. 201–214.
186. **Dickinson B. W., Gavin H. P.** Parametric Statistical Generalization of Uniform-Hazard Earthquake Ground Motions // *Journal of Structural Engineering*. — 2011. — Vol. 137, No. 3. — P. 410–422.
187. **Dinesh S., Ahmad Fadzil M. H.** Characterization of the Size Distribution of Mountains Extracted from Multiscale Digital Elevation Models // *Journal of Applied Sciences*. — 2007. — Vol. 7, No. 10. — P. 1410–1415.
188. **Dong P., Chen Q.** *LiDAR remote sensing and applications*. — Boca Raton London New York: CRC Press, 2018. — 220 p.
189. **Douglas D. H., Peucker T. K.** Algorithms for the reduction of the number of points required to represent a digitized line or its caricature // *The Canadian Cartographer*. — 1973. — P. 112–122.
190. **Dowson K.** *Towards extracting artistic sketches and maps from digital elevation models*: PhD thesis. — UK: University of Hull, 1994.
191. **Drăguț L., Blaschke T.** Automated classification of landform elements using object-based image analysis // *Geomorphology*. — 2006. — Vol. 81, No. 3–4. — P. 330–344.
192. **Dressler G., Groeneveld J., Hetzer J., Janischewski A., Nolzen H., Rödiger E., Schwarz N., Taubert F., Thober J., Will M., Williams T., Wirth S. B., Müller B.** Upscaling in socio-environmental systems modelling: Current challenges, promising strategies and insights from ecology // *Socio-Environmental Systems Modelling*. — 2022. — Vol. 4, No. 18112. — P. 1–25.
193. **Du J., Wu F., Yin J., Liu C., Gong X.** Polyline simplification based on the artificial neural network with constraints of generalization knowledge // *Cartography and Geographic Information Science*. — 2022. — Vol. 49, No. 4. — P. 313–337.
194. **Dubuisson M.-P., Jain A. K.** A Modified Hausdorff Distance for Object Matching // *Proceedings of 12th International Conference on Pattern Recognition*. — Long Beach, CA: IEEE Comput. Soc. Press, 1994. — P. 566–568.

195. **Duchaineau M., Wolinsky M., Sigeti D. E., Miller M. C., Aldrich C., Mineev-Weinstein M. B.** ROAMing terrain: Real-time Optimally Adapting Meshes // *Proceeding. Visualization '97 (Cat. No. 97CB36155)*. — Phoenix, AZ, USA: IEEE, 1997. — P. 81–88.
196. **Duchêne C., Baella B., Brewer C. A., Burghardt D., Battenfield B. P., Gaffuri J., Käuferle D., Lecordix F., Maugeais E., Nijhuis R., Pla M., Post M., Regnaud N., Stanislawski L. V., Stoter J., Tóth K., Urbanke S., Altena V. van, Wiedemann A.** Generalisation in Practice Within National Mapping Agencies // *Abstracting Geographic Information in a Data Rich World* / ed. by D. Burghardt, C. Duchêne, W. Mackaness. — Cham: Springer International Publishing, 2014. — P. 329–391.
197. **Dumont M., Touya G., Duchêne C.** Alternative transitions between existing representations in multi-scale maps // *Proceedings of the ICA*. — 2018. — Vol. 1. — P. 1–4.
198. **Dumont M., Touya G., Duchêne C.** Designing multi-scale maps: Lessons learned from existing practices // *International Journal of Cartography*. — 2020. — Vol. 6, No. 1. — P. 121–151.
199. **Dundas Data Visualization Inc.** *Dundas Dashboard* [Электронный ресурс]. URL: <https://www.dundas.com/support/> (дата обращения: 15.02.2025).
200. **Duo Z., Wang W., Wang H.** Oceanic Mesoscale Eddy Detection Method Based on Deep Learning // *Remote Sensing*. — 2019. — Vol. 11 (16), No. 1921. — P. 1–14.
201. **Dutton G.** Scale, Sinuosity, and Point Selection in Digital Line Generalization // *Cartography and Geographic Information Science*. — 1999. — Vol. 26, No. 1. — P. 33–54.
202. **Ebisch K.** Short note A correction to the Douglas – Peucker line generalization // *Computers & Geosciences*. — 2002. — Vol. 28. — P. 995–997.
203. **Eckert M.** *Die Kartenwissenschaft. Forschungen und grundlagenzueiner Kartographieals Wissenschaft*. — Berlin, Leipzig: Walter de Gruyter & Co, 1921. — 640 p.
204. **Edelsbrunner H., Kirkpatrick D., Seidel R.** On the shape of a set of points in the plane // *IEEE Transactions on Information Theory*. — 1983. — Vol. 29, No. 4. — P. 551–559.
205. **Eilander D., Van Verseveld W., Yamazaki D., Weerts A., Winsemius H. C., Ward P. J.** A hydrography upscaling method for scale-invariant parametrization of distributed hydrological models // *Hydrology and Earth System Sciences*. — 2021. — Vol. 25, No. 9. — P. 5287–5313.
206. **Eiter T., Mannila H.** *Computing Discrete Fréchet Distance*. — Vienna, Austria: Christian Doppler Laboratory for Expert Systems, TU Wien, 1994.

207. **Erdoğan S.** Modelling the spatial distribution of DEM error with geographically weighted regression: An experimental study // *Computers & Geosciences*. — 2010. — Vol. 36, No. 1. — P. 34–43.
208. **Eskandar G.** An Empirical Study of the Generalization Ability of Lidar 3D Object Detectors to Unseen Domains // *2024 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. — Seattle, WA, USA: IEEE, 2024. — P. 23815–23825.
209. **Esri Inc.** *What is ArcPy?* [Электронный ресурс]. URL: <https://desktop.arcgis.com/en/arcmap/latest/analyze/arcpy/what-is-arcpy-.htm> (дата обращения: 15.02.2025).
210. **European Yearbook 1994.** *Annuaire Européen, Volume 42*. — London: Martinus Nijhoff Publishers, 1996. — 1256 p.
211. **Evans I. S.** Scale-Specific Landforms and Aspects of the Land Surface // *Concepts and Modelling in Geomorphology: International Perspectives* / ed. by I. S. Evans, R. Dikau, E. Tokunaga, H. Ohmori, M. Hirano. — Tokyo: Terrapub, 2003. — P. 61–84.
212. **Fan Q., Yang L., Hu P.** DEM generalization based on analysis of geometry and landscape context // *International Symposium on Multispectral Image Processing and Pattern Recognition*. — 2007. — Vol. 6790. — P. 679035–679038.
213. **Fang Y., Fan L.** Comparisons of Eight Simplification Methods for Data Reduction of Terrain Point Cloud // *2021 4th International Conference on Control and Computer Vision*. — Macau China: ACM, 2021. — P. 135–141.
214. **Fei L. F.** Experiments of Group-generalization of Contour Lines on Topographic Maps // *Journal of Wuhan Technical University of Surveying and Mapping*. — 1993. — Vol. 18, No. S1. — P. 6–22.
215. **Fei L., Huang L., He J.** An indirect generalization of contour lines based on DEM generalization using the 3D Douglas-Peucker algorithm // *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. — 2008. — Vol. XXXVII. — P. 751–756.
216. **Filin S., Abo Akel N., Kremeike K., Sester M., Doytsher Y.** Interpretation and Generalization of 3D Landscapes from LiDAR Data // *Cartography and Geographic Information Science*. — 2007. — Vol. 34, No. 3. — P. 231–243.
217. **Floriani L. de, Marzano P., Puppo E.** Multiresolution models for topographic surface description // *The Visual Computer*. — 1996. — Vol. 12, No. 7. — P. 317–345.
218. **Florinsky I. V.** Accuracy of local topographic variables derived from digital elevation models // *International Journal of Geographical Information Science*. — 1998. — Vol. 12, No. 1. — P. 47–62.

219. **Florinsky I. V.** Spheroidal equal angular DEMs: The specificity of morphometric treatment // *Transactions in GIS*. — 2017. — Vol. 21, No. 6. — P. 1115–1129.
220. **Florinsky I. V., Pankratov A. N.** A universal spectral analytical method for digital terrain modeling // *International Journal of Geographical Information Science*. — 2016. — Vol. 30, No. 12. — P. 2506–2528.
221. **Foerster T., Stoter J., Lemmens R.** Towards automatic web-based generalisation processing: A case study // *Proceedings of 11th ICA Workshop on Generalisation and Multiple Representation*. — Moscow: ICA, 2007. — P. 1–10.
222. **Ford C. M., Hu Y., Ghosh C., Fry L. M., Malakpour-Estalaki S., Mason L., Fitzpatrick L., Mazrooei A., Goering D. C.** Generalization of Runoff Risk Prediction at Field Scales to a Continental-Scale Region Using Cluster Analysis and Hybrid Modeling // *Geophysical Research Letters*. — 2022. — Vol. 49 (e2022GL100667), No. 17. — P. 1–9.
223. **Fotheringham A. S., Wong D. W. S.** The Modifiable Areal Unit Problem in Multivariate Statistical Analysis // *Environment and Planning A: Economy and Space*. — 1991. — Vol. 23, No. 7. — P. 1025–1044.
224. **Fowler R. J., Little J. J.** Automatic extraction of Irregular Network digital terrain models // *ACM SIGGRAPH Computer Graphics*. — 1979. — Vol. 13, No. 2. — P. 199–207.
225. **Fréchet M.** Sur quelques points du calcul fonctionnel // *Rendiconti del Circolo Matematico di Palermo*. — 1906. — Vol. 22, No. 1. — P. 1–74.
226. **Friis C., Hernández-Morcillo M., Baumann M., Coral C., Frommen T., Ghoddousi A., Loibl D., Rufin P.** Enabling spaces for bridging scales: Scanning solutions for interdisciplinary human-environment research // *Sustainability Science*. — 2023. — Vol. 18, No. 3. — P. 1251–1269.
227. **Frye C.** A Product Driven Approach to Designing a Multi-Purpose Multi-Scale GIS Base Map Database that Supports High Quality Mapping. — Vancouver, Washington: USA, 2006. — P. 1–16.
228. **Frye C., Eicher C. L.** Modeling active database-driven cartography within GIS databases // *Proceedings of the 21st International Cartographic Conference, 10-16 August*. — Durban: South Africa, 2003. — P. 1–7.
229. **Gaffuri J.** Field Deformation in an Agent-Based Generalisation Model: The GAEL Model // *GI-days 2007 — young researchers forum*. — 2007. — Vol. 30. — P. 1–24.
230. **Gaffuri J.** Three reuse examples of a generic deformation model in map generalisation // *Proceedings of 24th International Cartographic Conference ICC, 15–21 November 2009*. — Santiago, Chile: ICA, 2009. — P. 1–10.

231. **Gafurov A., Mukharamova S., Saveliev A., Yermolaev O.** Advancing Agricultural Crop Recognition: The Application of LSTM Networks and Spatial Generalization in Satellite Data Analysis // *Agriculture*. — 2023. — Vol. 13, No. 9. — P. 1672.
232. **Gallant J. C., Hutchinson M. F.** Scale dependence in terrain analysis // *Mathematics and Computers in Simulation*. — 1997. — Vol. 43, No. 3–6. — P. 313–321.
233. **Gamache M.** Free and Low Cost Datasets for International Mountain Cartography // *Proceedings of Mountain Cartography Workshop*. — Núria, Spain: ICA, 2004. — P. 1–34.
234. **Garcia Balboa J. L., Ariza López F. J.** Generalization-oriented Road Line Classification by Means of an Artificial Neural Network // *GeoInformatica*. — 2007. — Vol. 12, No. 3. — P. 289–312.
235. **Garuma G. F.** Review of urban surface parameterizations for numerical climate models // *Urban Climate*. — 2018. — Vol. 24. — P. 830–851.
236. **Garvert M. F., Smull B., Mass C.** Multiscale Mountain Waves Influencing a Major Orographic Precipitation Event // *Journal of the Atmospheric Sciences*. — 2007. — Vol. 64, No. 3. — P. 711–737.
237. **GEBCO.** *Global ocean & land terrain models* [Электронный ресурс]. URL: [https://www.gebco.net/data\\_and\\_products/gridded\\_bathymetry\\_data/](https://www.gebco.net/data_and_products/gridded_bathymetry_data/) (дата обращения: 15.02.2025).
238. **Geländedarstellung. Richtlinien 6.7.** — Wabern, Switzerland: Federal Office of Topography swisstopo, 2008.
239. **Getirana A. C. V., Bonnet M.-P., Rotunno Filho O. C., Mansur W. J.** Improving hydrological information acquisition from DEM processing in floodplains // *Hydrological Processes*. — 2009. — Vol. 23, No. 3. — P. 502–514.
240. **Gheibi A., Maheshwari A., Sack J.-R.** Minimizing Walking Length in Map Matching // *Topics in Theoretical Computer Science* / ed. by M. T. Hajiaghayi, M. R. Mousavi. — Cham: Springer International Publishing, 2016. — P. 105–120.
241. **Goethem A. van, Meulemans W., Speckmann B., Wood J.** Exploring Curved Schematization of Territorial Outlines // *IEEE Transactions on Visualization and Computer Graphics*. — 2015. — Vol. 21, No. 8. — P. 889–902.
242. **Gökgöz T.** Generalization of Contours Using Deviation Angles and Error Bands // *The Cartographic Journal*. — 2005. — Vol. 42, No. 2. — P. 145–156.
243. **Gomez F., Guzman A.** Digital model for three dimensional surface representation // *Geoprocessing*. — 1979. — Vol. 1. — P. 53–70.

244. **Gonzalez R. C., Woods R. E.** *Digital image processing*. — Fourth ed. — New York: Pearson, 2017. — 1192 p.
245. **Goodchild M. F., Li L.** Assuring the quality of volunteered geographic information // *Spatial Statistics*. — 2012. — Vol. 1. — P. 110–120.
246. **Gribov A., Bodansky E.** Reconstruction of Orthogonal Polygonal Lines // *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*. — Berlin, Heidelberg: Springer Berlin Heidelberg, 2006. — P. 462–473.
247. **Gridasov M. V., Kiseleva S. V., Nefedova L. V., Popel' O. S., Frid S. E.** Development of the geoinformation system “Renewable sources of Russia”: Statement of the problem and choice of solution methods // *Thermal Engineering*. — 2011. — Vol. 58, No. 11. — P. 924–931.
248. **Gröbe M., Burghardt D.** Scale-Dependent Point Selection Methods for Web Maps // *KN - Journal of Cartography and Geographic Information*. — 2021. — Vol. 71, No. 3. — P. 143–154.
249. **Grohmann C. H., Smith M. J., Riccomini C.** Multiscale Analysis of Topographic Surface Roughness in the Midland Valley, Scotland // *IEEE Transactions on Geoscience and Remote Sensing*. — 2011. — Vol. 49, No. 4. — P. 1200–1213.
250. **Gross M. H., Gatti R., Stadt O.** Fast multiresolution surface meshing // *Proceedings Visualization '95*. — Atlanta, GA, USA: IEEE Comput. Soc. Press, 1995. — P. 135–142.
251. **Guezic A., Hummel R.** Exploiting triangulated surface extraction using tetrahedral decomposition // *IEEE Transactions on Visualization and Computer Graphics*. — 1995. — Vol. 1, No. 4. — P. 328–342.
252. **Guilbert E., Gaffuri J., Jenny B.** Terrain Generalisation // *Abstracting Geographic Information in a Data Rich World* / ed. by D. Burghardt, C. Duchêne, W. Mackaness. — Cham: Springer International Publishing, 2014. — P. 227–258.
253. **Guillaume A. S., Leempoel K., Rochat E., Rogivue A., Kasser M., Gugerli F., Parisod C., Joost S.** Multiscale Very High Resolution Topographic Models in Alpine Ecology: Pros and Cons of Airborne LiDAR and Drone-Based Stereo-Photogrammetry Technologies // *Remote Sensing*. — 2021. — Vol. 13 (8), No. 1588. — P. 1–18.
254. **Guo D.** Regionalization with dynamically constrained agglomerative clustering and partitioning (REDCAP) // *International Journal of Geographical Information Science*. — 2008. — Vol. 22, No. 7. — P. 801–823.

255. **Gupta H. V., Perrin C., Blöschl G., Montanari A., Kumar R., Clark M., Andréassian V.** Large-sample hydrology: A need to balance depth with breadth // *Hydrology and Earth System Sciences*. — 2014. — Vol. 18, No. 2. — P. 463–477.
256. **Guth P. L., Van Niekerk A., Grohmann C. H., Muller J.-P., Hawker L., Florinsky I. V., Gesch D., Reuter H. I., Herrera-Cruz V., Riazanoff S., López-Vázquez C., Carabajal C. C., Albinet C., Strobl P.** Digital Elevation Models: Terminology and Definitions // *Remote Sensing*. — 2021. — Vol. 13 (18), No. 3581. — P. 1–19.
257. **Hack J. T.** *Studies of longitudinal stream profiles in Virginia and Maryland*. — Washington: U.S. Geological Survey, 1957.
258. **Hae-Kyong K., Jung-Wook M., Ki-Joune L.** Data Update Across Multi-Scale Databases // *Proceedings of the 12th International Conference on Geoinformatics – Geospatial Information Research: Bridging the Pacific and Atlantic*. — Gavle: Sweden, 2004. — P. 749–756.
259. **Hampe M., Sester M., Harrie L.** Multiple representation databases to support visualisation on mobile devices // *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*. — 2004. — Vol. 35, No. 6. — P. 1–6.
260. **Han Y., Wang Z., Lu X., Hu B.** Application of AHP to Road Selection // *ISPRS International Journal of Geo-Information*. — 2020. — Vol. 9 (2), No. 86. — P. 1–21.
261. **Hangouët J.** Computation of the Hausdorff Distance Between Plane Vector Polylines. — Charlotte, NC: CaGIS, 1997. — P. 1–10.
262. **Harrie L.** *Generalization Methods for Propagating Updates Between Cartographic Data Sets*: Licentiate Thesis. — Lund, Sweden: Department of Surveying, Lund Institute of Technology, Lund University, 1998.
263. **Harrie L., Hellström A.-K.** A Prototype System for Propagating Updates Between Cartographic Data Sets // *The Cartographic Journal*. — 1999. — Vol. 36, No. 2. — P. 133–140.
264. **Harrie L., Sarjakoski T., Lehto L.** A mapping function for variable-scale maps in small-display cartography // *Journal of Geospatial Engineering*. — 2002. — Vol. 4, No. 2. — P. 111–123.
265. **Harrower M., Brewer C. A.** ColorBrewer.org: An Online Tool for Selecting Colour Schemes for Maps // *The Cartographic Journal*. — 2003. — Vol. 40, No. 1. — P. 27–37.
266. **Harvey L. D. D.** Upscaling In Global Change Research // *Climatic Change*. — 2000. — Vol. 44, No. 3. — P. 225–263.

267. **Hauert J.-H.** A symmetry detector for map generalization and urban-space analysis // *ISPRS Journal of Photogrammetry and Remote Sensing*. — 2012. — Vol. 74. — P. 66–77.
268. **Hauert J.-H., Wolff A.** Optimal and topologically safe simplification of building footprints. // *ACM SIGSPATIAL GIS* / ed. by D. Agrawal, P. Zhang, A. E. Abbadi, M. F. Mokbel. — San Jose, CA, USA: ACM, 2010a. — P. 192–201.
269. **Hauert J.-H., Wolff A.** Area aggregation in map generalisation by mixed-integer programming // *International Journal of Geographical Information Science*. — 2010b. — Vol. 24, No. 12. — P. 1871–1897.
270. **Hausdorff F.** *Grundzüge der Mengenlehre*. — Leipzig: Veit, 1914. — 500 p.
271. **Healey S. P., Cohen W. B., Yang Z., Kenneth Brewer C., Brooks E. B., Gorelick N., Hernandez A. J., Huang C., Joseph Hughes M., Kennedy R. E., Loveland T. R., Moisen G. G., Schroeder T. A., Stehman S. V., Vogelmann J. E., Woodcock C. E., Yang L., Zhu Z.** Mapping forest change using stacked generalization: An ensemble approach // *Remote Sensing of Environment*. — 2018. — Vol. 204. — P. 717–728.
272. **Heer J., Robertson G.** Animated Transitions in Statistical Data Graphics // *IEEE Transactions on Visualization and Computer Graphics*. — 2007. — Vol. 13, No. 6. — P. 1240–1247.
273. **Heinzle F., Anders K.-H.** Characterising Space via Pattern Recognition Techniques // *Generalisation of Geographic Information: Cartographic Modelling and Applications* / ed. by W. A. Mackaness, A. Ruas, L. T. Sarjakoski. — Netherlands: Elsevier, 2007. — P. 233–253.
274. **Heitzler M., Gkonos C., Tsorlini A., Hurni L.** A modular process to improve the georeferencing of the Siegfried map // *e-Perimetron*. — 2018. — Vol. 13, No. 2. — P. 85–100.
275. **Hemer M. A., Zieger S., Durrant T., O’Grady J., Hoeke R. K., McInnes K. L., Rosebrock U.** A revised assessment of Australia’s national wave energy resource // *Renewable Energy*. — 2017. — Vol. 114. — P. 85–107.
276. **Hemerijckx L.-M., Van Emelen S., Rymenants J., Davis J., Verburg P. H., Lwasa S., Van Rompaey A.** Upscaling Household Survey Data Using Remote Sensing to Map Socioeconomic Groups in Kampala, Uganda // *Remote Sensing*. — 2020. — Vol. 12 (20), No. 3468. — P. 1–21.
277. **Herring J. R.** *OpenGIS® Implementation Standard for Geographic information - Simple feature access - Part 1: Common architecture*. : Open Geospatial Consortium, 2010.

278. **Hesselbarth M. H. K., Sciaini M., With K. A., Wiegand K., Nowosad J.** *Landscapemetrics : An open-source R tool to calculate landscape metrics // Ecography.* — 2019. — Vol. 42, No. 10. — P. 1648–1657.
279. **Hodge R. A., Hoey T. B.** Upscaling from grain-scale processes to alluviation in bedrock channels using a cellular automaton model // *Journal of Geophysical Research: Earth Surface.* — 2012. — Vol. 117 (F01017). — P. 1–18.
280. **Höhle J., Höhle M.** Accuracy assessment of digital elevation models by means of robust statistical methods // *ISPRS Journal of Photogrammetry and Remote Sensing.* — 2009. — Vol. 64, No. 4. — P. 398–406.
281. **Holben B. N.** Characteristics of maximum-value composite images from temporal AVHRR data // *International Journal of Remote Sensing.* — 1986. — Vol. 7, No. 11. — P. 1417–1434.
282. **Hootenanny.** *Conflates multiple maps into a single seamless map* [Электронный ресурс]. URL: <https://github.com/ngageoint/hootenanny> (дата обращения: 16.02.2025).
283. **Howe N. R., Weinman J., Gouwar J., Shamji A.** Deformable Part Models for Automatically Georeferencing Historical Map Images // *Proceedings of the 27th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems.* — Chicago IL USA: ACM, 2019. — P. 540–543.
284. **Hu A., Xu K., Pedrycz W., Xing M.** LiDAR point cloud simplification algorithm with fuzzy encoding-decoding mechanism // *Applied Soft Computing.* — 2024. — Vol. 162, No. 111852. — P. 1–9.
285. **Hu G., Wang C., Li S., Dai W., Xiong L., Tang G., Strobl J.** Using vertices of a triangular irregular network to calculate slope and aspect // *International Journal of Geographical Information Science.* — 2022. — Vol. 36, No. 2. — P. 382–404.
286. **Hurni L., Dahinden T., Hutzler E.** Digital Cliff Drawing for Topographic Maps: Traditional Representations by Means of New Technologies // *Cartographica: The International Journal for Geographic Information and Geovisualization.* — 2001. — Vol. 38, No. 1–2. — P. 55–65.
287. **Hutchinson M. F.** A new procedure for gridding elevation and stream line data with automatic removal of spurious pits // *Journal of Hydrology.* — 1989. — Vol. 106, No. 3–4. — P. 211–232.
288. **Hypsography.** *U.S. Geological Survey National Mapping Program technical instructions, Part 7.* — Reston, VA: U.S. Geological Survey, 2001. — 69 p.
289. **Imhof E.** *Cartographic Relief Presentation.* — Berlin: Walter der Gruyter, 1982. — 416 p.

290. **Iwahashi J., Pike R. J.** Automated classifications of topography from DEMs by an unsupervised nested-means algorithm and a three-part geometric signature // *Geomorphology*. — 2007. — Vol. 86, No. 3–4. — P. 409–440.
291. **Jaara K., Lecordix F.** Extraction of Cartographic Contour Lines Using Digital Terrain Model (DTM) // *The Cartographic Journal*. — 2011. — Vol. 48, No. 2. — P. 131–137.
292. **Jasiewicz J., Stepinski T. F.** Geomorphons — a pattern recognition approach to classification and mapping of landforms // *Geomorphology*. — 2013. — Vol. 182. — P. 147–156.
293. **Jelinski D. E., Wu J.** The modifiable areal unit problem and implications for landscape ecology // *Landscape Ecology*. — 1996. — Vol. 11, No. 3. — P. 129–140.
294. **Jenks G. F.** The data model concept in statistical mapping // *International Yearbook of Cartography*. — 1967. — No. 7. — P. 186–190.
295. **Jenks G. F.** Lines, Computers, and Human Frailties // *Annals of the Association of American Geographers*. — 1981. — Vol. 71, No. 1. — P. 1–10.
296. **Jenny B., Gilgen J., Geithövel R., Marston B. E., Hurni L.** Design Principles for Swiss-style Rock Drawing // *The Cartographic Journal*. — 2014. — Vol. 51, No. 4. — P. 360–371.
297. **Jenson S. K., Domingue J. O.** Extracting topographic structure from digital elevation data for geographic information-system analysis // *Photogrammetric Engineering and Remote Sensing*. — 1988. — Vol. 54, No. 11. — P. 1593–1600.
298. **Ji C., Li Y., Fan J., Lan S.** A Novel Simplification Method for 3D Geometric Point Cloud Based on the Importance of Point // *IEEE Access*. — 2019. — Vol. 7. — P. 129029–129042.
299. **Jobard B., Lefer W.** Creating Evenly-Spaced Streamlines of Arbitrary Density // *Visualization in Scientific Computing '97* / ed. by W. Lefer, M. Grave. — Vienna: Springer Vienna, 1997. — P. 43–55.
300. **Jones C. B., Abraham I. M.** Design Considerations for a Scale-Independent Database // *Proceedings of Second International Symposium on Spatial Data Handling*. — Seattle, Washington: International Geographical Union, 1986. — P. 384–398.
301. **Jones C. B., Kidner D. B., Luo L. Q., Bundy G. L., Ware J. M.** Database design for a multi-scale spatial information system // *International Journal of Geographical Information Systems*. — 1996. — Vol. 10, No. 8. — P. 901–920.
302. **Jordan G.** Adaptive smoothing of valleys in DEMs using TIN interpolation from ridgeline elevations: An application to morphotectonic aspect analysis // *Computers and Geosciences*. — 2007. — Vol. 33, No. 4. — P. 573–585.

303. **Jörges C., Berkenbrink C., Stumpe B.** Prediction and reconstruction of ocean wave heights based on bathymetric data using LSTM neural networks // *Ocean Engineering*. — 2021. — Vol. 232, No. 109046. — P. 1–18.
304. **Kadmon N.** Automated Selection of Settlements in Map Generalisation // *The Cartographic Journal*. — 1972. — Vol. 9, No. 2. — P. 93–98.
305. **Kalbermatten M., Van de Ville D., Joost S., Unser M., Golay F.** Laplace-Gradient Wavelet Pyramid and Multiscale Tensor Structures Applied on High-Resolution DEMs // *Proceedings of Geomorphometry 2009. Zurich, Switzerland, 31 August – 2 September, 2009*. — Zurich, Switzerland: The International Society for Geomorphometry, 2009. — P. 124–132.
306. **Karsznia I., Adolf A., Leyk S., Weibel R.** Using machine learning and data enrichment in the selection of roads for small-scale maps // *Cartography and Geographic Information Science*. — 2024. — Vol. 51, No. 1. — P. 60–78.
307. **Karsznia I., Weibel R.** Improving settlement selection for small-scale maps using data enrichment and machine learning // *Cartography and Geographic Information Science*. — 2018. — Vol. 45, No. 2. — P. 111–127.
308. **Karsznia I., Wereszczyńska K., Weibel R.** Make It Simple: Effective Road Selection for Small-Scale Map Design Using Decision-Tree-Based Models // *ISPRS International Journal of Geo-Information*. — 2022. — Vol. 11 (8), No. 457. — P. 1–19.
309. **Keates J. S.** *Cartographic design and production*. — 2 ed. — Harlow: Longman Scientific & Technical [u.a.], 1989. — 261 p.
310. **Kennelly P. J., Kimerling A. J.** Desktop Hachure Maps from Digital Elevation Models // *Cartographic Perspectives*. — 2000. — No. 37. — P. 78–81.
311. **Kennelly P. J., Kimerling a. J.** Hillshading of Terrain Using Layer Tints with Aspect-Variant Luminosity // *Cartography and Geographic Information Science*. — 2004. — Vol. 31, No. 2. — P. 67–77.
312. **Kennelly P. J., Kimerling A. J.** Non-Photorealistic Rendering and Terrain Representation // *Cartographic Perspectives*. — 2006. — No. 54. — P. 35–54.
313. **Kennelly P. J., Patterson T., Jenny B., Huffman D. P., Marston B. E., Bell S., Tait A. M.** Elevation models for reproducible evaluation of terrain representation // *Cartography and Geographic Information Science*. — 2021. — Vol. 48, No. 1. — P. 63–77.
314. **Kenneth Field.** *Cartography*. — Redlands, California: Esri Press, 2018. — 576 p.

315. **Kettunen P., Koski C., Oksanen J.** A design of contour generation for topographic maps with adaptive DEM smoothing // *International Journal of Cartography*. — 2017. — Vol. 3, No. 1. — P. 19–30.
316. **Kilpelainen T.** Maintenance of multiple representation databases for topographic data // *Proceedings of International Workshop on Dynamic and Multi-Dimensional GIS*. — Hong Kong: The Hong Kong Polytechnic University, 1997. — P. 116–127.
317. **Kilpeläinen T.** Maintenance of Multiple Representation Databases for Topographic Data // *The Cartographic Journal*. — 2000. — Vol. 37, No. 2. — P. 101–107.
318. **Kim H., Kang Y., Oh C., Yoon K.-J.** Single Domain Generalization for LiDAR Semantic Segmentation // *2023 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. — Vancouver, BC, Canada: IEEE, 2023. — P. 17587–17598.
319. **Kim J., Im S.** Rethinking LiDAR Domain Generalization: Single Source as Multiple Density Domains // *Computer Vision – ECCV 2024* / ed. by J. Kim, E. Ricci, S. Roth, O. Russakovsky. — Cham: Springer Nature Switzerland, 2025. — P. 310–327.
320. **Kim Y., Heer J.** Gemini: A Grammar and Recommender System for Animated Transitions in Statistical Graphics // *IEEE Transactions on Visualization and Computer Graphics*. — 2021. — Vol. 27, No. 2. — P. 485–494.
321. **Kimerling A. J.** *Map Use: Reading, Analysis, Interpretation*. — 8 ed. — Redlands: Esri Press, 2016. — 666 p.
322. **Knapp N., Fischer R., Cazcarra-Bes V., Huth A.** Structure metrics to generalize biomass estimation from lidar across forest types from different continents // *Remote Sensing of Environment*. — 2020. — Vol. 237, No. 111597. — P. 1–14.
323. **Knura M.** Learning from vector data: Enhancing vector-based shape encoding and shape classification for map generalization purposes // *Cartography and Geographic Information Science*. — 2023. — P. 1–22.
324. **Knuth K. H.** Optimal Data-Based Binning for Histograms // *Arxiv preprint physics*. — 2013. — No. 0605197. — P. 1–22.
325. **Kobbelt L., Campagna S., Seidel H. P.** A general framework for mesh decimation // *Graphics Interface '98 Proceedings*. — Vancouver, British Columbia: Canada, 1998. — P. 43–50.
326. **Kohl J.** *Der Verkehr des Menschen in seiner Abhängigkeit von der Erdoberfläche*. — Dresden: Germany, 1841. — 602 p.

327. **Kolbe T. H., Gröger G., Plümer L.** CityGML: Interoperable Access to 3D City Models // *Geo-Information for Disaster Management* / ed. by P. van Oosterom, S. Zlatanova, E. M. Fendel. — Berlin, Heidelberg: Springer, 2005. — P. 883–899.
328. **Koshel S. M.** Algorithm for Topologically Correct Gridding of Contour Data // *Proceedings of Seventh International Conference on Geographic Information Science (GIScience 2012)*. — Columbus, Ohio: USA, 2012. — P. 1–5.
329. **Kosheleva N. E., Vlasov D. V., Timofeev I. V., Samsonov T. E., Kasimov N. S.** Benzo[a]pyrene in Moscow road dust: Pollution levels and health risks // *Environmental Geochemistry and Health*. — 2023. — Vol. 45, No. 5. — P. 1669–1694.
330. **Kraak M.-J., Ormeling F.** *Cartography: Visualization of geospatial data*. — 4 ed. — Boca Raton, London: CRC Press, 2021. — 261 p.
331. **Kratzert F., Klotz D., Shalev G., Klambauer G., Hochreiter S., Nearing G.** Towards learning universal, regional, and local hydrological behaviors via machine learning applied to large-sample datasets // *Hydrology and Earth System Sciences*. — 2019. — Vol. 23, No. 12. — P. 5089–5110.
332. **Kreveld M. van, Oosterom P. van, Snoeyink J.** Efficient settlement selection for interactive display // *Proceedings of AutoCarto XIII*. — Bethesda, MD: CaGIS, 1995. — P. 287–296.
333. **Krinitskiy M., Aleksandrova M., Verezemskaya P., Gulev S., Sinitsyn A., Kovaleva N., Gavrikov A.** On the Generalization Ability of Data-Driven Models in the Problem of Total Cloud Cover Retrieval // *Remote Sensing*. — 2021. — Vol. 13 (2), No. 326. — P. 1–28.
334. **Kruiger J. F., Kasalica V., Meerlo R., Lamprecht A., Nyamsuren E., Scheider S.** Loose programming of GIS workflows with geo-analytical concepts // *Transactions in GIS*. — 2021. — Vol. 25, No. 1. — P. 424–449.
335. **Krylov S., Dvornikov A., Zagrebin G., Petrov V., Plotnikov I.** The automation of technological processes for creating small-scale digital cartographic bases for general geographic and thematic mapping // *Proceedings of the 26th International Cartographic Conference, Dresden, Germany*. — Dresden, Germany: ICA, 2013. — P. 1–5.
336. **Kyriakidis P. C., Shortridge A. M., Goodchild M. F.** Geostatistics for conflation and accuracy assessment of digital elevation models // *International Journal of Geographical Information Science*. — 1999. — Vol. 13, No. 7. — P. 677–707.
337. **L3Harris Technologies Inc.** *MapMerger™ Geospatial Vector Conflation* [Электронный ресурс]. URL: <https://www.harris.com/solution/mapmerger-geospatial-vector-conflation> (дата обращения: 08.05.2020).

338. **Ladd C. J. T., Smeaton C., Skov M. W., Austin W. E. N.** Best practice for upscaling soil organic carbon stocks in salt marshes // *Geoderma*. — 2022. — Vol. 428, No. 116188. — P. 1–15.
339. **Langran C. E., Poiker T. K.** Integration of name selection and name placement // *Proceedings of 2nd International Symposium on Spatial Data Handling*. — Seattle, Washington: USA, 1986. — P. 50–64.
340. **Larsén X. G., Davis N., Hannesdóttir Á., Kelly M., Svenningsen L., Slot R., Imberger M., Olsen B. T., Floors R.** The Global Atlas for Siting Parameters project: Extreme wind, turbulence, and turbine classes // *Wind Energy*. — 2022. — Vol. 25, No. 11. — P. 1841–1859.
341. **Lee B., Satyanarayan A., Cordeil M., Prouzeau A., Jenny B., Dwyer T.** Deimos: A Grammar of Dynamic Embodied Immersive Visualisation Morphs and Transitions // *Proceedings of the 2023 CHI Conference on Human Factors in Computing Systems*. — Hamburg Germany: ACM, 2023. — P. 1–18.
342. **Lee D., Yang W., Ahmed N.** Conflation in Geoprocessing Framework – Case Studies Recent development and looking ahead // *Proceedings of Geoprocessing 2014*. — Barcelona: Spain, 2014. — P. 1–6.
343. **Lee J.** Comparison of existing methods for building triangular irregular network models of terrain from grid digital elevation models // *International Journal of Geographical Information Systems*. — 1991. — Vol. 5, No. 3. — P. 267–285.
344. **Lee S., Wolberg G., Shin S. Y.** Scattered data interpolation with multilevel B-splines // *IEEE Transactions on Visualization and Computer Graphics*. — 1997. — Vol. 3, No. 3. — P. 228–244.
345. **Lehmann J. G.** *Darstellung einer neuer theorie der bezeichnung der schiefen flachen*. — Leipzig, Germany: bei Johann Benjamin Georg Fleischer, 1799. — 172 p.
346. **Lei T., Lei Z.** Optimal spatial data matching for conflation: A network flow-based approach // *Transactions in GIS*. — 2019. — Vol. 23, No. 5. — P. 1152–1176.
347. **Lemmens M.** Quality of Geo-Information // *Geo-Information: Technologies, Applications and the Environment*. — Dordrecht: Springer Netherlands, 2011. — P. 211–227.
348. **Leonowicz A. M., Jenny B., Hurni L.** Automatic generation of hypsometric layers for small-scale maps // *Computers and Geosciences*. — 2009. — Vol. 35, No. 10. — P. 2074–2083.
349. **Lesage P.-L., Visvalingam M.** Towards sketch-based exploration of terrain // *Computers & Graphics*. — 2002. — Vol. 26, No. 2. — P. 309–328.

350. **Li L., Goodchild M. F.** An optimisation model for linear feature matching in geographical data conflation // *International Journal of Image and Data Fusion*. — 2011. — Vol. 2, No. 4. — P. 309–328.
351. **Li Q., Wang Z., Yang B.** Multi-resolution representation of digital terrain models with terrain features preservation // *Science in China Series E: Technological Sciences*. — 2008. — Vol. 51, No. S1. — P. 145–154.
352. **Li R., Shen Y., Dai W.** Cartographic Generalization of Islands Using Remote Sensing Images for Multiscale Representation // *Remote Sensing*. — 2024. — Vol. 16, No. 16. — P. 2971.
353. **Li X., Liu S., Li H., Ma Y., Wang J., Zhang Y., Xu Z., Xu T., Song L., Yang X., Lu Z., Wang Z., Guo Z.** Intercomparison of Six Upscaling Evapotranspiration Methods: From Site to the Satellite Pixel // *Journal of Geophysical Research: Atmospheres*. — 2018. — Vol. 123, No. 13. — P. 6777–6803.
354. **Li Y., Briggs R.** An Automated System for Image-to-Vector Georeferencing // *Cartography and Geographic Information Science*. — 2012. — Vol. 39, No. 4. — P. 199–217.
355. **Li Z.** *Algorithmic Foundation of Multi-Scale Spatial Representation*. — Boca Raton: CRC Press, 2006. — 310 p.
356. **Li Z., Choi Y. H.** Topographic Map Generalization: Association of Road Elimination with Thematic Attributes // *The Cartographic Journal*. — 2002. — Vol. 39, No. 2. — P. 153–166.
357. **Li Z., Ho A.** Design of Multi-Scale and Dynamic Maps for Land Vehicle Navigation // *The Cartographic Journal*. — 2004. — Vol. 41, No. 3. — P. 265–270.
358. **Li Z., Openshaw S.** Algorithms for Automated Line Generalization Based on a Natural Principle of Objective Generalization // *International Journal of Geographical Information Systems*. — 1992. — Vol. 6, No. 5. — P. 373–389.
359. **Li Z., Sui H.** An Integrated Technique for Automated Generalization of Contour Maps // *The Cartographic Journal*. — 2000. — Vol. 37, No. 1. — P. 29–37.
360. **Li Z., Zhou Q.** Integration of linear and areal hierarchies for continuous multi-scale representation of road networks // *International Journal of Geographical Information Science*. — 2012. — Vol. 26, No. 5. — P. 855–880.
361. **Lillesand T. M., Kiefer R. W., Chipman J. W.** *Remote sensing and image interpretation*. — Seventh ed. — Hoboken, N.J: John Wiley & Sons, Inc, 2015. — 720 p.

362. **Lin P., Pan M., Wood E. F., Yamazaki D., Allen G. H.** A new vector-based global river network dataset accounting for variable drainage density // *Scientific Data*. — 2021. — Vol. 8, No. 28. — P. 1–9.
363. **Lindsay J. B.** Efficient hybrid breaching-filling sink removal methods for flow path enforcement in digital elevation models: Efficient Hybrid Sink Removal Methods for Flow Path Enforcement // *Hydrological Processes*. — 2016a. — Vol. 30, No. 6. — P. 846–857.
364. **Lindsay J. B.** The practice of DEM stream burning revisited // *Earth Surface Processes and Landforms*. — 2016b. — Vol. 41, No. 5. — P. 658–668.
365. **Lindsay J. B., Seibert J.** Measuring the significance of a divide to local drainage patterns // *International Journal of Geographical Information Science*. — 2013. — Vol. 27, No. 7. — P. 1453–1468.
366. **Lindström P., Koller D., Ribarsky W., Hodges L. F., Faust N., Turner G.** Real-time continuous level of detail rendering of height fields // *Proceedings of SIGGRAPH August 4–9, 1996* / ed. by H. Rushmeier. — New Orleans, LA: ACM SIGGRAPH, 1996. — P. 109–118.
367. **Lipson M. J., Nazarian N., Hart M. A., Nice K. A., Conroy B.** A Transformation in City-Descriptive Input Data for Urban Climate Models // *Frontiers in Environmental Science*. — 2022. — Vol. 10, No. 866398. — P. 1–18.
368. **Lisiewicz M., Karsznia I.** Verifying and Exploring Settlement Selection Rules and Variables for Small-Scale Maps Using Decision Tree-Based Models // *Geographical Analysis*. — 2021. — Vol. 53, No. 3. — P. 560–587.
369. **Liu K., Paterson L., Wong P., Qi D.** A Sedimentological Approach to Upscaling // *Transport in Porous Media*. — 2002. — Vol. 46. — P. 285–310.
370. **Liu X., Zhan F. B., Ai T.** Road selection based on Voronoi diagrams and “strokes” in map generalization // *International Journal of Applied Earth Observation and Geoinformation*. — 2010. — Vol. 12S. — P. S194–S202.
371. **Loon J. C.** *Cartographic generalization of digital terrain models*. — Ann Arbor: University Microfilms International, 1977.
372. **Lorensen W. E., Cline H. E.** Marching cubes: A high resolution 3D surface construction algorithm // *ACM SIGGRAPH Computer Graphics*. — 1987. — Vol. 21, No. 4. — P. 163–169.

373. **Lou Y., Zhang C., Zheng Y., Xie X., Wang W., Huang Y.** Map-matching for low-sampling-rate GPS trajectories // *Proceedings of the 17th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems - GIS '09*. — Seattle, Washington: ACM Press, 2009. — P. 352–361.
374. **Lounsbery M., DeRose T. D., Warren J.** Multiresolution analysis for surfaces of arbitrary topological type // *ACM Transactions on Graphics*. — 1997. — Vol. 16, No. 1. — P. 34–73.
375. **Luft J.** Automatic Georeferencing of Historical Maps by Geocoding // *Automatic Vectorisation of Historical Maps: International workshop organized by the ICA Commission on Cartographic Heritage into the Digital. 13 March, 2020*. — Budapest, Hungary: Department of Cartography and Geoinformatics ELTE, 2020. — P. 77–89.
376. **Luft J., Schiewe J.** Automatic content-based georeferencing of historical topographic maps // *Transactions in GIS*. — 2021. — Vol. 25, No. 6. — P. 2888–2906.
377. **Lyu Z., Sun Q., Ma J., Xu Q., Li Y., Zhang F.** Road Network Generalization Method Constrained by Residential Areas // *ISPRS International Journal of Geo-Information*. — 2022. — Vol. 11, No. 3. — P. 159.
378. **Mackaness W. A., Beard K. M.** Use of Graph Theory to Support Map Generalization // *Cartography and Geographic Information Science*. — 1993. — Vol. 20, No. 4. — P. 210–221.
379. **Mackaness W. A., Ruas A.** *Generalisation of geographic information: Cartographic modelling and applications* / ed. by W. A. Mackaness, A. Ruas. — 1 ed. — Amsterdam: Elsevier, 2007a. — 370 p.
380. **Mackaness W. A., Ruas A.** Evaluation in the Map Generalisation Process // *Generalisation of Geographic Information: Cartographic Modelling and Applications*. — Amsterdam, Netherlands: Elsevier, 2007b. — P. 89–111.
381. **Mackaness W., Steven M.** An Algorithm for Localised Contour Removal over Steep Terrain // *The Cartographic Journal*. — 2006. — Vol. 43, No. 2. — P. 144–156.
382. **Marble D. F.** Geographic Information Systems: An overview // *Proceedings of Pecora 9: Spatial Information Technologies for Remote Sensing Today and Tomorrow*. — Sioux Falls, SD: NASA, 1984. — P. 18–24.
383. **Mark D. M.** Automated Detection Of Drainage Networks From Digital Elevation Models // *Cartographica: The International Journal for Geographic Information and Geovisualization*. — 1984. — Vol. 21, No. 2–3. — P. 168–178.

384. **Martin M. P., Orton T. G., Lacarce E., Meersmans J., Saby N. P. A., Paroissien J. B., Jolivet C., Boulonne L., Arrouays D.** Evaluation of modelling approaches for predicting the spatial distribution of soil organic carbon stocks at the national scale // *Geoderma*. — 2014. — Vol. 223–225. — P. 97–107.
385. **Masson V., Heldens W., Bocher E., Bonhomme M., Bucher B., Burmeister C., De Munck C., Esch T., Hidalgo J., Kanani-Sühring F., Kwok Y.-T., Lemonsu A., Lévy J.-P., Maronga B., Pavlik D., Petit G., See L., Schoetter R., Tornay N., Votsis A., Zeidler J.** City-descriptive input data for urban climate models: Model requirements, data sources and challenges // *Urban Climate*. — 2020. — Vol. 31, No. 100536. — P. 1–28.
386. **McMaster R. B.** A Statistical Analysis of Mathematical Measures for Linear Simplification // *Cartography and Geographic Information Science*. — 1986. — Vol. 13, No. 2. — P. 103–116.
387. **McMaster R. B.** The Geometric Properties of Numerical Generalization // *Geographical Analysis*. — 1987. — Vol. 19, No. 4. — P. 330–346.
388. **McMaster R. B., Shea K. S.** Cartographic Generalization in a Digital Environment: A Framework for Implementation in a Geographic Information System // *Proceedings of GIS/LIS'88, San Antonio, Texas*. — Falls Church, VA: American Society of Photogrammetry and Remote Sensing and American Congress on Surveying and Mapping, 1988. — P. 240–249.
389. **McMaster R. B., Shea K. S.** *Generalization in Digital Cartography*. — Washington, DC: Association of American Cartographers, 1992. — 133 p.
390. **McMaster R., Monmonier M.** A conceptual framework for quantitative and qualitative raster-mode generalization // *Proceedings of GIS/LIS'89 November 26-30*. — Orlando, Florida: AM / FM International, 1989. — P. 390–403.
391. **Medyńska-Gulij B., Forrest D., Cybulski P.** Modern Cartographic Forms of Expression: The Renaissance of Multimedia Cartography // *ISPRS International Journal of Geo-Information*. — 2021. — Vol. 10, No. 7. — P. 484.
392. **Meijer J. R., Huijbregts M. A. J., Schotten K. C. G. J., Schipper A. M.** Global patterns of current and future road infrastructure // *Environmental Research Letters*. — 2018. — Vol. 13, No. 064006. — P. 1–10.
393. **Meijers M., Oosterom P. van, Driel M., Šuba R.** Web-based dissemination of continuously generalized space-scale cube data for smooth user interaction // *International Journal of Cartography*. — 2020. — P. 1–25.

394. **Meijers M., Savino S., Van Oosterom P.** SPLITAREA: An algorithm for weighted splitting of faces in the context of a planar partition // *International Journal of Geographical Information Science*. — 2016. — Vol. 30, No. 8. — P. 1522–1551.
395. **Mekhedov I., Mestetskiy L.** Skeleton of a Multi-ribbon Surface // *Computational Science and Its Applications – ICCSA 2010* / ed. by D. Taniar, O. Gervasi, B. Murgante, E. Pardede, B. O. Apduhan. — Berlin, Heidelberg: Springer, 2010. — P. 557–573.
396. **Meng L., Forberg A.** 3D Building Generalisation // *Generalisation of Geographic Information: Cartographic Modelling and Applications* / ed. by W. A. Mackaness, A. Ruas, L. T. Sarjakoski. — Amsterdam, Netherlands: Elsevier, 2007. — P. 211–231.
397. **Mesa-Mingorance J. L., Ariza-López F. J.** Accuracy Assessment of Digital Elevation Models (DEMs): A Critical Review of Practices of the Past Three Decades // *Remote Sensing*. — 2020. — Vol. 12 (16), No. 2630. — P. 1–27.
398. **METI, NASA.** *Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) Global Digital Elevation Model Version 1* [Электронный ресурс]. URL: <https://asterweb.jpl.nasa.gov/gdem.asp> (дата обращения: 16.02.2025).
399. **Meulemans W., Renssen A. van, Speckmann B.** Area-Preserving Subdivision Schematization // *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*. — Berlin, Heidelberg: Springer, 2010. — P. 160–174.
400. **Miele V., Picard F., Dray S.** Spatially constrained clustering of ecological networks // *Methods in Ecology and Evolution*. — 2014. — Vol. 5, No. 8. — P. 771–779.
401. **Milad B., Ghosh S., Slatt R., Marfurt K., Fahes M.** Practical Aspects of Upscaling Geocellular Geological Models for Reservoir Fluid Flow Simulations: A Case Study in Integrating Geology, Geophysics, and Petroleum Engineering Multiscale Data from the Hunton Group // *Energies*. — 2020. — Vol. 13 (7), No. 1604. — P. 1–27.
402. **Mitropoulos V., Nakos B.** A Methodology on Natural Occurring Lines Segmentation and Generalization // *Proceedings of 14th ICA Workshop on Generalisation and Multiple Representation*. — Paris, France: ICA, 2011. — P. 1–7.
403. **Morrison J. L.** A theoretical framework for cartographic generalization with emphasis on the process of symbolization // *International Yearbook of Cartography*. — 1974. — Vol. 14. — P. 115–127.
404. **Mower J. E.** Automating Landscape Illustration with Pen and Ink Style Rendering // *Cartography and Geographic Information Science*. — 2009. — Vol. 36, No. 1. — P. 117–128.
405. **Müller J.-C., Lagrange J.-P., Weibel R.** *GIS and Generalization: Methodology and Practice*. — London: Routledge, 1995. — 276 p.

406. **Mustiere S.** Cartographic generalization of roads in a local and adaptive approach: A knowledge acquisition problem // *International Journal of Geographical Information Science*. — 2005. — Vol. 19, No. 8–9. — P. 937–955.
407. **Mustière S., Devogele T.** Matching networks with different levels of detail // *GeoInformatica*. — 2008. — Vol. 12, No. 4. — P. 435–453.
408. **Muthusamy M., Schellart A., Tait S., Heuvelink G. B. M.** Geostatistical upscaling of rain gauge data to support uncertainty analysis of lumped urban hydrological models // *Hydrology and Earth System Sciences*. — 2017. — Vol. 21, No. 2. — P. 1077–1091.
409. **Myslenkov S., Samsonov T., Shurygina A., Kiseleva S., Arkhipkin V.** Wind Waves Web Atlas of the Russian Seas // *Water*. — 2023. — Vol. 15, No. 2036. — P. 1–18.
410. **NASADEM.** *Merged DEM Global 1 arc second V001* [Электронный ресурс]. URL: [https://lpdaac.usgs.gov/products/nasadem\\_hgtv001](https://lpdaac.usgs.gov/products/nasadem_hgtv001) (дата обращения: 16.02.2025).
411. **Nickerson B. G.** Automated Cartographic Generalization For Linear Features // *Cartographica: The International Journal for Geographic Information and Geovisualization*. — 1988. — Vol. 25, No. 3. — P. 15–66.
412. **Nickerson B. G., Freeman H.** Development of a Rule-Based System for Automatic Map Generalization // *Proceedings of the Second International Symposium on Spatial Data Handling*. — Seattle, Washington: USA, 1986. — P. 537–556.
413. **Nie H., Huang Z.** A New Method Of Line Feature Generalization Based On Shape Characteristic Analysis // *Metrology and Measurement Systems*. — 2011. — Vol. XVIII, No. 4. — P. 597–606.
414. **Noetinger B., Artus V., Zargar G.** The future of stochastic and upscaling methods in hydrogeology // *Hydrogeology Journal*. — 2005. — Vol. 13, No. 1. — P. 184–201.
415. **Nowosad J., Stepinski T. F.** Information theory as a consistent framework for quantification and classification of landscape patterns // *Landscape Ecology*. — 2019. — Vol. 34, No. 9. — P. 2091–2101.
416. **O’Callaghan J. F., Mark D. M.** The extraction of drainage networks from digital elevation data // *Computer Vision, Graphics, and Image Processing*. — 1984. — Vol. 28, No. 3. — P. 323–344.
417. **Oberholzer C., Hurni L.** Visualization of change in the Interactive Multimedia Atlas of Switzerland // *Computers and Geosciences*. — 2000. — Vol. 26, No. 1. — P. 37–43.
418. **OGC 07-011r2.** (ISO 19123:2005) *The OpenGIS® Abstract Specification Topic 6: Schema for coverage geometry and functions*. — OGC: Open Geospatial Consortium Inc., 2024.

419. **Oke T. R.** *Boundary Layer Climates*. — 2 ed. — London: Routledge, 1987. — 464 p.
420. **Oliphant T. E.** *A guide to NumPy*. — USA: Tregol Publishing, 2006. — 378 p.
421. **Oliveira M., Crujeiras R. M., Rodríguez-Casal A.** A plug-in rule for bandwidth selection in circular density estimation // *Computational Statistics and Data Analysis*. — 2012. — Vol. 56, No. 12. — P. 3898–3908.
422. **Olivera F., Lear M. S., Famiglietti J. S., Asante K.** Extracting low-resolution river networks from high-resolution digital elevation models // *Water Resources Research*. — 2002. — Vol. 38 (11), No. 1231. — P. 1–8.
423. **Oosterom P. van, Meijers M.** Vario-scale data structures supporting smooth zoom and progressive transfer of 2D and 3D data // *International Journal of Geographical Information Science*. — 2014. — Vol. 28, No. 3. — P. 455–478.
424. **Openshaw S.** A Geographical Solution to Scale and Aggregation Problems in Region-Building, Partitioning and Spatial Modelling // *Transactions of the Institute of British Geographers*. — 1977. — Vol. 2, No. 4. — P. 459.
425. **Openshaw S.** *The modifiable areal unit problem*. — Norwich, UK: Geo Books, 1984. — 41 p.
426. **OpenStreetMap.** *A free, editable map of the whole world* [Электронный ресурс]. URL: <https://www.openstreetmap.org> (дата обращения: 10.02.2025).
427. **Pain C. F.** Multi-scale landscape mapping for catchment characterization // *Proceedings of 2nd International Salinity Forum Salinity, water and society—global issues, local action. 31 March – 3 April 2008*. — Adelaide: South Australia, 2008.
428. **Pallero G.** Robust Line Simplification On The Plane // *Computers & Geosciences*. — 2013. — Vol. 61. — P. 152–159.
429. **Pang Y., Räsänen A., Juselius-Rajamäki T., Aurela M., Juutinen S., Väiliranta M., Virtanen T.** Upscaling field-measured seasonal ground vegetation patterns with Sentinel-2 images in boreal ecosystems // *International Journal of Remote Sensing*. — 2023. — Vol. 44, No. 14. — P. 4239–4261.
430. **Park W., Yu K.** Hybrid line simplification for cartographic generalization // *Pattern Recognition Letters*. — 2011. — Vol. 32, No. 9. — P. 1267–1273.
431. **Parzen E.** On Estimation of a Probability Density Function and Mode // *The Annals of Mathematical Statistics*. — 1962. — Vol. 33, No. 3. — P. 1065–1076.

432. **Patterson T., Kelso N. V., et al.** *Natural Earth. Free vector and raster map data at 1:10m, 1:50m, and 1:110m scales* [Электронный ресурс]. URL: <https://www.naturalearthdata.com> (дата обращения: 25.03.2020).
433. **Paz A. R., Collischonn W., Lopes Da Silveira A. L.** Improvements in large-scale drainage networks derived from digital elevation models // *Water Resources Research*. — 2006. — Vol. 42, No. 8. — P. 2005WR004544.
434. **Pedrini H.** Multiresolution terrain modeling based on triangulated irregular networks // *Revista Brasileira de Geociencias*. — 2001. — Vol. 31, No. 2. — P. 117–122.
435. **Peled A., Loon J. C., Bossler J. D.** Producing Intermediate Contours from Digitized Contours // *Cartography and Geographic Information Science*. — 1989. — Vol. 16, No. 3. — P. 191–200.
436. **Peng W., Pilouk M., Tempfli K.** Generalizing Relief Representation Using Digitized Contours // *International Archives of Photogrammetry and Remote Sensing*. — 1996. — Vol. 31, No. B4. — P. 649–654.
437. **Peng Y., Xiong X., Adhikari K., Knadel M., Grunwald S., Greve M. H.** Modeling Soil Organic Carbon at Regional Scale by Combining Multi-Spectral Images with Laboratory Spectra // *PLOS ONE*. — 2015. — Vol. 10, No. 11. — P. 1–22 (e0142295).
438. **Petras V., Newcomb D. J., Mitasova H.** Generalized 3D fragmentation index derived from lidar point clouds // *Open Geospatial Data, Software and Standards*. — 2017. — Vol. 2, No. 9. — P. 1–14.
439. **Peucker T. K., Douglas D. H.** Detection of Surface-Specific Points by Local Parallel Processing of Discrete Terrain Elevation Data // *Computer Graphics and Image Processing*. — 1975. — Vol. 4, No. 4. — P. 375–387.
440. **Pewsey A., Neuhäuser M., Ruxton G. D.** *Circular statistics in R*. — Oxford New York: Oxford University Press, 2013. — 198 p.
441. **Plazanet C., Affholder J.-G., Fritsch E.** The Importance of Geometric Modeling in Linear Feature Generalization // *Cartography and Geographic Information Science*. — 1995. — Vol. 22, No. 4. — P. 291–305.
442. **Plazanet C., Bigolin N. M., Ruas A.** Experiments with Learning Techniques for Spatial Model Enrichment and Line Generalization // *GeoInformatica*. — 1998. — Vol. 2, No. 4. — P. 315–333.
443. **Podobnikar T.** Methods for visual quality assessment of a digital terrain model // *Surveys and Perspectives Integrating Environment and Society*. — 2008. — Vol. 1, No. 2. — P. 1–10.

444. **Polidori L., El Hage M.** Digital Elevation Model Quality Assessment Methods: A Critical Review // *Remote Sensing*. — 2020. — Vol. 12 (21), No. 3522. — P. 1–36.
445. **Preparata F. P., Shamos M. I.** *Computational Geometry*. — New York, NY: Springer New York, 1985. — 398 p.
446. **Pung J., D'Souza R. M., Ghosal D., Zhang M.** A road network simplification algorithm that preserves topological properties // *Applied Network Science*. — 2022. — Vol. 7, No. 79. — P. 1–16.
447. **Quental P., Almeida J. A., Simões M.** Construction of high-resolution stochastic geological models and optimal upscaling to a simplified layer-type hydrogeological model // *Advances in Water Resources*. — 2012. — Vol. 39. — P. 18–32.
448. **Rangayyan R. M., Guliato D., Carvalho J. D. de, Santiago S. A.** Polygonal approximation of contours based on the turning angle function // *Journal of Electronic Imaging*. — 2008. — Vol. 17 (2), No. 23016. — P. 1–14.
449. **Raposo P.** Variable DEM generalization using local entropy for terrain representation through scale // *International Journal of Cartography*. — 2020. — Vol. 6, No. 1. — P. 99–120.
450. **Rasaei M. R., Sahimi M.** Upscaling of the Geological Models of Large-Scale Porous Media Using Multiresolution Wavelet Transformations // *Journal of Heat Transfer*. — 2009. — Vol. 131 (10), No. 101007. — P. 1–12.
451. **Ratajski L.** Phénomènes des points de généralisation // *International Yearbook of Cartography*. — 1967. — Vol. 7. — P. 143–151.
452. **Reed S. M.** Deriving flow directions for coarse-resolution (1–4 km) gridded hydrologic modeling // *Water Resources Research*. — 2003. — Vol. 39 (9), No. 1238. — P. 1–11.
453. **Regnauld N., Mackaness W. A., Hart G.** Automated relief representation for visualisation of archaeological monuments and other anthropogenic forms // *Computers, Environment and Urban Systems*. — 2002. — Vol. 26, No. 2–3. — P. 219–239.
454. **Revell P., Regnauld N., Bulbrooke G.** OS VectorMap™ District: Automated generalisation, text placement and conflation in support of making Public Data Public // *Proceedings of the 25th International cartographic conference*. — Paris, France: ICA, 2011. — P. 1–13.
455. **Rhind D.** A GIS research agenda // *International Journal of Geographical Information Systems*. — 1988. — Vol. 2, No. 1. — P. 23–28.
456. **Riitters K. H., O'Neill R. V., Wickham J. D., Jones K. B.** A note on contagion indices for landscape analysis // *Landscape Ecology*. — 1996. — Vol. 11, No. 4. — P. 197–202.

457. **Riley S. J., De Gloria S. D., Elliot R.** A Terrain Ruggedness Index that Quantifies Topographic Heterogeneity // *Intermountain Journal of Science*. — 1999. — Vol. 5, No. 1–4. — P. 23–27.
458. **Rivillas-Ospina G., Casas D., Maza-Chamorro M. A., Bolívar M., Ruiz G., Guerrero R., Horrillo-Caraballo J. M., Guerrero M., Díaz K., Rio R. D., Campos E.** APPMAR 1.0: A Python application for downloading and analyzing of WAVEWATCH III® wave and wind data // *Computers & Geosciences*. — 2022. — Vol. 162, No. 105098. — P. 1–14.
459. **Rivin G. S., Rozinkina I. A., Vil'fand R. M., Kiktev D. B., Tudrii K. O., Blinov D. V., Varentsov M. I., Zakharchenko D. I., Samsonov T. E., Repina I. A., Artamonov A. Yu.** Development of the High-resolution Operational System for Numerical Prediction of Weather and Severe Weather Events for the Moscow Region // *Russian Meteorology and Hydrology*. — 2020. — Vol. 45, No. 7. — P. 455–465.
460. **Rivin G. S., Vil'fand R. M., Kiktev D. B., Rozinkina I. A., Tudriy K. O., Blinov D. V., Varentsov M. I., Samsonov T. E., Bundel' A. Yu., Kirsanov A. A., Zakharchenko D. I.** The System for Numerical Prediction of Weather Events (Including Severe Ones) for Moscow Megacity: The Prototype Development // *Russian Meteorology and Hydrology*. — 2019. — Vol. 44, No. 11. — P. 729–738.
461. **Robinson A. H., Morrison J. L., Muehrcke P. C., Kimerling J. A., Guptill S. C.** *Elements of cartography*. — 6 ed. — New York: John Wiley & Sons, 1995. — 688 p.
462. **Romstad B., Etzelmüller B.** Mean-curvature watersheds: A simple method for segmentation of a digital elevation model into terrain units // *Geomorphology*. — 2012. — Vol. 139–140. — P. 293–302.
463. **Rosenblatt M.** Remarks on Some Nonparametric Estimates of a Density Function // *The Annals of Mathematical Statistics*. — 1956. — Vol. 27, No. 3. — P. 832–837.
464. **Rosenfeld A., Johnston E.** Angle Detection on Digital Curves // *IEEE Transactions on Computers*. — 1973. — Vol. C-22, No. 9. — P. 875–878.
465. **Roth R. E., Brewer C. A., Stryker M. S.** A typology of operators for maintaining legible map designs at multiple scales // *Cartographic Perspectives*. — 2011. — No. 68. — P. 29–64.
466. **Roth R. E., Kelly M., Underwood N., Lally N., Vincent K., Sack C.** Interactive Multiscale Thematic Maps: A Preliminary Study // *Abstracts of the ICA*. — 2019. — Vol. 1. — P. 1–2.
467. **Ruas A., Bianchin A.** Echelle et Niveau de Detail // *Generalisation et Representation Multiple* / ed. by A. Ruas. — Paris: Hermes Lavoisier, 2002. — P. 25–44.

468. **Ruchikachorn P., Mueller K.** Learning Visualizations by Analogy: Promoting Visual Literacy through Visualization Morphing // *IEEE Transactions on Visualization and Computer Graphics*. — 2015. — Vol. 21, No. 9. — P. 1028–1044.
469. **Ryu Y.** Upscaling Land Surface Fluxes Through Hyper Resolution Remote Sensing in Space, Time, and the Spectrum // *Journal of Geophysical Research: Biogeosciences*. — 2024. — Vol. 129 (e2023JG007678), No. 10. — P. 1–19.
470. **Saalfeld A.** Conflation. Automated map compilation // *International Journal of Geographical Information Systems*. — 1988. — Vol. 2, No. 3. — P. 217–228.
471. **Sabins F. F. J., Ellis J. M.** *Remote sensing: Principles, Interpretation, and Applications*. — Fourth Edition ed. — Long Grove (Illinois): Waveland Press, Inc, 2020. — 524 p.
472. **Saga S. A.** Structural Knowledge to Support the Generalization of a Coast Line // *Proceedings of the 17th International Cartographic Conference*. — Barcelona, Spain: ICA, 1995. — P. 1491–1495.
473. **Samsonov T.** Multiscale Hypsometric Mapping // *Advances in Cartography and GIScience. Vol.1: Selection from ICC-2011, Paris*. Lecture Notes in Geoinformation and Cartography. / ed. by A. Ruas. — Berlin, Heidelberg: Springer, 2011. — P. 497–520.
474. **Samsonov T.** Morphometric Mapping of Topography by Flowline Hachures // *The Cartographic Journal*. — 2014. — Vol. 51, No. 1. — P. 63–74.
475. **Samsonov T.** Choropleth color scales for modifiable areal units // *Proceedings of XXVII International cartographic conference, 23–28 august 2015*. — Rio de Janeiro, Brazil: ICA, 2015. — P. 1–3.
476. **Samsonov T.** Generalize DEM: ArcGIS Python toolbox for automated structural generalization and conflation of digital elevation models // *Zenodo*. — 2020a.
477. **Samsonov T.** Optimal generalization of digital elevation model constrained by granularity measure // *Proceedings of the 30th International cartographic conference (ICC 2021)*. Abstracts of the international cartographic association. — Florence, Italy: Copernicus, 2021. — P. 255.
478. **Samsonov T.** Granularity of Digital Elevation Model and Optimal Level of Detail in Small-Scale Cartographic Relief Presentation // *Remote Sensing*. — 2022. — Vol. 14 (5), No. 1270. — P. 1–40.
479. **Samsonov T. E.** Automated Conflation of Digital Elevation Model with Reference Hydrographic Lines // *ISPRS International Journal of Geo-Information*. — 2020b. — Vol. 9 (5), No. 334. — P. 1–40.

480. **Samsonov T. E., Konstantinov P. I., Varentsov M. I.** Object-oriented approach to urban canyon analysis and its applications in meteorological modeling // *Urban Climate*. — 2015. — Vol. 13. — P. 122–139.
481. **Samsonov T. E., Varentsov M. I.** Computation of City-descriptive Parameters for High-resolution Numerical Weather Prediction in Moscow Megacity in the Framework of the COSMO Model // *Russian Meteorology and Hydrology*. — 2020. — Vol. 45, No. 7. — P. 515–521.
482. **Samsonov T. E., Yakimova O. P., Potemkin D. A., Guseva O. A.** Spatial variation of feature density in multiscale topographic data // *GEOGRAPHY, ENVIRONMENT, SUSTAINABILITY*. — 2023. — Vol. 16, No. 1. — P. 86–102.
483. **Samsonov T. E., Yakimova Olga. P.** Shape-Adaptive Geometric Simplification of Heterogeneous Line Datasets // *International Journal of Geographical Information Science*. — 2017. — Vol. 31, No. 8. — P. 1485–1520.
484. **Samsonov T., Konstantinov P.** OpenStreetMap data assessment for extraction of urban land cover and geometry parameters required by urban climate modeling // *Proceedings of GIScience 2014*. — Vienna, Austria: Vienna University of Technology, 2014. — P. 1–5.
485. **Samsonov T., Koshel S., Walther D., Jenny B.** Automated placement of supplementary contour lines // *International Journal of Geographical Information Science*. — 2019. — Vol. 33, No. 10. — P. 2072–2093.
486. **Samsonov T., Krivosheina A.** Joint generalization of city points and road network for small-scale mapping // *Proceedings of Seventh International Conference on Geographic Information Science GIScience*. — Columbus, Ohio: Ohio State University, 2012. — P. 1–7.
487. **Samsonov T., Walther D.** *Supplementary contours: ArcGIS Python toolbox for automated placement of supplementary contour lines* [Электронный ресурс]. URL: <https://zenodo.org/record/2647565> (дата обращения: 14.02.2025).
488. **Samsonov T., Yakimova O.** Regression modeling of reduction in spatial accuracy and detail for multiple geometric line simplification procedures // *International Journal of Cartography*. — 2020. — Vol. 6, No. 1. — P. 47–70.
489. **Sarjakoski T., Sarjakoski L. T.** A Real-Time Generalisation and Map Adaptation Approach for Location-Based Services // *Generalisation of Geographic Information: Cartographic Modelling and Applications*. Series of International Cartographic Association. / ed. by W. A. Mackaness, A. Ruas, L. T. Sarjakoski. — Amsterdam, The Netherlands: Elsevier, 2007. — P. 137–159.

490. **Saunders W.** Preparation of DEMs for use in environmental modeling analysis // *ESRI user conference*, 1999. — P. 24–30.
491. **Scarlato L. L.** An automatic critical line detector for digital elevation matrices // *Proceedings 1990 ACSM-ASPRS Annual Convention, Denver, Colorado*. — Denver, Colorado: ASPRS, 1990. — P. 43–52.
492. **Scarlato L., Pavlidis T.** Hierarchical triangulation using cartographic coherence // *CVGIP: Graphical Models and Image Processing*. — 1992. — Vol. 54, No. 2. — P. 147–161.
493. **Scheider S., Ballatore A.** Semantic typing of linked geoprocessing workflows // *International Journal of Digital Earth*. — 2018. — Vol. 11, No. 1. — P. 113–138.
494. **Scheider S., Gräler B., Pebesma E., Stasch C.** Modeling spatiotemporal information generation // *International Journal of Geographical Information Science*. — 2016. — Vol. 30, No. 10. — P. 1–29.
495. **Schmid F., Janatzek H. A.** A method for high-level street network extraction of OpenStreetMap data in OpenScienceMap // *Proceedings of the 26th International Cartographic Conference*. — Dresden, Germany: International Cartographic Association (ICA), 2013. — P. 1–14.
496. **Schmidt J., Hewitt A.** Fuzzy land element classification from DTMs based on geometry and terrain position // *Geoderma*. — 2004. — Vol. 121, No. 3–4. — P. 243–256.
497. **Schröder F., Roßbach P.** Managing the complexity of digital terrain models // *Computers & Graphics*. — 1994. — Vol. 18, No. 6. — P. 775–783.
498. **Serra J., Vincent L.** An overview of morphological filtering // *Circuits Systems and Signal Processing*. — 1992. — Vol. 11, No. 1. — P. 47–108.
499. **Sester M.** Optimization approaches for generalization and data abstraction // *International Journal of Geographical Information Science*. — 2005. — Vol. 19, No. 8–9. — P. 871–897.
500. **Shan J., Sampath A.** Urban Terrain and Building Extraction from Airborne LIDAR Data // *Urban Remote Sensing*. — Boca Raton: CRC Press, 2013. — P. 21–47.
501. **Shan J., Toth C. K.** *Topographic laser ranging and scanning: Principles and processing*. — Secon ed. — Boca Raton, FL: CRC Press, 2018. — 654 p.
502. **Shi W., Cheung C.** Performance Evaluation of Line Simplification Algorithms for Vector Generalization // *The Cartographic Journal*. — 2006. — Vol. 43, No. 1. — P. 27–44.
503. **Sibson R.** A brief description of natural neighbour interpolation // *Interpreting multivariate data* / ed. by V. Barnett. — Chichester: Wiley, 1981. — P. 21–36.

504. **Slocum T. A., McMaster R. B., Kessler F. C., Howard Hugh. H.** *Thematic Cartography and Geovisualization*. — 4 ed. — Boca Raton: CRC Press, 2022. — 612 p.
505. **Smith M. J. de, Goodchild M. F., Longley P. A.** *Geospatial Analysis: A Comprehensive Guide to Principles, Techniques and Software Tools*. — 6 ed. — Winchelsea, UK: The Winchelsea Press, 2018. — 618 p.
506. **Soille P.** Optimal removal of spurious pits in grid digital elevation models // *Water Resources Research*. — 2004. — Vol. 40, No. 12509. — P. 1–9.
507. **Soille P., Vogt J., Colombo R.** Carving and adaptive drainage enforcement of grid digital elevation models // *Water Resources Research*. — 2003. — Vol. 39 (12), No. 1366. — P. 1–13.
508. **Song H., Feng H.-Y.** A progressive point cloud simplification algorithm with preserved sharp edge data // *The International Journal of Advanced Manufacturing Technology*. — 2009. — Vol. 45, No. 5–6. — P. 583–592.
509. **Song J., Miao R.** A Novel Evaluation Approach for Line Simplification Algorithms towards Vector Map Visualization // *ISPRS International Journal of Geo-Information*. — 2016. — Vol. 5 (12), No. 223. — P. 1–13.
510. **Sonneveld M., Lguensat R., Jones D. C., Dueben P. D., Brajard J., Balaji V.** Bridging observations, theory and numerical simulation of the ocean using machine learning // *Environmental Research Letters*. — 2021. — Vol. 16 (7), No. 073008. — P. 1–28.
511. **Stanislowski L. V., Buttenfield B. P., Bereuter P., Savino S., Brewer C. A.** Generalisation Operators // *Advances in Cartography and GIScience. Volume 1*. — Cham: Springer International Publishing, 2014. — P. 157–195.
512. **Stanislowski L. V., Buttenfield B. P., Doumbouya A.** A rapid approach for automated comparison of independently derived stream networks // *Cartography and Geographic Information Science*. — 2015. — Vol. 42, No. 5. — P. 435–448.
513. **Stanislowski L. V., Nelson C., Hamann M.** Automated Conflation of Reach Data for the National Hydrography Dataset // *Proceedings of 22nd Annual Esri International User Conference*. — San Diego, CA, USA: Esri, 2002. — P. 1–10.
514. **Stanislowski L. V., Survila K., Wendel J., Liu Y., Buttenfield B. P.** An open source high-performance solution to extract surface water drainage networks from diverse terrain conditions // *Cartography and Geographic Information Science*. — 2018. — Vol. 45, No. 4. — P. 319–328.
515. **Stasch C., Scheider S., Pebesma E., Kuhn W.** Meaningful spatial prediction and aggregation // *Environmental Modelling and Software*. — 2014. — Vol. 51. — P. 149–165.

516. **Staufenbiel W.** *Zur Automation der Generalisierung topographischer Karten mit besonderer Berücksichtigung großmaßstäbiger Gebäudedarstellungen.* : Universität Hannover, 1973. — 164 p.
517. **Stewart I. D., Oke T. R.** Local climate zones for urban temperature studies // *Bulletin of the American Meteorological Society.* — 2012. — Vol. 93, No. 12. — P. 1879–1900.
518. **Stolte C., Tang D., Hanrahan P.** Multiscale visualization using data cubes // *IEEE Transactions on Visualization and Computer Graphics.* — 2003. — Vol. 9, No. 2. — P. 176–187.
519. **Strand G.-H., Heggem E. S. F., Aune-Lundberg L., Hościło A., Waśniewski A.** Transformation of a Classified Image from Pixel Clutter to Land Cover Map Using Geometric Generalization and Thematic Self-Enrichment // *Geomatics.* — 2024. — Vol. 4, No. 3. — P. 271–285.
520. **Sun J., Shen Z., Wang Y., Bao H., Zhou X.** LoFTR: Detector-Free Local Feature Matching with Transformers // *2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).* — Nashville, TN, USA: IEEE, 2021. — P. 8918–8927.
521. **Sun X., Rosin P. L., Martin R., Langbein F.** Fast and Effective Feature-Preserving Mesh Denoising // *IEEE Transactions on Visualization and Computer Graphics.* — 2007. — Vol. 13, No. 5. — P. 925–938.
522. **Tableau Software I.** *Tableau Software* [Электронный ресурс]. URL: <http://www.tableau.com> (дата обращения: 16.02.2025).
523. **Tang J., Deng M., Peng J., Liu H., Yang X., Chen X.** Automatic road network selection method considering functional semantic features of roads with graph convolutional networks // *International Journal of Geographical Information Science.* — 2024. — Vol. 38, No. 11. — P. 2403–2432.
524. **Tang L.** Automatic Extraction of Specific Geomorphological Elements from Contours // *Proceedings of the 5th International Symposium on Spatial Data Handling.* — Charleston, S.C: USA, 1992. — P. 544–566.
525. **Tay L. T., Sagar B. S. D., Chuah H. K.** Analysis of Geophysical Networks Derived From Multiscale Digital Elevation Models: A Morphological Approach // *IEEE Geoscience and Remote Sensing Letters.* — 2005. — Vol. 2, No. 4. — P. 399–403.
526. **Thenkabail P. S.** *Remote Sensing Handbook, Volume II: Image Processing, Change Detection, GIS, and Spatial Data Analysis* / ed. by P. S. Thenkabail. — Second ed. — Boca Raton, FL: CRC Press, 2024. — 464 p.

527. **Thompson J. F., Soni B. K., Weatherill N. P.** *Handbook of grid generation*. — Boca Raton, Fla: CRC Press, 1999. — 1096 p.
528. **Thomson R., Brooks R.** Chapter 13 - Generalisation of Geographical Networks // *Generalization of Geographic Information: Cartographic Modelling and Applications* / ed. by W. Mackaness, A. Ruas, L. T. Sarjakoski. — Amsterdam: Elsevier Science B.V., 2007. — P. 255–267.
529. **Tobler W. R.** Automation in the Preparation of Thematic Maps // *The Cartographic Journal*. — 1965. — Vol. 2, No. 1. — P. 32–38.
530. **Tomlin D.** *GIS and Cartographic Modeling*. — Second ed. — Redlands, CA: ESRI Press, 2012. — 204 p.
531. **Töpfer F., Pillewizer W.** The Principles of Selection // *The Cartographic Journal*. — 1966. — Vol. 3, No. 1. — P. 10–16.
532. **Torun A., Köbben B., Lemmens R.** Processing Spatial Data On the Internet // *International Archives of Photogrammetry and Remote Sensing*. — 2000. — Vol. 33, No. B6. — P. 269–278.
533. **Touya G.** A Road Network Selection Process Based on Data Enrichment and Structure Detection: Road Network Selection // *Transactions in GIS*. — 2010. — Vol. 14, No. 5. — P. 595–614.
534. **Touya G., Baley M.** Level of Details Harmonization Operations in OpenStreetMap Based Large Scale Maps // *Citizen Empowered Mapping*. — Berlin: Springer International Publishing AG, 2017. — P. 3–25.
535. **Touya G., Brando-Escobar C.** Detecting Level-of-Detail Inconsistencies in Volunteered Geographic Information Data Sets // *Cartographica: The International Journal for Geographic Information and Geovisualization*. — 2013. — Vol. 48, No. 2. — P. 134–143.
536. **Touya G., Bucher B., Falquet G., Jaara K., Steiniger S.** Modelling Geographic Relationships in Automated Environments // *Abstracting Geographic Information in a Data Rich World: Methodologies and Applications of Map Generalisation* / ed. by D. Burghardt, C. Duchêne, W. Mackaness. — Cham: Springer International Publishing, 2014. — P. 53–82.
537. **Touya G., Duchêne C., Ruas A.** Collaborative Generalisation: Formalisation of Generalisation Knowledge to Orchestrate Different Cartographic Generalisation Processes // *Theories and Methods of Spatio-Temporal Reasoning in Geographic Space*. — Berlin, Heidelberg: Springer, 2010. — P. 264–278.

538. **Touya G., Gruget M., Muehlenhaus I.** Where Am I Now? Modelling Disorientation in Pan-Scalar Maps // *ISPRS International Journal of Geo-Information*. — 2023. — Vol. 12, No. 2. — P. 62.
539. **Touya G., Potié Q., Mackaness W. A.** Incorporating ideas of structure and meaning in interactive multi scale mapping environments // *International Journal of Cartography*. — 2023. — Vol. 9, No. 2. — P. 342–372.
540. **Touya G., Reimer A.** Inferring the Scale of OpenStreetMap Features // *OpenStreetMap in GIScience: Experiences, Research, Applications*. Lecture Notes in Geoinformation and Cartography. — Berlin, Heidelberg: Springer, 2015. — P. 81–99.
541. **Tripathi P., Patel N. R., Kushwaha S. P. S., Dadhwal V. K.** Upscaling of leaf area index in Terai forest plantations using fine- and moderate-resolution satellite data // *International Journal of Remote Sensing*. — 2014. — Vol. 35, No. 22. — P. 7749–7762.
542. **Tufte E. R.** *The visual display of quantitative Information*. — Second ed. — Cheshire, Conn: Graphics Press, 2001. — 197 p.
543. **Tufte E. R.** *Beautiful evidence*. — Cheshire, Conn: Graphics Press, 2006. — 213 p.
544. **Turk G., Banks D.** Image-guided streamline placement // *Proceedings of SIGGRAPH96 the 23rd annual conference on Computer graphics and interactive techniques*. — New Orleans, Louisiana: ACM, 1996. — P. 453–460.
545. **Tversky B.** Parts, paronomies, and taxonomies. // *Developmental Psychology*. — 1989. — Vol. 25, No. 6. — P. 983–995.
546. **USGS.** *GTOPO30* [Электронный ресурс]. URL: <https://www.usgs.gov/centers/eros/science/usgs-eros-archive-digital-elevation-global-30-arc-second-elevation-gtopo30> (дата обращения: 16.02.2025).
547. **USGS.** *Topographic Map Symbols*. — Reston, VA: U.S. Geological Survey, 2005. — 4 p.
548. **Van Oosterom P.** Variable-scale Topological Data Structures Suitable for Progressive Data Transfer: The GAP-face Tree and GAP-edge Forest // *Cartography and Geographic Information Science*. — 2005. — Vol. 32, No. 4. — P. 331–346.
549. **Vangenot C.** Multi-representation in spatial databases using the MADS conceptual model // *Proceedings of the 8th ICA Workshop on Generalisation and Multiple Representation*. — Leicester, UK: ICA, 2004. — P. 1–8.

550. **Varentsov M. I., Konstantinov P. I., Shartova N. V., Samsonov T. E., Kargashin P. E., Varentsov A. I., Fenner D., Meier F.** Urban heat island of the Moscow megacity: The long-term trends and new approaches for monitoring and research based on crowdsourcing data // *IOP Conference Series: Earth and Environmental Science*. — 2020. — Vol. 606, No. 012063. — P. 1–10.
551. **Varentsov M., Fenner D., Meier F., Samsonov T., Demuzere M.** Quantifying Local and Mesoscale Drivers of the Urban Heat Island of Moscow with Reference and Crowdsourced Observations // *Frontiers in Environmental Science*. — 2021. — Vol. 9, No. 716968. — P. 1–21.
552. **Varentsov M., Wouters H., Platonov V., Konstantinov P.** Megacity-Induced Mesoclimatic Effects in the Lower Atmosphere: A Modeling Study for Multiple Summers over Moscow, Russia // *Atmosphere*. — 2018. — Vol. 9 (2), No. 50. — P. 1–24.
553. **Veltman N.** *Flubber: Tools for smoother shape animations* [Электронный ресурс]. URL: <https://github.com/veltman/flubber> (дата обращения: 16.02.2025).
554. **Veregin H.** Quantifying positional error induced by line simplification // *International Journal of Geographical Information Science*. — 2000. — Vol. 14, No. 2. — P. 113–130.
555. **Villoslada M., Berner L. T., Juutinen S., Yläne H., Kumpula T.** Upscaling vascular aboveground biomass and topsoil moisture of subarctic fens from Unoccupied Aerial Vehicles (UAVs) to satellite level // *Science of The Total Environment*. — 2024. — Vol. 933, No. 173049. — P. 1–15.
556. **Visvalingam M.** The Visvalingam Algorithm: Metrics, Measures and Heuristics // *The Cartographic Journal*. — 2016. — Vol. 53, No. 3. — P. 242–252.
557. **Visvalingam M., Whyatt J. D.** Line Generalisation by Repeated Elimination of Points // *The Cartographic Journal*. — 1993. — Vol. 30, No. 1. — P. 46–51.
558. **Von Herzen B., Barr A. H.** Accurate triangulations of deformed, intersecting surfaces // *ACM SIGGRAPH Computer Graphics*. — 1987. — Vol. 21, No. 4. — P. 103–110.
559. **Voronoi G.** Nouvelles applications des paramètres continus à la théorie des formes quadratiques. Premier mémoire. Sur quelques propriétés des formes quadratiques positives parfaites. // *Journal für die reine und angewandte Mathematik (Crelles Journal)*. — 1908. — Vol. 1908, No. 133. — P. 97–102.
560. **Wallgrün J. O., Wolter D., Richter K.-F.** Qualitative matching of spatial information // *Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems - GIS '10*. — San Jose, California: ACM Press, 2010. — P. 300–309.

561. **Walter V., Fritsch D.** Matching spatial data sets: A statistical approach // *International Journal of Geographical Information Science*. — 1999. — Vol. 13, No. 5. — P. 445–473.
562. **Wang B., Chen G., Bu J., Yu Y.** Multiscale Visualization Of Relational Databases Using Layered Zoom Trees And Partial Data Cubes // *Proceedings of International Conference on Information Visualization Theory and Applications*. — Angers: France, 2010. — P. 101–111.
563. **Wang D., Tang Y.** A one-parameter Budyko model for water balance captures emergent behavior in darwinian hydrologic models // *Geophysical Research Letters*. — 2014. — Vol. 41, No. 13. — P. 4569–4577.
564. **Wang L., Liu H.** An efficient method for identifying and filling surface depressions in digital elevation models for hydrologic analysis and modelling // *International Journal of Geographical Information Science*. — 2006. — Vol. 20, No. 2. — P. 193–213.
565. **Wang Y. H., Chen J., Jian J.** Multi-Scale Conceptual Model for GIS Geographical Features // *Journal of China University of Mining Technology*. — 2003. — Vol. 32, No. 4. — P. 376–382.
566. **Wang Y., Di Q.** Modifiable areal unit problem and environmental factors of COVID-19 outbreak // *Science of The Total Environment*. — 2020. — Vol. 740, No. 139984. — P. 1–5.
567. **Wang Z., Müller J.-C.** Line Generalization Based on Analysis of Shape Characteristics // *Cartography and Geographic Information Science*. — 1998. — Vol. 25, No. 1. — P. 3–15.
568. **WAVEWATCH. Version III®** [Электронный ресурс]. URL: <https://github.com/noaa-emc/ww3> (дата обращения: 10.02.2025).
569. **Wechsler S. P.** Perceptions by Digital Elevation Model Users of DEM Uncertainty // *URISA Journal*. — 2003. — Vol. 15, No. 2. — P. 61–69.
570. **Weibel R.** An Adaptive Methodology for Automated Relief Generalization // *Proceedings of AutoCarto 8*. — Baltimore, Maryland: USA, 1987. — P. 42–49.
571. **Weibel R.** Models and Experiments for Adaptive Computer-Assisted Terrain Generalization // *Cartography and Geographic Information Science*. — 1992. — Vol. 19, No. 3. — P. 133–153.
572. **Weibel R.** A Typology of Constraints to Line Simplification // *Proceedings of 7th International Symposium on Spatial Data Handling, 12-16 August 1996*. — Delft: The Netherlands, 1996. — P. 533–546.
573. **Weidner L., Walton G., Kromer R.** Generalization considerations and solutions for point cloud hillslope classifiers // *Geomorphology*. — 2020. — Vol. 354, No. 107039. — P. 1–14.

574. **Weinman J.** Toponym Recognition in Historical Maps by Gazetteer Alignment // *2013 12th International Conference on Document Analysis and Recognition*. — Washington, DC, USA: IEEE, 2013. — P. 1044–1048.
575. **Weinman J.** Geographic and Style Models for Historical Map Alignment and Toponym Recognition // *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)*. — Kyoto: IEEE, 2017. — P. 957–964.
576. **Weiss R., Weibel R.** Road network selection for small-scale maps using an improved centrality-based algorithm // *Journal of Spatial Information Science*. — 2014. — No. 9. — P. 71–99.
577. **White E. R.** Assessment of Line-Generalization Algorithms Using Characteristic Points // *The American Cartographer*. — 1985. — Vol. 12, No. 1. — P. 17–28.
578. **White M. S., Griffin P.** Piecewise Linear Rubber-Sheet Map Transformation // *The American Cartographer*. — 1985. — Vol. 12, No. 2. — P. 123–131.
579. **Williams M., Bell R., Spadavecchia L., Street L. E., Van Wijk M. T.** Upscaling leaf area index in an Arctic landscape through multiscale observations // *Global Change Biology*. — 2008. — Vol. 14, No. 7. — P. 1517–1530.
580. **Wolpert D. H.** Stacked generalization // *Neural Networks*. — 1992. — Vol. 5, No. 2. — P. 241–259.
581. **Wolter D., Blank D., Henrich A.** Georeferencing River Networks Using Spatial Reasoning // *Proceedings of the 11th Workshop on Geographic Information Retrieval*. — Heidelberg Germany: ACM, 2017. — P. 1–2.
582. **Wood J.** Scale-based Characterization of Digital Elevation Models // *Innovations in GIS*. — Boca Raton: CRC Press, 1996. — P. 163–175.
583. **Wouters H., Demuzere M., Blahak U., Fortuniak K., Maiheu B., Camps J., Tielemans D., Van Lipzig N. P. M.** The efficient urban canopy dependency parametrization (SURY) v1.0 for atmospheric modelling: Description and application with the COSMO-CLM model for a Belgian summer // *Geoscientific Model Development*. — 2016. — Vol. 9, No. 9. — P. 3027–3054.
584. **Wu F.** A Model for Multi-scale Handling and Representation of Relief Based on Wavelet Analysis // *Proceedings of the 3th Surveying and mapping workshop across strait*. — Hong Kong: China, 2000. — P. 435–441.
585. **Wu H. H.** Prinzip und Methode der automatischen Generalisierung der Reliefformen // *Nachrichten aus dem Karten- und Vermessungswesen. Series 1*. — 1981. — Vol. 85. — P. 163–174.

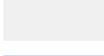
586. **Wu H., Kimball J. S., Li H., Huang M., Leung L. R., Adler R. F.** A new global river network database for macroscale hydrologic modeling // *Water Resources Research*. — 2012. — Vol. 48 (W09701), No. 9. — P. 1–5.
587. **Wu H., Kimball J. S., Mantua N., Stanford J.** Automated upscaling of river networks for macroscale hydrological modeling // *Water Resources Research*. — 2011. — Vol. 47 (W03517), No. 3. — P. 1–18.
588. **Wu T., Li J., Li T., Sivakumar B., Zhang G., Wang G.** High-efficient extraction of drainage networks from digital elevation models constrained by enhanced flow enforcement from known river maps // *Geomorphology*. — 2019. — Vol. 340. — P. 184–201.
589. **Xiao T., Ai T., Yu H., Yang M., Liu P.** A point selection method in map generalization using graph convolutional network model // *Cartography and Geographic Information Science*. — 2024. — Vol. 51, No. 1. — P. 20–40.
590. **Yadav B., Hatfield K.** Stream network conflation with topographic DEMs // *Environmental Modelling & Software*. — 2018. — Vol. 102. — P. 241–249.
591. **Yamazaki D., Ikeshima D., Tawatari R., Yamaguchi T., O’Loughlin F., Neal J. C., Sampson C. C., Kanae S., Bates P. D.** A high-accuracy map of global terrain elevations: Accurate Global Terrain Elevation map // *Geophysical Research Letters*. — 2017. — Vol. 44, No. 11. — P. 5844–5853.
592. **Yan H.** *Description Approaches and Automated Generalization Algorithms for Groups of Map Objects*. — Singapore: Springer Singapore, 2019. — 228 p.
593. **Yan H., Weibel R.** An algorithm for point cluster generalization based on the Voronoi diagram // *Computers & Geosciences*. — 2008. — Vol. 34, No. 8. — P. 939–954.
594. **Yan X., Yang M.** A deep learning approach for polyline and building simplification based on graph autoencoder with flexible constraints // *Cartography and Geographic Information Science*. — 2024. — Vol. 51, No. 1. — P. 79–96.
595. **Yang M., Huang H., Zhang Y., Yan X.** Pattern Recognition and Segmentation of Administrative Boundaries Using a One-Dimensional Convolutional Neural Network and Grid Shape Context Descriptor // *ISPRS International Journal of Geo-Information*. — 2022. — Vol. 11, No. 9. — P. 461.
596. **Yilmaz I., Gullu M.** Georeferencing of Historical Maps Using Back Propagation Artificial Neural Network // *Experimental Techniques*. — 2012. — Vol. 36, No. 5. — P. 15–19.

597. **Yin Z.** A multi-scale GIS database model based on Petri net // *Proceedings of ISPRS Workshop on Service and Application of Spatial Data Infrastructure*. — Hangzhou, China: ISPRS, 2005. — P. 271–276.
598. **Yoeli P.** Topographical Relief Depiction by Hachures with Computer and Plotter // *The Cartographic Journal*. — 1985. — Vol. 22, No. 2. — P. 111–124.
599. **Yoeli P.** Entwurf einer Methodologie für computergestütztes kartographisches Generalisieren topographischer Reliefs (Concept of a computer-assisted method for the cartographic generalisation of topographic relief) // *Kartographisches Generalisieren*. — Zürich: Switzerland, 1990. — P. 57–60.
600. **Yu W., Chen Y.** Data-driven polyline simplification using a stacked autoencoder-based deep neural network // *Transactions in GIS*. — 2022. — Vol. 26, No. 5. — P. 2302–2325.
601. **Yuan R., Kennedy D. M., Ierodiaconou D. A.** Spatial scaling in geomorphology: Extrapolating micro-erosion measurements to the landform scale on shore platforms using unoccupied aerial vehicles // *Geomorphology*. — 2022. — Vol. 408, No. 108252. — P. 1–13.
602. **Yue T.-X., Zhang L.-L., Zhao N., Zhao M.-W., Chen C.-F., Du Z.-P., Song D.-J., Fan Z.-M., Shi W.-J., Wang S.-H., Yan C.-Q., Li Q.-Q., Sun X.-F., Yang H., Wilson J., Xu B.** A review of recent developments in HASM // *Environmental Earth Sciences*. — 2015. — Vol. 74, No. 8. — P. 6541–6549.
603. **Zakšek K., Podobnikar T.** An effective DEM generalization with basic GIS operations // *Proceedings of 12th ICA Workshop on Generalisation and Multiple Representation, 20-21 June 2008*. — La Coruña, Spain: ICA, 2005. — P. 1–10.
604. **Zen M., Candiago S., Schirpke U., Egarter Vigl L., Giupponi C.** Upscaling ecosystem service maps to administrative levels: Beyond scale mismatches // *Science of The Total Environment*. — 2019. — Vol. 660. — P. 1565–1575.
605. **Zhang B., Fan Z., Du Z., Zheng J., Luo J., Wang N., Wang Q.** A Geomorphological Regionalization using the Upscaled DEM: The Beijing-Tianjin-Hebei Area, China Case Study // *Scientific Reports*. — 2020. — Vol. 10, No. 10532. — P. 1–12.
606. **Zhang C., Pan M., Wu H., Xu H.** Study on simplification of contour lines preserving topological coherence // *Acta Scientiarum Naturalium Universitatis Pekinensis*. — 2007. — Vol. 43, No. 2. — P. 216–222.
607. **Zhang D., Zhang W., Huang W., Hong Z., Meng L.** Upscaling of Surface Soil Moisture Using a Deep Learning Model with VIIRS RDR // *ISPRS International Journal of Geo-Information*. — 2017. — Vol. 6 (5), No. 130. — P. 1–20.

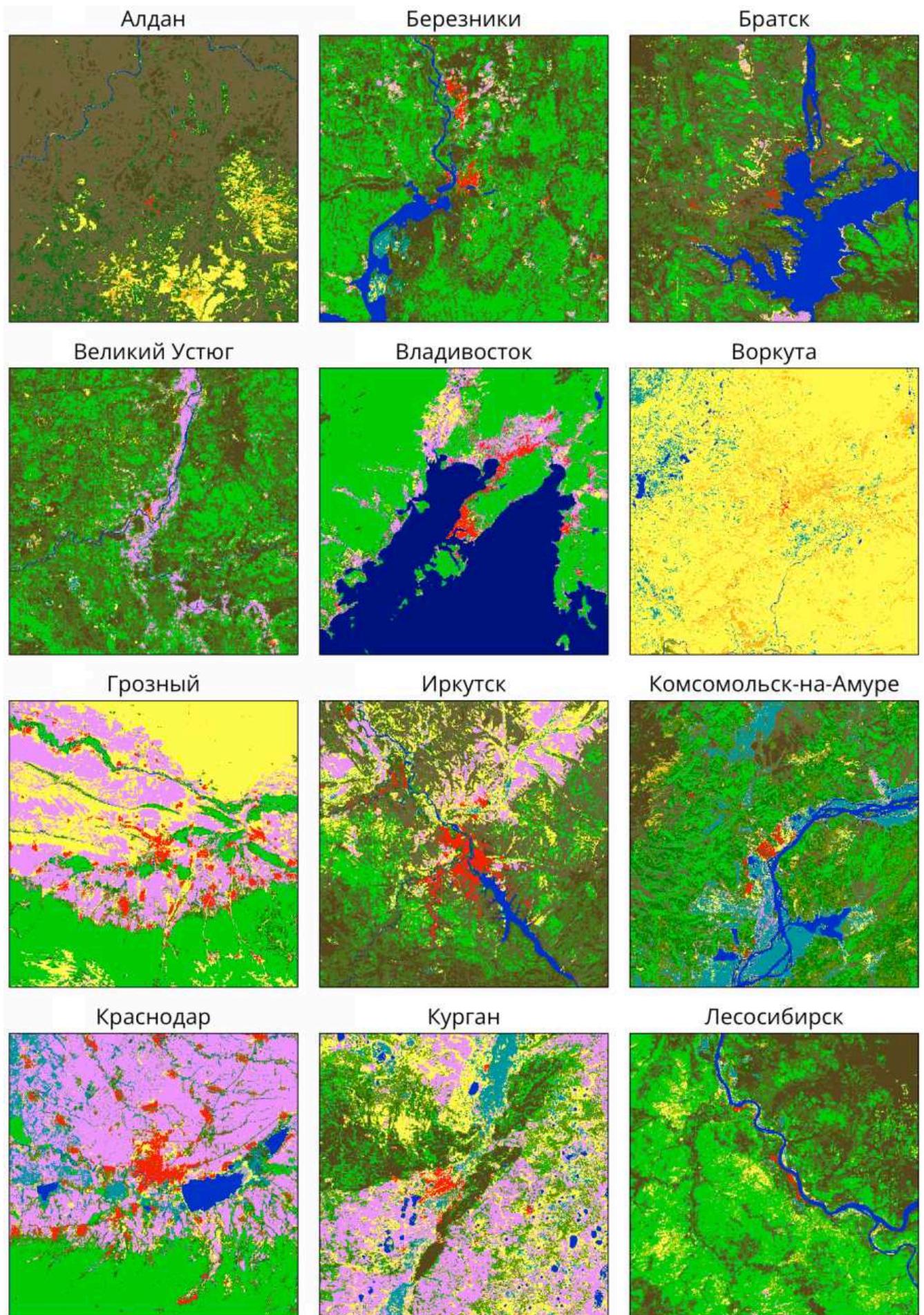
608. **Zhang M., Kukadia N.** Metrics of Urban Form and the Modifiable Areal Unit Problem // *Transportation Research Record: Journal of the Transportation Research Board.* — 2005. — No. 1902. — P. 71–79.
609. **Zhang Q., Huang Z., Xu Z., Shang H.** Study on Sampling Rule and Simplification of LiDAR Point Cloud Based on Terrain Complexity // *Journal of the Indian Society of Remote Sensing.* — 2018. — Vol. 46, No. 11. — P. 1773–1784.
610. **Zhang Z., Lin Y.** Data-Driven Seismic Waveform Inversion: A Study on the Robustness and Generalization // *IEEE Transactions on Geoscience and Remote Sensing.* — 2020. — Vol. 58, No. 10. — P. 6900–6913.
611. **Zhao Z., Saalfeld A.** Linear-time sleeve-fitting polyline // *Autocarto 13, ACSM/ASPRS'97 Technical Papers.* — Seattle, Washington: ASPRS, 1997. — P. 214–223.
612. **Zheng J., Gao Z., Ma J., Shen J., Zhang K.** Deep Graph Convolutional Networks for Accurate Automatic Road Network Selection // *ISPRS International Journal of Geo-Information.* — 2021. — Vol. 10 (11), No. 768. — P. 1–22.
613. **Zhi P., Jiang L., Yang X., Wang X., Li H.-W., Zhou Q., Li K.-C., Ivanović M.** Cross-Domain Generalization for LiDAR-Based 3D Object Detection in Infrastructure and Vehicle Environments // *Sensors.* — 2025. — Vol. 25 (3), No. 767. — P. 1–20.
614. **Zhou Q., Chen Y.** Generalization of DEM for terrain analysis using a compound method // *ISPRS Journal of Photogrammetry and Remote Sensing.* — 2011. — Vol. 66, No. 1. — P. 38–45.
615. **Zhou Q., Liu X., Sun Y.** Terrain complexity and uncertainties in grid-based digital terrain analysis // *International Journal of Geographical Information Science.* — 2006. — Vol. 20, No. 10. — P. 1137–1147.
616. **Zhou S., Jones C. B.** A Multi-representation Spatial Data Model // *Advances in Spatial and Temporal Databases* / ed. by T. Hadzilacos, Y. Manolopoulos, J. Roddick, Y. Theodoridis. — Berlin, Heidelberg: Springer, 2003. — P. 394–411.
617. **Zhu W., Hayashibe M.** A Hierarchical Deep Reinforcement Learning Framework With High Efficiency and Generalization for Fast and Safe Navigation // *IEEE Transactions on Industrial Electronics.* — 2023. — Vol. 70, No. 5. — P. 4962–4971.

# Приложение

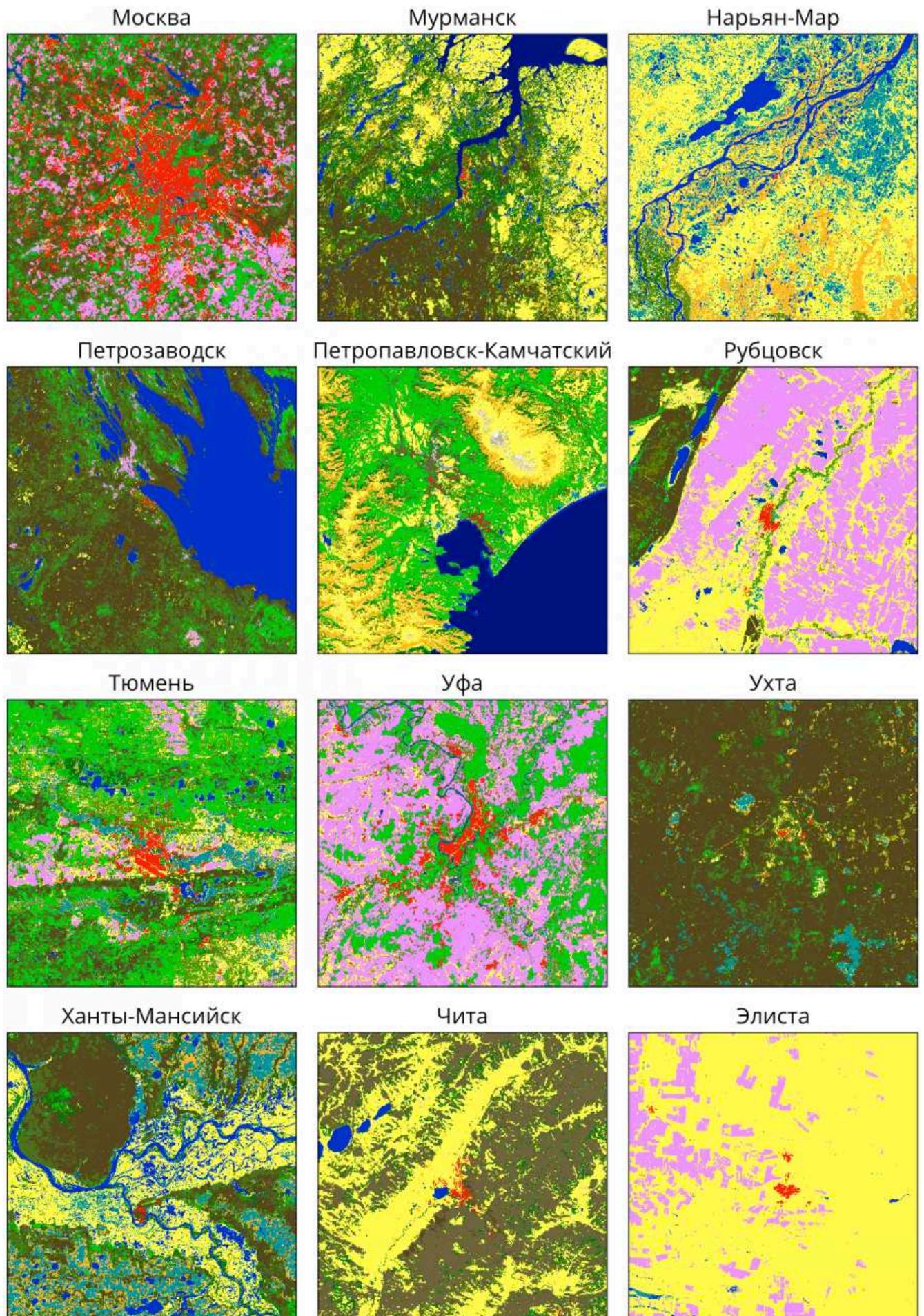
**Таблица А.1** – Типы наземного покрова по данным Copernicus Global Land Cover (CGLC)

Тип	Цвет	Описание
0		Входные данные отсутствуют
20		Кустарники
30		Травянистая растительность
40		Возделываемая и управляемая растительность/сельское хозяйство (пахотные земли)
50		Городской / застроенный
60		Голая / редкая растительность
70		Снег и лед
80		Постоянные водные объекты
90		Травянистые водно-болотные угодья
100		Мох и лишайник
111		Густой лес, вечнозеленый хвойный
112		Густой лес, вечнозеленый лиственный
113		Густой лес, листопадный хвойный
114		Густой лес, листопадный лиственный
115		Густой лес, смешанный
116		Густой лес, неизвестно
121		Редкий лес, вечнозеленый хвойный
122		Редкий лес, вечнозеленый лиственный
123		Редкий лес, листопадный хвойный
124		Редкий лес, листопадный лиственный
125		Редкий лес, смешанный

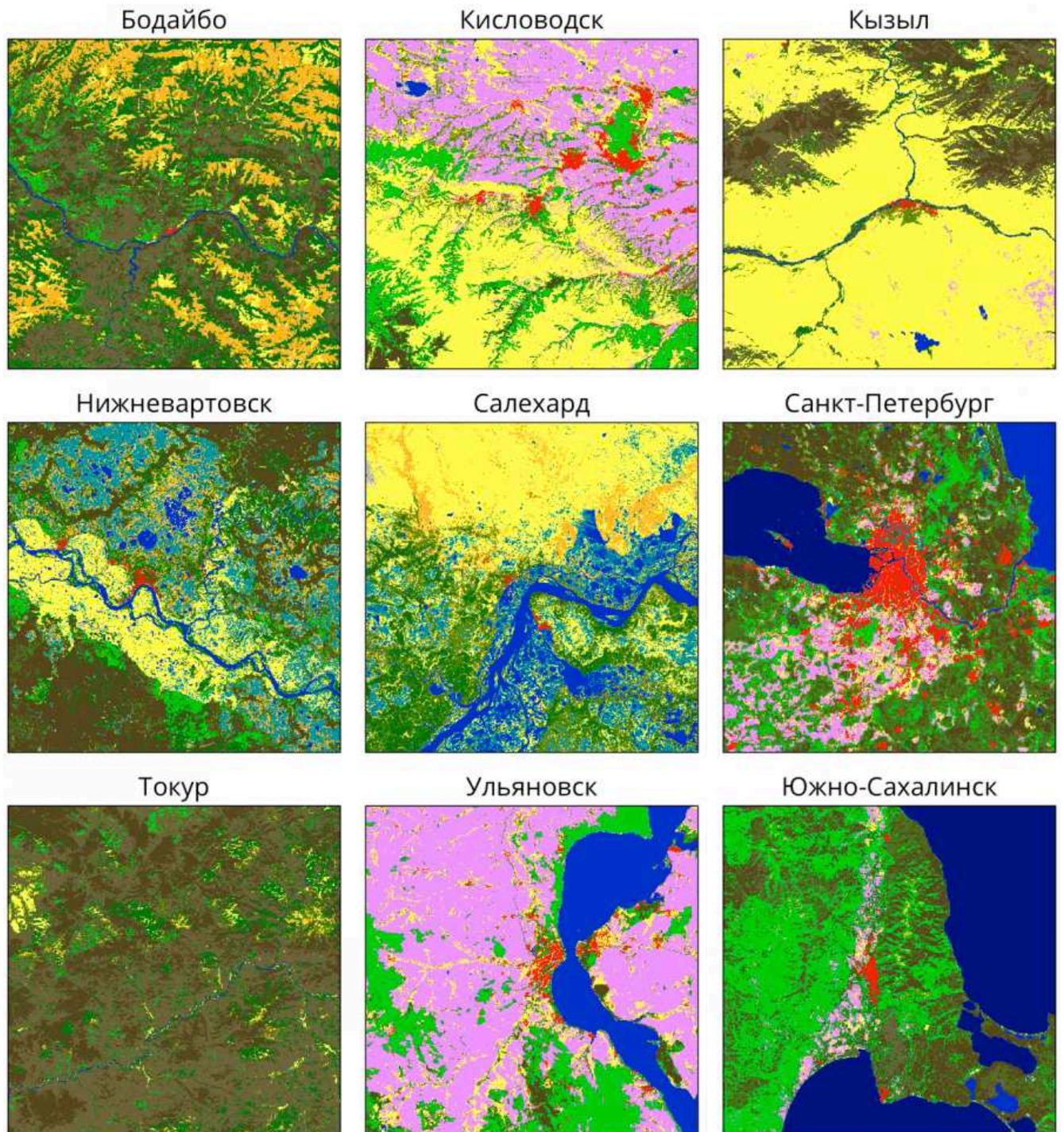
Тип	Цвет	Описание
126		Редкий лес, неизвестно
200		Открытое море



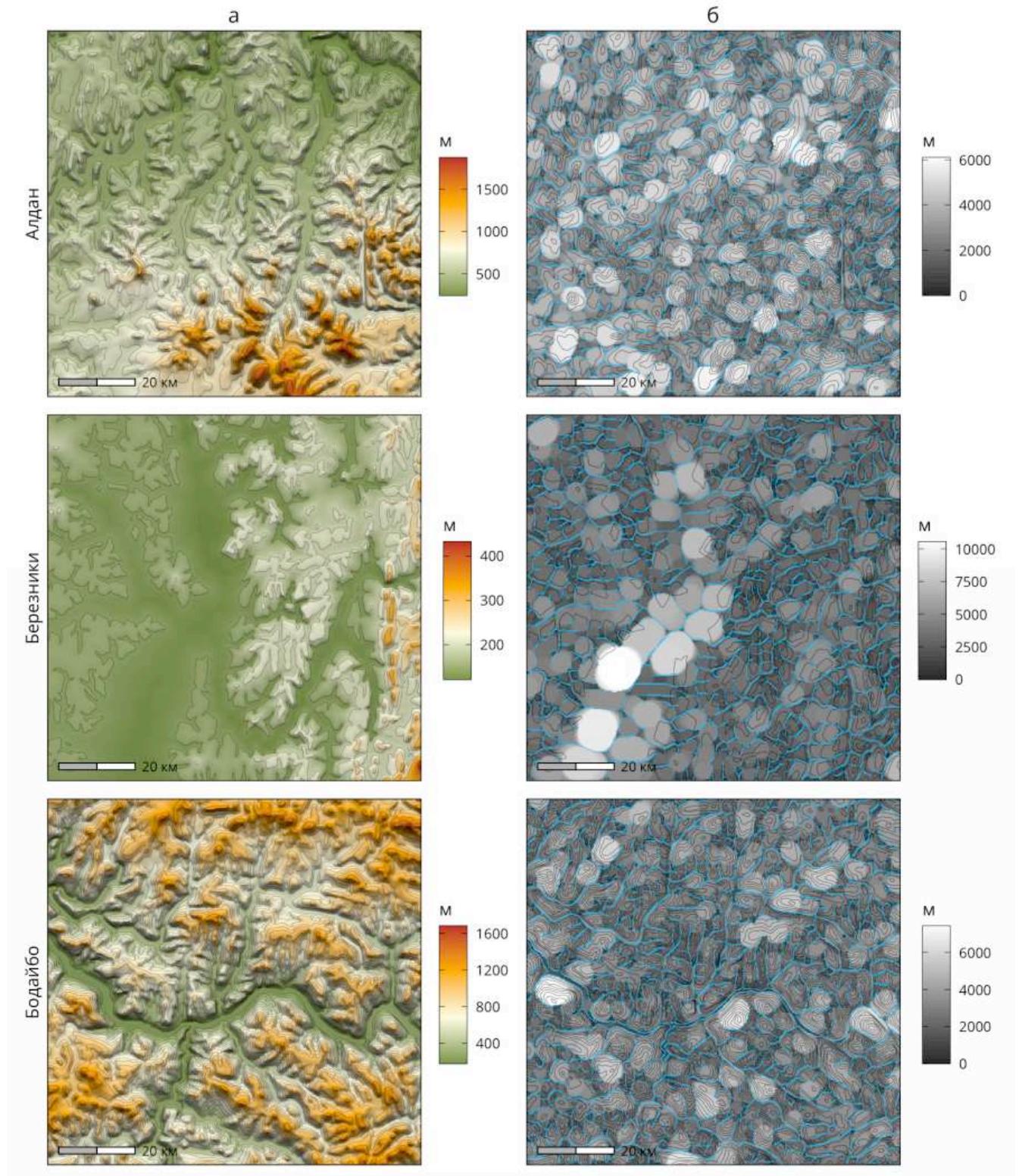
**Рисунок А.1** – Типы наземного покрова обучающих фрагментов 1-12 для исследования плотности пространственных данных



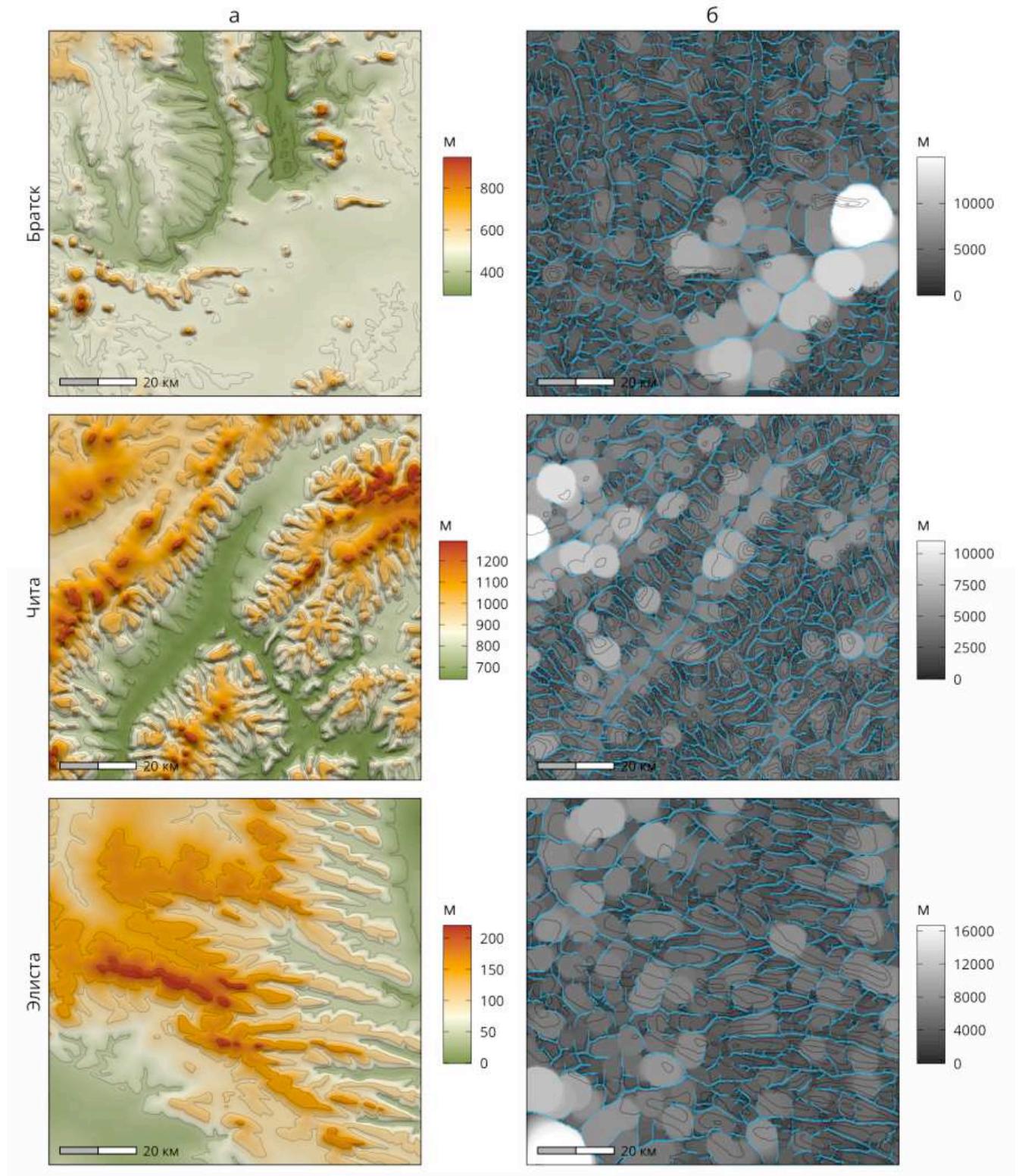
**Рисунок А.2** — Типы наземного покрова обучающих фрагментов 1-12 для исследования плотности пространственных данных



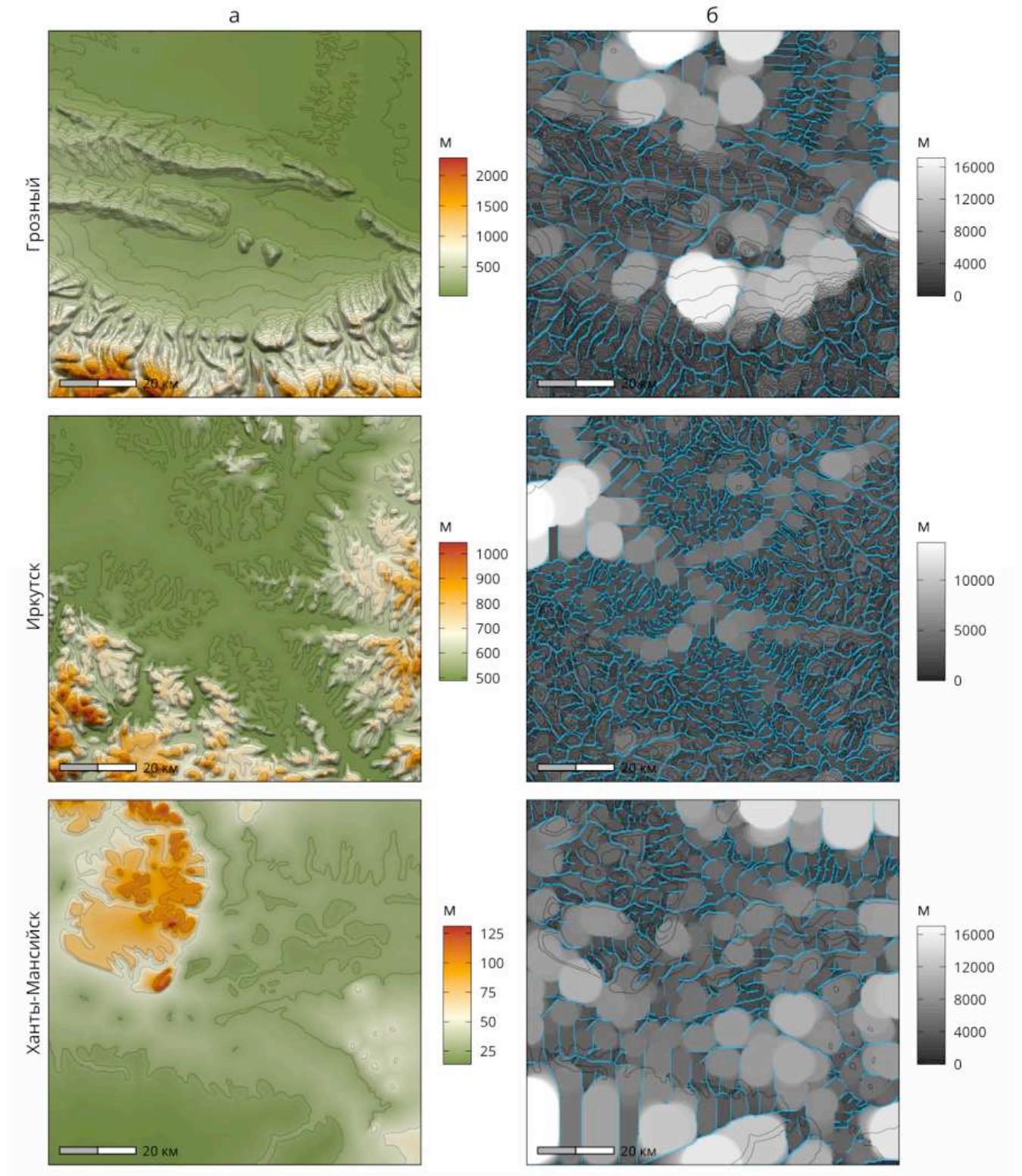
**Рисунок А.3** — Типы наземного покрова тестовых фрагментов 1-9 для исследования плотности пространственных данных



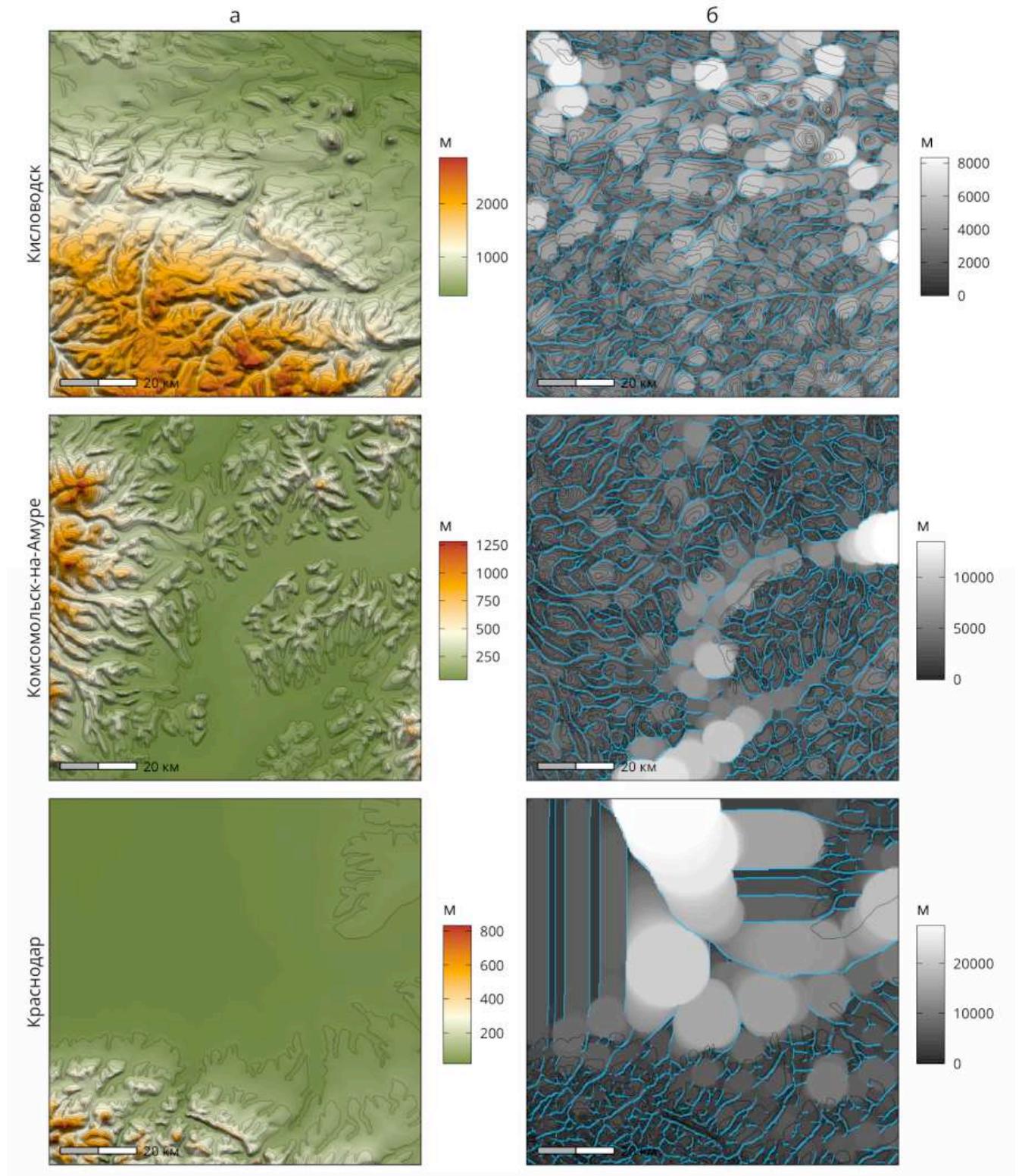
**Рисунок А.4** – Фрагменты ЦМР 1-3 для исследования гранулярности: (а) абсолютная высота; (б) ширина формы рельефа (уровень детализации 1000)



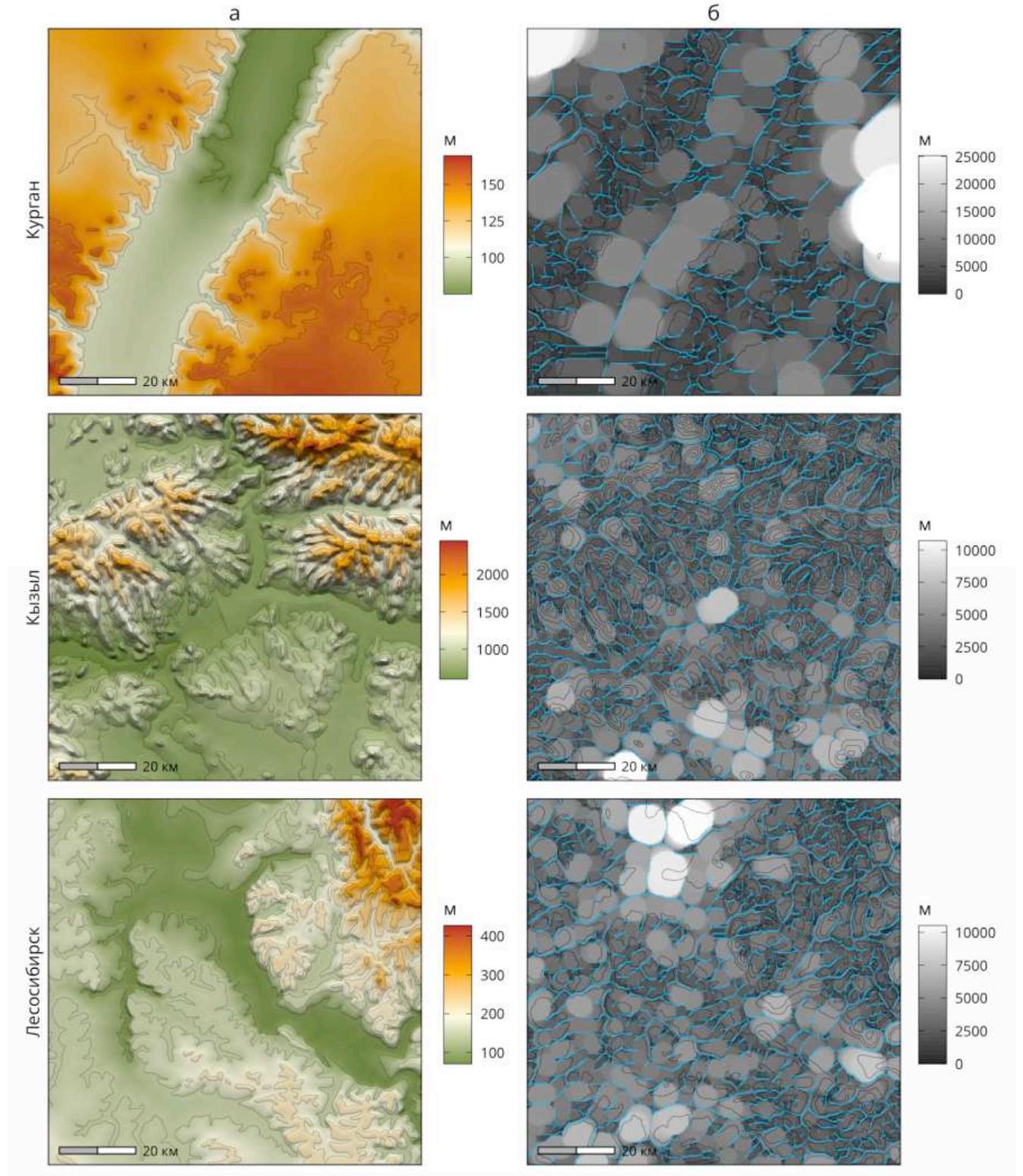
**Рисунок А.5** – Фрагменты ЦМР 4-6 для исследования гранулярности: (а) абсолютная высота; (б) ширина формы рельефа (уровень детализации 1000)



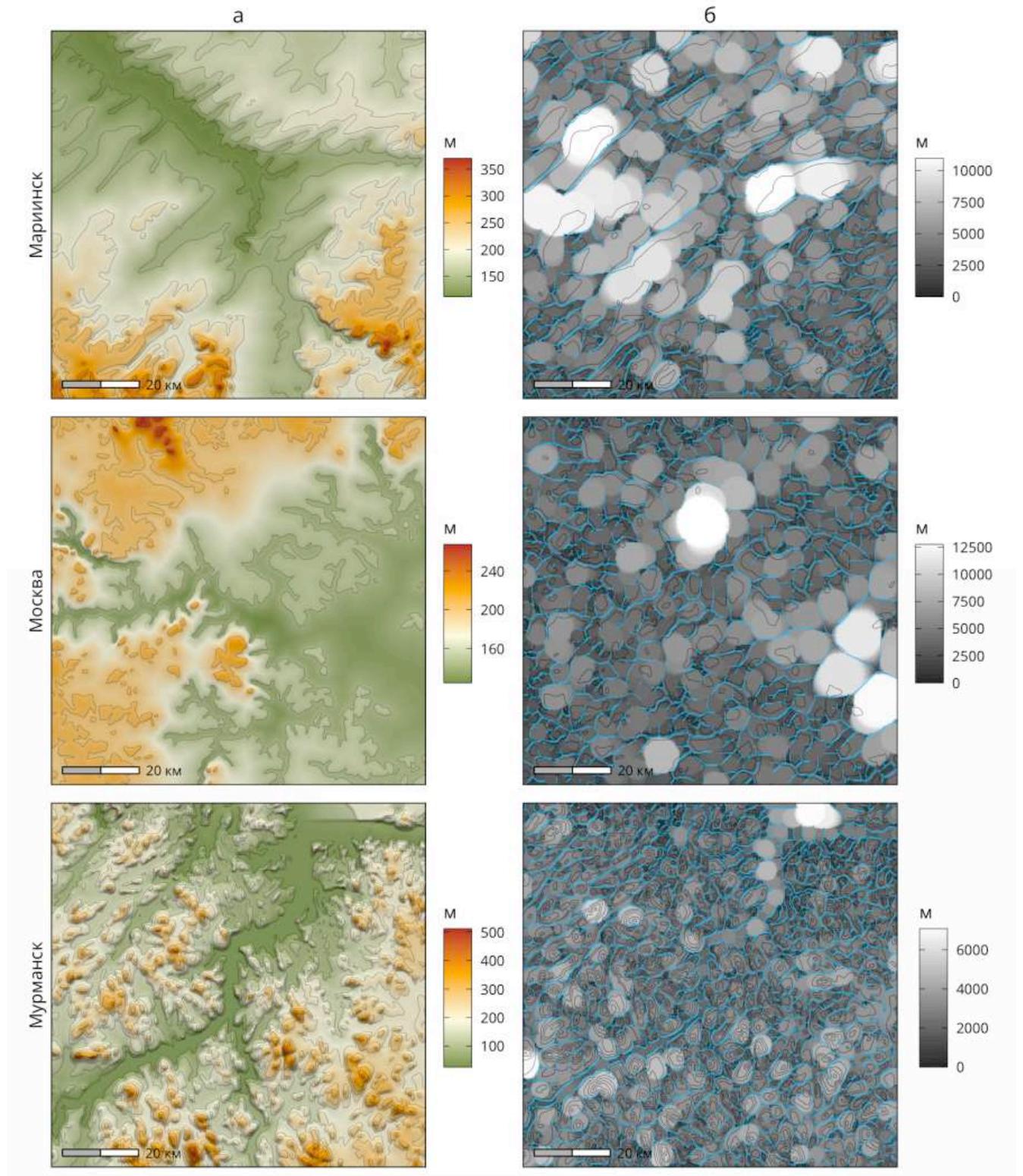
**Рисунок А.6** – Фрагменты ЦМР 7-9 для исследования гранулярности: (а) абсолютная высота; (б) ширина формы рельефа (уровень детализации 1000)



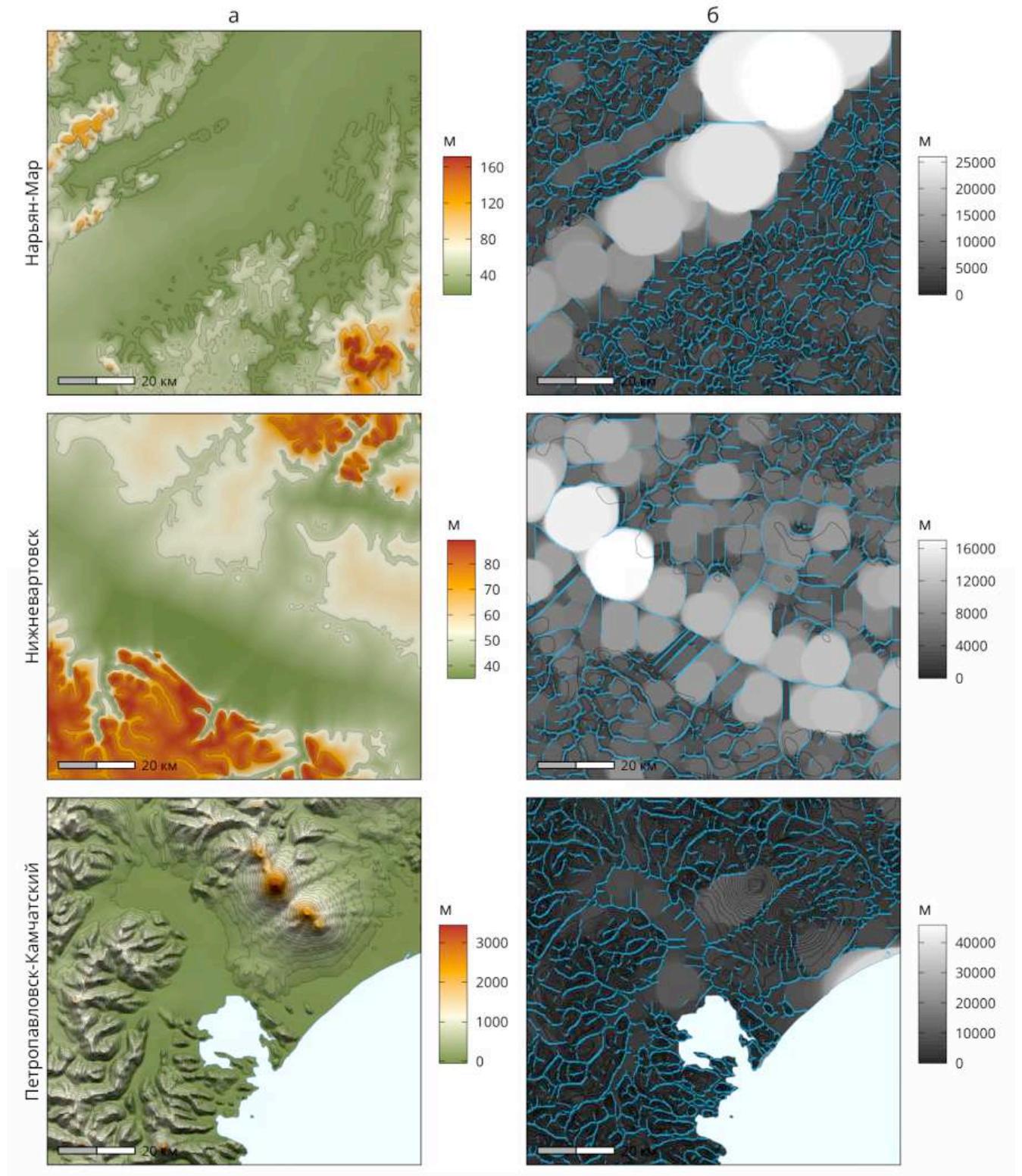
**Рисунок А.7** – Фрагменты ЦМР 10-12 для исследования гранулярности: (а) абсолютная высота; (б) ширина формы рельефа (уровень детализации 1000)



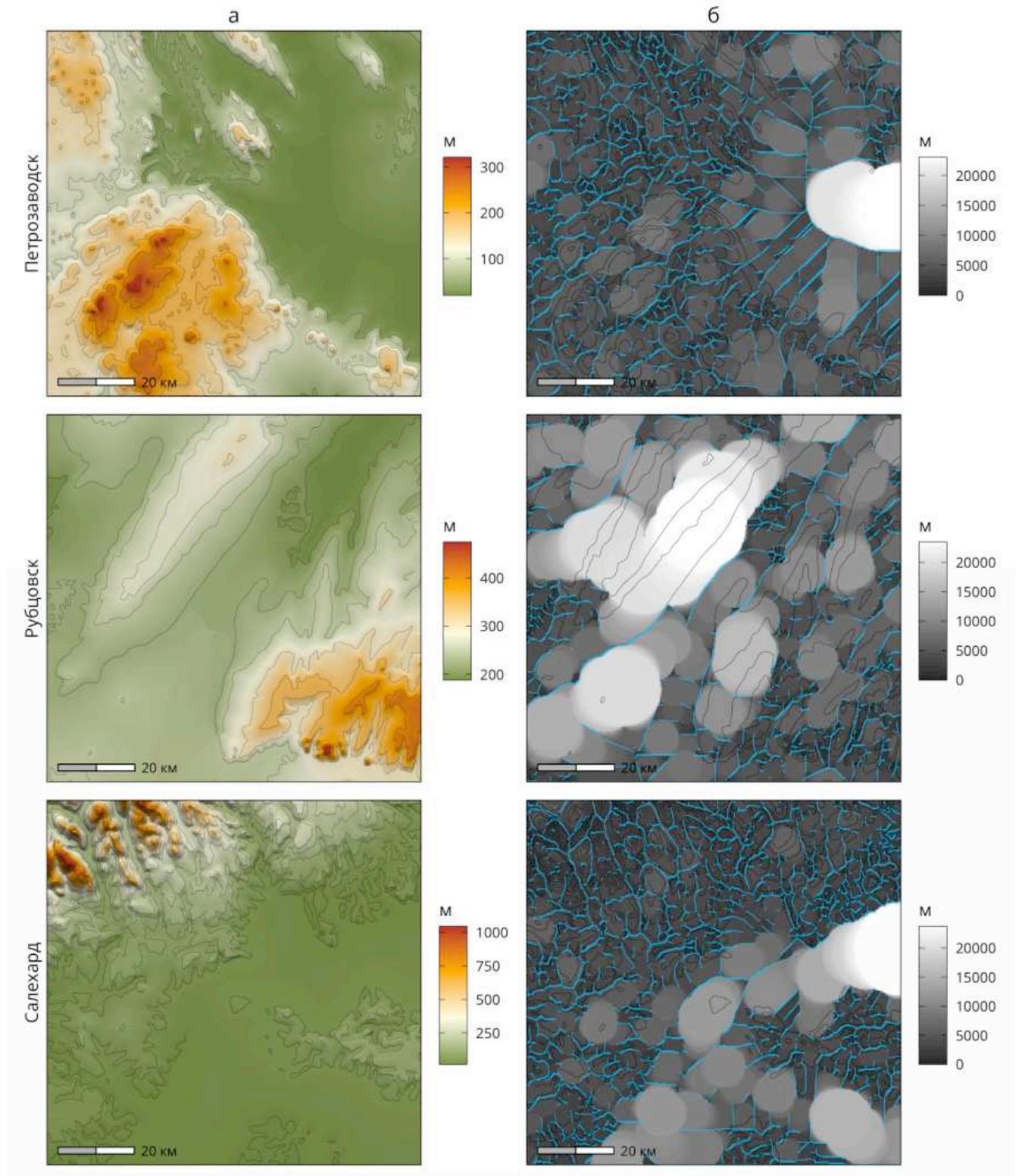
**Рисунок А.8** – Фрагменты ЦМР 13-15 для исследования гранулярности: (а) абсолютная высота; (б) ширина формы рельефа (уровень детализации 1000)



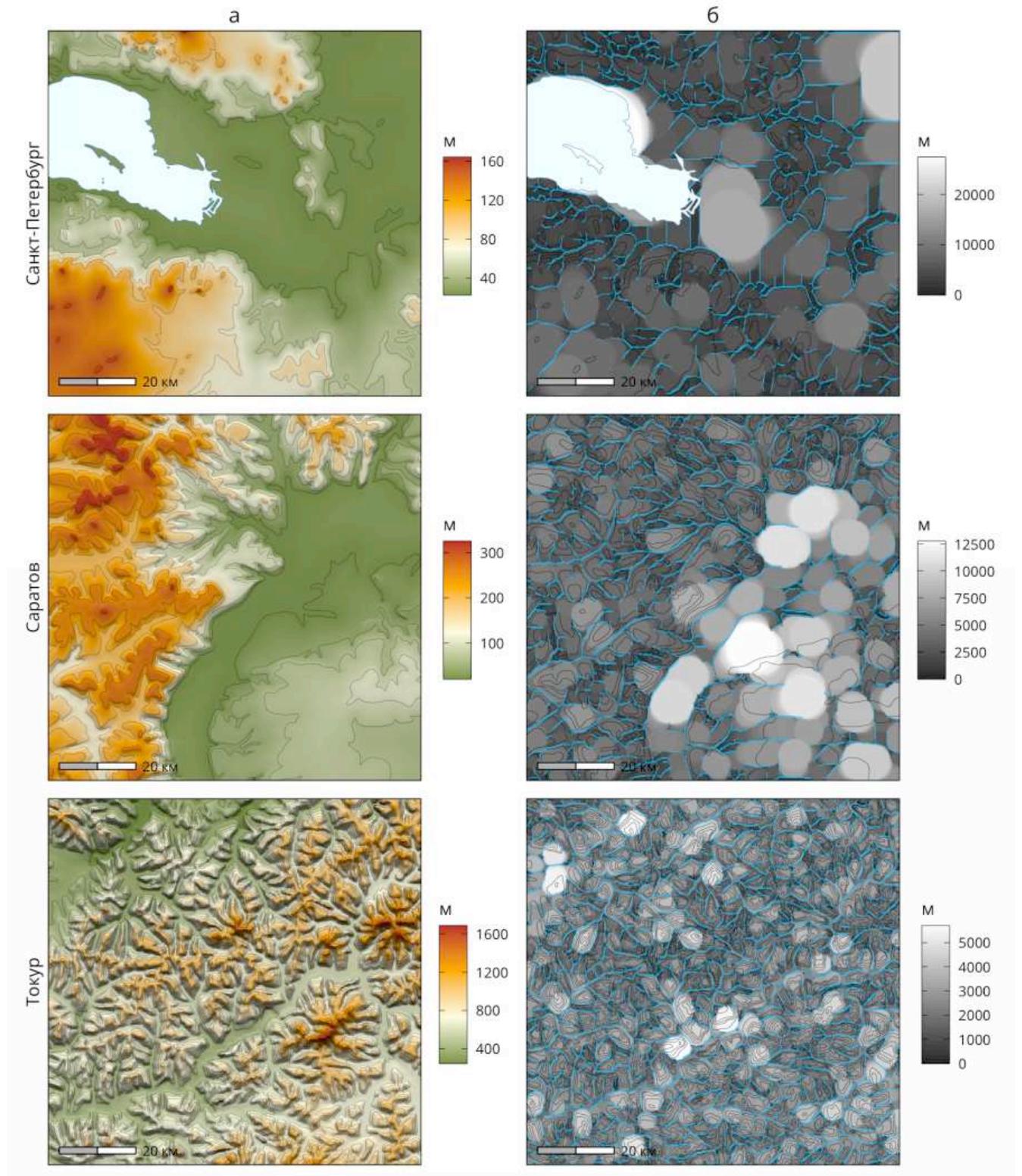
**Рисунок А.9** – Фрагменты ЦМР 16-18 для исследования гранулярности: (а) абсолютная высота; (б) ширина формы рельефа (уровень детализации 1000)



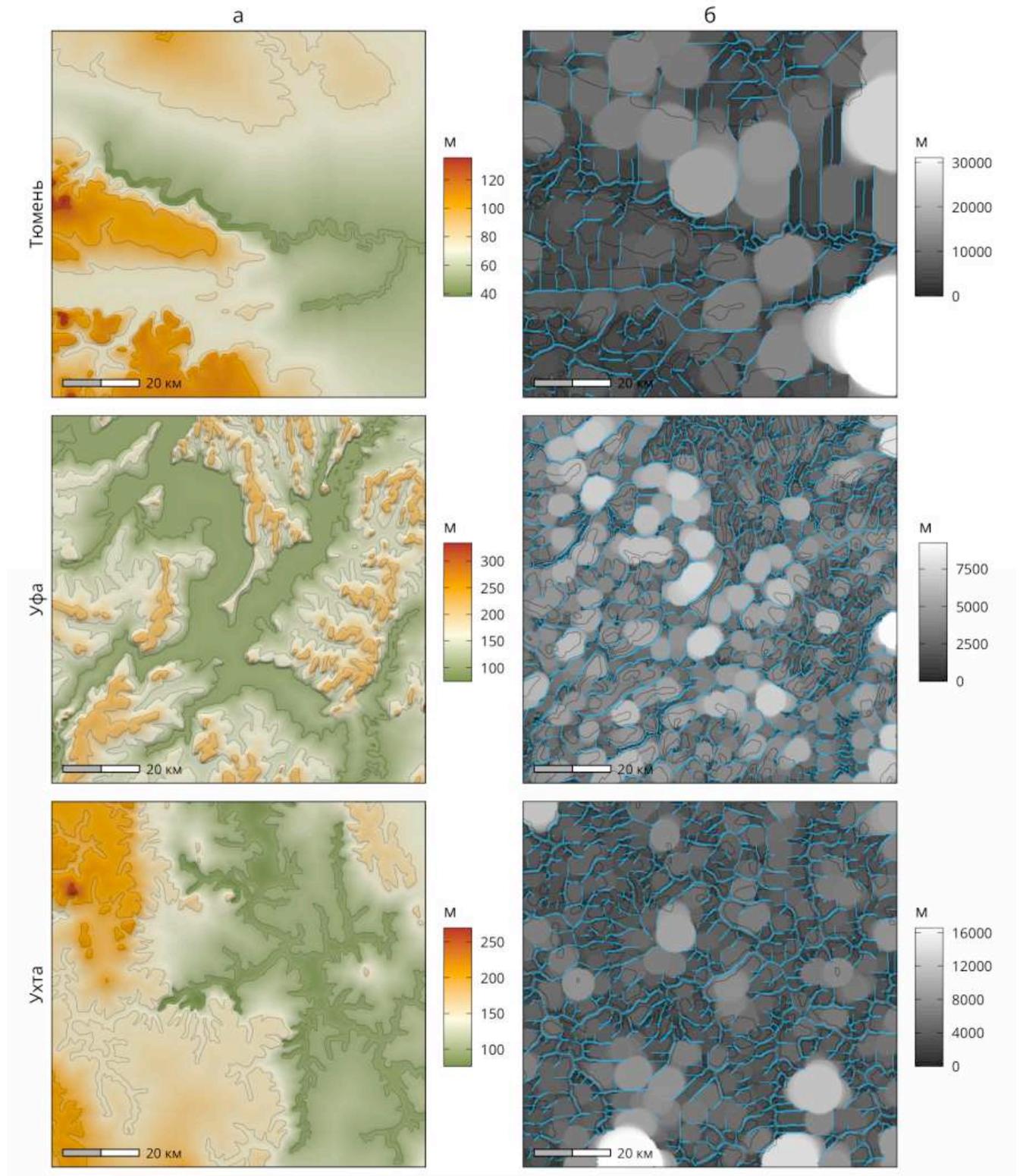
**Рисунок А.10** – Фрагменты ЦМР 19-21 для исследования гранулярности: (а) абсолютная высота; (б) ширина формы рельефа (уровень детализации 1000)



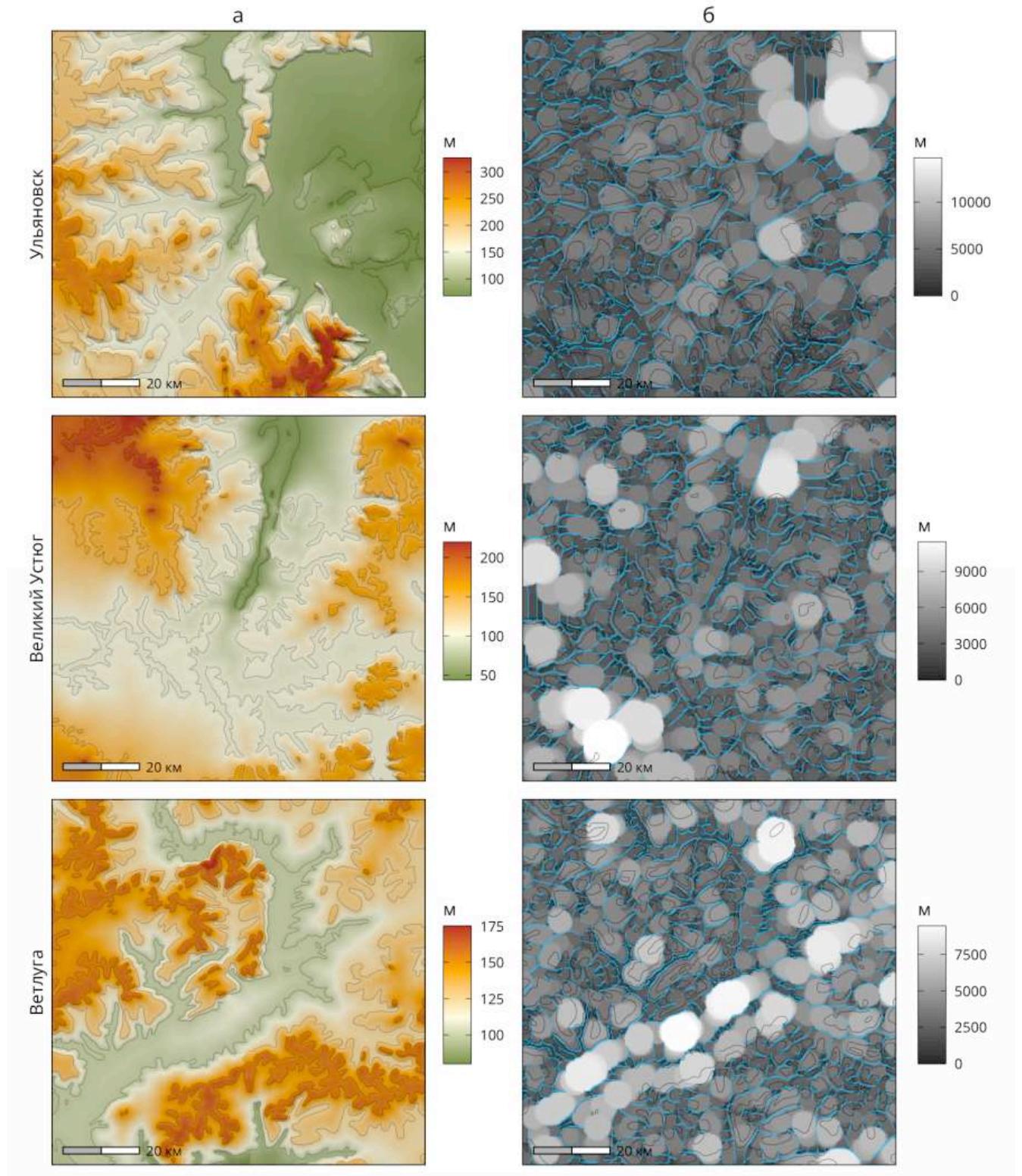
**Рисунок А.11** – Фрагменты ЦМР 22-24 для исследования гранулярности: (а) абсолютная высота; (б) ширина формы рельефа (уровень детализации 1000)



**Рисунок А.12** – Фрагменты ЦМР 25-27 для исследования гранулярности: (а) абсолютная высота; (б) ширина формы рельефа (уровень детализации 1000)



**Рисунок А.13** – Фрагменты ЦМР 28-30 для исследования гранулярности: (а) абсолютная высота; (б) ширина формы рельефа (уровень детализации 1000)



**Рисунок А.14** – Фрагменты ЦМР 31-33 для исследования гранулярности: (а) абсолютная высота; (б) ширина формы рельефа (уровень детализации 1000)