

ФЕДЕРАЛЬНОЕ ГОСУДАРСТВЕННОЕ БЮДЖЕТНОЕ
ОБРАЗОВАТЕЛЬНОЕ УЧРЕЖДЕНИЕ ВЫСШЕГО ОБРАЗОВАНИЯ
«МОСКОВСКИЙ ГОСУДАРСТВЕННЫЙ УНИВЕРСИТЕТ ИМЕНИ М.В.
ЛОМОНОСОВА»

На правах рукописи

ДОКУЧАЕВ

Павел Михайлович

**ПОСТРОЕНИЕ ЦИФРОВОЙ ПОЧВЕННОЙ КАРТЫ И
КАРТОГРАММЫ УГЛЕРОДА С ИСПОЛЬЗОВАНИЕМ МЕТОДОВ
ЦИФРОВОГО ПОЧВЕННОГО КАРТОГРАФИРОВАНИЯ (НА
ПРИМЕРЕ ВЯТСКО-КАМСКОЙ ПРОВИНЦИИ ДЕРНОВО-
ПОДЗОЛИСТЫХ ПОЧВ ЮЖНОЙ ТАЙГИ)**

Специальность 03.02.13 – Почвоведение

Диссертация
на соискание ученой степени
кандидата биологических наук

Научный руководитель:
кандидат сельскохозяйственных наук
Мешалкина Юлия Львовна

Москва 2017

ОГЛАВЛЕНИЕ

СПИСОК СОКРАЩЕНИЙ.....	6
ВВЕДЕНИЕ.....	7
ГЛАВА 1. ОБЗОР НАУЧНОЙ ЛИТЕРАТУРЫ	12
1.1 История изучения и картографирования почвенного покрова севера Удмуртской Республики	12
1.2 ЦПК и методы, используемые для построения картограмм	13
1.2.1 История развития цифровых методов в картографировании и их развитие в настоящее время.....	13
1.2.2 Что такое ЦПК.....	17
1.2.3 Источники получения данных для ЦПК.....	19
S (классификационное положение почвы или некоторое почвенное свойство в точке).....	19
C (климат)	20
O (организмы).....	21
R (рельеф)	21
P (почвообразующие породы)	22
A (временной фактор).....	23
N (пространственное расположение).....	23
1.2.4 Интерполяционные методы, используемые для построения картограмм отдельных свойств	24
1.2.5 Методы, используемые для построения почвенных карт.....	27
Метод множественной логистической регрессии	27
Метод деревьев решений	28
Метод случайного леса.....	30
Метод опорных векторов	32
Метод регрессионных деревьев.....	35
Другие методы.....	36
1.3 Применение ЦПК для картографирования пространственного распределения органического углерода и запасов гумуса	38

ГЛАВА 2. ОБЪЕКТЫ И МЕТОДЫ	41
2.1 Географическое положение, общая характеристика и история сельскохозяйственного использования ключевого участка	41
2.2 Агроэкологические условия.....	43
2.2.1 Агроклиматические ресурсы	43
2.2.2 Геоморфологические условия.....	45
2.2.3 Литологические условия	48
2.2.4 Гидрологические и гидрографические условия.....	49
2.2.5 Растительность	49
2.2.6 Почвенный покров	50
2.3 Характеристика материалов, имеющихся на район исследования.....	51
2.3.1 Архивные материалы хозяйства.....	51
2.3.2 Топографическая карта и ее производные	54
2.3.3 Анализ космического снимка	55
2.4 Полевой и камеральный этапы 2014-2016 гг.	56
2.4.1 Полевое обследование территории	56
2.4.2 Анализ почвенных образцов на содержание органического углерода и гранулометрический состав, а также создание базы данных	58
2.4.3 Создание традиционной почвенной карты	59
2.5 Цифровое почвенное картографирование.....	59
2.5.1 Описание используемого программного обеспечения	59
2.5.2 Отбор значимых ковариат для моделирования	60
2.5.3 Подбор размера пикселя для картографирования	62
2.5.4 Процесс создания почвенных карт на основе моделей.....	62
2.5.5 Сравнение карт между собой по 200-ам случайным точкам.....	67
2.5.6 Проверка карт по независимой случайной выборке	67
2.6 Построение картограммы гумуса	69
ГЛАВА 3. ПОСТРОЕНИЕ ПОЧВЕННОЙ КАРТЫ НА РАЙОН ИССЛЕДОВАНИЯ	71
3.1 Особенности почвенного покрова исследуемой территории.....	71

3.2 Создание индикационной таблицы по результатам полевого обследования	86
3.3 Создание картограммы пространственного распределения гранулометрического состава верхнего гумусого горизонта.....	88
3.4 Создание крупномасштабной почвенной карты.....	90
3.5 Анализ почвенной карты и сравнение с почвенной картой 1984 г.	92
3.5.1 Анализ изменения факторов почвообразования.....	92
3.5.2 Сравнение полученной почвенной карты с материалами обследования 1984 г.	93
3.6 Оценка точности почвенной карты.....	97
ГЛАВА 4. СОЗДАНИЕ ПОЧВЕННЫХ КАРТ СРЕДСТВАМИ ЦПК	100
4.1 Выбор ковариат и размера пикселя для моделирования	100
4.2 Моделирование на основе множественной логистической регрессии. 105	105
4.3 Моделирование на основе деревьев решений C5	109
4.4 Моделирование на основе метода случайного леса	112
4.5 Моделирование на основе регрессионных деревьев	116
4.6 Моделирование на основе метода опорных векторов.....	119
4.7 Сравнение и анализ результатов моделирования.....	122
4.7.1 Анализ схожести почвенных карт, построенных на основе одного и того же метода цифрового картографирования, но с разными коэффициентами каппа самовоспроизводимости моделей	122
4.7.2 Визуальный анализ степени схожести полученных карт с картой, построенной традиционными методами	124
4.7.3 Анализ общей точности и коэффициентов каппа при моделировании	124
4.7.4 Анализ таблиц точности производителя и пользователя	127
4.7.5 Анализ «устойчивых» пикселей на карте 2016 г. при моделировании	130
4.7.6 Анализ положения «устойчивых» пикселей по карте 1984 г.	135
ГЛАВА 5. СРАВНЕНИЕ ЦИФРОВЫХ КАРТ С ПОЧВЕННОЙ КАРТОЙ 2016 г. И ПРОВЕРКА ПО СЛУЧАЙНОЙ НЕЗАВИСИМОЙ ВЫБОРКЕ....	137
5.1 Сравнение карт почвенных типов, построенных различными методами, с картой 2016 г.....	137

5.2 Анализ схожести цифровых почвенных карт с картой 2016 г.	139
5.3 Проверка карт по случайной независимой выборке	142
ГЛАВА 6. ПОСТРОЕНИЕ КАРТОГРАММЫ СОДЕРЖАНИЯ ОРГАНИЧЕСКОГО УГЛЕРОДА В ВЕРХНЕМ ГУМУСОВОМ СЛОЕ ПОЛЕЙ, ПОДСЧЕТ ЗАПАСОВ ГУМУСА И ОЦЕНКА ТОЧНОСТИ ПРЕДСКАЗАНИЯ	146
6.1 Отбор ковариат для пространственного моделирования содержания органического углерода в верхнем гумусовом слое	146
6.2 Построение картограмм содержания органического углерода в верхнем гумусовом слое полей, оценка точности предсказания	148
6.3 Построение картограммы запасов гумуса	154
ВЫВОДЫ	159
СПИСОК ЛИТЕРАТУРЫ.....	161

СПИСОК СОКРАЩЕНИЙ

БД – база данных

ГИС – геоинформационные системы

ДДЗ – данные дистанционного зондирования

ПГБДР - Почвенно-географическая база данных России

ПО – программное обеспечение

ЦМР - цифровая модель рельефа

ЦПК – цифровое почвенное картографирование

DT – decision trees (деревья решений)

EVI - Enhanced Vegetation Index (улучшенный вегетационный индекс)

MNLR – multinomial logistic regression (множественная логистическая регрессия)

NDVI - Normalized Difference Vegetation Index (нормализованный относительный индекс растительности)

OK – ordinary kriging (ординарный кригинг)

RF – random forests (метод случайного леса)

RK – regression kriging (регрессионный кригинг)

RT – regression trees (регрессионные деревья)

SRTM - Shuttle Radar Topography Mission (радарная топографическая съемка Земли)

SVM – support vector machine (метод опорных векторов)

ВВЕДЕНИЕ

Актуальность работы

Почвенная карта является основным научным документом, на базе которого возможна грамотная оценка земельных фондов, а также разработка системы практических мероприятий, направленных на повышение плодородия почв. Важнейшим условием для ведения рационального сельского хозяйства является обновление устаревших почвенных карт и актуализация информации о состоянии и свойствах почвенного покрова. В то же время проведение крупномасштабного почвенного картографирования традиционным путем на всю территорию России не представляется возможным из-за больших трудозатрат на полевые исследования и ограниченного числа специалистов. Применение методов цифровой почвенной картографии (ЦПК) может стать одним из решений этой проблемы. Перспективным средством обновления старых почвенных карт хозяйств является сочетание современных компьютерных технологий (McBratney и др., 2003; Симакова, Савин, 1998; Козлов, Сорокина, 2012; Хитров, 2012) и грамотного полевого обследования изучаемой территории (Сорокина, 2006, Ульянова, Зборищук, 2005).

Таким образом, актуальность темы диссертационного исследования определяется объективной необходимостью разработки подходов к крупномасштабному картированию почвенных свойств и анализа возможностей моделирования почвенного покрова средствами цифровой почвенной картографии.

Цель работы - провести крупномасштабное почвенное картографирование территории левого берега реки Убыть (г. Глазов, Удмуртская Республика) методами ЦПК, а также оценить запасы гумуса в верхних 20 см сельскохозяйственных угодий.

Задачи исследования:

1. Построение традиционной почвенной карты на район исследования.

2. Создание цифровых почвенных карт на основе 5 алгоритмов и проверка моделей на самовоспроизводимость: 1) множественная логистическая регрессия, 2) деревья решений C5, 3) метод случайного леса, 4) регрессионные деревья, 5) метод опорных векторов.

3. Сравнение полученных цифровых карт с почвенной картой.

4. Проверка построенных карт по случайной независимой выборке.

5. Построение картограммы запасов гумуса сельскохозяйственных полей и оценка точности предсказания.

Научная новизна

Впервые в отечественном почвоведении проведена сравнительная характеристика результатов моделирования пространственного распределения почв на основе пяти алгоритмов ЦПК: 1) множественная логистическая регрессия, 2) деревья решений C5, 3) метод случайного леса, 4) регрессионные деревья, 5) метод опорных векторов, - созданы картографические модели, проверена их устойчивость и проведено сравнение с традиционной почвенной картой. Показано, что цифровые почвенные карты, построенные на основе моделей множественной логистической регрессии и деревьев решений, ближе соответствуют действительности, по сравнению с другими цифровыми картами.

Впервые в отечественной науке проиллюстрирован тот факт, что в результате применения алгоритмов ЦПК порождается равновероятный набор картографических моделей на одну территорию и показано, что в большинстве случаев лучшие результаты с точки зрения эксперта - почвоведа будут показывать средние по самовоспроизводимости (по результатам кросс-валидации) модели.

Впервые на исследуемую территорию создана картограмма запасов гумуса в верхнем слое 0-20 см (масштаба 1:50000) и обоснован алгоритм построения картограммы запасов гумуса, позволяющий снизить ошибку предсказания в два раза.

Впервые с 1985 г. была составлена почвенная карта масштаба 1:50 000 на территорию хозяйства «Чура» Глазовского района Удмурской Республики.

Практическая значимость

Впервые в отечественном почвоведении проведена сравнительная характеристика результатов моделирования пространственного распределения почв на основе пяти алгоритмов ЦПК: 1) множественная логистическая регрессия, 2) деревья решений C5, 3) метод случайного леса, 4) регрессионные деревья, 5) метод опорных векторов, - созданы картографические модели, проверена их устойчивость и проведено сравнение с традиционной почвенной картой. Показано, что цифровые почвенные карты, построенные на основе моделей деревьев решений и регрессионных деревьев, более схожи с почвенной картой, построенной традиционным методом.

Впервые в отечественной науке проиллюстрирован тот факт, что в результате применения алгоритмов ЦПК порождается равновероятный набор картографических моделей на одну территорию и показано, что в большинстве случаев лучшие результаты с точки зрения эксперта - почвоведа будут показывать средние по самовоспроизводимости (по результатам кросс-валидации) модели.

Впервые на исследуемую территорию создана картограмма запасов гумуса в верхнем слое 0-20 см (масштаба 1:50000) и обоснован алгоритм построения картограммы запасов гумуса, позволяющий снизить ошибку предсказания в два раза.

Впервые с 1984 г. была составлена почвенная карта масштаба 1:50 000 на территорию хозяйства «Чура» Глазовского района Удмуртской Республики.

Апробация работы

Основные положения и результаты исследования были представлены автором в виде устных докладов на 5 конференциях: на XVI и XVIII Докучаевских молодежных чтениях (Санкт-Петербург, 2013 и 2015 гг.), «Ломоносов» (Москва, 2013 и 2015 гг.), VII съезд Общества почвоведов им.

В.В. Докучаева и Всероссийской с международным участием научной конференции «Почвоведение-продовольственной и экологической безопасности страны» (Белгород, 2016 г.). Работа прошла апробацию на кафедре общего земледелия и агроэкологии МГУ им. М.В. Ломоносова.

Публикации

По результатам проведенных научных исследований опубликовано 7 печатных работ, в том числе 2 статьи в журналах списка ВАК.

Личный вклад автора

Автором было проведено обобщение литературных материалов, участие в организации и проведении полевых работ (2014-2016 гг.), пробоподготовка, статистический анализ полученных материалов, построение пространственных моделей, обобщение и интерпретация полученных результатов. Автор лично представлял результаты исследований на научных мероприятиях, перечисленных выше.

Структура и объем диссертации

Диссертация состоит из введения, 6 глав, выводов, списка литературы и 5 приложений. Она изложена на 186 страницах, содержит 28 таблиц и 48 рисунков. Список литературы включает 243 наименования, в том числе 162 на иностранных языках.

Благодарности

Автор глубоко признателен руководству кафедры общего земледелия и агроэкологии факультета почвоведения МГУ им. М.В. Ломоносова за предоставленную возможность проведения исследований по теме диссертации.

Автор благодарит научного руководителя к. с-х. н. Ю.Л. Мешалкину за внимательное отношение и помошь на всех этапах работы. Автор выражает искреннюю благодарность за ценные советы и плодотворную консультацию в различных вопросах своим учителям и коллегам В.П. Самсоновой, И.А. Мартыненко, Т.Ю. Ульяновой, а также А.М. Прокашеву.

Автор благодарен руководству ФГБНУ «Почвенного института имени В.В. Докучаева» за возможность освоения методов цифровой почвенной картографии в ходе Школы для молодых ученых; а также сотрудникам лаборатории Биологии и биохимии почв за помощь в проведении анализов почвенных образцов на содержание гумуса и А.В. Юдиной за помощь в проведении анализов гранулометрического состава образцов.

Особую благодарность автор выражает своим родным и близким за моральную поддержку и помощь в проведении работ, а также руководству хозяйства ООО «Чура» в лице директора А.М. Першина и сотрудникам хозяйства за предоставленную возможность работы на полях хозяйства и информацию об особенностях почвенного покрова и сельскохозяйственной деятельности. Исследование было частично поддержано грантом Российского Научного Фонда № 15-16-30007.

ГЛАВА 1. ОБЗОР НАУЧНОЙ ЛИТЕРАТУРЫ

1.1 История изучения и картографирования почвенного покрова севера Удмуртской Республики

Почвенный покров севера Удмуртской Республики имеет сложный генезис и является объектом неоднократных рекогносцировочных, почвенно-агрохимических и других научных исследований (Прокашев, 2003). Так, в начале XX века Р.В. Ризположенским был предпринят первый опыт систематического научного изучения почвенного покрова Вятской губернии. (В то время север Удмуртской Республики входил в состав Вятской губернии). Начиная со второй половины прошлого века были получены обобщающие сведения по номенклатуре, классификационному разнообразию, свойствам, режимам, качественной оценке, закономерностям географического распространения и структуре почвенного покрова севера Удмуртской Республики. Основное внимание в этих работах уделялось изучению генезиса зональных автоморфных компонентов почвенного покрова: подзолистых, дерново-подзолистых, серых лесных и других типов почв (Ногина, 1962; Жуйкова и др., 2000; Вологжанина, 1984), а также экологии этих почв (Владыкина, 2012).

Почвоведами в регионе исследования отмечено наличие почв, именуемых собирательным термином «почвы со сложным органопрофилем» (Тюлин, Россохина, 1967; Прокашев, 2003). Наличие вторых гумусовых горизонтов было обнаружено в следующих почвенных типах: серые лесные (Вологжанина, 1984; Добровольский, Урусевская, 1984), дерновые оглеенные, а также дерново-карбонатных почв (Прокашев, 2003).

В настоящее время большое внимание местными почвоведами уделяется разработке адаптивно-ландшафтных систем земледелия (Дзюин, Дзюин, 2013). Почва при этом рассматривается как двигатель социально-экономического развития региона (Кашин, 2015; Рысин, 2009).

Первые попытки картографирования почв территории Удмуртской Республики относятся к середине девятнадцатого века, когда в России бурно развивались работы по земельному кадастру. В 1851 г. под руководством академика К.С. Веселовского была создана первая общая почвенная карта Европейской части России в масштабе 200 верст в дюйме (1: 8 400 000), на которой был отражен почвенный покров Удмуртской Республики (Ковриго, 2004). В 1884-1898 гг. статистический отдел Вятского губернского земства составил печатные карты пахотных угодий Удмуртии в масштабе 10 верст в дюйме (1:420 000), классификация почв была представлена по гранулометрическому составу: суглинистые, песчаные, супесчаные и глинистые почвы (Ковриго, 2004).

Дальнейшие обследования почв Удмуртской Республики проводились под руководством профессора Н.П. Карпинского в период с 1926 по 1928 гг., в результате была составлена почвенная карта Удмуртии в масштабе 1:420 000 и краткий почвенный очерк к ней. Карта, в основном, верно отражала особенности почвенного покрова Удмуртии (Ковриго, 2004).

Почвенно-климатические условия Удмуртской Республики, природные факторы почвообразования и почвы Удмуртии описаны в работах по географическому районированию почв СССР (Ногина, 1962; Пермяков, 1972).

Итогом почвенных обследований с 1951 по 1990 г. стало составление подробной почвенной карты Удмуртской Республики масштаба 1:200 000, изданной Главным Управлением геодезии и картографии при Совете Министров СССР в 1990 г. (Почвенная карта ..., 1990).

1.2 ЦПК и методы, используемые для построения картограмм

1.2.1 История развития цифровых методов в картографировании и их развитие в настоящее время

Почвенная карта представляет собой уменьшенное и обобщенное в заданном масштабе изображение почвенного покрова на плоскости (топографической основе), построенное по математическим законам, в

определенной системе условных обозначений. Составление почвенных карт – это трудоемкий процесс, включающий в себя три этапа: 1) подготовительный камеральный - сбор и систематизацию имеющихся картографических, литературных, дистанционных и фондовых материалов по объекту съемки, анализ их полноты и качества; 2) полевая съемка - заложение точек опробования, уточнение на местности границ и форм почвенных контуров, морфологическое описание разрезов и отбор почвенных образцов; 3) заключительный камеральный - составление окончательного варианта карты с учетом результатов анализа отобранных образцов.

Сокращение людских и временных затрат при проведении почвенного картографирования стало возможным с развитием информационных технологий.

В работах 70- 80-ых годов предпринимались первые попытки описания пространственных закономерностей почвенных свойств путем интерполяции математическими методами известных значений на всю территорию обследования (Burrough и др., 1994), предпринимались попытки определения характерных масштабов пространственного варьирования почвенного покрова и выявления средних размеров почвенного индивидуума (Козловский, 1970), разрабатывались аэрокосмические методы изучения почв (Андроников, 1979). В российском почвоведении активно развивалась концепция структуры почвенного покрова (Фридланад, 1972), в последующем ставшая основой для цифрового почвенного картографирования.

В 1992 г. направление «педометрика» оформилось как отдельная ветвь почвоведения, занимающаяся применением математических и статистических методов для изучения распределения и генезиса почв (Мешалкина, 2012). Уже в 90-ых годах высказывались идеи о необходимости задания моделей для проведения почвенных обследований (Hewitt, 1993), в том числе моделей, основанных на экспертных знаниях (Bui, 2003).

К концу XX века были накоплены знания в области использования методов геостатистики в экологии и почвоведении (Isaaks, Srivastava, 1989;

Webster, Oliver, 2007). Геостатистические методы оказались полезным инструментом для количественной оценки изменчивости почв (Геостатистика..., 2010), хотя их использование подняло много вопросов: какое количество образцов необходимо и достаточно для обеспечения определенной точности предсказания, как моделировать вариограммы, как стратифицировать данные, когда и какие методы являются наиболее подходящими (Burrough и др., 1994).

В последние десятилетия идет разработка региональных и глобальных баз почвенных данных (Hartemink и др., 2013). Вот некоторые примеры: Всемирная Гармонизированная БД с масштабом карты 1:5 000 000, разработанная ФАО-ЮНЕСКО (Harmonized ..., 2009); Европейская БД с масштабом карты 1:1 000 000, являющейся частью европейской почвенно-информационной системы — EUSIS (Le Bas и др., 1998); SOTER для разных частей Африки в масштабе 1:2 000 000 (Dijkshoorn, 2003) и для Латинской Америки и островов карибского бассейна в масштабе 1:5 000 000 (Dijkshoorn и др., 2005). Разработаны национальные почвенные базы данных, например, такие как: Интерактивная БД Почвенной Службы Америки (American Web Soil Survey - WSS) и Географическая БД Почвенной Службы Америки (Soil Survey Geographic Data Base - SSURGO) (Soil Survey Staff, 2017); австралийская информационная система почвенных ресурсов (ASRIS) (ASRIS, 2014); канадская почвенная база данных (CANSIS) и национальная база данных почв Канады (NSDB) (AAFC, 2013); российская почвенная карта масштаба 1:2 500 000 (Программа ..., 1972, Stolbovoi, McCallum, 2002) и российская карта почвенно-географического районирования масштаба 1:15 000 000 (Добровольский и др., 2008). Кроме того, в России были разработаны почвенные карты мира 1:10 000 000 (Почвенная..., 1975) и 1:15 000 000 (Почвенная..., 1982).

Впервые в России Информационная база классификации была разработана более 25 лет назад в Почвенном институте имени В.В Докучаева под руководством В.А. Рожкова (Шишов и др., 1985), она была составной

частью автоматизированной информационной системы «Мерон» (Рожков, 2002). В настоящее время успешно развивается Почвенно-географическая база данных России (ПГБДР) масштаба 1:2 500 000 (Шоба и др., 2010). Развиваются БД, характеризующие отдельные регионы России (например, Белоусова, Мешалкина, 2012).

На современном этапе развития ЦПК большое внимание уделяется источникам первичных данных: старым почвенным картам и материалам обследования (Mayr, др., 2010; Dobos и др., 2010; Рухович и др., 2012), цифровым моделям рельефа (Shary и др., 2002; Advances in Digital Terrain Analysis, 2008; Florinsky, 2011) и ДДЗ (Кирьянова, Савин, 2012; Кренке, 2012; Самсонова, Мешалкина, 2012; Королюк, 2012; Куст и др., 2010) для нужд ЦПК. Быстрыми темпами развивается использование результатов наземной съемки для крупномасштабного картографирования (Viscarra Rossel и др., 2010, Gooley и др., 2014). Большое внимание также уделяется способам гармонизации почвенных данных (Dobos и др., 2010; Sulaeman и др., 2013; Белоусова, Мешалкина, 2012) и объединению материалов, полученных в разном масштабе (Behrens и др., 2010).

Активно развиваются такие направления, как: ранжирование почвообразующих факторов по степени их участия в формировании региональной неоднородности почвенного покрова (Пузаченко и др., 2006; Козлов, 2009; Пузаченко, 2009; Boettinger, 2010), исследования зависимости качества получаемых ковариатов для ЦПК при различных масштабах исходных картографических материалов (Roecker, Thompson, 2010), картографическое обоснование для внедрения новых методик выбора точек пробоотбора (Minasny, McBratney, 2006; Pahlavan Rad и др., 2014), а также их сравнение (Brungard, Boettinger, 2010).

Быстрыми темпами развивается тематическая цифровая картография. Примерами могут служить: картографирование содержания органического углерода (Carré и др., 2010; Zhao, Shi, 2010; Mendonça-Santos и др., 2010 и многие другие работы), тяжелых металлов (Ballabio, Comolli, 2010), состояния

экосистем и климатических условий (New и др., 2002), агроэкологическое картографирование земель (Козлов, Сорокина, 2012; Васенев, Васенева, 2012) и др.

Работы по внедрению цифровых технологий в производство показывают качественное снижение стоимости и повышение скорости картирования в условиях сокращения возможностей получения новых полевых материалов (MacMillan и др., 2007).

1.2.2 Что такое ЦПК

Для прогнозирования и картографирования почвенного покрова используются цифровые количественные техники, названные Ф. Лагашери и А. МакБратни, *цифровой почвенной картографией*. Она определяется ими как «создание и распространение пространственных почвенных информационных систем, численных моделей, описывающих пространственные и временные вариации почвенных типов и свойств на основе почвенных наблюдений и знаний, на основе связанных с ними переменных окружающей среды» (Lagacherie, McBratney, 2007). Или иными словами, ЦПК – это одновременно создание цифровых карт, то есть научная работа, и, в то же время, воспроизведение таких карт (Мешалкина, 2012). Задача ЦПК - на основе ограниченного количества почвенных данных и предварительных знаний о почвенно-ландшафтных связях обеспечить точное отображение пространственного распределения почв и почвенных свойств.

Создание цифровых почвенных карт осуществляется почвенными пространственными логическими системами, включающими почвенные пространственные информационные системы и построитель почвенных карт (рис.1.1) (Lagacherie, McBratney, 2007).

Предсказания почвенных классов и свойств в цифровой картографии основывается на взаимосвязях между почвами, факторами и процессами почвообразования, которые входят в уравнения в качестве переменных-предсказателей (Mendonça-Santos и др., 2010).

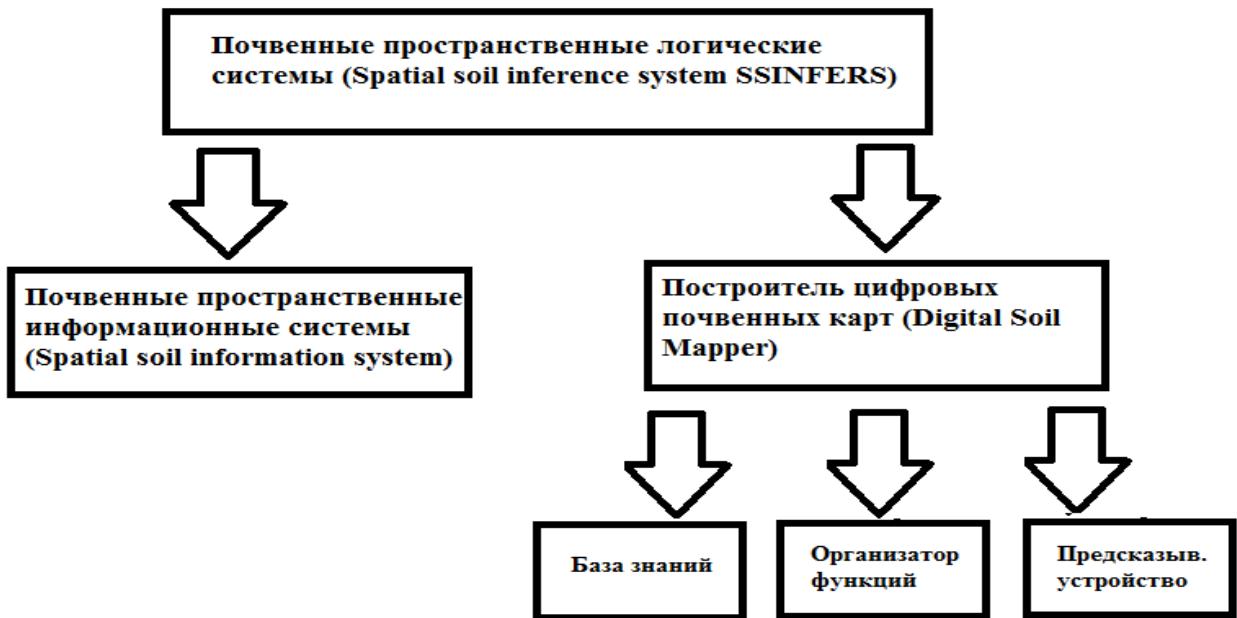


Рис. 1.1 Схема создания ЦПК (Lagacherie, McBratney, 2007)

Логика ЦПК основана на уравнении (1) (McBratney и др., 2003):

$$S_a[x, y, \sim t] \text{ или } S_c[x, y, \sim t] = f(s[x, y, \sim t], c[x, y, \sim t], o[x, y, \sim t], r[x, y, \sim t], p[x, y, \sim t], a[x, y, \sim t], n) \quad (1)$$

где S_a – количественная почвенная характеристика, S_c - почвенная таксономическая единица, s – почва или некоторое почвенное свойство в точке, c – климат (локальные климатические условия), o - организмы, включая человеческую деятельность, r – рельеф, включая морфометрические характеристики, p – почвообразующая порода, литологическое сложение, a – временной фактор, n - пространственное расположение, t - время (McBratney и др., 2003).

Эта формула является результатом трудов многих поколений ученых-почвоведов (например, Захарова, Shaw, Jenny), которые развивали основной закон почвоведения, сформулированный В.В. Докучаевым (Флоринский, 2012).

1.2.3 Источники получения данных для ЦПК

Согласно уравнению (1) почва в данной точке пространства является функцией от 7 переменных, входящих в состав уравнения. Поэтому, зная факторы почвообразования и/или ковариаты в конкретной точке пространства, мы можем предсказать почву в этой точке.

S (классификационное положение почвы или некоторое почвенное свойство в точке)

Источниками информации о классификационном положении почвы или о некоторых почвенных свойствах в точке могут служить старые почвенные карты, базы данных, результаты полевых обследований и другие материалы (Pásztor и др., 2016; Odeh и др., 2012), в том числе результаты дистанционного и наземного зондирования земли.

Различные спутниковые изображения используются в качестве источников для ЦПК: LANDSAT, Advanced Spaceborne Thermal Emission и Reflection Radiometer (ASTER), Quickbird, IKONOS и другие (Grunwald, 2009; Симакова, Савин, 1998; Кравцова, 2005; Konyushkova, 2010). Эти спутниковые изображения отличаются друг от друга пространственным разрешением, диапазонами спектральной съемки и спектральными разрешениями. Гиперспектральная съемка позволяет получить информацию о минерологических особенностях почв, таких как содержание оксидов железа (King и др., 1995), карбонатов (Lagacherie и др., 2010), сульфатов (Crowley, 1993). Кроме того, на основе гиперспектральной съемки можно получать информацию о содержании органического вещества (Gomez и др., 2008).

Данные наземной съемки отражательной способности почв в полевых условиях в видимой, ближней инфракрасной и средней инфракрасной области спектра широко используются для определения различных почвенных свойств (Viscarra Rossel и др., 2010; Reeves, 2010; Докучаев, Мешалкина, 2016а; Savin и др., 2016). Определение в полевых условиях электропроводности позволяет комплексно охарактеризовать почвенный покров, выявить его структуру

(Поздняков, 2001), диагностировать минералогический состав почв (Viscarra Rossel и др, 2009), содержание солей (Aly и др., 2007), влаги (Robinson и др., 2012) и другое. Этот метод также удачно используется при картографирования почвенных типов (Bishop, McBratney, 2001; Anderson-Cook и др., 2002; Samsonova и др., 2007; Huang и др., 2014). Измерение магнитных свойств почв дает информацию о составе и уровне загрязнения почвенного покрова (Jordanovaa и др., 2008).

Для прогнозирования свойств почвы успешно разрабатывается использование педотрансферных функций для пересчета одних почвенных свойств в другие (McBratney и др., 2002).

C (климат)

Климат играет важную роль в формировании и эволюции почвенного покрова. Широко используются следующие климатические параметры: минимальные и максимальные температуры, сумма температур выше 10 $^{\circ}\text{C}$, средняя температура за период, количество осадков, интенсивность испарения, уровень солнечной радиации, глубина снежного покрова и другие.

Получение пространственно-распределенных климатических данных возможно двумя путями: или наземные измерения в конкретных точках с дальнейшей интерполяцией на всю исследуемую территорию (Genovese, 2001) или посредством использования ДДЗ (New и др., 2001). ДДЗ могут служить источником пространственных данных по оценке коэффициентов испарения (Boegh и др., 2002), альбедо поверхностей (Brest и Goward, 1987). Температурные данные можно получить со спутников Geostationary Orbiting Earth Satellite - GOES (Diak и др., 1998) и TIROS (Susskind и др., 1997).

Для получения пространственно распределенных климатических данных на конкретный регион используются специальные базы данных. Создание таких баз возможно путем интерполяции данных с метеорологических станций (Howell, Kim, 2009).

Микроклимат территории формируется под влиянием рельефа (уклон, экспозиция и так далее) и может быть количественно определен путем оценки солнечной радиации, рассчитанной на основе ЦМР. Например, крутые северные склоны получают гораздо меньше солнечного излучения, чем соответствующие южные склоны водоразделов, что приводит к снижению испарения и большему увлажнению северных склонов (Boettinger, 2010).

Существуют попытки классификации почв на основе данных по почвенной влаге и температурным режимам почв (Donatelli и др., 2002).

O (организмы)

Для получения данных по растительности в ЦПК широко используются ДДЗ и ГИС-технологии (Mulder и др., 2011; Ballabio и др., 2012; Johnson и др., 2012). Для пространственного оценивания биомассы и картографирования растительного покрова используются различные вегетационные индексы (Grunwald и др., 2011; Mulder и др., 2011), наиболее распространенным из которых является NDVI (Meirik и др., 2010; Wang и др., 2007; Трифонова, Мищенко, 2016) и EVI (Peng и др., 2015; Huete и др., 1997), а также мониторинг урожайности на уборочных машинах (Bishop, McBratney, 2001) или анализ неоднородности почвенного покрова (Кутузова и др., 2015).

В отечественном почвоведении разработаны подходы, описывающие корреляцию между фитомассой основных растительных сообществ и почвенными свойствами (Родина, Базилевич, 1965), запасами гумуса (Кононова, 1963) и индексом NDVI (Савин и др., 2015).

R (рельеф)

Важнейшим фактором почвообразования является рельеф. В отечественном почвоведении были разработаны различные классификации элементов рельефа (Гедымин, 1990; Фридланд, 1984).

Традиционно информацию о рельефе местности почвоведы получали путем визуальной интерпретации аэрофотоснимков или топографических карт. С развитием компьютерной техники и цифровых источников данных

описание особенностей рельефа стало возможно в цифровом виде. Под термином «цифровая модель рельефа» обычно подразумеваю трехмерную модель превышений. Источниками ЦМР являются: топографические карты, съемка с помощью приемников GPS, стереопары аэро- и фотосъемки, радарная и топографическая съемки. Наиболее широко используемыми источниками для получения ЦМР являются данные радарной съемки LIDAR, SAR или SRTM (Roecker, Thompson, 2010; Woolard, Colby, 2002; Farr, 2000).

Такие характеристики как уклон, экспозиция, горизонтальная и вертикальная кривизны и другие являются производными ЦМР (Florinsky, 1998) и называются обычно его морфометрическими характеристиками. Точность информации о морфометрических характеристиках рельефа зависит от размера пикселя исходных данных, используемых для построения ЦМР (Roecker, Thompson, 2010). Морфометрические характеристики рельефа часто имеют тесную связь со свойствами почв, такими как содержание органического углерода (Maynard, Johnson, 2014; McKenzie, Ryan, 1999), гранулометрический состав (Bishop, Minasny, 2006; Levi, Rasmussen, 2014) и мощность почвенных горизонтов (Park и др., 2001; Sinowski, Auerswald, 1999).

Одним из способов визуализации картограмм пространственного распределения различных почвенных свойств является наложение результатов интерполяции на трехмерную модель рельефа: например, картографирование содержания гумуса (Смирнова и др., 2011; Adhikari и др., 2012), pH верхнего почвенного слоя (Castrignano и др., 2011).

P (почвообразующие породы)

Информация о почвообразующих породах для ЦПК может быть получена из оцифрованных геологических карт. Измерение естественных электрических, гравитационных, магнитных (Hansen и др., 2005) и электромагнитных (Huang и др., 2014; Buchanan, Triantafyllis, 2009) полей земли также может дать информацию о литологическом строении изучаемой территории. Другим важным инструментом получения информации о факторе

почвообразования p является гамма-спектрометрия (Mulder, 2011; Huang и др., 2014; Buchanan, Triantafilis, 2009).

Использование спутниковых ДДЗ в диапазонах VNIR (видимые и ближние инфракрасные длины волн) – TIR (тепловые инфракрасные длины волн) с высокими спектральным и пространственным разрешениями является полезным инструментом в получении информации о почвообразующих породах (Mulder, 2011).

Разработано несколько надежных индексов (Quartz Index, Carbonate Index и Mafic Index), использование которых помогает при среднемасштабном и крупномасштабном картографировании почвообразующих пород для засушливых регионов на основе данных ASTER (Ninomiya и др., 2005).

A (временной фактор)

Фактор А показывает время, в течение которого происходило почвообразование. Этот фактор относительно редко используется для целей ЦПК (Grunwald, 2009). Временной фактор может быть охарактеризован путем определения возраста горных пород, на которых развивалась почва (McBratney и др., 2003). Существуют методики датирования почвенного возраста на основе содержания в почвах ^{14}C и термолюминисценции (Matt, Johnson, 1996), отношения $^{40}\text{Ar}/^{39}\text{Ar}$ (Van Niekerk и др., 1999). Наземная электромагнитная съемка также используется для определения возраста почвенного покрова и стратиграфического картографирования (Sinha, 1990).

N (пространственное расположение)

Практически все почвенные профили, описанные до 2005 г. не имеют точных географических координат, описания снабжались лишь примерной привязкой на местности с указанием ближайшего населенного пункта. Однако при среднемасштабном и мелкомасштабном картографировании точное местоположение почвенного разреза не имеет большого значения, важным является наличие точного описания тех экологических условий и факторов почвообразования, которым соответствует данная почвенная единица.

Фактор N может быть описан некоторыми линейными или нелинейными преобразованиями исходных пространственных координат, например, может быть введена новая координата, такая как расстояние до водораздела, или расстояние до реки (Bui, Moran, 2001), или расстояние между солонцзовыми пятнами в зоне каштановых почв (Konyushkova, 2010).

1.2.4 Интерполяционные методы, используемые для построения картограмм отдельных свойств

Принцип действия *детерминированных интерполяционных методов* (таких как обратные взвешенные расстояния, триангуляция Делоне, локальная и глобальная полиномиальная интерполяция и другие) заключается в прогнозировании неизвестных значений свойств в заданной на основе математических формул, сформированных путем определения средневзвешенных значений свойства в соседних с ней точках (Геостатистика..., 2010).

Визуализация пространственного распределения почвенных свойств посредством применения *геостатистических методов* широко применяется в современном ЦПК для различных природных зон, климатических условий и ландшафтов (Aksoy и др., 2012; Сидорова, 2011; Самсонова, 2008).

Простейшие модели описания пространственного варьирования почвенных свойств - это хороплеты (однородные, согласно некоторым принятым критериям, участки территории) и тренды (пространственные модели, описывающие постепенные изменения почвенного свойства) (Геостатистика..., 2010). Существуют работы по предсказанию почвенных свойств на основе использования хороплет (Hoosbeek, 1998) и трендов (Сидорова, Соломатова, 2000).

Кригинг – метод пространственной интерполяции, основанный на модели вариограммы, по которой определяются веса значений переменной в окрестностях искомой точки или области (Webster, 2008). Термин «кригинг» объединяет целое семейство алгоритмов пространственной регрессии

(Геостатистика..., 2010). Кригинг бывает однофакторный, например, ординарный, блоковый, индикаторный и так далее и многофакторный, например, регрессионный, кокригинг или кригинг главных компонент (Grunwald, 2009).

Если корреляция между переменными является значимой, то методы многофакторного кригинга (регрессионный кригинг и др.) показывают значительно лучшие результаты по сравнению с ординарным кригингом или другими методами интерполяции (Minasny, McBratney, 2007б). Это преимущество тем больше, чем теснее связь между основными и второстепенными переменными (Zhao, Shi, 2010). В настоящее время регрессионный кригинг получил наибольшее распространение в почвоведении (Zhao, Shi, 2010; Marchetti и др., 2010; Mendonça-Santos и др., 2010).

Особенностью методов *интеллектуального анализа данных* является поиск неизвестных корреляций между ковариатами окружающей среды и прогнозируемой переменной S - почвой или почвенным свойством (Digital..., 2006). Методы интеллектуального анализа позволяют быстро автоматически извлечь информацию из больших наборов данных (Behrens и др., 2005). Наиболее часто используемыми иерархическими моделями интеллектуального анализа данных в ЦПК являются: деревья классификации (Bou Kheir и др., 2010; Stoorvogel и др., 2009), регрессионные деревья (Vasques и др., 2008; Stoorvogel и др., 2009), методы случайного леса (Grimm и др., 2008; Wiesmeier и др., 2011), деревья принятия решений (Häring и др., 2012; Subburayalu и др., 2014). Эти методы реализованы во многих статистических программах и сочетаются с ГИС (Digital..., 2006).

Принцип действия *нейронных сетей* заключается в извлечении знаний и правил из обучающих выборок на основе имитации механизмов организации и функционирования биологических нейронных сетей — сетей нервных клеток живого организма (Behrens и др., 2005). Нейронные сети часто применяются для описания пространственного варьирования: содержания

органического вещества почв (Zhao, Shi, 2010), отношения C/N (Carré и др., 2010), pH и гранулометрического состава (Boruvka, Penizek, 2007), гидрологических характеристик (Minasny, McBratney, 2007a), антропогенного влияния (Докучаев, 2013, 2014, 2016). Существуют примеры использования нейронных сетей в сочетании с кригингом для картографирования пространственного распределения органического углерода в почвах (Zhao, Shi, 2010).

Группа *описательных методов* описывает связи почвы с факторами почвообразования и индикаторами, основываясь на опыте экспертов-почвоведов, работающих в данной регионе исследования (Digital..., 2006). Примерами таких методов являются: набор правил типа «*ковариатам окружающей среды Q соответствует почва S*» (McKenzie и др., 1999), байсовские сети (Mayg и др., 2010; Самсонова и др., 2017), системы нечеткой логики (Qi , Zhu и др., 2006; Jordanova и др., 2008), системы условных вероятностей, полученных на основе существующих почвенных карт (Hewitt и др., 2010).

Существует группа методов картографирования пространственного распределения почвенных свойств, *основанная на построении регрессионных зависимостей* между измеренными значениями в точках и неизвестными значениями (Zhao, Shi, 2010; Vasques и др., 2008; Gomez и др., 2008).

Классификационные и дискrimинационные методы широко применяются для пространственного описания содержания Ca, Mg, K и pH верхнего горизонта (Awiti и др., 2008), содержания органического углерода (Awiti и др., 2008) или литологических свойств территорий (Sinowski, Auerswald, 1999).

Анализ главных компонент используется в ЦПК для картографирования гранулометрического состава почвенного покрова (Hengl и др., 2007), микробных сообществ (Allison и др., 2007) и питательных элементов (Awiti и др., 2008).

Для визуализации пространственного распределения почвенных свойств часто используются *методы кластерного анализа* (Genxu и др., 2008; Zhang и др., 2008б), в том числе классификации с использованием четких или нечетких методов кластеризации (Nauman и др., 2012; Schmidt, Hewitt, 2004).

1.2.5 Методы, используемые для построения почвенных карт

Метод множественной логистической регрессии

Существует множество работ по ЦПК, где в качестве модели для картографирования почвенных типов используется метод множественной логистической регрессии (Kempen и др., 2009; Brungard и др., 2015; Marchetti и др., 2011). Множественная логистическая регрессия является обобщением логистического регрессионного анализа, применяемого к качественным переменным (Venables, Ripley, 2002).

Предположим, что переменная Y_i представляет определенный тип почвы в месте отбора проб, где $i = 1, \dots, n$, а n - это количество почвенных типов, встречающихся на исследуемой территории. В случае, если n равно 2, то у переменной Y могут быть два случая: Y_1 и Y_2 . Вероятность появления Y_1 будет равняться p_1 , а вероятность появления Y_2 будет равняться p_2 . Логистическая регрессия относит вероятность p_1 к набору ковариат, используемых для предсказания, используя логистическую связывающую функцию:

$$\text{logit}(p_1) = \ln\left(\frac{p_1}{p_2}\right) = \ln\left(\frac{p_1}{1-p_1}\right) = x_1\beta \quad (2),$$

где x_1 – это вектор переменных-предсказателей, а β – вектор коэффициентов модели, который, как правило, оцениваются методом максимального правдоподобия.

Уравнение (2) может быть переписано в следующем виде:

$$\frac{p_1}{1-p_1} = \exp(x_1\beta) = \exp(\eta) \quad (3),$$

Частное в уравнении (3) называется вероятностью. Из этого уравнения следует:

$$p_1 = \frac{\exp(\eta)}{1 + \exp(\eta)} \quad (4).$$

Бинарная модель логистической регрессии легко преобразуется в множественную. Если у нас имеется n почвенных типов на исследуемой территории, соответственно будет n переменных Y_1, Y_2, \dots, Y_n с соответствующими вероятностями появления p_1, p_2, \dots, p_n . По аналогии с двумя исходной логистической регрессией вероятности будут равняться $p_1/p_n, p_2/p_n, \dots, p_{n-1}/p_n$ и будут описываться средними значениями $\exp(\eta_1), \exp(\eta_2), \dots, \exp(\eta_{n-1})$. Для $\sum_{i=1}^n \pi_i = 1$:

$$\pi_i = \frac{\exp(\eta_i)}{\exp(\eta_1) + \exp(\eta_2) + \dots + \exp(\eta_n)} \quad (5)$$

где $\eta_n=0$. Такое построение модели обеспечивает, что все вероятности находятся в интервале от 0 до 1, а сумма всех вероятностей будет равна 1.

В процессе настройки модели на обучающей выборке формируется таблица, в которой для каждого типа или класса почвы создается вектор с коэффициентами для каждой переменной-предсказателя (ковариаты). При проведении предсказания для всего растра вычисляются вероятности возникновения в этой точке того или иного типа или класса почвы (Kempen и др., 2009). Точка относится к тому классу, вероятность которого для данной точки получается наибольшей. Тем самым, в ходе картографирования для каждой точки растра составляется вектор, состоящий из вероятностей нахождения в этой точке того или иного типа или класса почвы.

Метод деревьев решений

Метод деревьев решений широко применяется для моделирования пространственного распределения почвенных классов (Moonjun и др., 2010; Connolly и др. 2007; Brungard и др., 2015), для дезагрегации старых почвенных карт (Subburayalu и др., 2014; Häring и др., 2012) или картографирования различных почвенных свойств и характеристик (Bou Kheir и др., 2010). В этом методе используется древовидная структура для построения модели классификации. Во время обучения набор данных делится на подмножества.

Одна ветвь представляет собой одно решение. Деревья решений формируются с учетом особенностей выборки. В каждом узле формируется классифицирующее правило, в соответствии с которым дерево решений относит ту или иную точку к тому или иному классу. Классификация начинается с корневого узла, в котором заключено первое правило классификации, то есть то правило, которое четко отделяет один из классов от всех других классов (Quinlan, 1993).

Пусть имеется T — обучающая выборка примеров, состоящая из точек, для которых известен почвенный тип или класс в этой точке и набор ковариат, тогда C — множество почвенных классов или типов, состоящее из k элементов. Для каждого примера из T известна его принадлежность к какому-либо из классов C_1, C_2, \dots, C_k .

На первом шаге имеется узел и ассоциированное с ним множество T , которое необходимо разбить на подмножества. Для этого выбирается одна из ковариат, которая наилучшим образом разделит выборку T . Выбранная ковариата A имеет n значений, что даёт разбиение на n подмножеств. Далее от этого узла создается n ветвей, каждой из которых поставлено в соответствие своё подмножество, полученное при разбиении T . Процедура подбора атрибута и разбиения по нему рекурсивно применяется ко всем n подмножествам и останавливается в двух случаях:

1. после очередного ветвления в вершине оказываются примеры из одного класса; тогда она становится листом, а класс, которому принадлежат её примеры, будет решением листа;

2. вершина оказалась ассоциированной с пустым множеством; тогда она становится листом, а в качестве решения выбирается наиболее часто встречающийся класс у непосредственного предка этой вершины.

В результате получается дерево решений, в соответствии с которым происходит отнесение каждого пикселя растрового изображения к тому или иному почвенному типу или классу.

Метод случайного леса

Метод случайного леса активно используется в цифровой почвенной картографии для описания пространственного распределения почвенных типов (Barthold и др., 2013; Brungard и др., 2015) и для улучшения старых почвенных карт (Pahlavan Rad и др., 2014). Этот метод является статистическим методом, и он развивает большую группу методов, так называемых, классификационных и регрессионных деревьев (CART) (Breiman и др., 1984). Классификационные и регрессионные деревья - это непараметрические методы интеллектуального анализа данных, использующие для моделирования рекурсивное разбиение набора данных для изучения взаимозависимостей между зависимой (определяемой) переменной и переменными-предсказателями. Другими словами, это методы, решающие задачи классификации и регрессии путем построения дерева решений. Для этого набор данных из зависимых переменных разбивается последовательно на деревья все более и более мелкие, и на все более и более однородные подмножества. Разбиение основано на том значении переменной-предсказателя, которое наилучшим образом уменьшает «засоренность» в узлах. Важная особенность этих методов в том, что они часто дают четкое представление о структуре данных, то есть о природе отношений между переменной-отклика и переменными-предсказателями (Prasad и др., 2006). Однако использование этих методов связано с высокой чувствительностью структуры получаемых деревьев к выбору набора данных, на основе которых может происходить переобучение (Prasad и др., 2006).

Метод случайного леса был разработан с целью улучшения прогнозирования при моделировании. Для метода случайного леса процесс построения модели такой же, как и для классификационных и регрессионных деревьев (CART) с той разницей, что строится множество деревьев, то есть лес моделей. Для каждого дерева, используется только подмножество переменных-предсказателей. Число прогнозирующих переменных является

изменяемым пользователем параметром и перестановка переменных между деревьями происходит случайным образом. Этот процесс реализован в пакете «RandomForest» (Liaw, Wiener, 2002). Случайный выбор переменных-предсказателей, среди которых лучшее разбиение достигается в каждом дереве, увеличивает разнообразие леса моделей и, следовательно, уменьшает корреляцию между отдельными деревьями. Это предотвращает переобучение (Breiman, 2001).

Каждое дерево строится на основе бутстреп-выборки входного набора данных, который позволяет на основе оставшейся после изъятия части выборки для построение дерева провести надежную оценку погрешности. Погрешность предсказания вычисляется для каждого дерева, а затем обобщается по всем классификаторам. Результатом предсказания на основе метода случайного леса является один единственный прогноз, который получается из набора предсказаний путем агрегации предсказаний по всем деревьям на основе большинства голосов. Это означает, что самый часто предсказываемый класс на основе каждого дерева из набора деревьев считается результирующим (Liaw, Wiener, 2002).

Преимущества процедуры, заложенной в основу метода случайного леса (Diaz-Uriarte, de Andres, 2006):

- 1) выше производительность прогнозирования;
- 2) отсутствует переобучение;
- 3) низкая корреляция отдельных деревьев между собой, поскольку разнообразие лесов увеличивается за счет использования ограниченного числа переменных-предсказателей;
- 4) низкое смещение и дисперсия из-за усреднения по большому количеству деревьев.

Одним из основных недостатков метода случайного леса является внутренняя сложность получаемого леса моделей (Prasad и др., 2006), которая затрудняет интерпретацию взаимозависимостей между зависимыми

переменными и переменными-предсказателями, поскольку невозможно исследовать строение всех деревьев в лесу.

Во время проведения процедур по моделированию путем метода случайного леса на основе обучающей выборки создается большое количество деревьев. На основе полученных деревьев происходит классификация всех точек растра и отнесение их к одному из классов. То есть каждая точка классифицируется всем множеством деревьев, полученных в результате обучения. Конечным результатом классификации для данной точки являлся тот класс, за который проголосовало большинство деревьев.

Метод опорных векторов

Существует множество работ по ЦПК, где в качестве модели для картографирования почвенных типов и почвенных свойств используется метод опорных векторов (Kovačević и др., 2010; Simin и др., 2010; Lamorski и др., 2008). Метод опорных векторов - это набор непараметрических методов машинного обучения, работающий с бинарными классификационными моделями. Основная идея метода — перевод исходных векторов в пространство более высокой размерности и поиск разделяющей гиперплоскости с максимальным зазором в этом пространстве. Две параллельных гиперплоскости строятся по обеим сторонам гиперплоскости, разделяющей классы. Разделяющей гиперплоскостью будет гиперплоскость, максимизирующая расстояние до двух параллельных гиперплоскостей. Алгоритм работает в предположении, что чем больше разница или расстояние между этими параллельными гиперплоскостями, тем меньше будет средняя ошибка классификатора.

Идею метода удобно проиллюстрировать на примере: даны точки на плоскости, разбитые на два класса (рис. 1.1). Проведем линию, разделяющую эти два класса (красная линия на рис. 1.1). Далее, все новые точки (не из обучающей выборки) автоматически классифицируются следующим образом: точка выше прямой попадает в класс А, точка ниже прямой — в класс В.

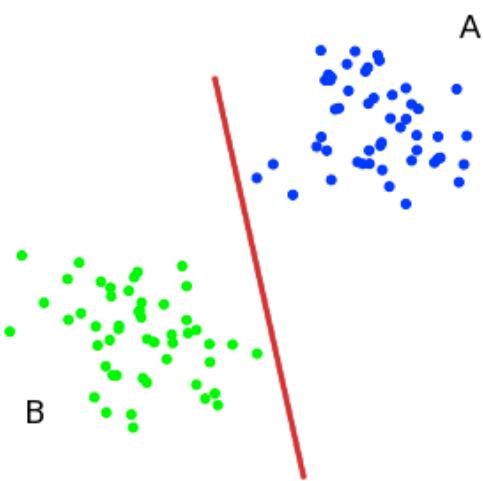


Рис. 1.1 Разделение 2 классов прямой

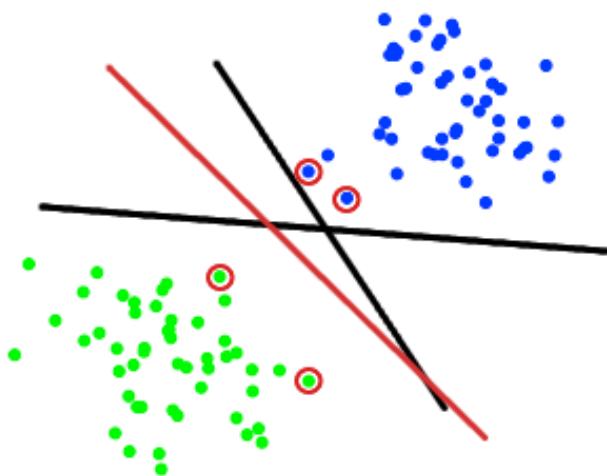


Рис. 1.2 Несколько прямых, разделяющих 2 класса

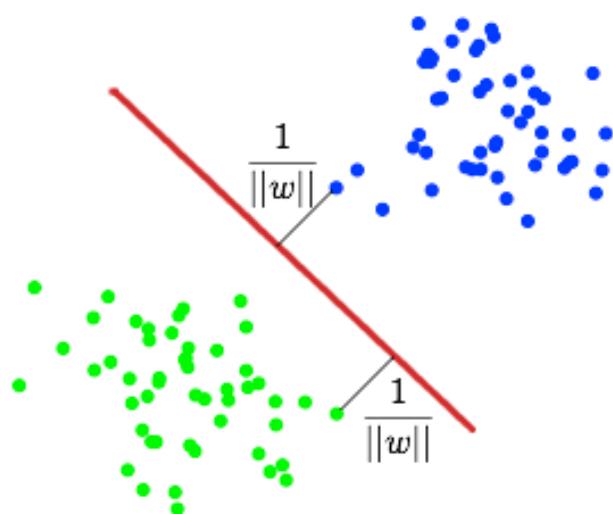


Рис. 1.3 Максимизация расстояния до каждого класса

Такую прямую назовем разделяющей прямой. Однако, в пространствах высоких размерностей прямая уже не будет разделять наши классы, так как понятие «ниже прямой» или «выше прямой» теряет всякий смысл. Поэтому вместо прямых необходимо рассматривать гиперплоскости — пространства, размерность которых на единицу меньше, чем размерность исходного пространства. В трехмерном пространстве, например, гиперплоскость — это обычная двумерная плоскость. В нашем примере существует несколько прямых, разделяющих два класса (рис. 1.2):

С точки зрения точности классификации лучше всего выбрать прямую, расстояние от которой до каждого класса максимально. Другими словами, выберем ту прямую, которая разделяет классы наилучшим образом (красная прямая на рис.1.2). Такая прямая, а в общем случае — гиперплоскость, называется оптимальной разделяющей гиперплоскостью. Вектора, лежащие ближе всех к разделяющей гиперплоскости, называются опорными векторами (support vectors). На рисунке 1.2 они помечены красным.

Пусть имеется обучающая выборка: $(x_1, y_1), \dots, (x_m, y_m)$, $x_i \in R^n$, $y_i \in \{-1; 1\}$. Метод опорных векторов строит классифицирующую функцию F в виде

$$F(x) = \text{sign} ([w, x] + b) \quad (6),$$

где $[,]$ — скалярное произведение, w — нормальный вектор к разделяющей гиперплоскости, b — вспомогательный параметр. Те объекты, для которых $F(x) = 1$ попадают в один класс, а объекты с $F(x) = -1$ — в другой. Выбор именно такой функции неслучаен: любая гиперплоскость может быть задана в виде $[w, x] + b = 0$ для некоторых w и b (рис. 1.3).

Далее, мы хотим выбрать такие w и b которые максимизируют расстояние до каждого класса (рис. 1.3). Можно подсчитать, что данное расстояние равно $\frac{1}{\|w\|}$. Проблема нахождения максимума $\frac{1}{\|w\|}$ эквивалентна проблеме нахождения минимума $\|w\|^2$. Запишем все это в виде задачи оптимизации:

$$\begin{cases} \arg \min_{w,b} \|w\|^2, \\ y_i ([w, x] + b) \geq 1, i = 1, \dots, m. \end{cases} \quad (7)$$

которая является стандартной задачей квадратичного программирования и решается с помощью множителей Лагранжа.

На практике случаи, когда данные можно разделить гиперплоскостью довольно редки. В этом случае все элементы обучающей выборки вкладываются в пространство X более высокой размерности с помощью специального отображения $\varphi: R^n \rightarrow X$. При этом отображение φ выбирается так, чтобы в новом пространстве X выборка была линейно разделима. Классифицирующая функция F принимает вид $F(x) = \text{sign}([w, \varphi(x)] + b)$.

При бинарном моделировании предполагается, что почва принадлежит только к одному классу, при этом существует всего два класса ($C = \{c_1, c_2\}$). Каждая классификационная задача с n количеством классов может быть смоделирована как последовательность $(n/2)$ бинарных задач с использованием подхода один-против-одного, при котором происходит возникновение $n^*(n-1)/2$ бинарных классификаторов по одному для каждой пары классов. Окончательное решение принимается путем голосования, т. е. наиболее часто прогнозируемый класс выбирается в качестве результирующего.

Метод регрессионных деревьев

Другим распространенным методом ЦПК являются регрессионные деревья, которые часто используются для пространственного моделирования почвенного покрова (Lemercier и др., 2012; Schmidt и др., 2008; Scull и др., 2005). Алгоритм построение регрессионных деревьев, в целом, схож с алгоритмом построения деревьев решений. Однако во время классификации на основе регрессионных деревьев строится множество деревьев, и существует второй этап, который заключается в использовании кросс-валидации для создания конечного дерева решений. Проводится кросс-валидационная оценка

риска для всех образовавшихся деревьев, и конечной моделью будет то дерево, которое будет иметь самую низкую оценку риска.

При моделировании пространственного распределения почвенного покрова на основе метода регрессионных деревьев реализуется разбиение обучающей почвенной выборки на непересекающиеся подмножества на основе использования наиболее подходящей для этого переменной-предсказателя, что приводит к получению подмножеств, содержащих минимальное количество посторонних включений. Далее проводится поиск переменной-предсказателя, которая наилучшим образом разобьет получившиеся подмножества и так далее до того момента, когда получающееся подмножество будет содержать минимальное количество компонентов. Правила разбиения записываются в узлы деревьев. На основе полученного регрессионного дерева происходит дальнейшая классификация всего растрового изображения и отнесение каждого пикселя к тому или иному классу.

Другие методы

В современном ЦПК существует множество других методов, применяемых для описания, моделирования или предсказания пространственного распределения почвенных типов, классов и даже разрядов. Например, активно используются различные дискриминационные и классификационные методы для описания пространственного распределения почвенных типов и классов: например, *метод главных компонент* (Du и др. 2008; Hengl и др. 2007), *дискриминантный анализ* (Du и др. 2008; Konyushkova, 2010), *кластеризация и группировка* (Hengl и др. 2007), *классификация на основе максимального подобия* (Hengl и др. 2007; Кириллова и др., 2014), *канонический анализ* (Liu и др. 2008). Часто хорошие результаты моделирования дают *искусственные нейронные сети* (Moonjun и др., 2010; Du и др., 2008; Taghizadeh-Mehrjardi и др., 2015).

Быстрое развитие методов нечеткой логики для нужд ЦПК связано со способностью этого метода выявлять скрытые закономерности и описывать непрерывный характер пространственной изменчивости почв (McBratney, Odeh, 1997; Zhu, 1997). В работах по картографированию пространственного распределения почвенных классов методом нечеткой логики используются два различных 2 алгоритма: fuzzy c-means (FCM) (De Gruijter и др., 1997) и Semantic Import Model (SI) (Burrough и др., 1992). Примером успешного применения этих методов может служить работа по пространственному моделированию агроэкологических группировок почвенных типов (Vitharana и др., 2008; MacMillan и др., 2007).

Другим направлением картографирования почвенного покрова, используемым в ЦПК является ГИС-моделирование, используемое, например, при описании распределения почвенных типов на территории Антарктики (Bockheim, McLeod, 2008), болотных почв в Ирландии (Connolly и др. 2007). Кроме того, ГИС-системы используются для мониторинга окружающей среды (Мясникова, Васенев, 2015; Мищенко, Трифонова, 2007).

Для нужд ЦПК часто используются существующие почвенные карты, материалы почвенных и геоботанических обследований территорий (Du и др. 2008; Nauman, Thompson, 2014; Pahlavan-Rad и др., 2014). Для этих случаев используется алгоритм дезагрегации и гармонизации почвенных типов на основе классификационных деревьев DSMART (Disaggregation and Harmonisation of Soil Map Units Through Resampled Classification Trees), который дает возможность быстрого определения наиболее и наименее вероятного для данной точки почвенного типа на основе старых почвенных карт (Odgers и др., 2014).

В ЦПК методы, основанные на экспертных знаниях, используются реже по сравнению со стохастическими и детерминированными моделями для прогнозирования почвенного покрова (Grunwald, 2009). В то же время, использование комплексных моделей для предсказания почв чаще точнее

отражает реальную ситуацию, чем предсказание на основе простых моделей (Brungard и др., 2015).

Методы, используемые для построения цифровых почвенных карт, можно ранжировать следующим образом (Grunwald, 2009):

1-е место - регрессии (41,1%),

2-е место - классификационные дискриминационные методы (32,2%),

3-е место - методы однофакторного кригинга (18,9 %),

4-е место - методы на основе деревьев, например, классификационные деревья, деревья регрессии и случайного леса) (13,3%).

1.3 Применение ЦПК для картографирования пространственного распределения органического углерода и запасов гумуса

Изучение пространственного распределения запасов гумуса в почвах имеет важное значение для решения целого ряда научных и практических вопросов: отслеживание изменения климата, изучение почвенного плодородия и пространственных закономерностей распределения органического вещества в различных ландшафтах (Mendonça-Santos и др., 2010; Grimm и др., 2008; Sidorova, Krasilnikov, 2007; Кузяков, Кузякова, 1997).

Первые работы по предсказанию содержания органического углерода на основе моделирования описаны для лесных почв умеренных широт Франции. Предсказание проводилось на основе модели линейной регрессии, в качестве ковариатов использовались данные по почвообразующим породам, абсолютным высотам местности и углам наклона (Arrouays и др., 1998)

В работе китайских ученых Yong-Cun Zhao и Xue-Zheng Shi сравнивалось пять моделей, построенных на основе 359 точек измерений, на основе которых проводилось картографирование в провинции Хэбэй (Zhao, Shi, 2010): множественная линейная регрессия; универсальный кригинг; регрессионный кригинг; нейронные сети в сочетании с кригингом; деревья регрессии. Наилучшие результаты показали методы регрессионных деревьев и нейронных сетей (точность предсказания - 67%), в то время как метод

множественной линейной регрессии давал правильные предсказания лишь в 19,5% случаев (Zhao, Shi, 2010).

В работах чилийских ученых проводилось сравнение картограмм пространственного распределения гумуса, построенных на основе четырех методов: ординарного кригинга, кокригинга, ординарного кригинга с применением локальных семивариограмм и нейронных сетей. Точность предсказания различными методами составила от 67 до 69% (Padarian и др., 2012).

Показано хорошее качество предсказания содержания органического углерода в неизвестных точках на территориях полузасушливых степных экосистем Китая и Монголии, площадью около 3600 км², на основе классификационных и регрессионных деревьев, а также методов случайного леса (Wiesmeier и др., 2011).

Целый ряд различных моделей применялся для описания пространственного варьирования гумуса в штате Флорида (США): множественная логистическая регрессия, регрессионные деревья и деревья решений, различные регрессии (Vasques и др. 2009). Американскими учеными также была проделана работа по картографированию содержания органического углерода на основе анализа ДДЗ, используя метод нейронных сетей, в качестве ковариат использовались данные по растительному покрову (Chen и др., 2008).

Существуют работы по картографированию содержания органического углерода на основе регрессионных методов, где в качестве ковариат применялись ДДЗ: например, для пустынных территорий Австралии (Gomez и др., 2008), для сельскохозяйственных полей Европы (Stevens и др., 2008).

Путем использования лабораторных результатов спектрального анализа образцов и ДДЗ, моделируя на основе дискриминантного анализа, метода главных компонент и регрессии было описано пространственное распределение органического углерода на территориях недавно вовлеченных

в сельскохозяйственную деятельность к югу от пустыни Сахара (Awiti и др., 2008).

Результаты картографирования содержания органического углерода, получаемые на небольшие по площади территории, могут быть экстраполированы на большие территории путем применения экспертных оценок и геостатистических методов (Yigini, Panagos, 2014).

ГЛАВА 2. ОБЪЕКТЫ И МЕТОДЫ

2.1 Географическое положение, общая характеристика и история сельскохозяйственного использования ключевого участка

Ключевой участок охватывает примерно половину площади хозяйства ООО «Чура», которое располагается в Глазовском районе севера Удмуртской Республики в 9 км юго-западнее г. Глазова. Протяженность участка с севера на юг – 4,5 км, а с запада на восток – 5,2 км (рис. 2.1).



Рис. 2.1 Космический снимок территории исследования WorldView-1. Дата съемки 15 сентября 2015 г.

Общая площадь его составляет 19 км². В центральной части исследуемого района расположена д. Кыпка, и протекает одноименная река. В восточной части ключевого района протекает река Убыть, а северо-западной его части – река Варсемашур.

Транспортная связь между производственными участками хозяйства осуществляется по грунтовым дорогам и дорогам с твердым покрытием.

Основное направление деятельности данного хозяйства – мясное животноводство. Севернее д. Кыпка располагается комплекс по выращиванию молодых бычков. На полях хозяйства выращивается пшеница яровая, ячмень яровой и овес, рапс, корнеплодные кормовые культуры, а также горох. Часть полей засевается однолетними и многолетними травами. Поля, засеянные многолетними травами, используются в качестве сенокосов и в качестве пастбищ для крупного рогатого скота. Часть угодий, засеянных многолетними травами, используются для выращивания их на семена (люцерна, клевер).

Антропогенное влияние на исследуемой территории значительно: 65 % территории занимают пашни, 5 % - сенокосы, 6 % - приусадебные участки д. Кыпка (Докучаев, 2015б).

Пахотные земли хозяйства представлены участками, занимающими водораздельные пространства между балками и лощинами. В хозяйстве на территории ключевого участка выделяется 6 полей и 35 рабочих участков. Площади рабочих участков варьируются от 1,3 га до 123,9 га (данные взяты из проектной документации). Также имеются обширные лесные массивы и небольшое количество сенокосных площадей.

На территории ключевого участка используются различные приемы обработки почв и внесения удобрений. Например, лущение стерни чистого пара, августовская зяблевая вспашка, культивация и боронование при появлении сорняков и другие приемы агротехники. Семена высеваются поперек склона с последующим их прикатыванием. На поля с многолетними и однолетними травами регулярно вносятся рассчитанные дозы минеральных и азотных удобрений. Кроме того, проводятся мероприятия по снегозадержанию и регулированию снеготаяния.

Хозяйство характеризуется высоким природно-ресурсным потенциалом и значительной обеспеченностью производственными ресурсами, а также высокой культурой землепользования (Докучаев, 2015а).

2.2 Агроэкологические условия

2.2.1 Агроклиматические ресурсы

Климат района исследования умеренно-континентальный с продолжительной холодной многоснежной зимой и довольно жарким летом. Большую часть года преобладают юго-западные ветры.

В зимнее время вторжение арктического воздуха с севера приносит сильные морозы (до -40°C и ниже), а весной и осенью - частые ночные заморозки. Зимние потепления, доходящие до оттепели, вызываются вхождением морского воздуха с запада, с Атлантики. Летом такая воздушная масса приносит прохладную влажную погоду. В летнее время с юга и юго-востока приходят очень теплые воздушные массы, с которыми связана засушливая, жаркая погода. Средняя температура самого холодного месяца - $14^{\circ} - 17^{\circ}\text{C}$, самого теплого - $17,2^{\circ} - 18,8^{\circ}\text{C}$ (табл.2.1).

Осадки составляют 525-560 мм, коэффициент увлажнения 1,00-1,33 (табл. 2.1). По характеру увлажнения территория относится к зоне с неустойчивым увлажнением, где в теплое время испарение нередко превышает количество выпавших осадков.

Продолжительность периода с температурой воздуха выше 0°C равна 190-210 дням, а с температурой выше 10°C – 111-115 дням. Сумма среднесуточных температур за период с температурой выше 10°C около 2100.

Зима холодная и продолжается шесть с половиной месяцев. Продолжительность периода с устойчивым снежным покровом составляет 170 дней. Глубина промерзания почв составляет 90 см. Запасы продуктивной влаги весной в метровом слое почв на зяби составляют 175-200 мм (Ковриго, 2004).

Из неблагоприятных климатических явлений для ведения сельского хозяйства следует отметить:

1) холодные, многоснежные и продолжительные зимы, после которых наступает резкое потепление, соответственно, происходит резкое высвобождение большой массы талой воды.

Таблица 2.1

Характеристика агроклиматических условий

<i>Показатель</i>	<i>Характеристика</i>
Среднегодовая температура воздуха, °C	+ 3
Годовое количество осадков, мм	556
Средний из абсолютных максимумов температуры, °C	+ 37,5
Средний из абсолютных минимумов температуры, °C	- 40,2
Средняя температура самого теплого месяца, °C	17,2-18,8
Средняя температура самого холодного месяца, °C	- 15,5-17
Сумма активных (выше 10 °C) температур	2100
Сумма осадков за период активной вегетации (со среднесуточными температурами выше 10 °C), мм	150...220
Гидротермический коэффициент	1,0...1,33
Весенняя дата перехода среднесуточной температуры через 0 °C	05 апреля
через + 10 °C	03 мая
Осенняя дата перехода среднесуточной температуры через + 10 °C	26 сентября
через 0 °C	15 октября
Продолжительность периода (в днях) со среднесуточной температурой выше:	
0 °C	208
+ 5 °C	174
+ 10 °C	119
+ 15 °C	89
Дата наступления заморозков: последних весенних	03 мая
первых осенних	21 сентября
Продолжительность безморозного периода, дни	126
Дата образования устойчивого снежного покрова	09 октября
Число дней в году с устойчивым снежным покровом	203
Средняя высота снежного покрова, см	36,5
Максимальная высота снежного покрова, см	92
Дата схода устойчивого снежного покрова	24 апреля

Продолжение таблицы 2.1

<i>Показатель</i>	<i>Характеристика</i>
Средняя температура (°C):	
Апреля	+ 5,6
Июля	+ 20,1
Октября	+ 1,5
Января	- 14,0
Количество осадков (мм) за:	
Весну	120...380
Лето	140...420
Осень	170...300
Зиму	100...220
Направление ветра по месяцам:	
В апреле	З, СВ, ЮЗ
В июле	З, СВ
В октябре	ЮЗ, З, В
В январе	В, ЮВ
Выращиваемые сельскохозяйственные культуры, и культуры, которые можно выращивать	Озимая рожь, ячмень, многолетние травы, пшеница яровая, овес, горох, бобовые

2) Часть влаги не успевает впитываться в почву и стекает в гидрографическую сеть. Это приводит к сильной водной эрозии почв в этом регионе. В отдельные годы наблюдаются засухи. В среднем, каждый 3 – 4 год является засушливым.

2.2.2 Геоморфологические условия

Территория ключевого участка сильно расчленена овражно-балочной сетью, между которой располагаются пологие увалы (Подсосова, 1972). Современные очертания рельеф принял в четвертичный период.

Во время плейстоцена происходило выравнивание территории вследствие водно-ледниковой аккумуляции отложений, а в период голоцена

стали активно проявляться эрозионные процессы (плоскостная и линейная эрозия), внесшие существенные дополнения в очертания рельефа Глазовского района (Ковриго, 2004). Максимальный перепад высот в пределах участка составляет 65 м. Геоморфологические условия землепользования ключевого участка «Чура» приведены в таблице 2.2.

Таблица 2.2

Геоморфологические условия землепользования ключевого участка «Чура»

<i>Показатель</i>	<i>Характеристика</i>
Геоморфологическая область	Возвышенная эрозионно-денудационная равнина
Абсолютные отметки высот водоразделов	Более 220 м
Геоморфологический район	Верхнекамская возвышенность
Тип рельефа	Неоген-четвертичные сильно расчлененные эрозионно-денудационные возвышенные равнины вне ледниковой области
Типы морфоскульптуры (формы рельефа, созданные экзогенными процессами)	Азональная, созданная эрозионно-денудационными процессами; увалистый рельеф
Тип эрозионных систем	Глубоковрезанные овражно-балочные и лощинно-балочные
Степень вертикального расчленения	65 м
Коэффициент горизонтальной расчлененности территории	0,6...1,5
Коэффициент расчлененности территории оврагами, км/км ²	более 2,0
Неблагоприятные экзогенные процессы	Растущие овраги и промоины, донные овраги в балках, активные процессы водной эрозии

Для разделения почвенного покрова по степени эродированности и для выявления закономерностей распределения почв по рельефу вся исследуемая территория была условно разделена нами на 10 элементов мезорельефа согласно предложению А.В. Гедымина (Гедымин, 1990). Были выделены ровные поверхности крутизной до 1 град., слабые склоны 1-3 град., средние по крутизне склоны 3-5 град., крутые склоны 5-8 град. На изучаемой территории

встретились также склоны более 8 град. Кроме того, были определены склоны выпуклой и вогнутой формы (рис. 2.2, табл.2.3). На территории ключевого участка преимущественно распространены склоны от 1 градуса до 3 градусов (44,6 %) (табл. 2.3).

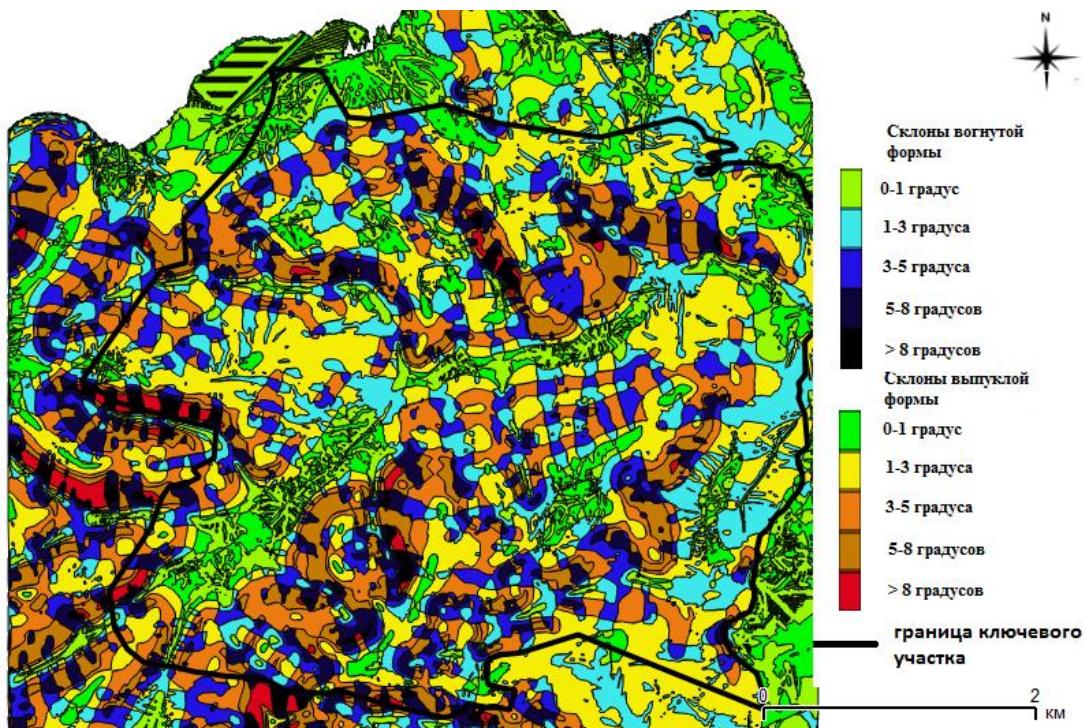


Рис.2.2 Карта элементарных поверхностей мезорельефа на исследуемую территорию. Масштаб 1: 50 000

В тоже время практически треть территории занята склонами от 3 до 5 град. (31,5 %). Выровненные поверхности и крутые склоны занимают по 11,1 % и 11,3 %, соответственно.

Таким образом, изрезанность территории балками, оврагами, лощинами и ложбинами стока создают неудобства при обработке пахотных угодий. Поля в результате часто имеют небольшой размер: менее 50 га. Преобладание склоновых территорий в пахотных землях приводит к распространению эрозионных процессов и потерям влаги с поверхностным и внутрипочвенным стоком в гидрографическую сеть. Стоит отметить, что работами местных почвоведов (Ковриго, 2004; Колосков, 1939; Тайчинов, 1973) было показано, что рельеф в сочетании с особенностями почвообразующих пород, оказывает

решающее значение на формирование почвенного покрова исследуемой территории.

Таблица 2.3

Распределение элементов мезорельефа на территории ключевого участка

Слоны крутизной		Общая площадь, км ²	% от общей площади	
0-1 град.	вогнутые	1,1	5,8	11,1
	выпуклые	1	5,3	
1-3 град.	вогнутые	3,27	17,2	44,6
	выпуклые	5,22	27,4	
3-5 град.	вогнутые	2,73	14,3	31,5
	выпуклые	3,27	17,2	
5-8 град.	вогнутые	0,96	5	11,3
	выпуклые	1,19	6,3	
> 8 град.	вогнутые	0,13	0,7	1,5
	выпуклые	0,15	0,8	

2.2.3 Литологические условия

Согласно карте почвообразующих пород Удмуртской Республики (Ковриго, 2004) территория ключевого участка однородна по распространению почвообразующих пород, которыми являются тяжелосуглинистые покровные красновато-бурые суглинки водно-ледникового происхождения, подстилаемые моренными отложениями. Эти отложения бескарбонатные, хорошо отсортированные, по валовому содержанию SiO₂ – кислые. Местами на крутых склонах моренные отложения, содержащие гальку, выходят наружу.

В поймах рек Убыть, Кыпка и Варсемашур почвообразующими породами являются современные аллювиальные отложения.

Почвы, сформированные на тяжелосуглинистых покровных суглинках, не вполне благоприятны в агрономическом отношении: следует отметить малую мощность гумусового слоя, низкую обеспеченность подвижным

фосфором и азотом, недостаток микроэлементов, кислую реакцию среды верхних гумусовых горизонтов и тяжелый гранулометрический состав пахотного слоя.

2.2.4 Гидрологические и гидрографические условия

Главным элементом гидрографической сети ключевого участка является река Убыть, расположенная в его восточной части ключевого участка. В центральной части участка протекает река Кыпка, которая является притоком реки Убыть. В северо-западной части участка протекает река Варсемашур, которая в период полевых сезонов 2014-2015 гг. была пересохшей.

Ручьи, протекающие по днищу оврагов и балок и часто пересыхающие летом, имеют сток к рекам Убыть и Кыпка. Помимо оврагов и балок к гидрографической сети относятся лощины и ложбины стока, широко распространенные на территории хозяйства.

Грунтовые воды на водораздельных территориях залегают обычно глубоко и не оказывают существенного влияния на водный режим почв. По днищам балок, поймам рек Убыть, Кыпка и Варсемашур грунтовые воды залегают близко к поверхности, что приводит к их заболачиванию и оглеению.

Таким образом, поверхностные воды, стекая по ложбинам стока, приводят к эрозии. А после ливней и весенних паводков по ложбинам стока происходит смыв почв и погребение свежими наносами посевов. В то же время грунтовые воды на водораздельных территориях находятся глубоко и не оказывают влияния на урожай.

2.2.5 Растительность

Растительный покров исследуемой территории относится к южно-таежной лесной зоне. Естественная растительность севера Удмуртии входит в северо-западную геоботаническую зону и представлена в основном пихтово-еловыми и темными сосновыми лесами со значительной примесью березы. Соотношение хвойных и лиственных пород в лесах составляет 57% и 43%, соответственно (Ефимова и др., 1972).

Лесистость исследуемого участка составляет 24 %. Наиболее облесенными элементами рельефа являются крутые склоны, а также балки и овраги. Можно выделить 3 крупных массива леса: на северо-востоке участка - площадью 62 га, на востоке – 92 га и на юго-востоке – 85 га. Травяный ярус лесов представлен мятым, зверобоем, костяникой, клевером, заразихой, папоротником, кислицей, снытью.

Пойменные луга занимают около 5 % территории. Преобладают злаково-разнотравные травостоя, в состав которых входят: тимофеевка, лисохвост, овсянница красная, люцерна, клевер. Днища балок покрыты крапивой, малиной, папоротником и хвошами, а почвенный покров представлен дерново-глеевым почвами.

2.2.6 Почвенный покров

Согласно почвенно-географическому районированию ключевой участок принадлежит Вятско-Камской провинции дерново-подзолистых почв южной тайги (Добровольский, Урусовская, 2004). В соответствии с этим самыми распространенными почвами на исследуемой территории являются дерново-неглубокоподзолистые и дерново-глубокоподзолистые почвы. Дерново-неглубокоподзолистые целинные лесные несмытые почвы имеют сплошной подзолистый горизонт до 15 см, а у дерново-глубокоподзолистых почв его мощность более 15 см. На пашне мощность подзолистого горизонта у несмытых дерново-неглубокоподзолистых почв доходит до 10 см. Чаще всего в профилях дерново-подзолистых почв отсутствуют признаки глееватости, границы генетических горизонтов четкие (Ковриго, 2004).

Большая часть дерново-подзолистых почв на пашне представлена смытыми категориями. Встречаются также дерново-подзолистые почвы со вторым гумусовым горизонтом.

Небольшие участки заняты светло-серыми лесными почвами. Дерново-глеевые и аллювиальные почвы являются интразональными почвами для исследуемого участка.

Гранулометрический состав почв на исследуемой территории средне- и тяжелосуглинистый. В связи с активной сельскохозяйственной деятельностью и процессами водной эрозии, протекающими в данном регионе, большая часть почвенных профилей имеет черты нарушенного состояния.

2.3 Характеристика материалов, имеющихся на район исследования

2.3.1 Архивные материалы хозяйства

В ходе сбора материала на район исследования от главного агронома хозяйства нами была получена проектная документация «Система земледелия и землеустройство колхоза «Чура» Глазовского района Удмуртской АССР», разработанная Удмуртским филиалом института Уралгипрозема, совместно со специалистами хозяйства и при участии Удмуртской проектно-изыскательской станции химизации сельского хозяйства и Удмуртской республиканской станции по защите растений 1984 г. выпуска.

Одним из основных материалов, используемых нами в дальнейшем, была почвенная карта 1984 г. (рис. 2.3).

Согласно почвенной карте 1984 г. (рис. 2.3) на территории исследования представлены четыре типа почв и 19 их разновидностей. При этом примерно 62 % территории занимают дерново-подзолистые почвы разной степени смытости и оподзоленности (Классификация..., 1977).

Наибольшее распространение получили дерново-подзолистые слабосмытые почвы (45,7 % от общей площади территории, рис. 2.3). Несмытые дерново-подзолистые почвы занимают лишь 8,7 % от общей площади (рис. 2.3), среднесмытые – 10,9 % (рис. 2.3), сильносмытые – 2,0 % (рис. 2.3). Большое количество дерново-глеевых почв встречается на территории исследования (19,3% от общей площади, рис. 2.3). Аллювиальные почвы распространены в долине рек Кыпка и Убыть (рис. 2.3) и занимают 3,7 % территории. Светло-серые и серые лесные почвы встречаются в

подчиненных позициях рельефа: в нижних расположенных частях и подножиях склонов (рис. 2.3) (Ковриго, 2004).

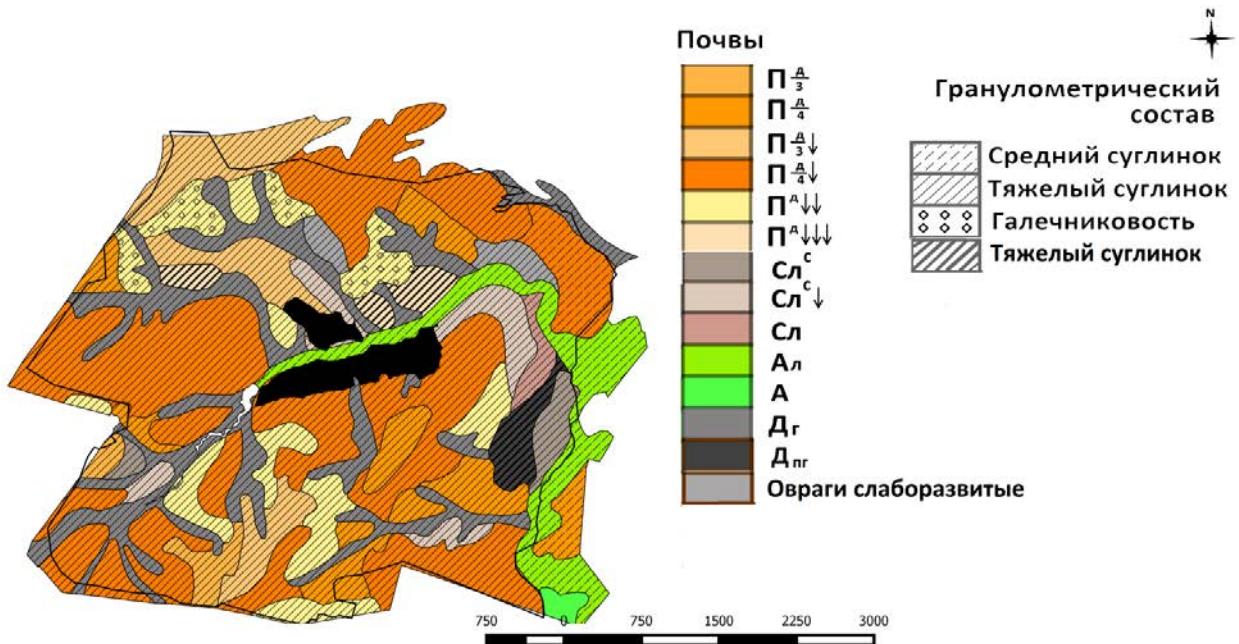


Рис. 2.3 Почвенная карта 1984 г. (масштаб 1:50000): Ал - аллювиальная иловато-торфяная, А – аллювиальная дерновая кислая слоистая, Дг - дерново-глеевая, $\text{П}_{\frac{A}{3}}$ – дерново-неглубокоподзолистая, $\text{П}_{\frac{A}{4}}$ – дерново-глубокоподзолистая, $\text{П}_{\frac{A}{3}}\downarrow$ – дерново-неглубокоподзолистая слабосмытая, $\text{П}_{\frac{A}{4}}\downarrow$ – дерново-глубокоподзолистая слабосмытая, $\text{П}^{\downarrow\downarrow}$ – дерново-подзолистая среднесмытая, $\text{П}^{\downarrow\downarrow\downarrow}$ – дерново-подзолистая сильносмытая, СЛ – серая лесная, СЛ^c – светло-серая лесная, СЛ^c↓ – светло-серая лесная слабосмытая

Эти типы почв, согласно карте 1984 г., занимают незначительную часть территории, около 7 % от общей площади. Гранулометрический состав почв на данной территории преимущественно тяжелосуглинистый.

В хозяйстве на территории ключевого участка выделяется 6 полей и 35 рабочих участков, имеются обширные лесные массивы и небольшое количество сенокосных площадей (рис. 2.4). В тоже время, территория исследования подвержена воздействию водной эрозии (рис. 2.5).

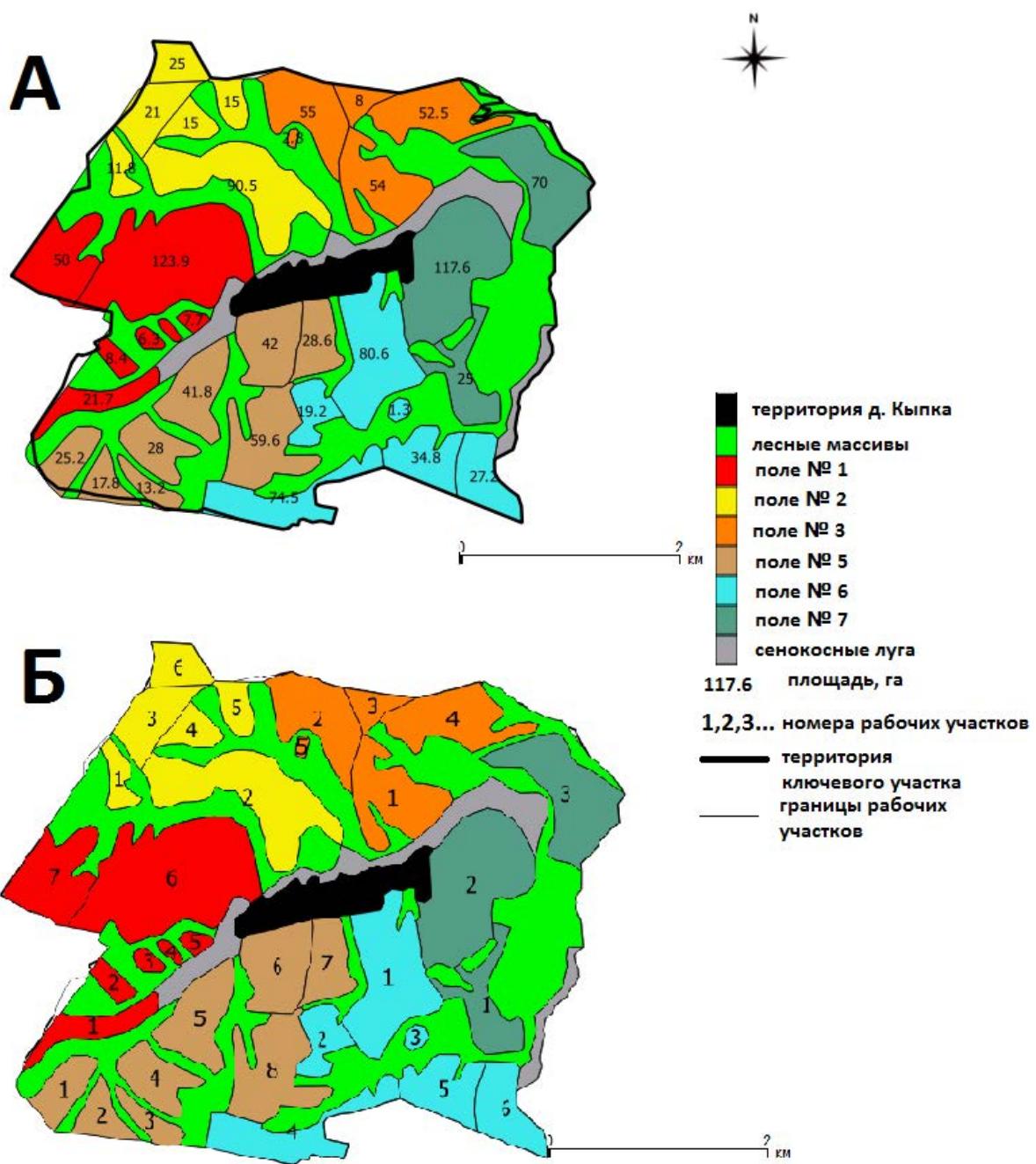


Рис. 2.4 Карта организации пашни. Масштаб 1: 50 000. А- площади рабочих участков. Б- номера рабочих участков полей

Площади рабочих участков различны и варьируются от 1,3 га до 123,9 га (площади взяты из проектной документации). Площадь пашни составляет 67 % от общей площади территории. Сопоставление карты организации пашни (рис. 2.4) и карты элементарных поверхностей мезорельефа (рис. 2.2) показало, что наиболее крупные рабочие участки занимают самые крупные и наиболее выровненные участки территории исследования.

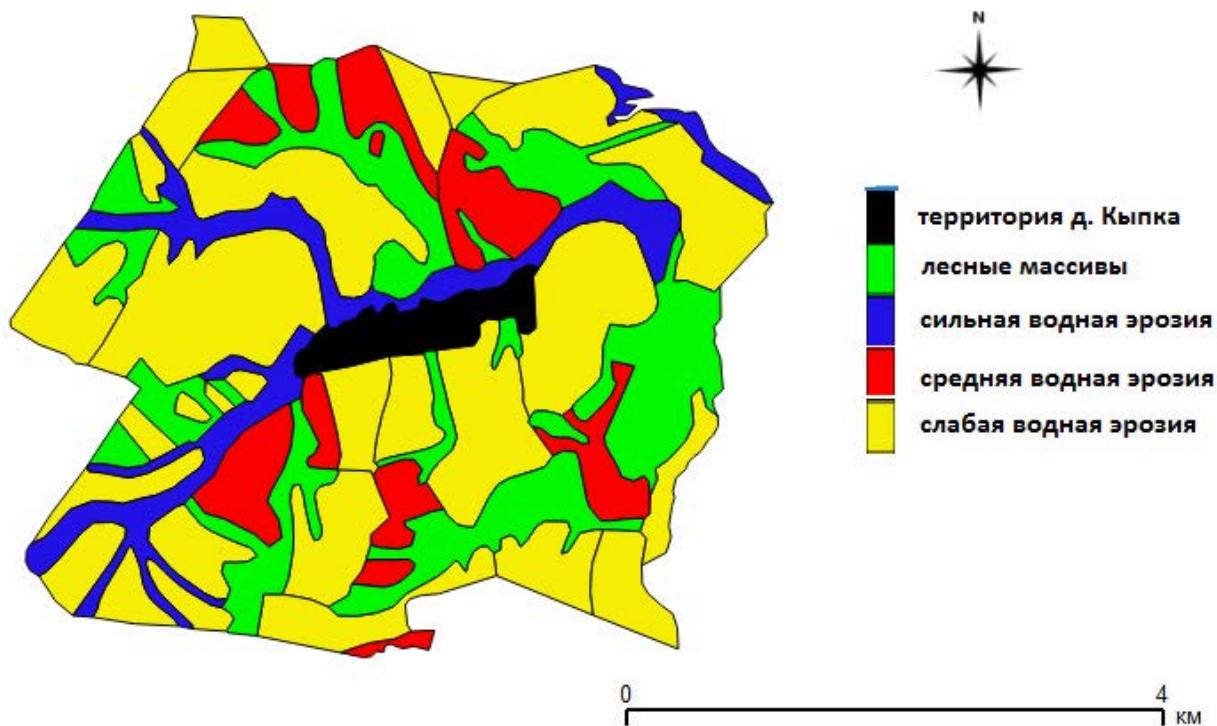


Рис. 2.5 Карта эрозии 1984 г. Масштаб 1: 50 000

Специалистами Уралгипрозема выделяются области сильной водной эрозии в поймах рек Кыпка, Варсемашур и в наиболее глубоких балках. Территории, подверженные сильной водной эрозии, занимают около 10% общей площади. Около 12% территории заняты почвами, подверженными водной эрозии средней степени (рис. 2.5).

По почвенной карте 1984 г. были отрисованы границы почвенных контуров и созданы файлы с векторными слоями, включающими следующие условные почвенные группы: дерново-подзолистые почвы разной степени смытости, аллювиальные, серые лесные и дерново-глеевые почвы.

2.3.2 Топографическая карта и ее производные

Для создания ЦМР была оцифрована топографическая карта масштаба 1:50 000 (рис. 2.6). В результате проведения оцифровки топографической карты был создан векторный слой с отрисованными изолиниями рельефа. Полученный файл с векторным слоем, содержащим изолинии рельефа, был обработан. На основе универсального кригинга были построены ЦМР с разрешением 5, 20 и 50 м. В дальнейшем на основе этих моделей рельефа были

рассчитаны морфометрические характеристики рельефа и построены их растровые изображения.

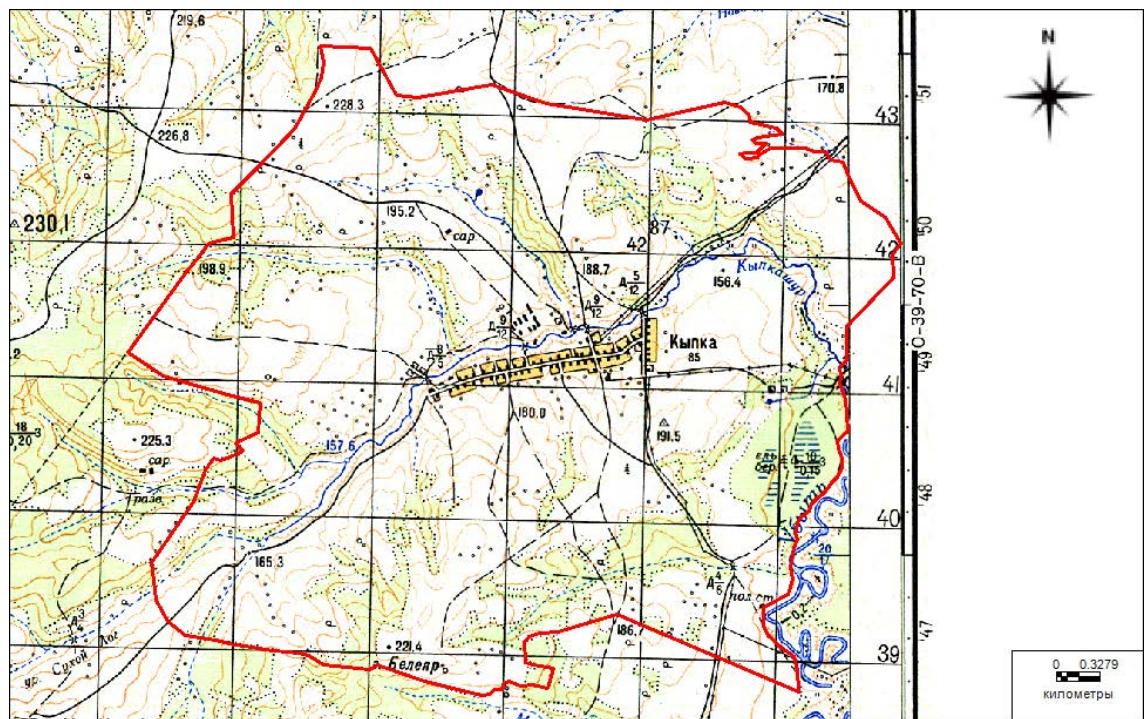


Рис. 2.6. Топографическая карта на исследуемую территорию масштаба 1:50 000. Красной линией показаны границы участка

2.3.3 Анализ космического снимка

Был использован космический снимок высокого разрешения WorldView-1. Дата съемки 15 сентября 2015 г. (рис. 2.1). Снимок состоял из 4 каналов, характеризующих синюю, зеленую, красную и инфракрасную часть электромагнитного спектра видимого излучения. Пространственное разрешение космического снимка 2 м.

На основе этого космического снимка был рассчитан индекс NDVI и создано растровое изображение, характеризующее пространственное распределение индекса на территории исследования (рис. 2.7).

Космический снимок позволил выявить области избыточного увлажнения и использовался в планировании и реализации полевых исследований территории.



Рис. 2.7 Индекс NDVI для территории исследования

Созданное на его основе растровое изображение, характеризующее распределение индекса NDVI на территории исследования, использовалось в качестве ковариаты для цифрового почвенного картографирования.

2.4 Полевой и камеральный этапы 2014-2016 гг.

2.4.1 Полевое обследование территории

Площадь территории исследования составила около 19 км², категория сложности местности для почвенной съемки – 4. А это значит, что для составления почвенной карты масштаба 1:50000 необходим один разрез на 80 га, а соотношение между почвенными разрезами, полумирами и прикопками должно составлять 1:4:2 при работе с использованием материалов космической съемки (Общесоюзная инструкция..., 1973; Методика..., 1962).

В рамках исследования было проведено 3 полевых сезона (рис. 2.8):

1. Полевой этап 2014 г.: рекогносцировка, знакомство с местностью особенностями сельскохозяйственной деятельности, выявление закономерностей пространственного распределения почв в районе ключевого

участка, проведение рекогносцировочной съемки с заложением 8 полнопрофильных почвенных разреза.

2. Полевой этап 2015 г.: проведение основных почвенных изысканий, планирование маршрутов почвенных съемок, отбор почвенных образцов. Было заложено 15 полнопрофильных разреза, 93 полужамы и 50 прикопок.

3. Полевой этап 2016 г.: проведение обследования по случайной независимой выборке, заложение почвенных прикопок в 27 точках выборки.

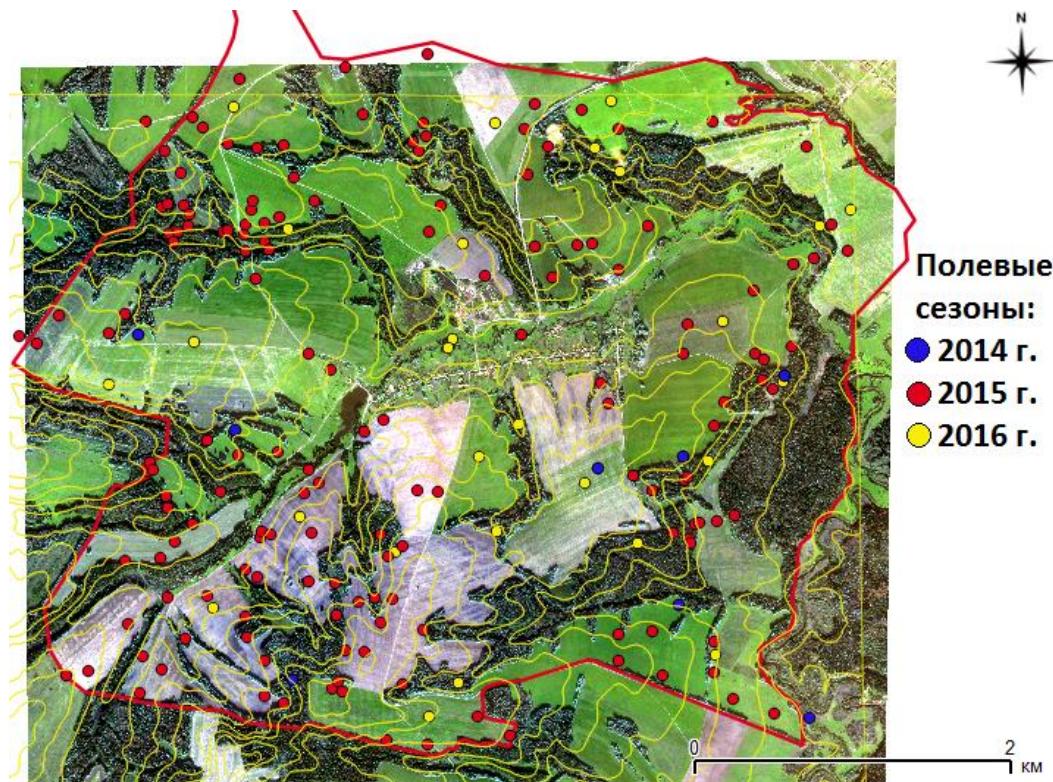


Рис 2.8 Распределение точек обследования полевых сезонов 2014-2016 гг.

Таким образом, в ходе полевых сезонов 2014-2015 гг. было обследовано 166 точек заложения почвенных профилей, из которых 7 относились к аллювиальным почвам, 6 – к дерново-глеевым почвам, 10 – светло-серым лесным почвам, 12 – к дерново-подзолистым несмытым почвам, 69 - к дерново-подзолистым слабосмытым почвам, 44 - к дерново-подзолистым среднесмытым почвам, 9 - к дерново-подзолистым сильносмытым почвам, 9 - к дерново-подзолистым намытым почвам (рис.2.8).

Случайная независимая почвенная выборка из 27 точек состояла из 2 аллювиальных почв, 2 дерново-глеевых почв, 2 дерново-подзолистых

несмытых почв, 2 светло-серых лесных почв, 11 дерново-подзолистых слабосмытых почв, 3 дерново-подзолистых среднесмытых почв, 2 дерново-подзолистых сильносмытых почв, 2 дерново-подзолистых намытых почвы.

Во время полевых исследований определялись точные географические координаты, делались цветные фотографии профилей и окружающей местности, проводилось полное морфологическое описание и отбирались почвенные образцы.

2.4.2 Анализ почвенных образцов на содержание органического углерода и гранулометрический состав, а также создание базы данных

В каждой точке наблюдения из верхнего гумусового горизонта (0-20 см) отбирались почвенные образцы. В части разрезов отбор образцов проводился по-горизонтно. В каждом почвенном образце определялось содержание органического углерода методом каталитического сжигания. Анализы проводились на элементном анализаторе VARIO EL III в лаборатории Биологии и биохимии почв ФГБНУ «Почвенного института имени В.В. Докучаева». Также в каждом почвенном образце определялось содержание физической глины и физического песка методом лазерной дифрактометрии. Анализ проводился сотрудником ФГБНУ «Почвенного института имени В.В. Докучаева» А.В. Юдиной.

В итоге, была создана база данных, содержащая 166 точек, в которой для каждого наблюдения записывались следующие характеристики: координаты, полное название почвы согласно классификации 1977 года (Классификация, 1977), гранулометрический состав верхнего горизонта, содержание органического углерода, степень смытости, индекс NDVI, посчитанные морфометрические характеристики рельефа, принадлежность к тому или иному поченному типу согласно карте 1984 г., а также цветные фотографии профилей. Для 27 точек случайной независимой выборки была также создана подобная таблица.

2.4.3 Создание традиционной почвенной карты

Для построения почвенной карты нами был создан проект в программе QGIS 2.8.2, который включал в себя несколько тематических слоев: построенную карту элементарных поверхностей мезорельефа, космический снимок высокого разрешения WorldView-1, карты производных морфометрических характеристик рельефа и слой с точками полевого обследования. На основе анализа всех слоев и руководствуясь индикационной таблицей, эксперты путем было проведено определение границ почвенных контуров для исследуемой территории.

2.5 Цифровое почвенное картографирование

2.5.1 Описание используемого программного обеспечения

В ходе исследования нами использовалось следующее программное обеспечение:

1. QuantumGIS 2.8.2 – свободная кроссплатформенная геоинформационная система, позволяющая выполнять полный набор действий, связанных с производством и визуализацией картографической информации (QGIS, 2017). Использовалась в работе для компоновки картографического материала, для редактирования и анализа растровых данных, для обрезки, визуализации и калькуляции с растровыми слоями, для работы с атрибутивной информацией векторных слоев, для автоматического проведения подсчетов внутри контуров, для создания буферных зон и других операций.

2. SAGA GIS 2.1.2 – система географического информационного анализа данных, позволяющая проводить различные операции с геоданными, например подготовка ДДЗ, анализ изображений, геостатистическая обработка и моделирование. В наших исследованиях использовалась в качестве основного программного обеспечения, посредством которого создавалась ЦМР и растровые изображения морфометрических характеристик рельефа (SAGA, 2017).

3. MultiSpec – это программа для компьютерной обработки многозональных снимков, разработанная учеными американского университета Пердью (MultiSpec, 2016). Использовалась нами для работы с космическим снимками Landsat и WorldView, для комбинирования и визуализации разных наборов каналов.

4. R 3.3.1 - свободная среда разработки программного обеспечения с открытым исходным кодом для языка программирования R, который предназначен для статистической обработки данных и работы с графикой (R Core Team, 2017). Были использованы следующие пакеты R: «gstat» для создания и подбора семивариограмм, проведения кригинга (Pebesma, 2004), «raster» для обработки растровых слоев, визуализации и отрисовки смоделированных почвенных карт (Hijmans, van Etten, 2013), «rgdal» для работы с растровыми слоями, для проведения операций по перепроектированию и записыванию растровых изображения в файл (Keitt и др., 2009), «rpart» для моделирования с помощью регрессионных деревьев (Breiman и др., 1984; Therneau и др., 2011), «Random Forest» для моделирования с помощью метода случайного леса (Liaw, Wiener, 2002), «C50» для моделирования на основе деревьев решений (Quinlan, 1993), «e1071» для моделирования на основе метода опорных векторов (Chang, Lin, 2001), «ithir» для вычисления коэффициента каппа (Congalton, 1991), «nnet» для моделирования на основе множественной логистической регрессии (Venables, Ripley, 2002), «stats» для проведения статистической обработки данных (Chambers, 1992), «sp» для трансформации пространственных данных, для перевода из одного формата данных в другой (Pebesma, Bivand, 2005).

2.5.2 Отбор значимых ковариат для моделирования

Для грамотного проведения моделирования пространственного распределения почвенного покрова необходим отбор ковариат (или факторно-индикационных переменных) на территорию исследования. В связи с тем, что важнейшую роль в дифференциации почвенного покрова территории в нашем

случае играет рельеф, так как климатические характеристики, растительность и почвообразующие породы не так контрастны на ключевом участке, то для его моделирования нами были получены следующие ковариаты:

1. Растворные изображения 35 морфометрических характеристик рельефа, полученные из карты рельефа масштаба 1:50000: длина склона, абсолютная высота, текстура, меридианальная кривизна, максимальный вес, длина линии потока, водосборная площадь, долины, относительная позиция склона, русловая сеть, кривизна в профиле, фактор длины склона, основная кривизна, тангенциальная кривизна, топографический индекс позиции, отмывка рельефа, индекс силы потока, максимальная кривизна, превышение, крутизна склонов водосбора, индекс расчленности рельефа, индекс сходимости, минимальная кривизна, кривизна в плане, крутизна склона, топографический индекс влажности, вектор всхолмленности рельефа, выпуклость поверхности, аспект, индекс баланса массы, модифицированная водосборная площадь, поперечная кривизна и другие.

2. Оцифрованные границы пашни, лесных массивов, поймы, лугов и балочных понижений, полученные по космическому снимку WorldView-1.

3. Вегетационный индекс NDVI и спектральные яркости каналов снимка WorldView-1.

4. Границы ареалов распространения условных почвенных групп по карте 1984 г.: дерново-подзолистых почв разной степени смытости, дерново-глеевых, светло-серых лесных и аллювиальных почв.

Использовался дисперсионный анализ сводной таблицы, состоящей из точек обследования и характеристик ковариат в этих точках, который показал какие именно ковариаты из общего массива ковариат являются значимыми для отнесения той или иной точки к изначально выбранным условным почвенным группам. Составлялись таблицы, отражающие степень влияния ковариат по формуле Миллса (Дмитриев, 2001).

2.5.3 Подбор размера пикселя для картографирования

Для проведения моделирования почвенных карт на основе цифровых методов были проведены исследования по подбору оптимального размера пикселя растровых изображений ковариат. В качестве алгоритма для тестирования размера пикселя была взята множественная логистическая регрессия. При моделировании брались значимые ковариаты, отобранные на основе дисперсионного анализа сводной таблицы, в которой для каждой точки были отображены значения ковариат и условный почвенный выдел в этой точке.

В результате моделирования было получено 600 моделей, по 100 моделей для каждого из 6 случаев: 1) размер пикселя растров ковариат 5 м, для обучения использовалась почвенная карта 1984 г.; 2) размер пикселя растров ковариат 5 м, для обучения не использовалась почвенная карта 1984 г.; 3) размер пикселя растров ковариат 20 м, для обучения использовалась почвенная карта 1984 г.; 4) размер пикселя растров ковариат 20 м, для обучения не использовалась почвенная карта 1984 г.; 5) размер пикселя растров ковариат 50 м, для обучения использовалась почвенная карта 1984 г.; 6) размер пикселя растров ковариат 50 м, для обучения не использовалась почвенная карта 1984 г.

2.5.4 Процесс создания почвенных карт на основе моделей

Для целей ЦПК нами использовались следующие методы: множественная логистическая регрессия, метод деревьев решений C5, опорных векторов и случайного леса, а также регрессионные деревья. Для создания почвенных карт методами ЦПК был создан файл формата .cvs, в котором была занесена информация по каждой точке наблюдения, включающая в себя весь набор отобранных ковариат и закодированное обозначение условной почвенной группы.

Устойчивость моделей определялась методом наибольшей беспристрастности (кросс-валидации). Суть метода состоит в том, что

подмножество наблюдений делится случайным образом на два поднабора: «обучающую» выборку и «проверяющую» выборку. По обучающей выборке происходит процесс настройки модели, а по второй выборке производится ее проверка. Этот процесс повторяется много раз. Мерой точности прогноза считают среднюю оценку, полученную по результатам оценки каждого значения «проверяющей» выборки.

При создании цифровых карт на основе пяти моделей проводилось по 100 итераций, в каждую из которых машина случайным образом из обучающей выборки отбирала 88% точек для создания модели, а на оставшихся 12% точек проводилась проверка самовоспроизводимости модели или, другими словами, проводилась проверка ее устойчивости (рис. 2.9). Для каждой проверяющей выборки была построена таблица сопряженности между реальными почвенными группами, выявленными в ходе полевых исследований, и почвенными группами, предсказанными в этих точках в результате моделирования.

Таблица сопряженности – эта матрица, по одной из осей которой записываются названия классов проверяемого набора данных, по второй – классы, используемые для проверки.

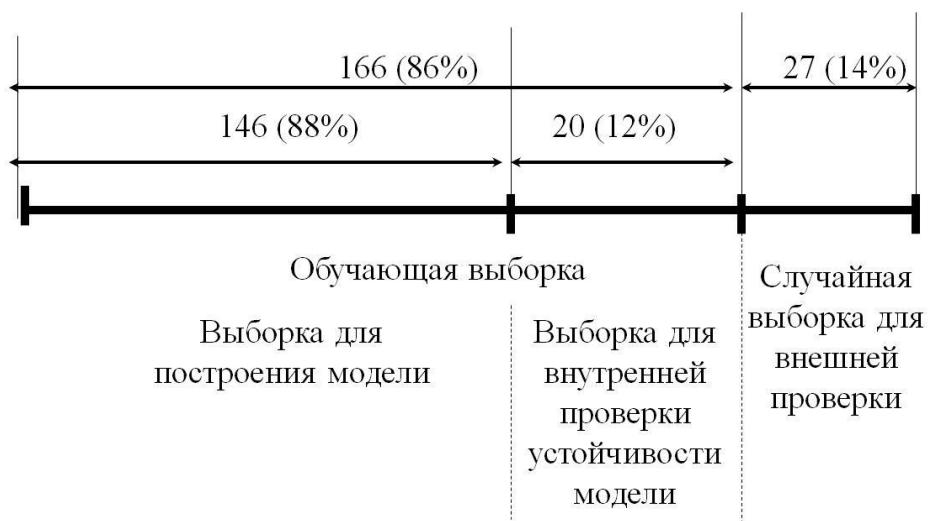


Рис. 2.9 Соотношение обучающей и проверочной выборок

На пересечении k -ой строки и l -того столбца расположен X_{kl} – это количество точек (разрезов, случаев), отнесенных одновременно по почвенной карте к классу k , а по результатам проверки к классу. Главная диагональ таблицы показывает случаи, где результаты, полученные в двух случаях, совпадают. Вне-диагональные элементы соответствуют случаям несовпадения.

При интерпретации результатов обычно полагается, что проверяемый результат потенциально является *неточным*, а проверочная выборка хорошо отражает реальную ситуацию. В противном случае, если проверочная выборка также не совершена, то говорят не об «ошибке», а о «разнице» между двумя наборами данных.

Числа по диагонали матрицы сопряженности показывают верность предсказания между наблюдаемым в реальности классом и соответствующим предсказанием по почвенной карте. Числа за пределами диагонали указывают на неверно классифицированные объекты или ошибки.

Сумма значений диагональных элементов (X_{ii}) отражает общее количество одинаково классифицированных точек, а отношение этого количества к общему количеству сравниваемых точек N называется *общей точностью* (или *коэффициентом общей точности*) и обычно выражается в процентах:

$$A_o = \theta_1 = \frac{\sum_{i=1}^r x_{ii}}{N} = \sum_{i=1}^r p_{ii}; \text{ где } p_{kl} = \frac{x_{kl}}{N}$$

Точность предсказания отдельных классов рассчитывается аналогично. Однако, есть выбор деления количества правильных предсказаний каждого класса либо на общее количество наблюдений, либо на общее количество предсказаний по соответствующим колонкам или строкам.

Традиционно, делят на общее количество предсказаний по классу (то есть на сумму по колонке). Эта мера точности предсказания указывает на вероятность того, что наблюдение было классифицировано правильно и на

самом деле является мерой «ошибки пропуска», или «точности модели», или «точности производителя» (producer's accuracy), потому что создатель модели интересуется, насколько хорошо определенный класс может быть предсказан.

Для класса i :

$$producer's_accuracy_i = \frac{x_{ii}}{x_{+1}}$$

Если общее количество правильных предсказаний одного класса поделить на общее количество разрезов, описанных в поле, то этот результат является мерой «ошибок выполнения», т.е. «точности определения», или «точности пользователя» (user's accuracy), так как она указывает на вероятность того, что предсказанные почвы по карте, на самом деле, представлены той же категорией в реальности в поле.

Для класса i :

$$user's_accuracy_i = \frac{x_{ii}}{x_{1+}}$$

Коэффициент каппа или просто каппа является еще одной статистической мерой совпадения между наблюдаемыми и предсказанными классами. Расчет основан на разнице между тем, сколько совпадений реально присутствует («наблюдаемые» совпадения), и тем, насколько можно ожидать совпадения, получившегося просто случайным образом. Наблюдаемое совпадение – это процент общей точности предсказания.

Индекс каппа учитывает, что совпадение двух картосхем может быть обусловлено случаем. Например, на некоторой территории распространено 2 типа почв, и один из них занимает 80% площади, а другой - 20%. Тогда вероятность того, что на карте и при полевом обследовании случайно обнаружат в одном и том же месте почву 1 равна $0,8 \cdot 0,8 = 0,64$, а почву 2 – $0,2 \cdot 0,2 = 0,04$. То есть, вероятность того, что обе группы обнаружат в одних и тех же точках одинаковые почвы, если будут называть почвы по игре случая, равна $0,64 + 0,04 = 0,68$. Обозначим ее как θ_2 :

$$\theta_2 = \sum_{i=1}^r \frac{x_{i+}x_{+i}}{N^2};$$

Если предположить, что совпадение результатов на карте и при полевом обследовании двумя группами может быть или случайным, или неслучайным, то $1 - \theta_2$ равно вероятности неслучайного совпадения двух картосхем.

Индекс *kappa* задается как $\hat{\kappa} = \frac{\theta_1 - \theta_2}{1 - \theta_2}$

Каппа равно 1 при полном совпадении двух картосхем или при полном совпадении результатов полевого обследования и почвенной карты (отсутствие случайности). Каппа равно нулю, если θ_1 совпадает с θ_2 и наблюдается чисто случайное совпадение двух картограмм или при случайном совпадении почвенной карты и результатов полевого обследования. В большинстве случаев каппа лежит от 0 до 1 и показывает, насколько похожи картосхемы или насколько совпали почвенная карта и результаты обследования. Если каппа меньше нуля, то несогласие больше, чем могло бы возникнуть случайно.

В результате кросс-валидации получалась таблица, в которой для каждой итерации отображались следующие показатели модели, полученной в результате этой итерации: коэффициент каппа, общая точность моделирования, точность производителя и пользователя (Самсонова, Мешалкина, 2011; Hengl и др., 2007; Докучаев, Мешалкина, 2016б).

Для каждой модели были рассчитаны средние значения, медианы и стандартные отклонения и границы вышеперечисленных характеристик. Строились кривые распределения каппа для каждого алгоритма, и полученный набор данных, состоящий из 100 коэффициентов капп для каждой модели, проверялся на нормальность распределения.

Для каждой из моделей для дальнейшей работы были выбраны две конкретные реализации моделей: для среднего и для максимального значения каппы.

На основе построенных картографических моделей, в которых заключены правила классификации, предсказывались все точки растра.

2.5.5 Сравнение карт между собой по 200-ам случайным точкам

Для сравнения карт между собой была построена случайная выборка, состоящая из 200 случайных точек, брошенных на карту 2016 г. Количество точек для каждой условной почвенной группы пропорционально площадям, занимаемым этими группами на карте 2016 г. (рис.2.10). Точки были равномерно распределены по всей территории исследования. Далее составлялась база данных, в которой для каждой из 200 точек соответствовало 12 колонок, в которых были отражены наименования условных почвенных групп для всех видов карт, построенных нами в ходе исследования:

1. Почвенная карта 2016 г., построенная традиционным методом с элементами ГИС-анализа.

2. Почвенная карта 1984 г.

3. Десять почвенных карт, созданных цифровыми методами почвенного картографирования. Для каждого вида моделирования было создано по 2 карты. Одна карта на основе модели со средней крапкой самовоспроизводимости, а другая – на основе модели с максимальной крапкой.

Из этой таблицы составлялись таблицы сопряженности цифровых почвенных карт с картой 2016 г. и с картой 1984 г., которые в дальнейшем анализировались.

Проводился подсчет характеристик, отражающих степень сходимости картографических материалов, а именно коэффициента кappa и общей точности.

2.5.6 Проверка карт по независимой случайной выборке

На следующем этапе исследования нами было проведено полевое обследование 2016 г. в 27 случайных независимых точках (рис.2.11).

Количество точек, относимых к тому или иному условному поченному выделу, было пропорционально площади этих почв согласно карте 2016 г., созданной нами традиционным методом с использованием элементов ГИС-анализа.

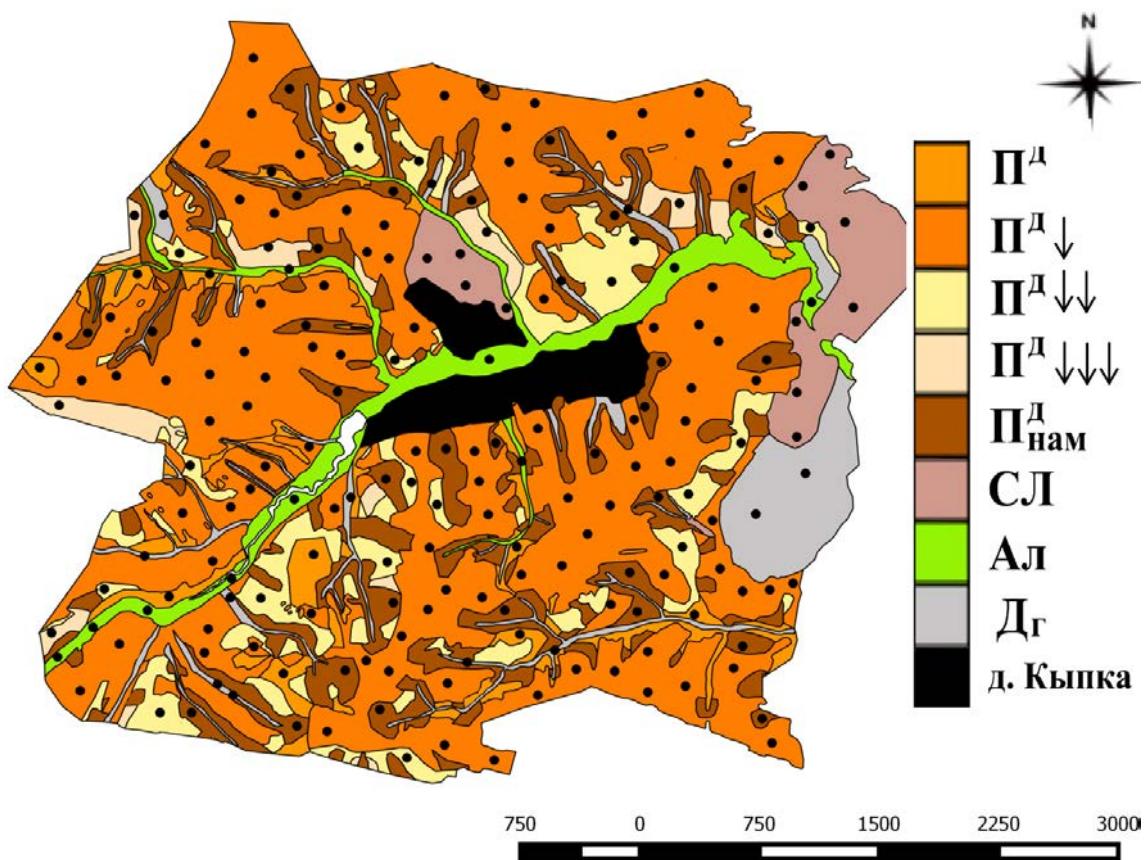


Рис. 2.10 Распределение 200 проверочных точек на территории исследования

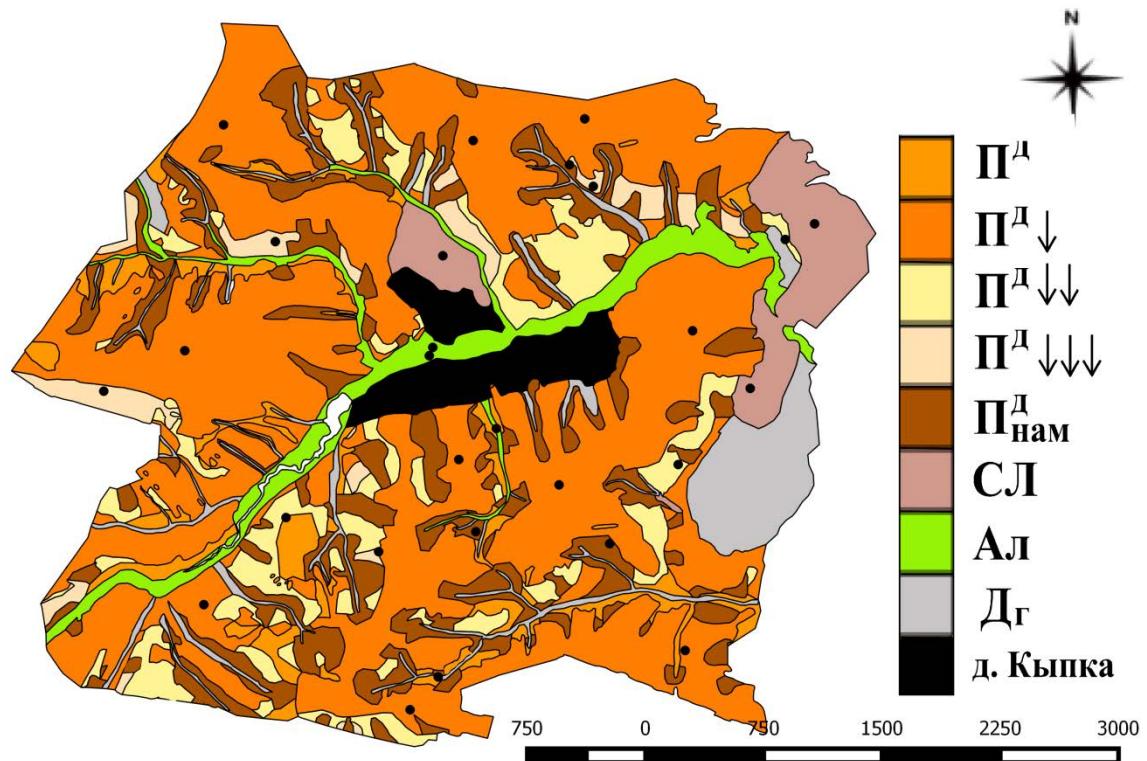


Рис. 2.11 Распределение 27 точек обследования случайной независимой выборки

Для каждой точки записывались следующие характеристики: оординаты, полное название почвы согласно классификации 1977 года, почвенно-морфологическое описание разреза, степень смытости и др., а также цветные фотографии профилей. Эти точки были нанесены на карты, построенные различными методами цифровой почвенной картографии, и на карту 2016 г., построенную традиционным методом с использованием элементов ГИС-анализа.

Была создана база данных, в которой каждой из 27 точек обследования соответствовало 13 наименований условных почвенных групп: принадлежность к условному поченному выделу, выявленная в данной точке, в ходе обследования 2016 г., условный почвенный выдел, согласно карте 2016 г., построенной традиционным методом, и согласно карте 1984 г., а также десять почвенных выделов согласно картам, построенным на основе 5 моделей со средними и максимальными каппами самовоспроизводимости этих моделей (см. приложение 4). Проводилось создание таблиц сопряженности и их анализ для выявление карты наиболее приближенной к действительности.

2.6 Построение картограммы гумуса

Для построения картограммы содержания органического углерода в верхнем гумусовом слое из всего массива данных нами была отобрана 101 точка обследования, находящиеся на полях исследуемого участка (рис. 2.12). Проводился анализ коэффициентов корреляции между содержанием органического углерода в верхнем гумусовом слое и ковариатами.

Для описания пространственного распределения органического углерода в верхнем гумусовом слое полей использовались ординарный и регрессионный кригинг. На основе картограммы, построенной методом регрессионного кригинга, была создана карта запасов гумуса и карта ошибок предсказания запасов гумуса. Был проведен анализ запасов гумуса по полям и рабочим участкам, выделенным согласно проектной документации хозяйства.

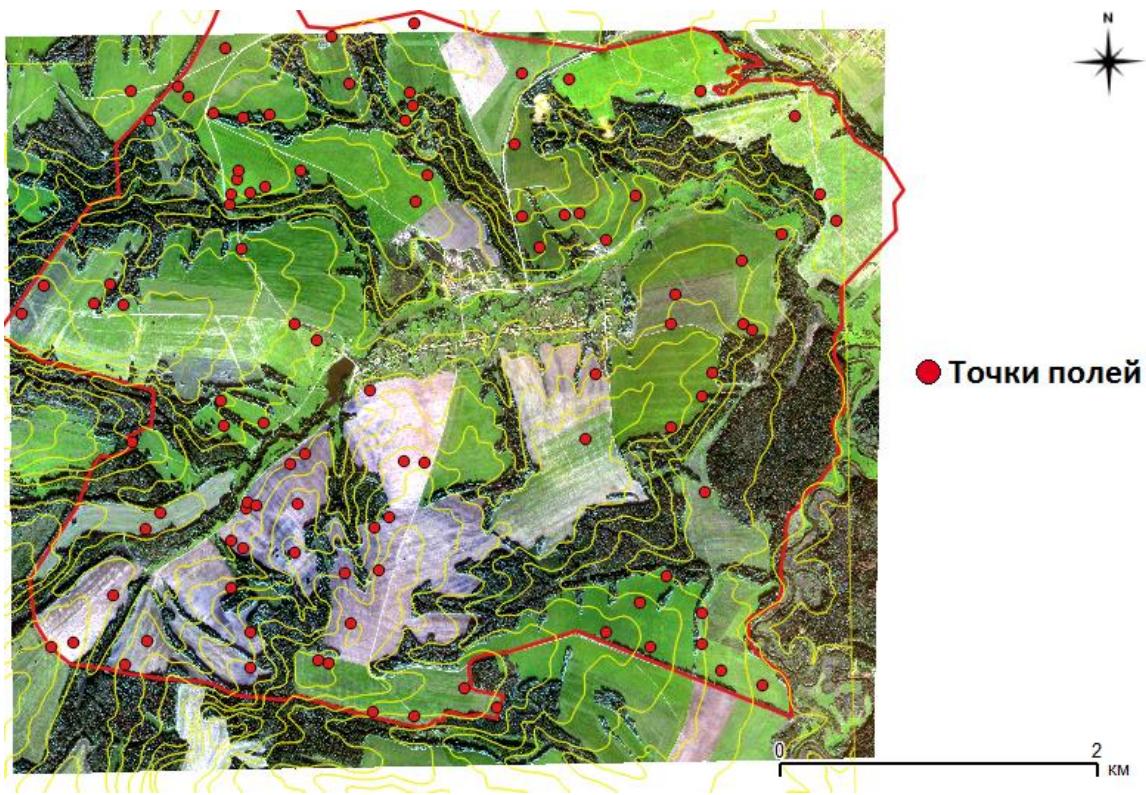


Рис.2.12 Точки обследования, находящиеся на полях и используемые для построения картограмм гумуса

Были вычислены средние запасы гумуса для каждой условной почвенной группы по почвенной карте, созданной традиционным методом обследования с элементами ГИС-анализа.

ГЛАВА 3. ПОСТРОЕНИЕ ПОЧВЕННОЙ КАРТЫ НА РАЙОН ИССЛЕДОВАНИЯ

3.1 Особенности почвенного покрова исследуемой территории

В результате полевых обследований 2014-2016 гг. было установлено, что основной фон в структуре почвенного покрова исследуемой территории составляют дерново-подзолистые почвы разной степени подзолистости и смытости. Встречаются также светло-серые лесные, дерново-глеевые и аллювиальные почвы.

В наших исследованиях мы пользовались Классификацией почв СССР 1977 г. (Классификация..., 1977) из-за двух объективных причин:

1. Для сравнения почвенных карт необходимо иметь одинаковые классификации, на основе которых эти почвенные карты созданы. Так как в нашем распоряжении была почвенная карта 1984 г., в основе которой была Классификация почв СССР 1977 г., то нашу почвенную карту мы также решили создавать, руководствуясь этой классификацией. Перевод из одной почвенной классификации в другую не всегда корректен, потому что для проведения такого перевода необходимо знать исходные данные, которых для карты 1984 г. у нас не было.

2. С момента почвенного обследования 1984 г. произошли значительные изменения в теории и практике отечественного почвоведения, нашедшие отражение в Классификации почв России 2004 (Классификация..., 2004). Тип светло-серые лесные почвы, который прописан в Классификации почв СССР 1977 г., отсутствует в Классификации почв России 2004 г. Этот тип почв согласно российской классификации относится к дерново-подзолистым почвам. Поэтому использование Классификации почв России 2004 г. в нашей работе приводило бы к сокращению числа почвенных типов, которые мы изначально наметили выделять для описания почвенного покрова исследуемой территории.

Нами был проведен перевод названий почв из Классификации почв СССР 1977 г. в Классификацию почв России 2004 г. (Классификация..., 2004) и Реферативную Базу Почв Мира – «WRB» (табл. 3.1). Названия почв брались в соответствии с работой П.В. Красильникова (1999).

Таблица 3.1

Таблица корреляций классификаций для почвенных групп территории исследования

Почвенные группы, согласно Классификации почв СССР 1977 г.	Почвенные группы, согласно Классификации почв России 2004 г.	Почвенные группы, согласно WRB.
Дерново-подзолистые несмытые	Дерново-подзолистые и агродерново-подзолистые	Umbric Albeluvisols Abruptic
Дерново-подзолистые слабосмытые	Дерново-подзолистые и агродерново-подзолистые	Umbric Albeluvisols Abruptic
Дерново-подзолистые среднесмытые	Агрообраземы, агроземы текстурно-дифференцированные	Anthraquic Cambic Cutanic, Anthraquic Cambic Clayic, Anthraquic Cambic Skeletic
Дерново-подзолистые сильносмытые	Агрообраземы, агроземы текстурно-дифференцированные	Anthraquic Cambic Clayic, Anthraquic Cambic Skeletic
Дерново-подзолистые намытые	Дерново-подзолистые и агродерново-подзолистые	Umbric Albeluvisols Abruptic
Светло-серые лесные	Дерново-подзолистые и агродерново-подзолистые	Greyic Phaeozems Albic
Дерново-глеевые	Темногумусово-глеевые, перегнойно-гумусовые глеевые	Haplic Gleysols Humic
Аллювиальные луговые кислые	Аллювиальные темногумусовые и серогумусовые глеевые	Umbric Fluvisols Oxyaquic

Была составлена специальная таблица, характеризующая особенности основных почвенных выделов на территории исследования (табл.3.2). Символьные обозначения: П-0-1 – выровненные участки (плакоры с крутизной склона от 0 до 1^0); П-1-3 - слабые склоны выпуклой формы крутизной от 1 до 3^0 ; П-3-5 - средние склоны выпуклой формы крутизной от 3 до 5^0 ; П >5 - крутые склоны выпуклой формы крутизной $> 5^0$; У-1-3 - слабые склоны вогнутой формы крутизной от 1 до 3^0 ; У-3-5 - средние склоны вогнутой формы крутизной от 3 до 5^0 ; У >5 - крутые склоны вогнутой формы крутизной $> 5^0$

Выровненные поверхности занимают дерново-подзолистые несмытые почвы (рис.3.1, А; табл. 3.2), встречающиеся как на пашне, так и под лесной растительностью. Однако доля площадей, занимаемых такими почвами относительно невелика. Наиболее распространеными на данной территории являются дерново-подзолистые слабосмытые почвы, которые составляют основу почвенного покрова территории и занимают склоны различной крутизны, которые также располагаются как под лесной растительностью, так и под пашней (рис.3.1, Б; табл. 3.2). Как уже отмечалось в разделе 2 территории ключевого участка сильно рассечена овражно-балочной сетью, поэтому часто в почвенном покрове встречаются дерново-подзолистые среднесмытые (рис.3.2, А; табл. 3.2) и сильносмытые (рис.3.2, Б; табл. 3.2) почвы. Депрессивные формы рельефа чаще всего заняты дерново-подзолистыми намытыми почвами (рис.3.3, А; табл. 3.2). В расположенных частях склонов можно встретить светло-серые лесные почвы занимают (рис.3.3, Б; табл. 3.2), доля которых в общей площади территории невелика. Днища оврагов и балок, а также участки с близким залеганием грунтовых вод заняты дерново-глеевыми почвами (рис.3.3, В; табл. 3.2).

Дерново-неглубоко- и глубокоподзолистые несмытые почвы на исследуемой территории преимущественно распаханы или были когда-то распаханы. Эти почвы занимают выровненные участки территории преимущественно на водоразделах, встречаются как на пашне, так и под

лесной растительностью. Однако доля площадей, занимаемых такими почвами относительно невелика.

Морфологическое строение профиля данной условной почвенной группы имеет следующие характерные особенности: максимальную мощность и степень выраженности подзолистого горизонта, четкую языковатую форму границы этого горизонта. Содержание физической глины в верхнем горизонте таких почв колеблется от 38,0 до 47,0% при сравнительно однородном содержании ила по профилю. Среднее содержание физической глины в верхнем горизонте таких почв составляет 42%.

Содержание гумуса в таких почвах варьируется от 0,9 до 2,2%, при среднем содержании гумуса в верхнем горизонте 1,7%. В почвах, подверженных эрозии, содержание гумуса уменьшается. По содержанию гумуса характеризуемые почвы относятся к малогумусированным видам (Классификация..., 2004).

Дерново-неглубоко- и глубокоподзолистые слабосмытые почвы занимают большую часть исследуемой территории. Эти почвы также в основном находятся в распаханном состоянии. Этот условный почвенный выдел занимает склоны небольшой крутизны

Морфологическое строение профиля дерново-неглубоко- и глубокоподзолистых слабосмытых почв имеет следующие характерные особенности: сокращенную мощность и фрагментарность подзолистого горизонта, распаханные почвы имеют четко выраженный пахотный горизонт, имеющий наиболее темную окраску по профилю.

Содержание физической глины в верхнем горизонте таких почв колеблется от 35,0 до 46,0% при сравнительно однородном содержании ила по профилю. Среднее содержание физической глины – 40%.

Содержание гумуса в таких почвах варьируется от 0,9 до 3,6% при среднем значении 1,7%. По содержанию гумуса характеризуемые почвы относятся к малогумусированным видам (Классификация..., 2004).

Дерново-подзолистые среднесмытые почвы занимают на ключевом участке территории с крутыми склонами. Так как исследуемая территория сильно расчленена овражно-балочной сетью, то значительная часть территории занята склоновыми формами рельефа и дерново-подзолистые среднесмытые почвы, соответственно, встречаются на территории исследования часто. Большая часть из этих почв находится под лесной растительностью, однако часть из них вовлечена в пашню.

Морфологическое строение профиля таких почв имеет следующие характерные особенности: отсутствие признаков подзолистого горизонта, постепенная смена окраски от верхних к нижним горизонтам, она становится более желтоватой, местами в нижней части профиля присутствует галька, слабая гумусированность верхнего горизонта, большая уплотненность верхнего горизонта по сравнению с дерново-подзолистыми несмытыми и слабосмытыми почвами.

Содержание физической глины в верхнем горизонте таких почв колеблется в широком диапазоне от 37,0 до 50,0% при сравнительно однородном содержании ила по профилю. Среднее содержание физической глины – 43,0%.

Содержание гумуса в верхнем слое в таких почвах варьируется от 0,5 до 2,2% при среднем значении 1,4%. По содержанию гумуса характеризуемые почвы относятся к малогумусированным видам (Классификация..., 2004).

Небольшие площади и самые крутые склоны на территории заняты *дерново-подзолистыми сильносмытыми почвами*. Данный условный почвенный выдел встречается под лесной растительностью.

Особенности морфологического строения таких почв: размытость границ и сложность выделения генетических горизонтов, желтоватые и красноватые оттенки верхнего гумусированного горизонта, слабая гумусированность верхнего горизонта, на глубине 10-15 см залегание почвообразующих пород глинистого и суглинистого состава, иногда в

профиле наблюдается обильное содержание гальки, высокая уплотненность как верхних, так и нижних горизонтов.

Диапазон содержания физической глины в верхнем горизонте дерново-подзолистых сильносмытых почв колеблется от 45,0 до 60,0% при сравнительно однородном содержании ила по профилю. Среднее содержание физической глины – 49,0%.

Содержание гумуса в верхнем горизонте дерново-подзолистых сильносмытых почв варьируется от 0,5 до 1,7% при среднем значении 1,2%. По содержанию гумуса характеризуемые почвы относятся к слабогумусированным и малогумусированным видам (Классификация..., 2004).

Так как в исследуемом регионе идут активные процессы водной эрозии, то некоторая часть территории занята *дерново-подзолистыми намытыми почвами*. Данный условный почвенный выдел занимает вогнутые позиции рельефа, в которых происходит накопление смываемого из почв других позиций рельефа верхнего слоя. Такие почвы встречаются как на пашне, так и под лесной растительностью.

Характерные черты морфологического строения профиля: темная окраска верхнего гумусого горизонта (темнее, чем у других дерново-подзолистых почв, выделяемых на исследуемой территории), слоистость гумусого горизонта, мощность этого горизонта доходит до 40-50 см.

Содержание физической глины в верхнем горизонте дерново-подзолистых намытых почв колеблется от 36,0 до 46,0% при сравнительно однородном содержании ила по профилю. Среднее содержание физической глины – 40,2%.

Содержание гумуса в верхнем слое данной условной почвенной группы варьируется от 1,7 до 2,0% при среднем значении 1,9%. По содержанию гумуса дерново-подзолистые намытые почвы также относятся к малогумусированным видам (Классификация..., 2004).

Светло-серые лесные почвы занимают выположенные нижние части склонов (Ковриго, 2004). Часть из этих почв распахана, часть встречается под лесной растительностью. Эти почвы характеризуются лучшими гидрофизическими условиями почвообразования, чем дерново-подзолистые почвы этой территории, в связи с близким залеганием грунтовых вод, а, соответственно, и большим содержанием гумуса верхнем слое.

Морфологическое строение профиля данного типа почв имеет следующие характерные особенности: большая мощность, темно-серая окраска и мелкокомковатая структура верхнего гумусого горизонта, наличие переходного гумусово-элювиального горизонта с ярко выраженными потеками гумуса и обильной белесой присыпкой, комковатая структура этого горизонта, иногда в профиле встречаются признаки глееватости.

Содержание физической глины в верхнем горизонте светло-серых лесных почв колеблется от 33,0 до 44,0%. Гранулометрический состав характеризуется сравнительно однородным содержании ила по профилю. Среднее содержание физической глины – 39,4%.

Содержание гумуса для светло-серых лесных почв региона исследования невелико для данного типа почв и варьируется от 1,7 до 6,4% при среднем значении 3,3%. По мощности гумусого горизонта среди этого типа почв региона исследования встречаются как маломощные, так и среднемощные виды этих почв. По содержанию гумуса согласно классификации и диагностике почв России светло-серые почвы этого региона относятся к среднегумусированным (Классификация..., 2004).

Аллювиальные почвы на территории исследования представлены 2 типами почв: аллювиальные луговые кислые и аллювиальные дерновые кислые. В свою очередь аллювиальные луговые кислые встречаются 2 родов: обычные и ожелезненные. Для всех аллювиальных почв характерно, что они занимают поймы рек Убыть и Кыпка, а также часть аллювиальных почв диагностирована в наиболее крупных овражно-балочных комплексах, где

весенние потоки вод, возникающие в результате снеготаяния, формируют характерный профиль аллювиальной дерновой кислой почвы.

Морфологическое строение профиля данного типа почв имеет следующие характерные особенности: большая мощность гумусого горизонта, темно-серая окраска всего профиля, наличие оглеения, близкое залегание грунтовых вод.

Содержание физической глины в верхнем горизонте аллювиальных почв колеблется от 27,0 до 40,0%. Гранулометрический состав преимущественно среднесуглинистый. Среднее содержание физической глины – 35%.

Содержание органического углерода в верхних слоях для аллювиальных почв региона исследования невелико для данного типа почв и варьируется от 1,0 до 1,5% при среднем значении 0,9% (Классификация..., 2004).

Другим видом интразональных почв ключевого участка являются **дерново-глеевые почвы**, которые представлены двумя подтипами: дерново-грунтово-глеевые и перегнойно-грунтово-глеевые. Эти почвы занимают наиболее низкие позиции рельефа. Например, данные почвы встречаются в днищах балок и оврагов.

Морфологическое строение профиля дерново-глеевых почв имеет следующие особенности: часто наблюдается слоистость профиля, оглеение по всему профилю, иногда встречается глеевый горизонт.

Диапазон содержания физической глины в верхнем горизонте дерново-глеевых почв колеблется от 37,0 до 50,0% при сравнительно однородном содержании ила по профилю. Среднее содержание физической глины в верхнем горизонте – 43,0%.

Содержание гумуса в верхнем горизонте сильно варьируется и находится в диапазоне от 1,2 до 5,2% при среднем значении 2,5%. Эти почвы в основе своей относятся к малогумусированным (Классификация..., 2004).

Таблица 3.2

Характеристика особенностей основных почвенных выделов на территории исследования

Почвенные выделы	Положение в рельефе	Растительный покров	Характерная черта морфологического строения профиля
Дерново-подзолистые несмытые	П-0-1; П-1-3	Лиственница, ель, береза, сосна, осина, хвощ, клевер	Четко выраженный элювиальный горизонт белесого цвета
Дерново-подзолистые слабосмытые	П-0-1; П-1-3; П-3-5; У-1-3	Клевер, одуванчик, подорожник, люцерна	Мраморовидность почвенного профиля, залегающего под гумусовым горизонтом для пахотных почв, смытость части верхнего гумусового горизонта для почв под лесом
Дерново-подзолистые среднесмытые	П-3-5; П>5	Лисохвост, мятлик, костер, овсянница, зверобой, красный клевер	Включения материала из горизонта В в верхнем пахотном горизонте, отсутствие мраморовидности, четкое разделение верхнего гумусового горизонта и нижележащей почвенной массы по цвету для пахотных почв, смытость верхнего гумусового горизонта и части элювиального горизонта для почв под лесом

Продолжение таблицы 3.2

Почвенные выделы	Положение в рельефе	Растительный покров	Характерная черта морфологического строения профиля
Дерново-подзолистые сильносмытые	Π>5	Клевер, сосна, береза, ель	Слабое различие по цвету верхнего гумусого горизонта и нижележащей почвенной массы, наличие гальки в верхнем горизонте, буроватые, желтоватые и красноватые оттенки пахотного слоя, смытость части или полностью горизонта В для почв под лесом
Дерново-подзолистые намытые	U-1-3; U-3-5; U>5	Клевер, одуванчик, подорожник, люцерна, лиственница, ель, береза, сосна, осина	Большая мощность гумусового горизонта (от 25 см и выше), темно-серая окраска и слоистость для пахотных почв, слоистость верхнего гумусового горизонта и его большая мощность для почв под лесом
Светло-серые лесные	Π-0-1; Π-1-3; U-1-3	Лиственница, ель, береза, сосна, осина, копытень, сныть, хвощ	Наличие переходного гумусово-элювиального горизонта, серая окраска верхнего гумусового слоя
Дерново-глеевые	Π-0-1; днища оврагов и балок	Крапива, хвощ, папоротник	Слоистость почвенного профиля, оглеение наличие горизонта G

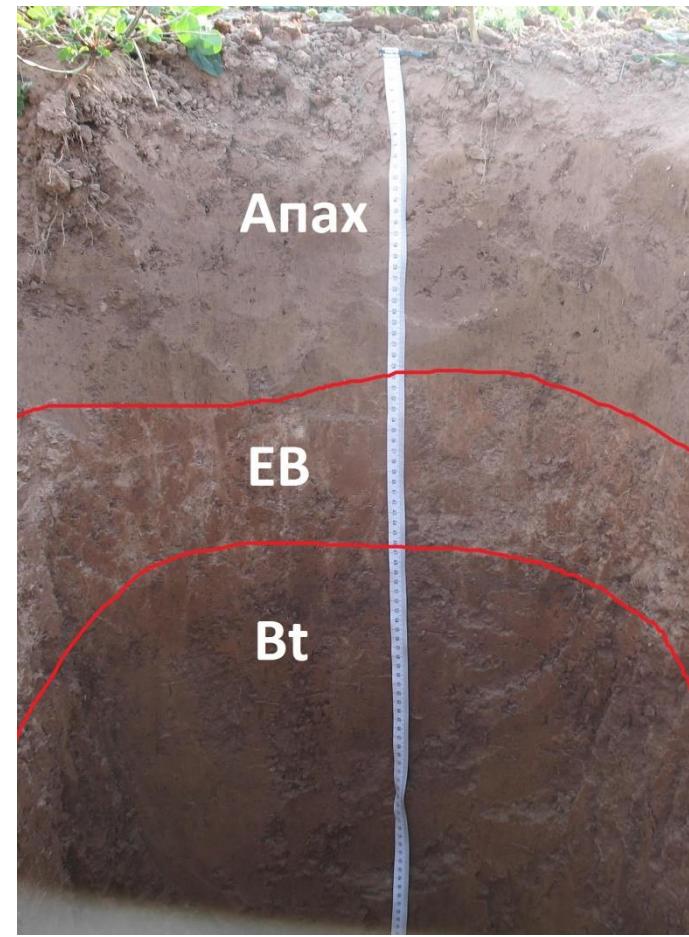
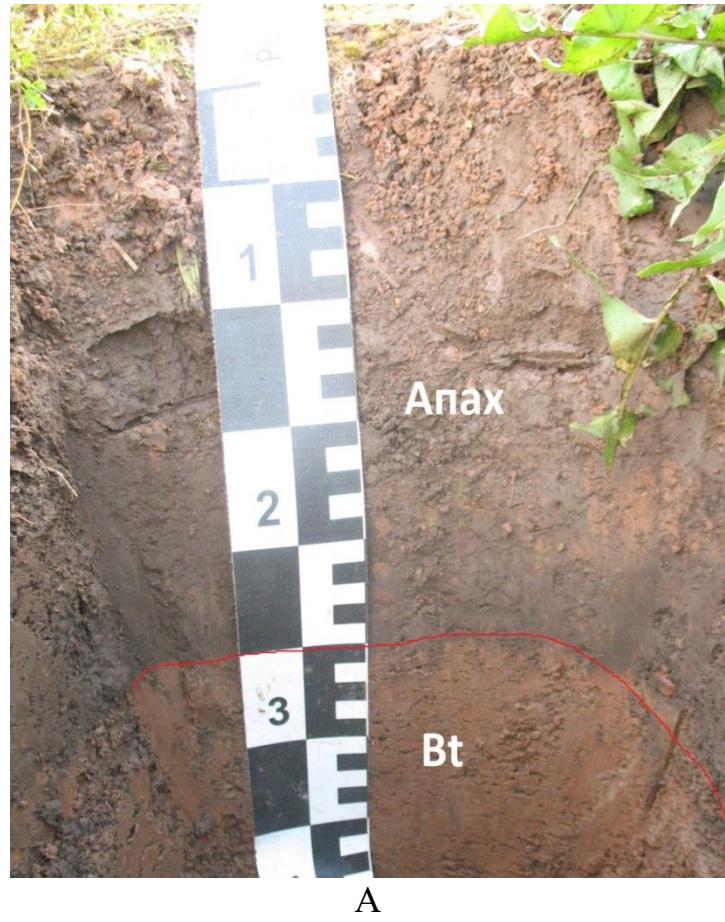
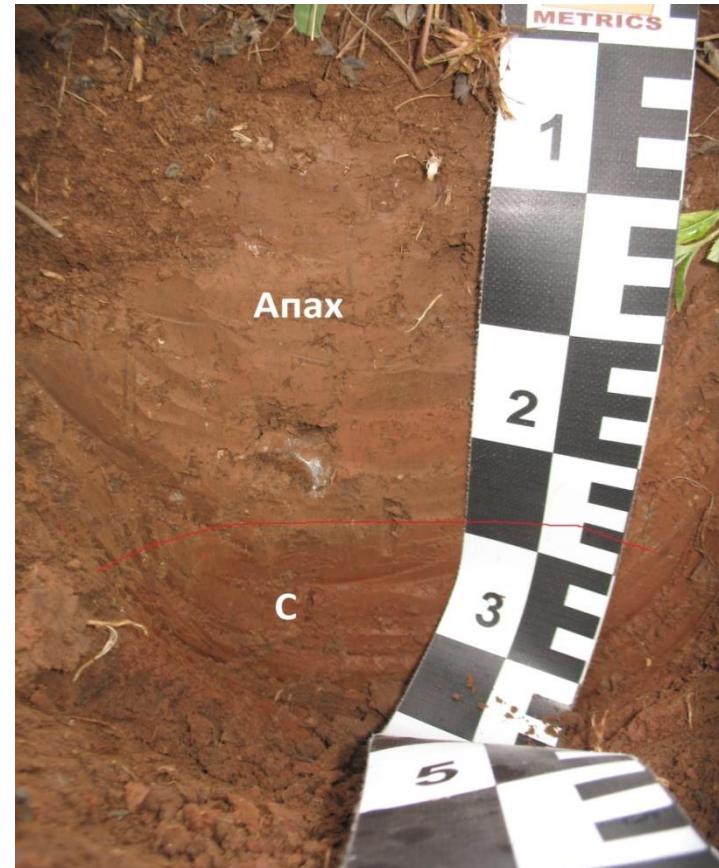


Рис. 3.1 Примеры дерново-подзолистых почв. А: дерново-неглубокоподзолистая старопахотная тяжелосуглинистая почва. Б: дерново-неглубокоподзолистая освоенная среднепахотная тяжелосуглинистая почва

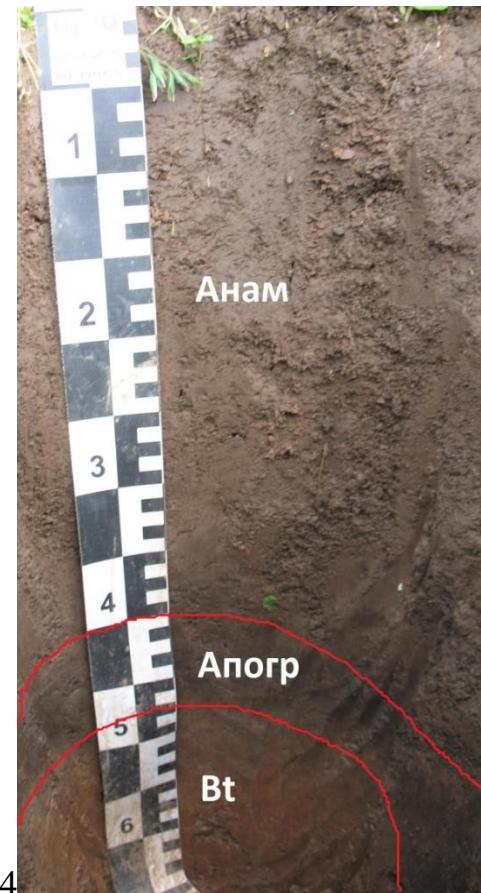


А

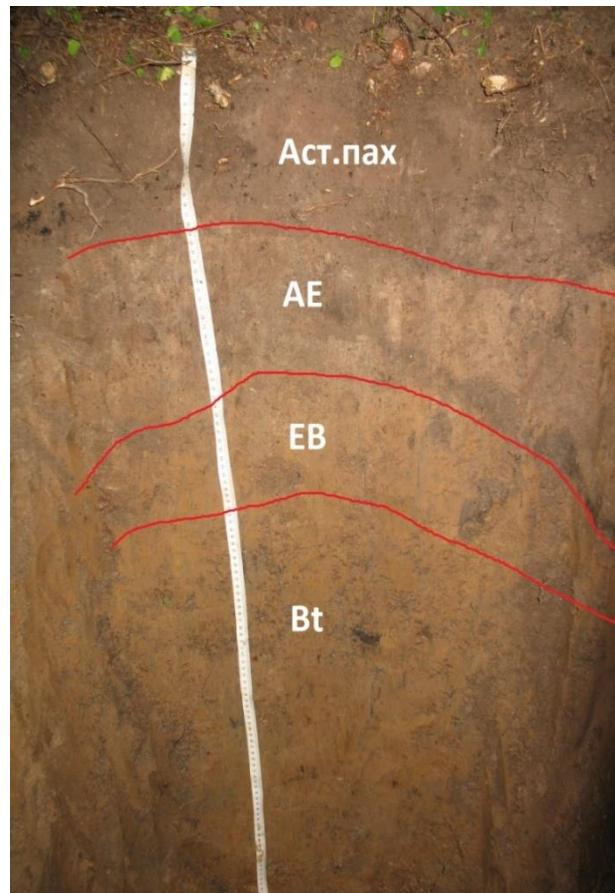


Б

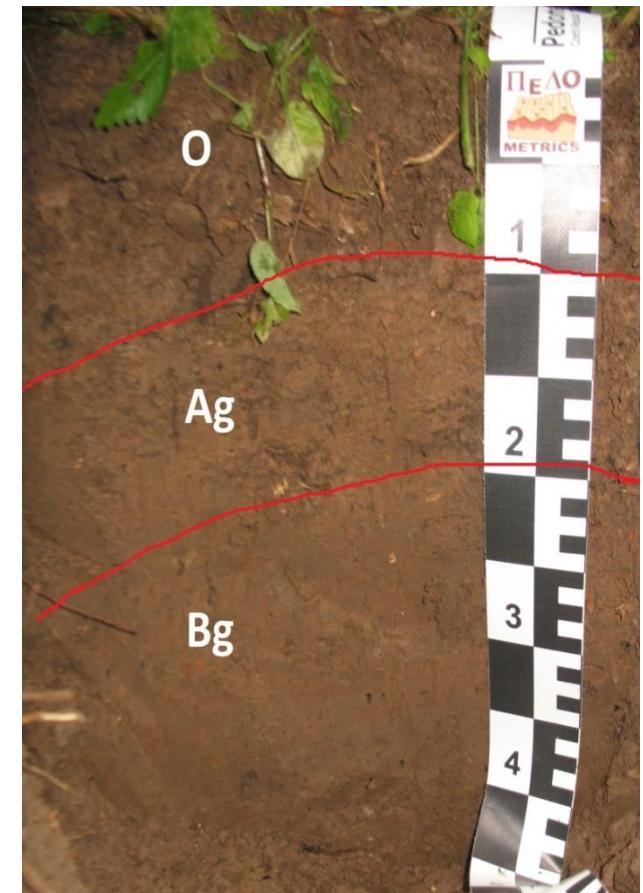
Рис.3.2 Примеры дерново-подзолистых почв. А: дерново-подзолистая освоенная среднепахотная среднесмытая среднесуглинистая почва. Б: дерново-подзолистая освоенная среднепахотная сильносмытая глинистая почва



А



Б



В

Рис.3.3 Примеры почв региона исследования. А: дерново-неглубокоподзолистая освоенная среднепахотная тяжелосуглинистая почва. Б: светло-серая лесная старопахотная грунтово-глееватая среднесуглинистая почва. В: Дерново-поверхностно-глееватая

Таблица 3.3

Индикационная таблица

Тип землепользования	Пашня							Лес							Пойма	Овраги и балки
	Π_3^{Δ} ; $\Pi_4^{\Delta} - 3$	$\Pi_3^{\Delta} \downarrow$; $\Pi_4^{\Delta} \downarrow - \mathcal{Z}$	$\Pi^{\Delta} \downarrow$	$\Pi^{\Delta} \downarrow \downarrow - \mathcal{E}$	СЛ - 3	$\Pi_3^{\Delta} \text{нам}; \Pi_4^{\Delta} \text{нам-ЭА}$	Π_3^{Δ} ; $\Pi_4^{\Delta} - 3$	$\Pi_3^{\Delta} \downarrow$; $\Pi_4^{\Delta} \downarrow - \mathcal{Z}$	$\Pi^{\Delta} \downarrow$	$\Pi^{\Delta} \downarrow \downarrow - \mathcal{E}$	СЛ - 3	$\Pi_3^{\Delta} \text{нам}; \Pi_4^{\Delta} \text{нам-ЭА}$				
1	2	3	4	5	6	7	8	9	10	11	12	13	14	15		
Поймы р. Убыть и р. Кылка															1,0	
Выровненные участки (плакоры с крутизной склона от 0 до 1 градуса)	0,2	0,3			0,05		0,15	0,1			0,05				0,15	
Слабые склоны выпуклой формы крутизной от 1 до 3°	0,1	0,6			0,1		0,05	0,05			0,1					
Средние склоны выпуклой формы крутизной от 3 до 5°		0,1	0,4					0,3	0,2							

Продолжение таблицы 3.3

1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
Крутые склоны выпуклой формы крутизной > 5°			0,3	0,2					0,2	0,3				
Слабые склоны вогнутой формы крутизной от 1 до 3°		0,3			0,1	0,2					0,2	0,2		
Средние склоны вогнутой формы крутизной от 3 до 5°						0,3						0,7		
Крутые склоны вогнутой формы крутизной > 5°						0,6						0,4		
Днища оврагов														1,0
Днища глубоких оврагов (глубина >5 м)													0,3	0,7

* Π_3^D - дерново-неглубокоподзолистые почвы, Π_4^D - дерново-глубокоподзолистые почвы, $\Pi_3^D \downarrow$ - дерново-неглубокоподзолистые слабосмытые почвы, $\Pi_4^D \downarrow$ - дерново-глубокоподзолистые слабосмытые почвы, $\Pi^D \downarrow\downarrow$ - дерново-подзолистые среднесмытые почвы, $\Pi^D \downarrow\downarrow\downarrow$ - дерново-подзолистые сильносмытые почвы, СЛ- светло-серые лесные почвы, Π_3^D нам- дерново-неглубокоподзолистые намытые почвы, Π_4^D нам- дерново-глубокоподзолистые намытые почвы, Дг- дерново-глеевые почвы, А- аллювиальные дерновые слоистые почвы, Ал - аллювиальные луговые кислые почвы, A_{Fe} - аллювиальные луговые кислые ожелезненные почвы, З- зональные почвы, ЭЗ- эрозионно-зональные почвы, Э- эрозионные почвы, ЭА- эрозионно-аккумулятивные, Г- гидроморфные почвы, ПП- полугидроморфные подчиненных элементов ландшафта.

3.2 Создание индикационной таблицы по результатам полевого обследования

Основываясь на собранном материале в результате полевых сезонов 2014-2015 гг., а также карте элементарных поверхностей мезорельефа, нами была составлена индикационная таблица для выделения почвенных контуров (табл. 3.3).

Согласно индикационной таблице, выровненные участки территории (то есть плакоры, крутизна склона которых составляла не более 1^0 с различной формой склона) заняты зональными и эрозионно-зональными почвенными структурами: 20% приходится на пахотные дерново-подзолистые почвы разной степени подзолистости, 30% - слабосмытые пахотные дерново-подзолистые почвы разной степени подзолистости, 5% - светло-серые лесные пахотные почвы, 15% - дерново-подзолистые почвы разной степени подзолистости под лесом, 10% - слабосмытые дерново-подзолистые почвы разной степени подзолистости под лесом, 5% - серые лесные непахотные почвы и 15% - перегнойно-грунтово-глеевые почвы. Почвенный покров слабых склонов выпуклой формы крутизной от 1 до 3^0 представлен зональными и эрозионно-зональными почвенными структурами: 10% приходится на пахотные дерново-подзолистые почвы разной степени подзолистости, 60% - слабосмытые пахотные дерново-подзолистые почвы разной степени подзолистости, 10% - светло-серые лесные пахотные почвы, 5% - дерново-подзолистые почвы разной степени подзолистости под лесом, 5% - слабосмытые дерново-подзолистые почвы разной степени подзолистости под лесом, 10% - серые лесные непахотные почвы. На средних склонах выпуклой формы крутизной от 3 до 5^0 расположены эрозионно-зональные и эрозионные почвенные структуры: 10% - пахотные слабосмытые дерново-подзолистые почвы разной степени подзолистости, 40% - пахотные дерново-подзолистые среднесмытые почвы, 30% - слабосмытые дерново-подзолистые почвы разной степени подзолистости под лесом, 20% - дерново-подзолистые

среднесмытые почвы под лесом. Крутые склоны выпуклой формы с крутизной более 5^0 заняты только эрозионными почвенными структурами: 30% территории занимают дерново-подзолистые пахотные среднесмытые почвы, 20% - дерново-подзолистые пахотные сильносмытые почвы, 20% - дерново-подзолистые среднесмытые почвы под лесом, 30% - дерново-подзолистые сильносмытые почвы под лесом. Составленная нами индикационная таблица показала также, что слабые склоны вогнутой формы крутизной от 1 до 3^0 заняты различными почвенными структурами такими, как эрозионно-зональные, зональные и эрозионно-аккумулятивные: 30% территории таких склонов заняты пахотными slabosmytymi дерново-подзолистыми почвами разной степени подзолистости, 10% - серыми лесными пахотными почвами, 20% - пахотными дерново-подзолистыми почвами разной степени подзолистости и намытости, 20% - серыми лесными почвами под лесом, 20% - дерново-подзолистыми почвами разной степени подзолистости и намытости под лесными массивами. Средние склоны вогнутой формы крутизной от 3 до 5^0 заняты эрозионно-аккумулятивными почвенными структурами: 30% занимают пахотные дерново-подзолистые почвы разной степени подзолистости и намытости, а 70% - дерново-подзолистые почвы разной степени подзолистости и намытости под лесными массивами. Почвенный покров крутых вогнутых склонов представлен также эрозионно-аккумулятивными почвенными структурами и состоит из пахотных дерново-подзолистых почв разной степени подзолистости и намытости и таких же почв только под лесом (60% и 40%, соответственно) (табл.3.3). Анализ этой таблицы показал, что днища оврагов (глубиной менее 5 м) изучаемой территории на 100% заняты дерново-глеевыми почвами, в то время как 70% оврагов глубиной более 5 метров заняты дерново-глеевыми почвами. 30% почвенного покрова таких оврагов занимают аллювиальные дерновые слоистые почвы. Причиной этому являются мощные водные потоки талой воды, протекающие по этим оврагам в весенний период. Кроме того, 15% территории выровненных участков также занимают полугидроморфные

почвы подчиненных элементов ландшафта, а именно перегнойно-грунтово-глеевые почвы. Почвенный покров деревенского поселения на 100% был отнесен к антропогенно-преобразованным почвам приусадебных участков. Поймы рек Убыти и Кыпки на 100% заняты аллювиальными луговыми почвами.

3.3 Создание картограммы пространственного распределения гранулометрического состава верхнего гумусого горизонта

Среднее содержание физической глины в верхних гумусовых горизонтах обследованных точек составило 41 % при стандартном отклонении 4,5%, минимальное значение составило 27,3%, а максимальное – 60,6% (см. приложение 5). Наибольшие значения содержания физической глины характерны для точек, располагающихся в тальвегах оврагов, а также в участках, где происходит намыв почв и в верхних горизонтах среднесмытых и сильносмытых почв. Наименьшие значения содержания физической глины наблюдались в поймах рек, а также на территориях, расположенных на выровненных частях рельефа или территории со слабыми углами наклона поверхности.

Анализ коэффициентов корреляции между полученными данными по гранулометрическому составу и ковариатами окружающей среды показал зависимость этого показателя от некоторых морфометрических характеристик, таких как крутизна склона, абсолютная высота, фактор длины склона, аспект, продольная кривизна и другие. Последующий регрессионный анализ выявил ковариаты, оказывающие большее влияние, и позволил составить уравнение для интерполяции данных по содержанию физической глины на всю территорию исследования (табл. 3.4). Коэффициент детерминации R^2 получился равным 0,31, а скорректированный R^2 оказался равным 0,29.

В итоге было получено регрессионное уравнение, коэффициенты которого приведены в таблице 3.4. Не указанные в таблицы ковариаты не являются значимыми.

Таблица 3.4

Результаты регрессионного анализа между содержанием физической глины и морфометрическими характеристиками рельефа

Морфометрические характеристики	Коэффициент	Ошибка предсказания
Свободный член	29,00	3,40
Крутизна склона	-274,08	84,87
Русловая сеть	0,049	0,020
Продольная кривизна	1795,74	493,00
Угол наклона	0,14	0,05
Число мельтон неровностей	3,15	1,00
Индекс пересеченности местности	20,00	6,09

Была построена картограмма пространственного распределения гранулометрического состава верхнего гумусого горизонта на исследуемой территории (рис.3.4). В результате вычислений содержание физической глины на территории исследования находилось в диапазоне от 36% до 52%. Так как, на исследуемой территории распространен подзолистый тип почвообразования, то, согласно классификации почв по гранулометрическому составу, предложенной Н.А. Качинским, почвы имеют 3 градации гранулометрического состава: средний, тяжелый суглинок и глина легкая (Качинский, 1958).

Территория ключевого участка преимущественно имеет тяжелосуглинистый гранулометрический состав, в то время как аллювиальные почвы, распространенные в поймах рек Убыть и Кыпка, имеют среднесуглинистый гранулометрический состав верхнего горизонта. Наложение изолиний рельефа показывает, что среднесуглинистый гранулометрический состав имеют почвы, залегающие на относительно выровненных территориях с углами наклона поверхности до 3⁰ (рис. 3.4).

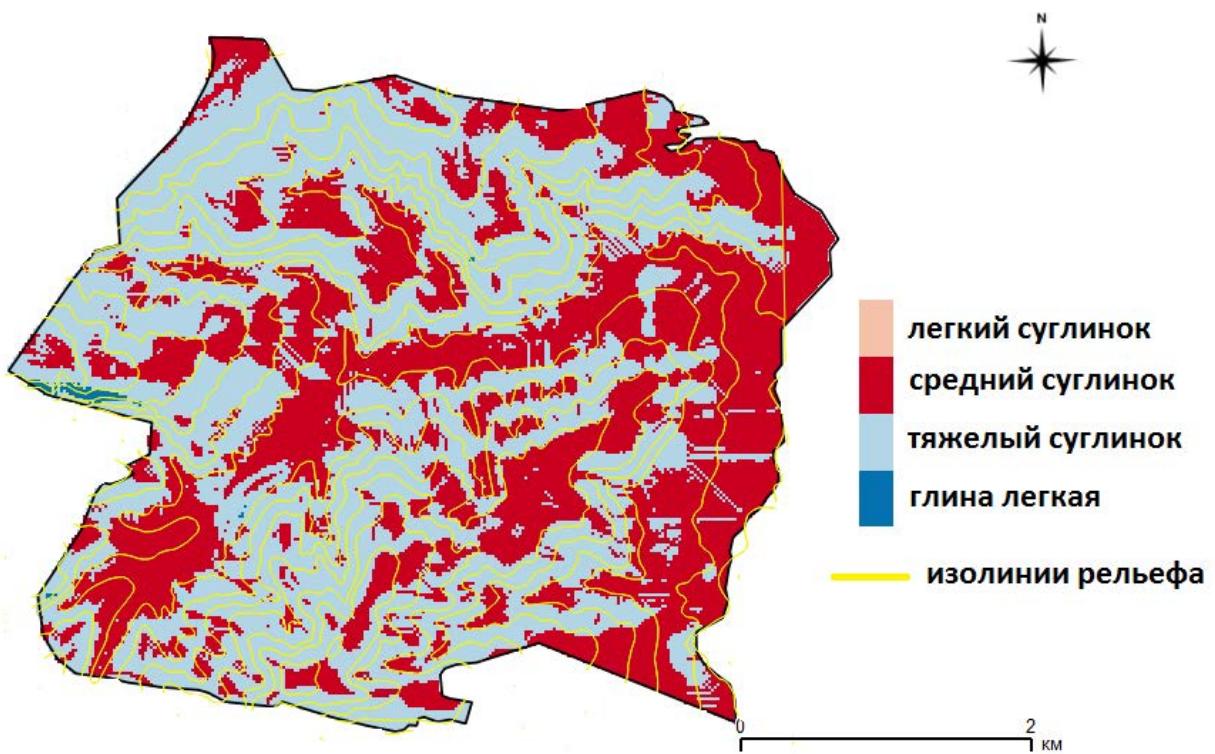


Рис. 3.4 Картограмма пространственного распределения гранулометрического состава верхнего гумусого горизонта на исследуемой территории

3.4 Создание крупномасштабной почвенной карты

На основе созданной ГИС-системы, картограммы пространственного распределения гранулометрического состава верхнего гумусого горизонта и индикационной таблицы была построена почвенная карта на исследуемую территорию (рис.3.5). На ней отображены следующие почвенные выделы: А_{Fe} - аллювиальная луговая кислая ожелезненная, Ал - аллювиальная луговая кислая, А – аллювиальная дерновая кислая слоистая, Дг - дерново-глеевая, П $\frac{d}{3}$ – дерново-неглубокоподзолистая, П $\frac{d}{4}$ – дерново-глубокоподзолистая, П $\frac{d}{3} \downarrow$ – дерново-неглубокоподзолистая слабосмытая, П $\frac{d}{4} \downarrow$ – дерново-подзолистая среднесмытая, П $d \downarrow\downarrow$ – дерново-подзолистая сильносмытая, П $\frac{d}{3}$ нам – дерново-неглубокоподзолистая намытая, П $\frac{d}{4}$ нам – дерново-глубокоподзолистая намытая, СЛ – светло-серая лесная , СЛ Γ – светло-серая лесная глееватая.

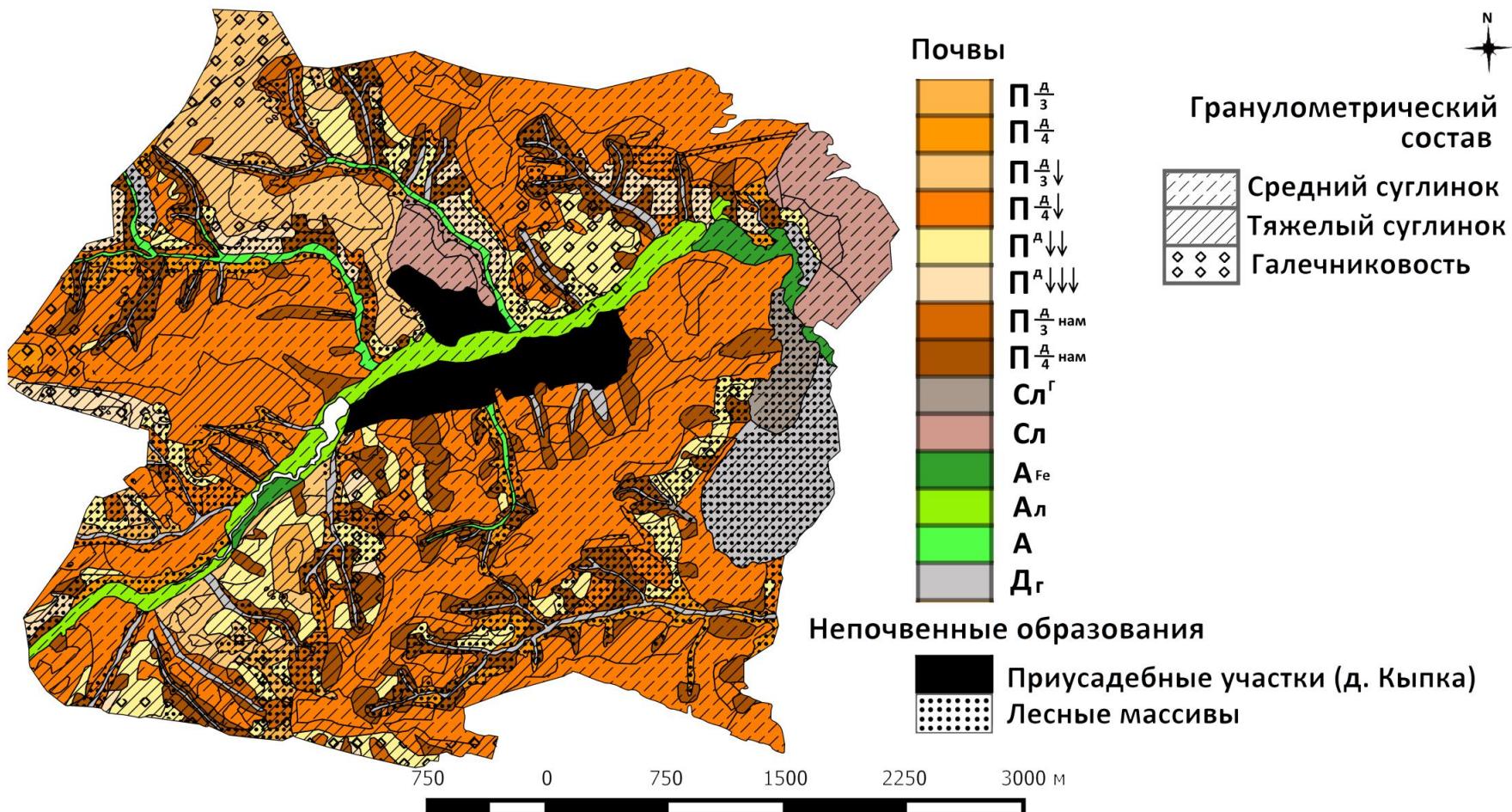


Рис. 3.5 Почвенная карта, составленная традиционными методами почвенного обследования с элементами ГИС-анализа. Обозначения почв приведены в тексте

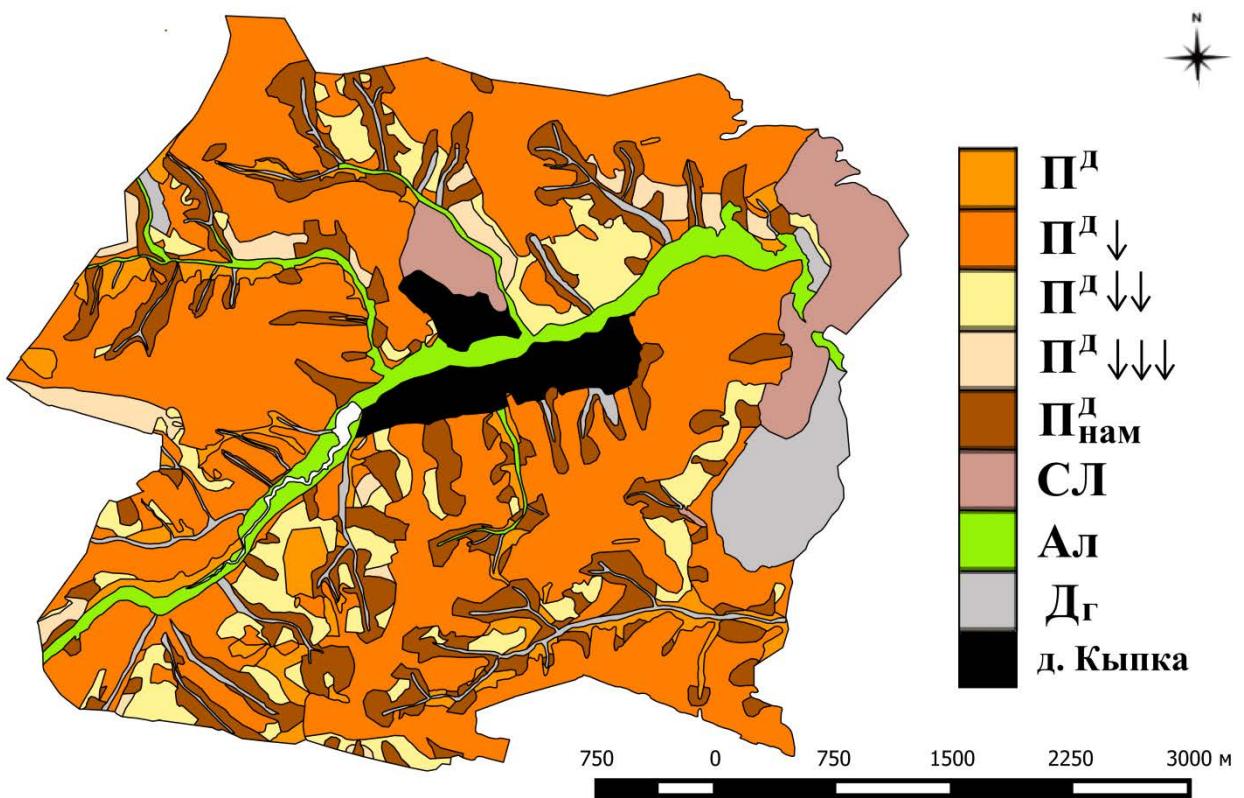


Рис. 3.6 Почвенная карта 8 условных почвенных групп. Обозначения почв приведены в тексте

Карта включает 41 разновидность почв, которые относятся к четырем различным типам почв.

Для дальнейшего сравнения почвенных карт использовалась упрощенная версия почвенной карты, в которой были отражены 8 условных почвенных групп: дерново-подзолистые несмытые, дерново-подзолистые слабосмытые, дерново-подзолистые среднесмытые, дерново-подзолистые сильносмытые, дерново-подзолистые намытые, аллювиальные, дерново-глеевые и светло-серые лесные почвы (рис.3.6).

3.5 Анализ почвенной карты и сравнение с почвенной картой 1984 г.

3.5.1 Анализ изменения факторов почвообразования

Был проведен анализ изменения факторов почвообразования на исследуемой территории с 1984 по 2016 г. Изменение антропогенного влияния было оценено путем сравнения карты организации пашни и космического

снимка разрешения WorldView-1 (дата съемки – 15 сентября 2015 года). Оказалось, что контура полей в хозяйстве не претерпели значительных изменений. Дополнительно, из каталога снимков Landsat (5, 7 и 8) были отобраны весенние и летние сцены 1984 - 2015 гг. с низкой облачностью и контрастными границами полей. Всего было найдено 25 сцен, анализ которых показал, что состав растительного покрова территории, контура полей и интенсивность их использования на протяжении последних 30 лет оставались неизменными.

Полевое обследование показало, что на исследуемой территории повсеместно распространенной почвообразующей породой являются покровные красновато-бурые суглинки, подстилаемые галечниковой мореной. Можно считать, что за последние 30 лет не произошли перемены в составе почвообразующих пород данного региона и существенного изменения климата.

Визуальное сравнение созданной нами ЦМР и карты эрозии 1984 г. позволяет предположить, что рельеф данной местности существенных изменений не претерпел, в то время как интенсивность эрозионных процессов могла поменяться.

Неизменность направленности и интенсивности хозяйственной деятельности (антропогенного воздействия), состава растительного покрова, рельефа, климата и почвообразующих пород, распространённых на обследованной территории позволяет предположить, что почвенная карта 1984 г. отражает близкий к современному почвенный покров и составлена специалистами Гипрозема с высокой точностью.

3.5.2 Сравнение полученной почвенной карты с материалами обследования 1984 г.

В результате анализа построенной почвенной карты было выявлено, что менее 4,9% или $0,95 \text{ км}^2$ территории занято зональными несмытыми дерново-подзолистыми почвами, а 5,4% (1 км^2) территории занимают светло-серые

лесные почвы. Большую часть территории занимают эрозионно-зональные, эрозионные и эрозионно-аккумулятивные почвенные структуры: дерново-подзолистые почвы различной степени смытости – 60% (11,5 км²), дерново-подзолистые намытые почвы – 15,4% (2,95 км²). Гидроморфный ряд почв представлен аллювиальными почвами, которые согласно построенной карте занимают 0,96 км² (около 5% от общей площади территории). Дерново-глеевыми полугидроморфными почвами занято 6,6% территории или 1,2 км² (рис.3.6, табл.3.5).

Таблица 3.5
Площади основных почвенных выделов на исследуемой территории
согласно картам 1984 г. и 2016 г.

Почвенные выделы	1984		2016	
	Общая площадь (км ²)	% от площади	Общая площадь (км ²)	% от площади
Аллювиальные почвы	0,71	3,7	0,96	5,0
Дерново-подзолистые несмытые почвы	1,67	8,7	0,95	4,9
Дерново-подзолистые слабосмытые почвы	8,8	45,7	9,3	48,4
Дерново-подзолистые среднесмытые почвы	2,1	10,9	1,7	8,6
Дерново-подзолистые сильносмытые почвы	0,38	2,0	0,6	3,0
Дерново-подзолистые намытые почвы	-	-	2,95	15,4
Светло-серые лесные почвы	1,3	7,0	1,0	5,4
Дерново-глеевые почвы	3,7	19,3	1,2	6,6
Почвы приусадебных участков	0,5	2,5	0,5	2,5
Водная поверхность	0,04	0,2	0,2	0,2
Итого	19,2		19,2	

Сравнение почвенной карты 2016 г. и карты 1984 г. показало незначительное увеличение доли слабосмытых и сильносмытых дерново-подзолистых почв при одновременном уменьшении доли среднесмытых почв на карте 2016 г. по сравнению с картой 1984 г. (табл.3.5). В отличие от

почвенной карты 1984 г., на данной территории были вскрыты контуры намытых почв ($2,95 \text{ км}^2$ или 15,4% от площади исследуемого участка). Сравнение площадей, занимаемых основными почвенными группами по картам 2016 и 1984 гг., показало значительное уменьшение доли дерново-глеевых почв (карта 1984 г. – 19,3%, карта 2016 г. – 6,6%), светло-серых лесных почв (карта 1984 г. – 8,7%, карта 2016 г. – 4,9%) и дерново-подзолистых несмытых почв (карта 1984 г. – 7,0%, карта 2016 г. – 5,4%) (табл.3.5). Кроме того, около 2,5% территории приходится на антропогенно-преобразованные почвы приусадебных участков, расположенные непосредственно в районе д. Кыпка.

Визуальное сравнение двух карт показало, что обе карты одинаково отражают почвенный покров исследуемой территории. На обеих картах в одних и тех же местах выделяются области распространения одинаковых почв. Точное выделение ареалов является результатом использования данных дистанционного зондирования и цифровой модели рельефа, так как появилась возможность более точной отрисовки оврагов, выделения ареалов дерново-подзолистых среднесмытых и сильносмытых почв.

Анализ таблиц сопряженности, отражающих гранулометрический состав почв, согласно составленной карте и карте 1984 г., показал, что сходимость этих результатов нулевая. Коэффициент каппа равен 0,003. Это объясняется субъективностью в определении гранулометрического состава почвы при полевых обследованиях. Решающую роль в получении конечного результата играет человеческий фактор. Все это в конечном итоге, приводит к различиям в полученной карте и карте 1984 г. Практически все почвенные контуры карты 1984 г. тяжелосуглинистые, в то время как у нас встречались как среднесуглинистые почвы, так и тяжелосуглинистые. Несомненно, гранулометрический состав почвенного покрова данной территории тяжелый, что является следствием почвообразующих пород.

Была изучена точность отнесения 166 точек обследования. На карте 2016 г. 144 точки исследования - это дерново-подзолистые почвы, в то время как

всего лишь 109 точек из них соответствуют дерново-подзолистым почвам на карте 1984 г. Составленная таблица сопряженности между почвенной группой в точках обследования по карте, созданной нами на основе традиционной почвенной съемки с использованием элементов ГИС-анализа, и почвенной группой по карте 1984 г. показала, что общая точность составила 47%, а каппа – 0,26.

Эти результаты объясняются тем, что на карте 1984 г. овраги были выделены более тонкими. Часть диагностированных нами дерново-подзолистых почв попала на почвы овражно-балочной сети по карте 1984 г. Некоторые почвенные разрезы закладывались в непосредственной близости от оврагов на расстоянии 10-20 м, и эти точки, в которых нами вскрывались дерново-подзолистые почвы, по карте 1984 г. соответствовали овражно-балочным почвам, что и объясняется «широкими» отрисованными ареалами овражных почв. Обследование территории также показало, что часть почвенных ареалов, отнесенных по карте 1984 г. к серым лесным почвам, нами были отнесены к дерново-подзолистым почвам с разной степенью смытости. Поймы рек Убыть и Кыпка на карте 1984 г. занимают аллювиальные дерновые кислые почвы, в то время как нами в поймах этих рек были нами вскрыты аллювиальные луговые кислые почвы. А почвы некоторых самых глубоких оврагов на исследуемой территории нами были отнесены к аллювиальным луговым слоистым.

Анализ таблиц сопряженности, составленных по степени смытости для дерново-подзолистых почв, показал, что общая точность, характеризующая сходимость результатов, равна 49%. То есть степень смытости в любой точке нашей карты будет совпадать со степенью смытости в этой же точке по карте 1984 г. в 49 процентах случаев. Коэффициент каппа получился равным 0,25. Возможно такой низкий коэффициент каппа в этом случае является результатом совокупного влияния трех факторов: выделение контуров намытых почв на карте 2016 г. (1), субъективность оценки при определении степени смытости в полевых условиях (2), а также возможное нарастание

интенсивности эрозионных процессов в регионе исследования за последние 30 лет (3).

Детальное выделение почвенных контуров не являлось для авторов карты 1984 г. первостепенной задачей. Это являлось одной из причин низких уровней сходимости при анализе таблиц сопряженности. Кроме того, немаловажной причиной, являлось то, что при определении степени смытости, оподзоленности или гранулометрического состава присутствует высокая доля субъективности, что также приводит к снижению степени сходимости результатов.

3.6 Оценка точности почвенной карты

Оценка созданной нами почвенной карты проверялась двумя способами.

Таблица 3.6

Таблица сопряженности между построенной нами почвенной картой и почвами, диагностированными в точках независимой случайной выборки

		Почвы независимой случайной выборки								
		Ал	Дг	СЛ	Π^d	$\Pi^d \downarrow$	$\Pi^d \downarrow\downarrow$	$\Pi^d \downarrow\downarrow\downarrow$	$\Pi_{нам}^d$	Сумма
Почвы согласно почвенной карте 2016 г.	Ал	2			1					3
	Дг		1							1
	СЛ			2		1				3
	Π^d				1					1
	$\Pi^d \downarrow$				1	10				11
	$\Pi^d \downarrow\downarrow$						2			2
	$\Pi^d \downarrow\downarrow\downarrow$						1	2		3
	$\Pi_{нам}^d$		1						2	3
	Сумма	2	2	2	3	11	3	2	2	27

*Примечание: Обозначения почв приведены в разделе 3.4.

Таблица 3.7

Распределение почвенных разновидностей и условных вероятностей их определения

Способ диагностики ЭПС	Средние оценки условных вероятностей	Количество почвенных разновидностей	Наименования почвенных разновидностей*
полное почвенное опробование (3- 5 точек на один ареал)	0,96-1	2	Светло-серые лесные средне- и тяжелосуглинистые и глееватые
неполное почвенное опробование (2 точки + индикация)	0,85-0,95	6	Дерново-неглубоко- и глубокоподзолистые освоенные несмытые и слабосмытые, светло-серые лесные среднесуглинистые, дерново-подзолистые сильносмытые освоенные
слабое почвенное опробование (небольшое количество точек + индикация)	0,71-0,84	13	Все аллювиальные почвы, дерново-подзолистые среднесмытые освоенные, дерново-неглубоко- и глубокоподзолистые освоенные слабосмытые, светло-серые лесные намытые, дерново-подзолистые сильносмытые освоенные тяжелосуглинистые, дерново-подзолистые среднесмытые среднесуглинистые
только индикация (разрезы практически отсутствуют)	0,67-0,70	20	Все дерново-подзолистые намытые, все дерново-глеевые почвы, дерново-подзолистые среднесмытые тяжелосуглинистые, дерново-неглубоко- и глубокоподзолистые слабосмытые, дерново-неглубоко- и глубокоподзолистые несмытые, дерново-подзолистые сильносмытые освоенные глинистые, дерново-подзолистые сильносмытые тяжелосуглинистые, светло-серые лесные

* там где не указан гранулометрический состав, то он средне- и тяжелосуглинистый

На основе случайной независимой проверочной выборки, состоящей из 27 точек, была составлена таблица сопряженности между построенной нами почвенной картой и почвами, выявленными в этих точках (табл. 3.6). Анализ этой таблицы показал, что общая точность созданной нами почвенной карты составила 82%, а коэффициент каппа показал почти полное совпадение «реальных» почв в точках с почвами, отраженными на почвенной карте и составил 0,76.

Для оценки созданной нами почвенной карты была составлена таблица 3.7, описывающая распределение оценок условных вероятностей среди почвенных разновидностей территории исследования (Сорокина, 2006). Средняя оценка условных вероятностей для созданной нами почвенной карты составила 0,74, что согласно критериям, предложенным Сорокиной Н.П., соответствует категории точности – «вероятная».

ГЛАВА 4. СОЗДАНИЕ ПОЧВЕННЫХ КАРТ СРЕДСТВАМИ ЦПК

4.1 Выбор ковариат и размера пикселя для моделирования

В результате дисперсионного анализа было отобрано 25 факторно-индикационных переменных, 20 из которых представлены в таблице (табл.4.1).

Таблица 4.1

Минимальные, максимальные и средние значения ковариат,
отобранных для моделирования

Ковариата	Минимальное значение	Среднее значение	Максимальное значение
абсолютная высота	131,83	182,99	236,92
водосборная площадь	5	110	377495
фактор длины склона	0	0,39	4,78
тангенциальная кривизна долины	-0,011	0,00003	0,0087
долины	0	3,72	40,7
индекс расчленности рельефа	0	0,145	1,09
минимальная кривизна	-0,0163	-0,00023	0,0051
топографический индекс влажности	4,43	7,82	23,13
выпуклость поверхности	0	0,55	1
модифицированная водосборная площадь	66,58	1786	183796
максимальный вес	132,56	191,49	236,96
относительная позиция склона	0	0,62	1
основная кривизна	-0,032	0,000044	0,035
топографический индекс позиции	-5,1	0,033	5,21
индекс сходимости	-73,20	0,15	69,11
крутизна склонов водосбора	0	2,19	12,75
крутизна склона	0	0,04	0,3
суточное освещение	-0,28	-0,0016	0,23
индекс баланса массы	-0,61	0,0028	1,63
поперечная кривизна	-0,023	0,000048	0,015

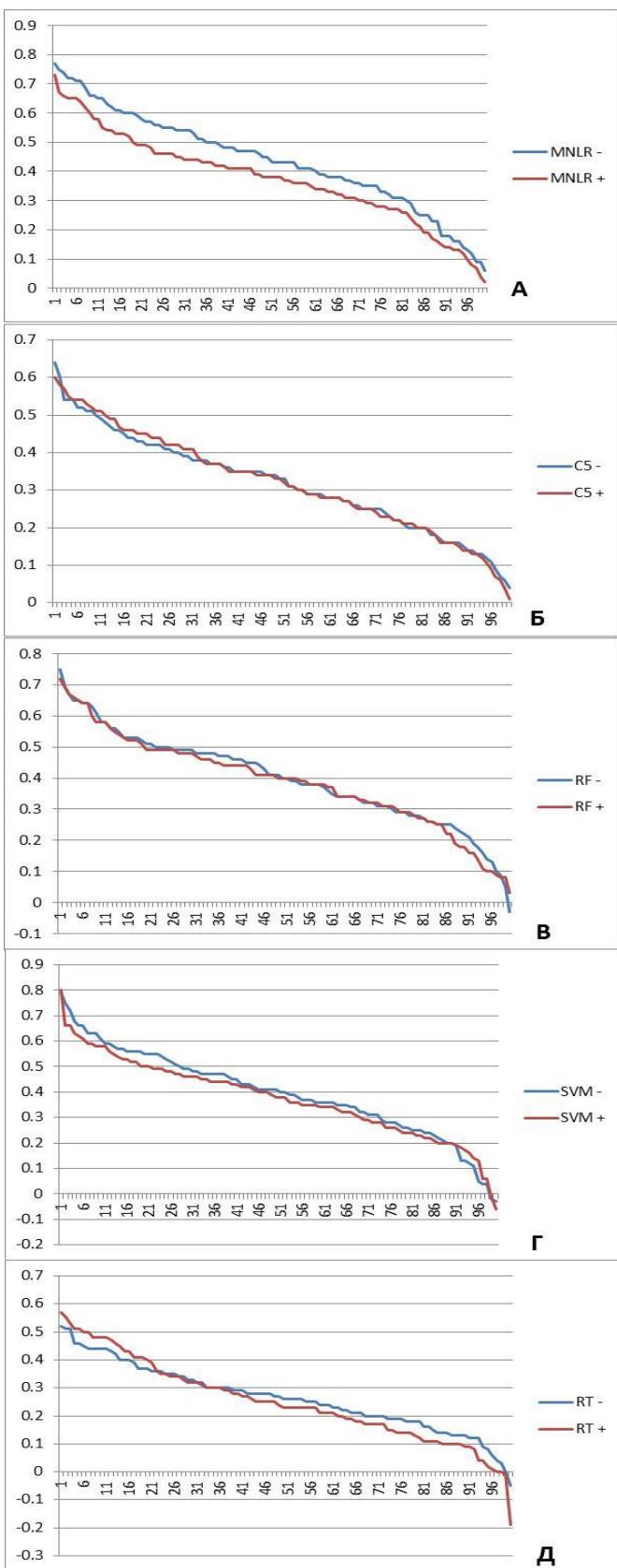


Рис. 4.1 Ранжированные ряды для капп самовоспроизводимости при моделировании с использованием в качестве ковариат почвенной карты 1984 г. (+) и без них (-).^{*}Примечания: MNLR- множественная логистическая регрессия, C5 – дерево решений C5, RF – метод случайного леса, SVM – метод опорных векторов, RT – регрессионные деревья.

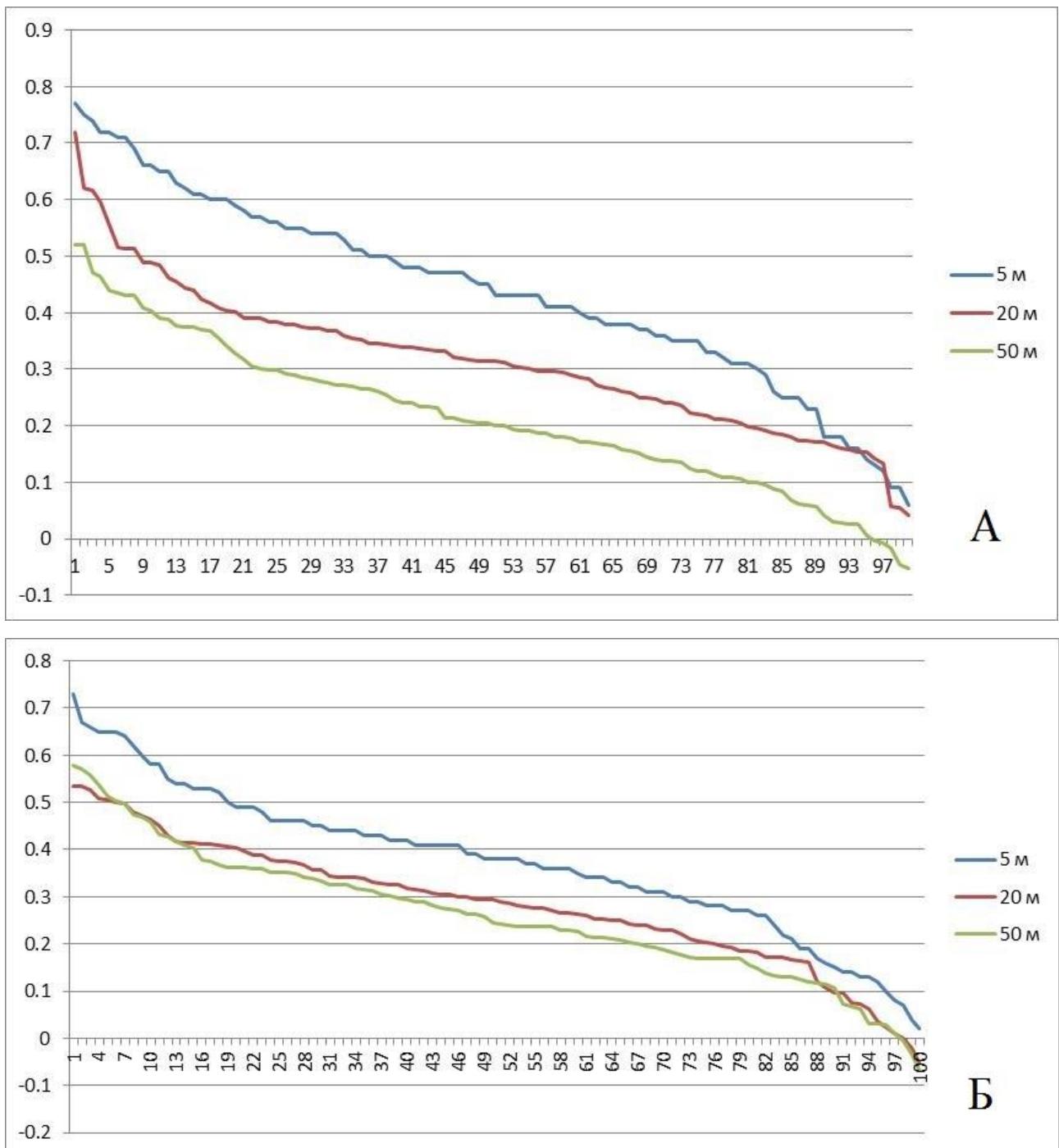


Рис. 4.2 Ранжированные ряды для капп самовоспроизводимости при моделировании на основе множественной логистической регрессии для разных размеров пикселей исходных растрков ковариат. А – без использования для обучения карты 1984 г. Б – с использованием для обучения карты 1984 г.

Значимыми переменными оказались оцифрованные границы пашни, лесных массивов, поймы, лугов и балочных понижений. Далее проводился процесс моделирования пятью алгоритмами с использованием отобранных

ковариат, включая в процесс моделирования почвенную карту 1984 г. и без нее.

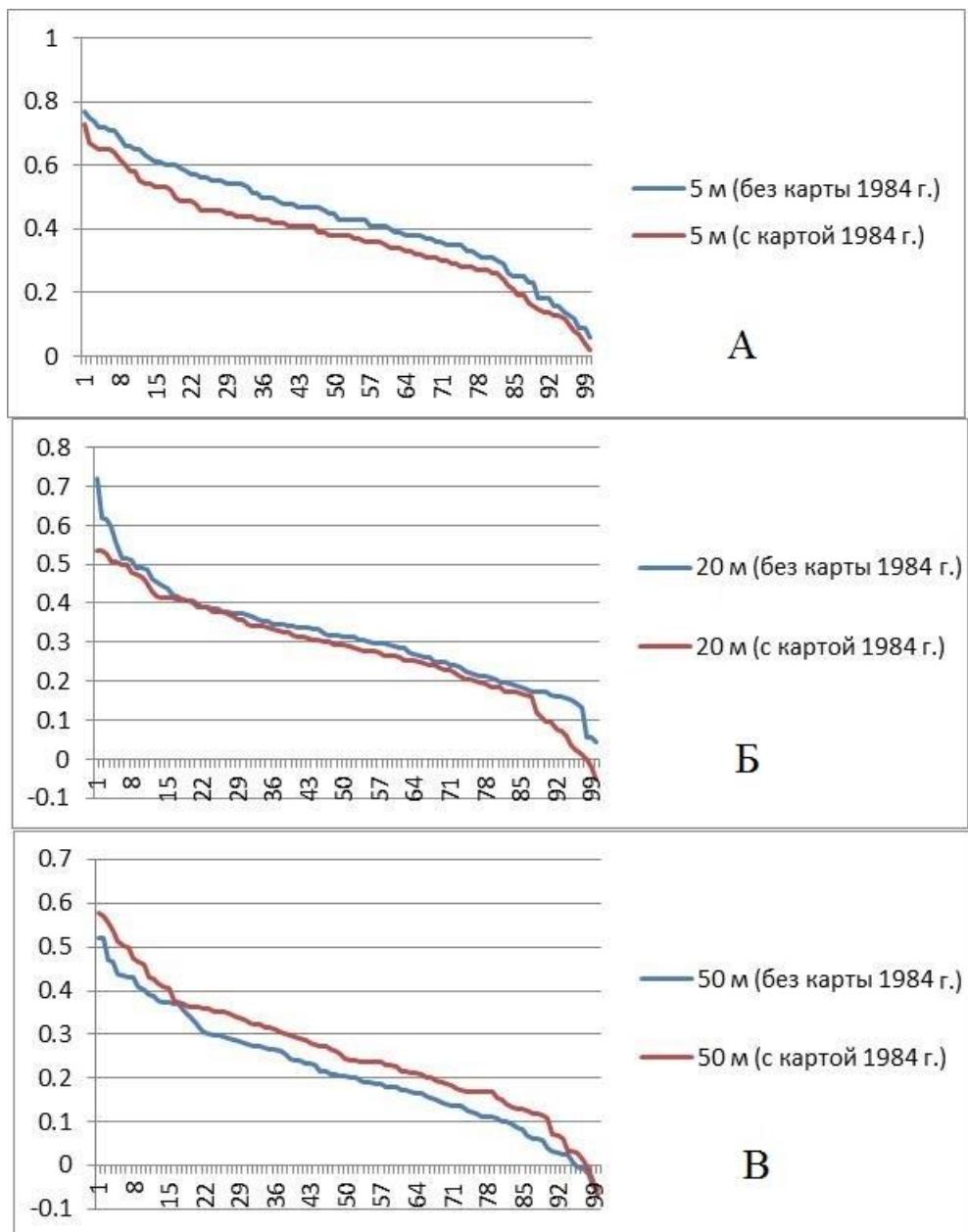


Рис. 4.3 Ранжированные ряды для каппа самовоспроизводимости при моделировании на основе множественной логистической регрессии для разных размеров пикселей исходных растров ковариат. А – размер пикселя 5 м. Б – размер пикселя 20 м. В - размер пикселя 50 м.

Процесс моделирования почвенных карт происходил 2 путями: с использованием карты 1984 г., и без нее. Внедрение в процесс моделирования старых почвенных материалов дало схожие результаты по средней общей точности и среднему коэффициенту каппа при моделировании с

использованием карты 1984 г. и при моделировании без нее (см. приложение 1,2, рис.4.1).

Визуальное сравнение карт, построенных с использованием и без использования в процессе моделирования почвенной карты 1984 г., показало, что различия в этих картах невелики: соотношение площадей почвенных выделов и рисунок распределения их на территории исследования по картам, построенным одним и тем же методом, схож.

В ходе подбора оптимального размера пикселя для цифрового почвенного картографирования были построены кривые, отображающие ранжирование коэффициента каппа моделей всех вариантов размера пикселей и при условии использования для обучения почвенной карты 1984 г. и без нее (рис. 4.2, 4.3; см. приложение 3). В качестве модели для картографирования была взята множественная логистическая регрессия.

Таким образом, распределение коэффициента каппа моделей множественной логистической регрессии для разных размеров пикселей показало, что при снижении размера пикселя значения капп моделей увеличиваются, соответственно увеличивается и точность моделирования (рис. 4.2, 4.3). Добавление в процесс обучения почвенной карты 1984 г. приводит к уменьшению разброса коэффициентов каппа между моделями при размере пикселя 5 метров и 50 метров (рис.4.2). Добавление в процесс обучения почвенной карты 1984 г. при размере пикселя 20 м практически не влияет на величину коэффициента каппа, и кривые распределения в этих случаях практически совпадают (рис.4.3, Б), а для пикселя размером 50 м наблюдается обратная картина, кривая, характеризующая ранжированный ряд коэффициентов каппа при моделировании с использованием карты 1984 г., находится выше, чем кривая, описывающая распределение капп при моделировании без нее (рис.4.3, В). В итоге, моделирование почвенного покрова и обсуждение результатов моделирования в дальнейшем основывалось на почвенных картах, построенных без использования в качестве ковариат почвенной карты 1984 г.,

а в процессе моделирования нами использовались раstry ковариат с размером пикселей 5 м.

4.2 Моделирование на основе множественной логистической регрессии

Общая точность при моделировании пространственного распределения почв на основе множественной логистической регрессии без использования в качестве ковариат карты 1984 г. составила от 30 до 85 % при средней общей точности 61,1 %, а каппа - от 0,06 до 0,77 при среднем каппе, равном 0,44 (см. приложение 1,2).

В ходе исследования было выявлено, что распределение каппы и общей точности для каждой модели подчиняются закону нормального распределения (см. приложение 1,2, рис.4.4). Поэтому в дальнейших исследованиях мы использовали модели со средней и максимальной каппой самовоспроизводимости (по нашему предположению, карта, построенная на основе модели с максимальной каппой, должна была быть более приближенной к реальности).

Визуальный анализ полученных цифровых почвенных карт и анализ долей площадей почвенных выделов показал, что карты с максимальной и средней каппой, в целом, похожи и достаточно верно отражают почвенный покров исследуемой территории (рис.4.5, табл.4.2).

Таблица 4.2

Доля площадей почвенных выделов (в %) для карт, построенных на основе множественной логистической регрессии: Карта А- на основе модели со средней каппой, $k = 0,44$; Карта Б - на основе модели с максимальной каппой, $k=0,77$ (рис.4.5)

Почвенные выделы	A	ДГ	СЛ	П ^д	П ^д ↓	П ^д ↓↓	П ^д ↓↓↓	П ^д _{нам}
Карта А	4,3	6,1	2,4	6,6	53,2	21,4	0,2	5,8
Карта Б	4,5	5,8	2,0	7,6	55,7	19,6	0,4	4,4

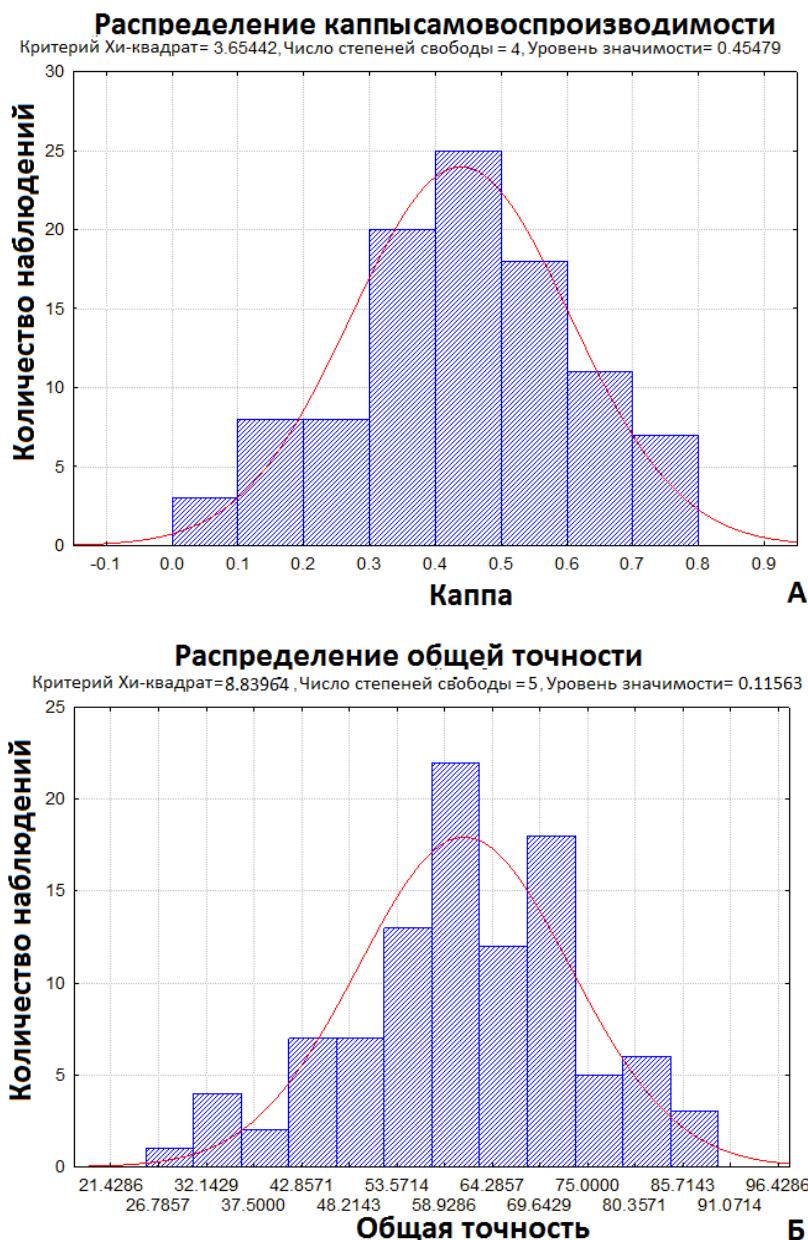


Рис.4.4 Распределение каппы самовоспроизводимости моделей (А) и общей точности (Б) при моделировании на основе множественной логистической регрессии

Этот анализ выявил незначительные различия (около 2 %) в доле площадей дерново-подзолистых несмытых и слабосмытых почв при классификации на основании моделей множественной логистической регрессии с максимальной и средней кappой самовоспроизводимости.

По картам на основе этих моделей четко отделяются балки и овраги, занятые ареалами дерново-глеевых почв (рис. 4.5). Множественная логистическая регрессия лучше всего предсказывает дерново-подзолистые

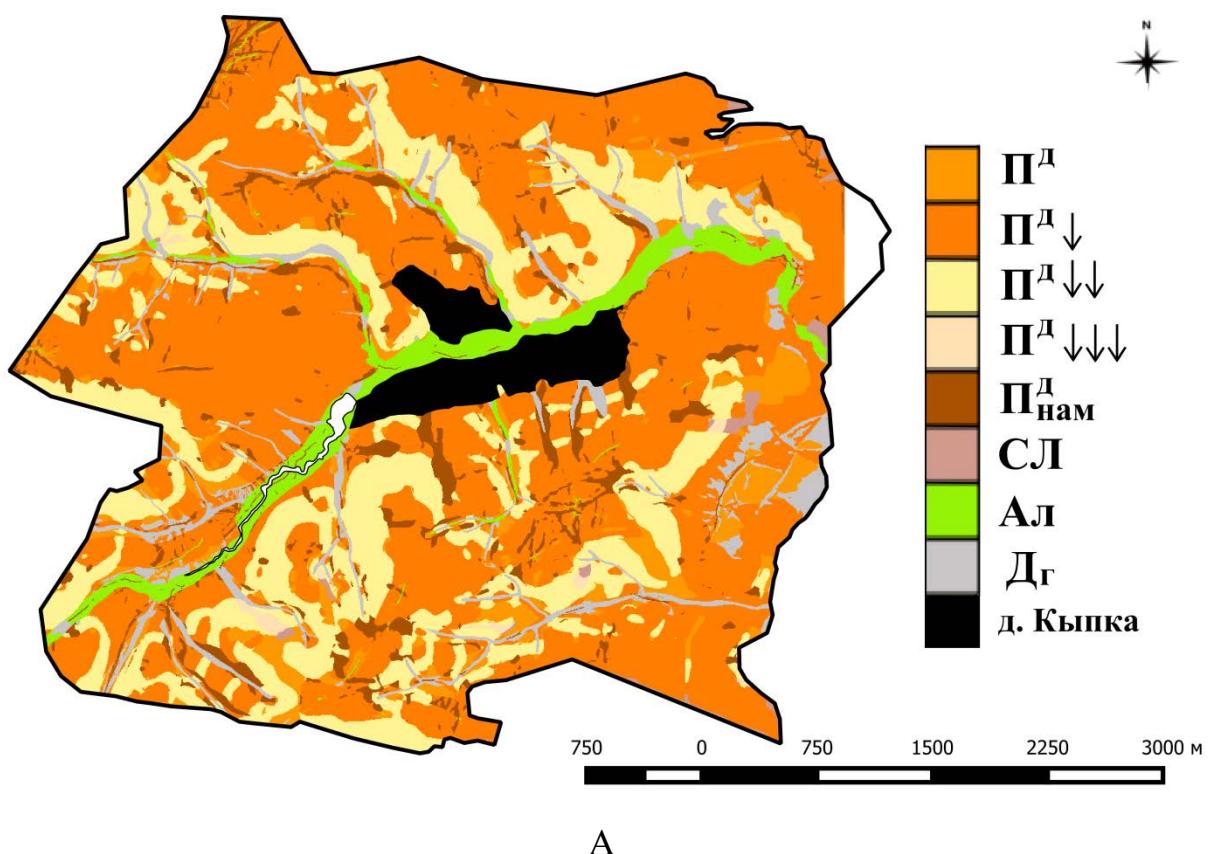
несмытые и слабосмытые почвы, а также аллювиальные почвы (точность производителя 71,7 %, 78,4 % и 95,4 %, соответственно) (табл. 4.3).

Таблица 4.3

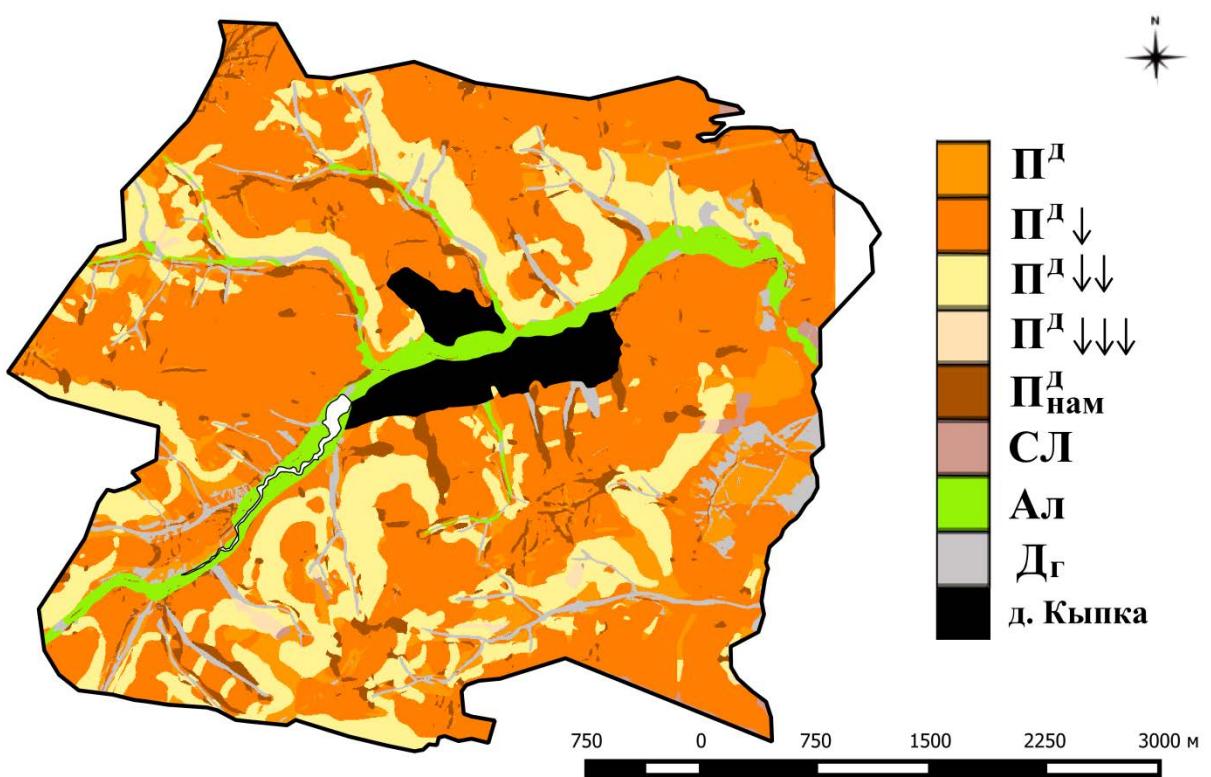
Точность производителя и пользователя при моделировании на основе множественной логистической регрессии.

Почвы	Точность производителя, %	Точность пользователя, %
Дерново-подзолистые несмытые	71,7 (от 0 до 100 %)	71,3 (от 0 до 100 %)
Дерново-подзолистые слабосмытые	78,4 (от 40 до 100 %)	57,9 (от 25 до 91 %)
Дерново-подзолистые среднесмытые	60,9 (от 0 до 100 %)	66,0 (от 0 до 100 %)
Дерново-подзолистые сильносмытые	19,1 (от 0 до 100 %)	63,5 (от 0 до 100 %)
Дерново-подзолистые намытые	45,4 (от 0 до 100 %)	65,6 (от 0 до 100 %)
Аллювиальные	95,4 (от 0 до 100 %)	85,7 (от 0 до 100 %)
Дерново-глеевые	43,7 (от 0 до 100 %)	56,6 (от 0 до 100 %)
Светло-серые лесные	0	0

Модель плохо предсказывает дерново-подзолистые сильносмытые и светло-серые лесные почвы (точность производителя около 19,1 % и 0 %, соответственно). Полученные результаты показали, что использование карт, построенных на основе модели множественной логистической регрессии, с более чем 50 %-ой долей вероятности позволяет в контурах всех почвенных выделов, кроме светло-серых лесных, обнаружить именно эти почвенные выделы (табл. 4.3).



А



Б

Рис.4.5 Почвенные карты, построенные на основе множественной логистической регрессии. А: каппа самовоспроизводимости модели = 0,44, Б: каппа самовоспроизводимости модели = 0,77

Таким образом, использование модели множественной логистической регрессии при пространственном предсказании распределения почвенных выделов позволяет достоверно провести почвенное картографирование исследуемого участка территории. Эта модель имеет ряд преимуществ: уверенное выделение ареалов распространения дерново-подзолистых, дерново-глеевых и аллювиальных почв, слабая зависимость качества картографического материала от каппы самовоспроизводимости модели при проведении классификации.

4.3 Моделирование на основе деревьев решений С5

Общая точность при моделировании пространственного распределения почв на основе деревьев решений С5 составляет от 35 до 75 % при средней общей точности 55,7 %, а каппа от 0,04 до 0,64 при среднем каппа, равном 0,32 (см. приложение 1,2).

Визуальное сравнение карт, построенных на основе моделей деревьев решений с максимальной и средней кappой, показало их существенное различие и слабую устойчивость алгоритма, так как отнесение к тому или иному поченному выделу одной и той же точки растра для двух этих моделей может быть различное в связи с тем, что при каждой новой итерации построения модели формируются различные классификационные правила (рис. 4.6, табл. 4.4). Анализ долей площадей, занятых почвенными группами по картам, построенным на основе моделей со средним и максимальным значением кappы, также показал слабую устойчивость алгоритма. Модель со средней кappой не выделяла на территории дерново-глеевые почвы, в то время как модель с максимальной кappой предсказала их присутствие на 11,3 % территории. Доля же дерново-подзолистых несмытых почв по карте, созданной на основе модели со средней кappой практически в 2 раза больше, чем по карте, созданной на основе модели деревьев решений с максимальной кappой (табл. 4.4). Возможно, набор ковариат и наша база данных не дают четкого разделения этих почвенных выделов методом деревьев решений С5.

Моделирование на основе деревьев решений с использованием, созданной нами базы данных, не выделяет контура светло-серых лесных и дерново-подзолистых сильносмытых почв.

Таблица 4.4

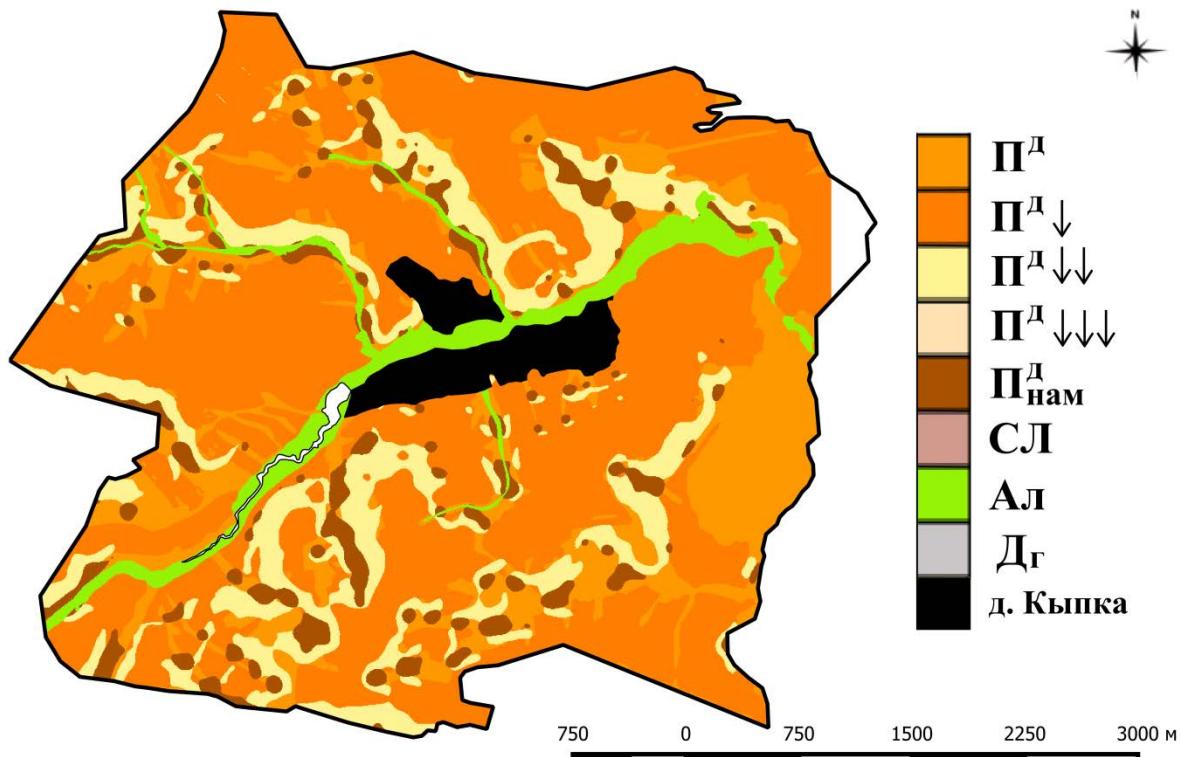
Доля площадей почвенных выделов (в %) для карт, построенных на основе метода деревьев решений С5: Карта А- на основе модели со средней каппой, $k=0,32$; Карта Б - на основе модели с максимальной каппой, $k=0,64$

Почвенные выделы	A	ДГ	СЛ	Π^d	$\Pi^d \downarrow$	$\Pi^d \downarrow\downarrow$	$\Pi^d \downarrow\downarrow\downarrow$	$\Pi_{нам}^d$
Карта А (рис. 4.6)	5,3	0	0	20,0	54,6	12,9	0	7,2
Карта Б (рис. 4.6)	5,3	11,3	0	11,9	51,5	13,0	0	7,1

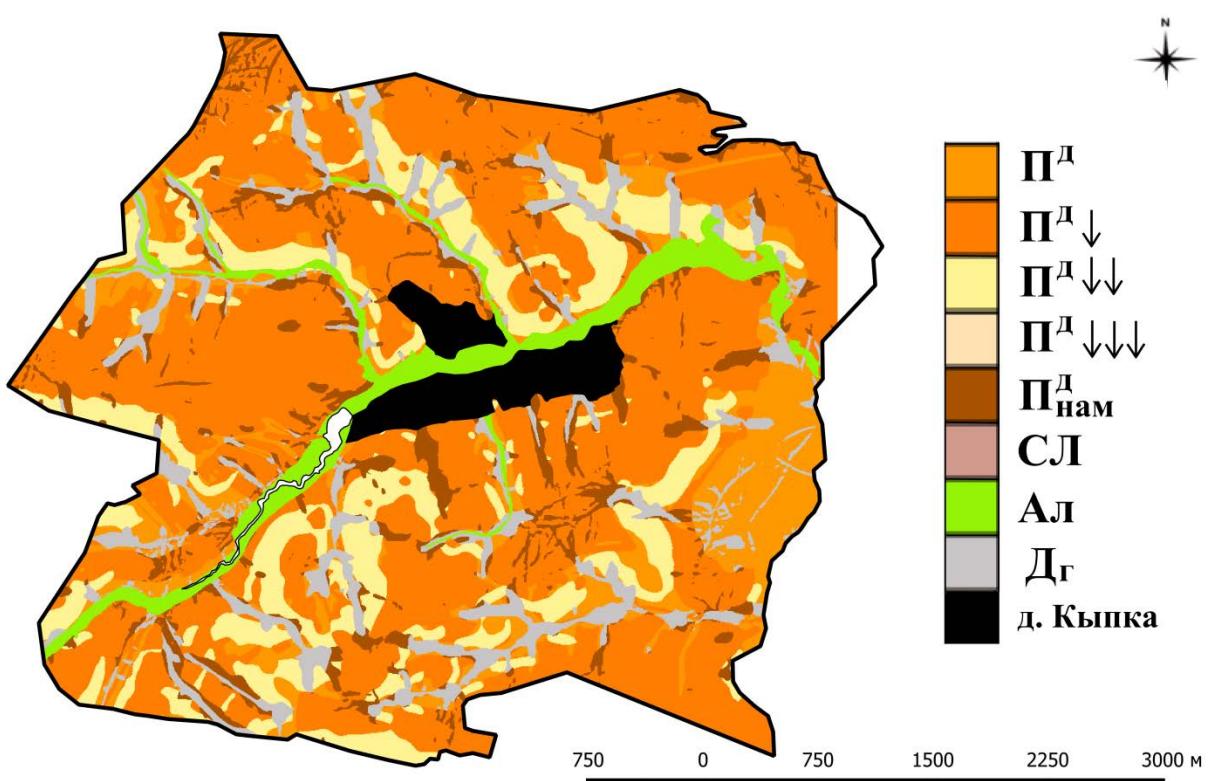
Таблица 4.5

Точность производителя и пользователя при моделировании на основе деревьев решений С5

Почвы	Точность производителя, %	Точность пользователя, %
Дерново-подзолистые несмытые	26,0 (от 0 до 100 %)	40,9 (от 0 до 100 %)
Дерново-подзолистые слабосмытые	86,2 (от 38 до 100 %)	54,4 (от 24 до 80 %)
Дерново-подзолистые среднесмытые	55,3 (от 0 до 100 %)	66,2 (от 17 до 100 %)
Дерново-подзолистые сильносмытые	0	0
Дерново-подзолистые намытые	5,2 (от 0 до 100 %)	16,1 (от 0 до 100 %)
Аллювиальные	97,3 (от 0 до 100 %)	87,3 (от 0 до 100 %)
Дерново-глеевые	8,2 (от 0 до 100 %)	28,4 (от 0 до 100 %)
Светло-серые лесные	0	0



А



Б

Рис. 4.6 Почвенные карты, построенные на основе деревьев решений С5. А: каппа самовоспроизводимости модели = 0,32, Б: каппа самовоспроизводимости модели = 0,64

Деревья решений с наибольшей точностью предсказывают дерново-подзолистые слабосмытые, среднесмытые и аллювиальные почвы (точность производителя 86,2 %, 55,3 % и 97,3 %, соответственно) (табл.4.5). Моделирование на основе деревьев решений C5 плохо предсказывает дерново-подзолистые сильносмытые и намытые почвы, а также светло-серые лесные и дерново-глеевые почвы. Полученные результаты также показали, что модель – деревья решений C5 хорошо отделяет от других классов следующие почвенные выделы: дерново-подзолистые слабосмытые, среднесмытые и аллювиальные почвы (точность пользователя 54,4 %, 66,2 %, и 87,3 %, соответственно) (табл.4.5). Поэтому пользователь карт, построенных на основе деревьев решений, с высокой долей вероятности в контурах дерново-подзолистых слабосмытых, среднесмытых и аллювиальных почв встретит именно эти почвы.

Таким образом, использование модели деревьев решений C5 при пространственном предсказании распределения почвенных выделов помогает выявить лишь некоторые закономерности распределения почвенных выделов на территории исследования в связи со слабой устойчивостью алгоритма. Однако эта модель имеет также ряд преимуществ: высокая точность производителя при предсказании дерново-подзолистых слабосмытых, среднесмытых и аллювиальных почв и высокая точность пользователя для дерново-подзолистых среднесмытых почв.

4.4 Моделирование на основе метода случайного леса

Общая точность при моделировании пространственного распределения почв на основе метода случайного леса составляет от 35 до 85 % при средней общей точности 59,3 %, а каппа от -0,03 до 0,75 при среднем каппа равном 0,40 (см. приложение 1,2).

Визуальное сравнение карт, построенных на основе метода случайного леса с максимальной и средней каппой, показало их небольшое различие: карта, построенная на основе модели с максимальной каппой, является более

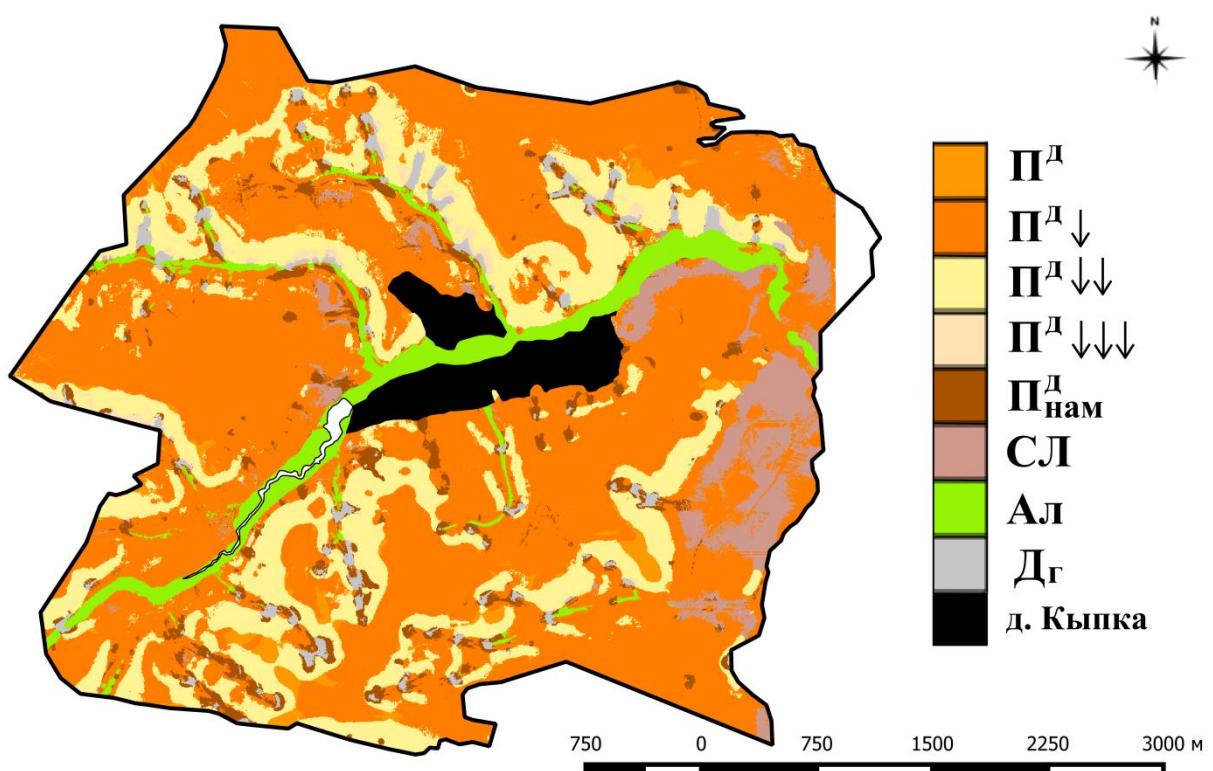
дискретной, чем карта, построенная на основе модели со средней каппой. Встречаются участки, где на общем фоне дерново-подзолистых слабосмытых почв встречаются пиксели намытых, несмытых и дерново-глеевых почв (рис. 4.7). Визуальная интерпретация карт, построенных на основе метода случного леса, показала также, что контура светло-серых лесных почв выделяются этой моделью в местах, где выделялись нами при проведении полевых обследований. В то же время, контура дерново-глеевых почв определяются этой моделью неверно (рис. 4.7). Анализ долей площадей, занятых почвенными группами по картам, построенным на основе моделей со средним и максимальным значением кппы, также показал хорошую устойчивость алгоритма. Были выявлены незначительные различия (не более 3 %) между картами, построенными на основе моделей с максимальной и средней кппой самовоспроизводимости (табл. 4.6).

Таблица 4.6

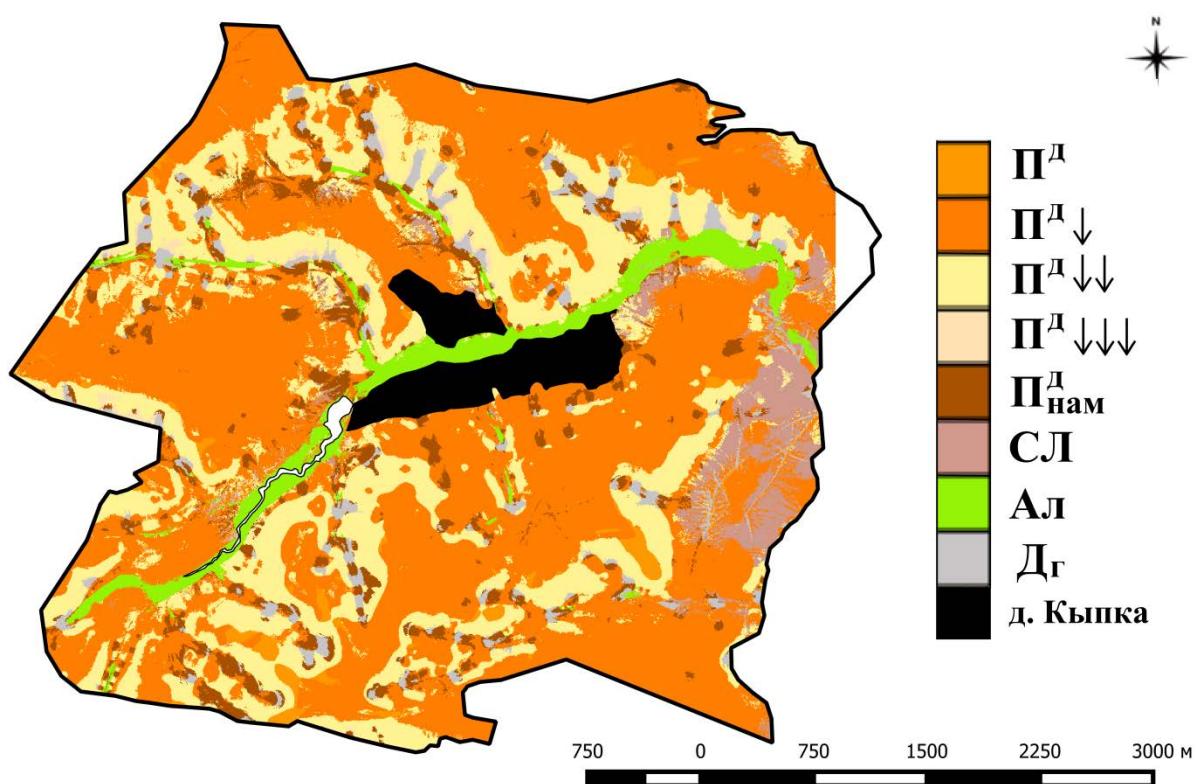
Доля площадей почвенных выделов (в %) для карт, построенных на основе метода случного леса: Карта А- на основе модели со средней кппой, $k=0,40$; Карта Б - на основе модели с максимальной кппой, $k=0,75$

Почвенные выделы	A	ДГ	СЛ	Π^d	$\Pi^d \downarrow$	$\Pi^d \downarrow\downarrow$	$\Pi^d \downarrow\downarrow\downarrow$	$\Pi_{нам}^d$
Карта А (рис. 4.7)	6,1	2,6	5,1	2,1	59,7	19,3	0,9	4,2
Карта Б (рис. 4.7)	4,8	4,0	3,4	2,0	56,4	22,7	0,8	5,9

Метод случного леса с наибольшей точностью предсказывает дерново-подзолистые слабо- и среднесмытые и аллювиальные почвы (точность производителя от 67,1 до 91,2 %) (табл. 4.7). Плохо предсказываются дерново-подзолистые сильносмытые и светло-серые лесные почвы. Как и предыдущие модели, пользователь карт, построенных на основе метода случного леса, с высокой долей вероятности в контурах дерново-подзолистых слабосмытых, среднесмытых и аллювиальных почв встретит именно эти почвы (точность пользователя 58,8 %, 63,9 %, и 87,6 %, соответственно) (табл. 4.7).



А



Б

Рис.4.7 Почвенная карта, построенная на основе метода случайного леса. А: каппа самовоспроизводимости модели = 0,40, Б: каппа самовоспроизводимости модели = 0,75

В 82,7 % случаев в контурах дерново-подзолистых несмытых почв также будут обнаружены эти почвы. Это свидетельствует о том, что модель хорошо распознает и отделяет перечисленные выше почвенные выделы по созданной нами базе данных.

Таблица 4.7

Точность производителя и пользователя при моделировании на основе метода случайного леса

Почвы	Точность производителя, %	Точность пользователя, %
Дерново-подзолистые несмытые	43,7 (от 0 до 100 %)	82,7 (от 0 до 100 %)
Дерново-подзолистые слабосмытые	77,6 (от 34 до 100 %)	58,8 (от 20 до 100 %)
Дерново-подзолистые среднесмытые	67,1 (от 20 до 100 %)	63,9 (от 17 до 100 %)
Дерново-подзолистые сильносмытые	0	0
Дерново-подзолистые намытые	23,5 (от 0 до 100 %)	38,0 (от 0 до 100 %)
Аллювиальные	91,2 (от 0 до 100 %)	87,6 (от 0 до 100 %)
Дерново-глеевые	53,4 (от 0 до 100 %)	56,7 (от 0 до 100 %)
Светло-серые лесные	0	0

Таким образом, использование метода случайного леса при пространственном предсказании распределения почвенных выделов помогает достоверно провести почвенное картографирование исследуемого участка территории. Преимуществами этой модели являются: правильное выделение ареалов распространения светло-серых лесных почв, высокие точности производителя при картографировании дерново-подзолистых почв разной степени смытости, слабая зависимость качества картографического материала от каппы самовоспроизводимости модели при проведении классификации.

4.5 Моделирование на основе регрессионных деревьев

Общая точность при моделировании пространственного распределения почв на основе регрессионных деревьев составляет от 30 до 70 % при средней общей точности 48,5 %, а каппа от -0,05 до 0,52 при среднем каппа равном 0,26 (см. приложение 1,2).

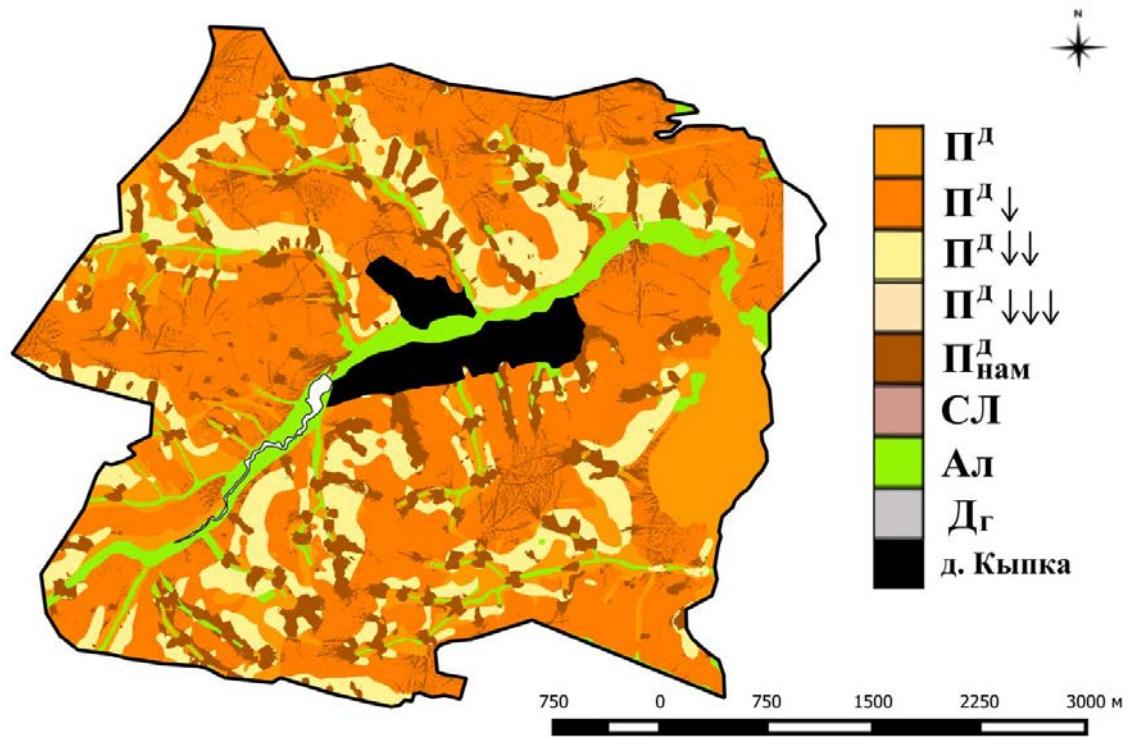
Картографирование на основе моделей регрессионных деревьев со средней и максимальной каппой самовоспроизводимости моделей показало неустойчивость этого метода цифрового почвенного картографирования. Визуальное сравнение карт, построенных на основе регрессионных деревьев с максимальной и средней каппой, и анализ таблиц показали их различия: увеличение доли дерново-подзолистых среднесмытых и аллювиальных почв по карте с максимальной каппой более чем в 2 раза, в то время как площади, занимаемые дерново-подзолистыми намытыми почвами, сократились (рис. 4.8, табл. 4.8). Отмечено, что доля площадей дерново-подзолистых несмытых и намытых почв для карты, построенной на основе модели со средней каппой самовоспроизводимости более чем в 2 раза больше, чем для карты, созданной путем предсказания на основе модели с максимальной каппой. Однако сравнение карты, построенной традиционными методами и карты, созданной на основе модели регрессионных деревьев со средней каппой

Таблица 4.8

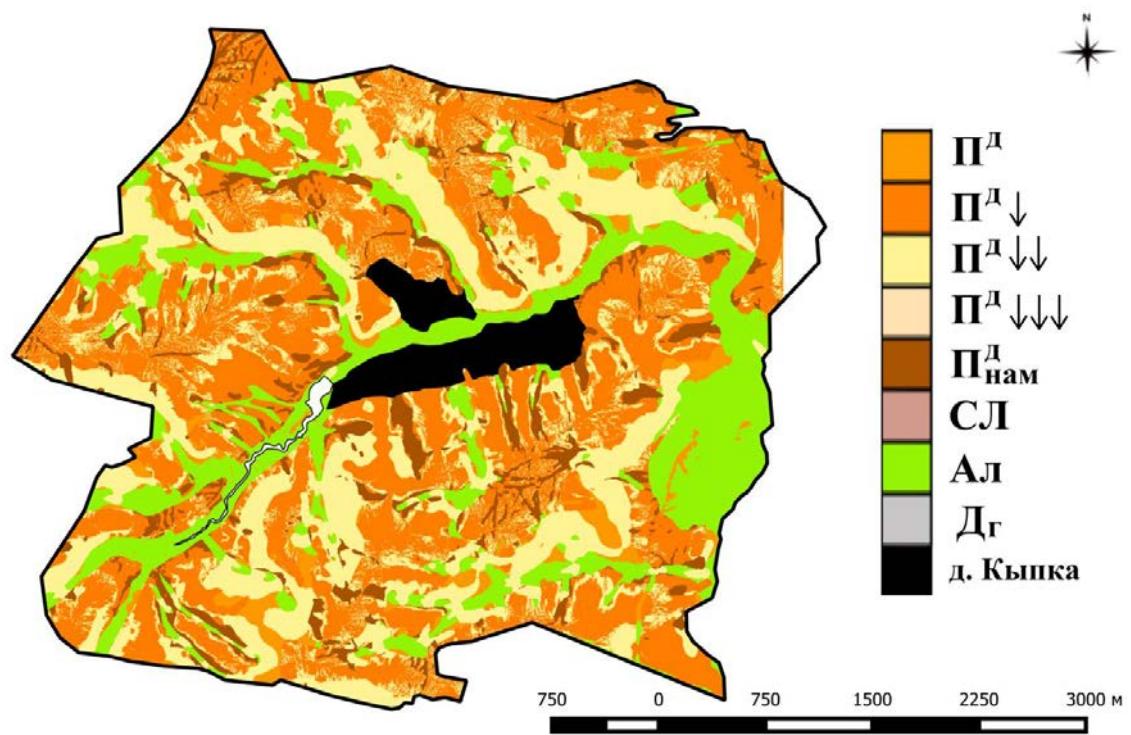
Доля площадей почвенных выделов (в %) для карт, построенных на основе регрессионных деревьев: Карта А- на основе модели со средней каппой, $k=0,26$; Карта Б - на основе модели с максимальной каппой, $k=0,52$

Почвенные выделы	A	ДГ	СЛ	П ^д	П ^д ↓	П ^д ↓↓	П ^д ↓↓↓	П ^д _{нам}
Карта А (рис. 4.8)	8,2	0	0	16,2	46,2	14,8	0	14,6
Карта Б (рис. 4.8)	18,7	0	0	5,5	40,6	28,1	0	7,1

самовоспроизводимости показало, что рисунок пространственного распределения дерново-подзолистых намытых почв по карте, построенной на



А



Б

Рис.4.8 Почвенная карта, построенная на основе деревьев регрессии. А: каппа самовоспроизводимости модели = 0,26, Б: каппа самовоспроизводимости модели = 0,52

основе регрессионных деревьев, наиболее схож с рисунком этого же почвенной группы согласно карте, построенной нами традиционными методами с элементами ГИС-анализа (рис. 3.6).

Регрессионные деревья с наибольшей точностью предсказывают дерново-подзолистые несмытые, слабо- и среднесмытые почвы (точность производителя в этих случаях составляет 68,3 %, 68,8 % и 54,7 %, соответственно). Все остальные почвенные выделы плохо предсказываются данной моделью (табл. 4.9).

Таблица 4.9

Точность производителя и пользователя при моделировании на основе деревьев регрессии

Почвы	Точность производителя, %	Точность пользователя, %
Дерново-подзолистые несмытые	68,3 (от 0 до 100 %)	39,6 (от 0 до 100 %)
Дерново-подзолистые слабосмытые	68,8 (от 30 до 100 %)	54,8 (от 17 до 100 %)
Дерново-подзолистые среднесмытые	54,7 (от 0 до 100 %)	60,8 (от 0 до 100 %)
Дерново-подзолистые сильносмытые	0	0
Дерново-подзолистые намытые	18,6 (от 0 до 100 %)	16,2 (от 0 до 100 %)
Аллювиальные	11,8 (от 0 до 100 %)	23,9 (от 0 до 100 %)
Дерново-глеевые	0	0
Светло-серые лесные	0	0

Деревьями регрессии хорошо отделяются дерново-подзолистые слабо- и среднесмытые почвы. В среднем более 50 % всех точек, которые были отнесены к дерново-подзолистым слабо- и среднесмытым почвам, определены верно, и пользователь полученной карты более чем с 54 %-ой вероятностью в

ареалах дерново-подзолистых слабосмытых почв встретит именно эти почвы, а в ареалах дерново-подзолистых среднесмытых почв с 60 %-ой вероятностью встретит именно их (табл. 4.9).

Таким образом, использование деревьев регрессий при пространственном предсказании распределения почвенных выделов плохо помогает выявить закономерности распределения этих выделов на территории исследования в связи со слабой устойчивостью алгоритма, а, соответственно, и сильными различиями в результатах картографирования. Однако эта модель имеет одно преимущество: правильное выделение контуров намытых дерново-подзолистых почв.

4.6 Моделирование на основе метода опорных векторов

Общая точность при моделировании пространственного распределения почв на основе метода опорных векторов составляет от 30 до 85 % при средней общей точности 59,4 %, а каппа от -0,03 до 0,79 при среднем каппа равном 0,40 (см. приложение 1,2).

Таблица 4.10

Доля площадей почвенных выделов (в %) для карт, построенных на основе метода опорных векторов: Карта 1 - на основе модели со средней каппой, $k=0,40$; Карта 2 - на основе модели с максимальной каппой, $k=0,79$

Почвенные выделы	A	Дг	СЛ	Π^d	$\Pi^d \downarrow$	$\Pi^d \downarrow\downarrow$	$\Pi^d \downarrow\downarrow\downarrow$	$\Pi_{нам}^d$
Карта А (рис. 4.9)	4,1	8,1	3,1	5,3	59,4	16,8	0,9	2,4
Карта Б (рис. 4.9)	4,3	5,8	1,6	2,3	65,1	15,3	0,6	5,1

Визуально отмечается высокая степень схожести между картами, построенными на основе метода опорных векторов с максимальной и средней каппой самовоспроизводимости. В то же время эти карты схожи и с картами, построенными на основе множественной логистической регрессии (рис. 4.9). Сравнение площадей почв, занимаемых различными группами на картах, построенных на основе моделей со средней и максимальной каппой, показало

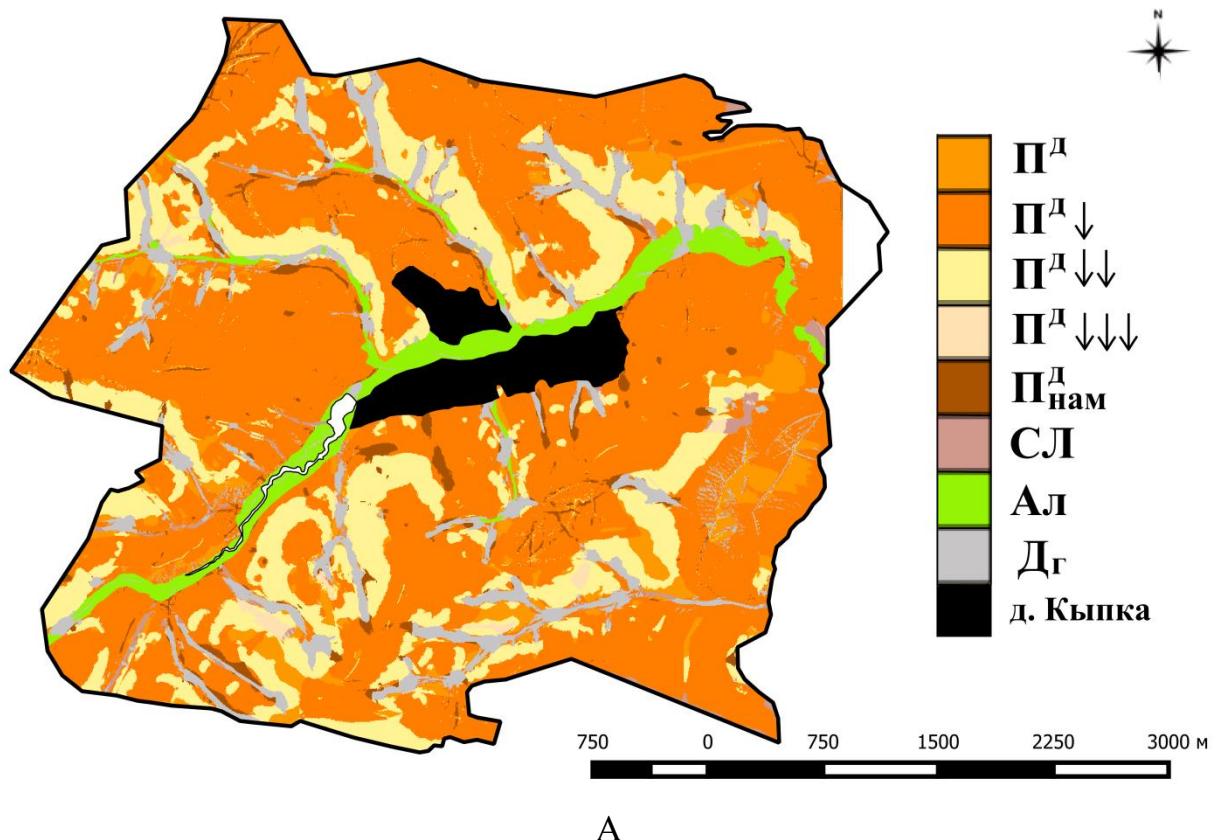
также их небольшое различие. Отмечено небольшое увеличение доли дерново-глеевых, серых лесных и дерново-подзолистых несмытых и среднесмытых почв на карте, построенной на основе модели со средней каппой, при одновременном незначительном уменьшении процентного содержания дерново-подзолистых слабосмытых почв (табл. 4.10).

Таблица 4. 11

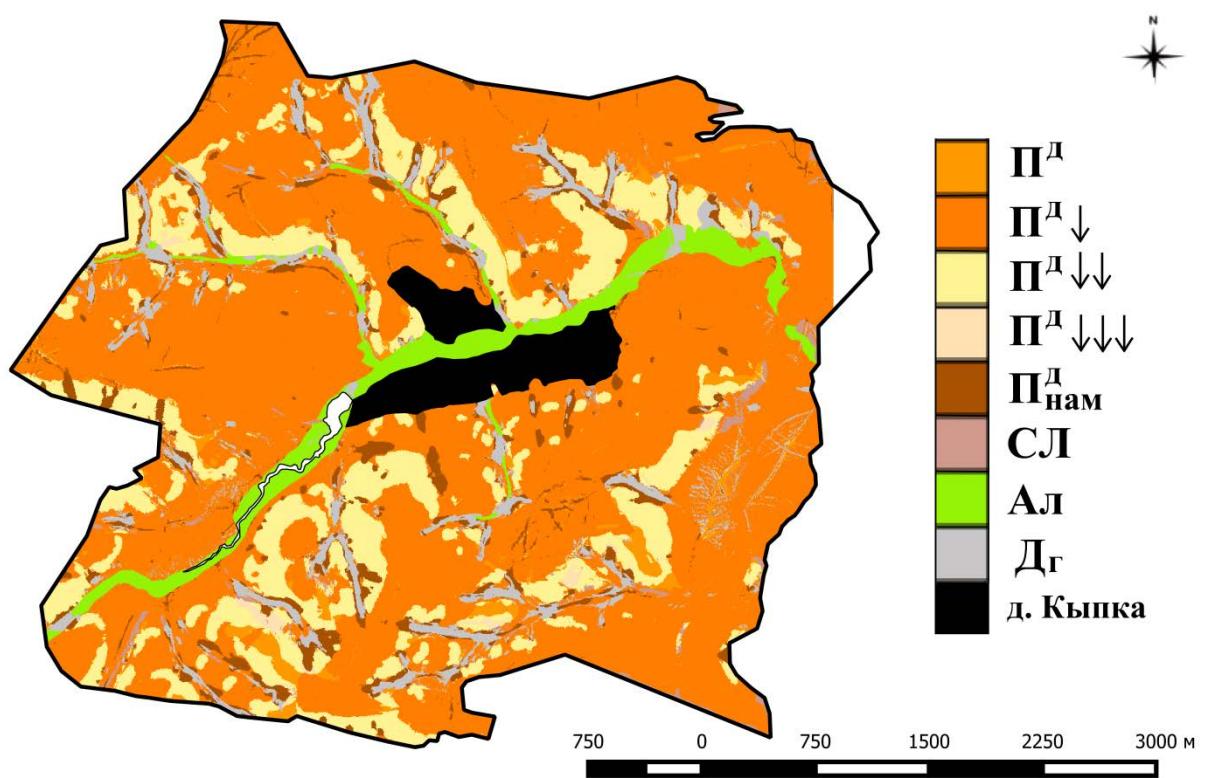
Точность производителя и пользователя при моделировании на основе метода опорных векторов

Почвы	Точность производителя, %	Точность пользователя, %
Дерново-подзолистые несмытые	46,2 (от 0 до 100 %)	69,9 (от 0 до 100 %)
Дерново-подзолистые слабосмытые	75,9 (от 17 до 100 %)	56,2 (от 15 до 90 %)
Дерново-подзолистые среднесмытые	57,6 (от 0 до 100 %)	61,9 (от 0 до 100 %)
Дерново-подзолистые сильносмытые	9,1 (от 0 до 100 %)	56,3 (от 0 до 100 %)
Дерново-подзолистые намытые	42,1 (от 0 до 100 %)	64,3 (от 0 до 100 %)
Аллювиальные	100	86,4 (от 0 до 100 %)
Дерново-глеевые	63,3 (от 0 до 100 %)	57,2 (от 0 до 100 %)
Светло-серые лесные	0	0

Метод опорных векторов дает наибольшую точность производителя для аллювиальных, дерново-подзолистых слабосмытых и среднесмытых почв (100 %, 75,9 % и 57,6 %, соответственно). Отмечена наибольшая точность производителя среди других моделей для дерново-подзолистых намытых почв (42,1 %) (табл. 4.11).



А



Б

Рис.4.9 Почвенная карта, построенная на основе метода опорных векторов. А: каппа самовоспроизводимости модели = 0,40. Б: каппа самовоспроизводимости модели = 0,79

Более чем с 60 %-ой вероятностью в ареалах дерново-подзолистых несмытых, среднесмытых, намытых и аллювиальных почв пользователь карты встретит именно эти почвы (точность пользователя 69,9 %, 61,9 %, 64,3 % и 86,4 %, соответственно) (табл. 4.11). В то же время отмечается достаточно высокая точность при разделении дерново-подзолистых слабосмытых и дерново-глеевых почв от других почвенных выделов.

Таким образом, использование метода опорных векторов при пространственном предсказании распределения почвенных выделов помогает достаточно достоверно провести почвенное картографирование исследуемого участка территории. Преимуществами этой модели являются: уверенное выделение ареалов распространения дерново-подзолистых, дерново-глеевых и аллювиальных почв, слабая зависимость качества картографического материала от каппы самовоспроизводимости модели при проведении классификации, наилучшее среди всех моделей точность пользователя и производителя для дерново-подзолистых намытых почв.

4.7 Сравнение и анализ результатов моделирования

4.7.1 Анализ схожести почвенных карт, построенных на основе одного и того же метода цифрового картографирования, но с разными коэффициентами каппа самовоспроизводимости моделей

В рамках исследования было проведено сравнение карт, построенных на основе одного и того же метода цифрового картографирования, но с двумя разными коэффициентами каппа самовоспроизводимости моделей: средней каппой и максимальной каппой. Сравнение проводилось по 200 проверочным, случайно брошенным точкам (рис. 2.14). В результате сравнения оказалось, что наиболее схожими между собой оказались карты, построенные на основе метода множественной логистической регрессии. Коэффициент каппа, равный 0,91, и общая точность, составляющая 95%, показывали почти полное совпадение двух карт (табл. 4.12).

Худшие результаты сравнения получились у карт, созданных на основе регрессионных деревьев. Коэффициент каппа показал лишь существенной совпадение этих карт и составил 0,63, а общая точность составила 74% (табл. 4.12). Сравнение трех других почвенных карт, созданных на основе деревьев решений C5, метода случайного леса и опорных векторов, показали также существенной совпадение для каждого из алгоритмов. Коэффициент каппа для этих случаев составил от 0,70 до 0,74, а общая точность составила около 83%, что говорит нам о том, что около 166 точек на этих картах оказались единообразно классифицированными каждым из алгоритмов с моделями, характеризующимися средней и максимальной каппой самовоспроизводимости (табл. 4.12).

Таблица 4.12

Сравнение почвенных карт, построенных на основе одного и того же метода цифрового картографирования, между собой

Почвенные карты	Каппа, характеризующая самовоспроизводимость цифровой модели	Каппа	Общая точность, %
Множественная логистическая регрессия	средняя $k= 0,44$	0,91	95
	максимальная $k= 0,77$		
Деревья решений C5	средняя $k= 0,32$	0,73	83
	максимальная $k= 0,64$		
Метод случайного леса	средняя $k= 0,40$	0,70	83
	максимальная $k= 0,75$		
Регрессионные деревья	средняя $k= 0,26$	0,63	74
	максимальная $k= 0,52$		
Метод опорных векторов	средняя $k= 0,40$	0,74	85
	максимальная $k= 0,79$		

Таким образом, модели множественной логистической регрессии, характеризующиеся различными коэффициентами каппа, будут разделять исследуемую территорию схожим образом, то есть уровень единообразия получаемых картографических материалов на основе этой модели будет высок.

В то же время, карты, построенные на основе моделей регрессионных деревьев с различными каппами самовоспроизводимости моделей, будут достаточно сильно отличаться друг от друга. Такие особенности картографирования разными цифровыми моделями, по-видимому, связаны с внутренними процессами формирования классификационных правил.

4.7.2 Визуальный анализ степени схожести полученных карт с картой, построенной традиционными методами

На основании визуального анализа наиболее схожей с почвенной картой, построенной традиционными методами, оказалась карта, построенная методом множественной логистической регрессии. Большое сходство также наблюдается у нее с картами, смоделированными на основе методов случайного леса и опорных векторов. Визуальное сравнение показало, что наименьшая степень схожести традиционной почвенной карты, построенной нами, наблюдается с картой, созданной на основе регрессионных деревьев.

4.7.3 Анализ общей точности и коэффициентов каппа при моделировании

В ходе проведения исследования для каждого алгоритма цифрового почвенного моделирования было создано по 200 моделей. 100 моделей обучались с использованием ковариат, взятых из карты 1984 г., и 100 моделей обучались без использования ковариат. Были построены ранжированные ряды для кпп (см. приложение 1, рис.4.10).

Наибольшая средняя кпп самовоспроизводимости модели 0,44 наблюдается в случае множественной логистической регрессии, в то время как наименьшая средняя кпп равняется 0,26 для модели регрессионных деревьев (см. приложение 1,2, рис. 4.11).

Наибольшая средняя общая точность моделирования равняется 61,1 % и наблюдается в случае множественной логистической регрессии, в то время как наименьшая средняя общая точность моделирования равняется 48,5 % для модели регрессионных деревьев (см. приложение 1,2, рис. 4.12).

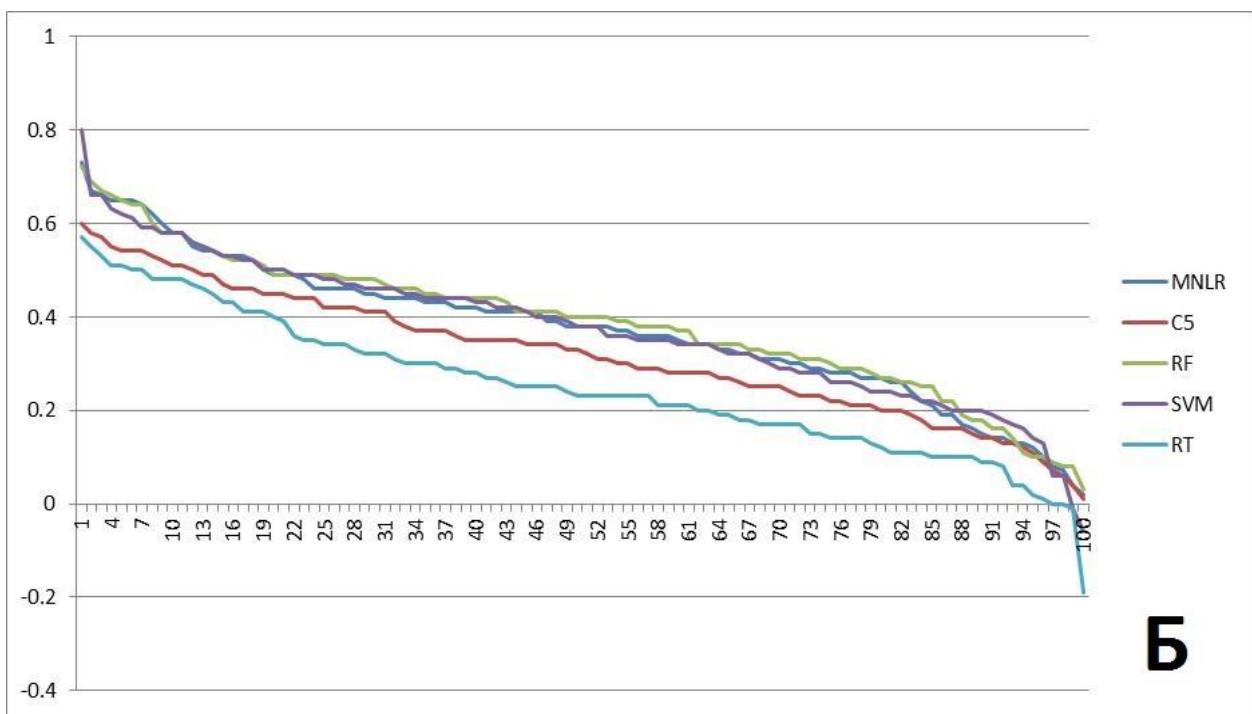
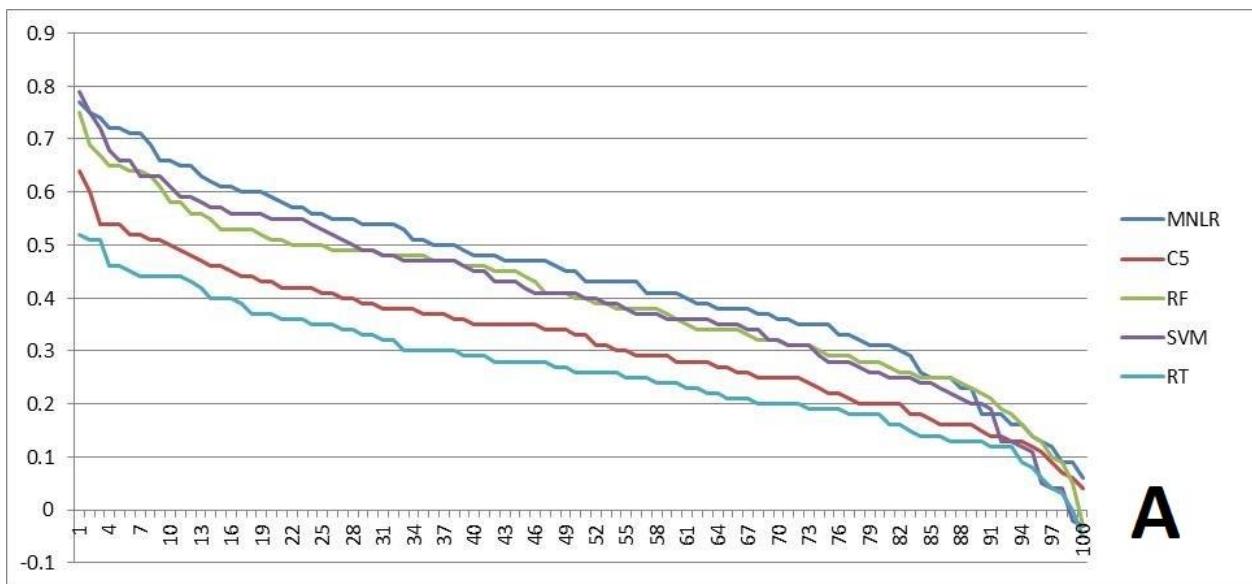


Рис. 4.10 Ранжированный ряд для коэффициента каппа при моделировании без использования карты 1984 г. (А) и с использованием карты 1984 г. (Б).
*Примечания: MNLR - множественная логистическая регрессия, C5 – дерево решений C5, RF – метод случайного леса, SVM – метод опорных векторов, RT – регрессионные деревья

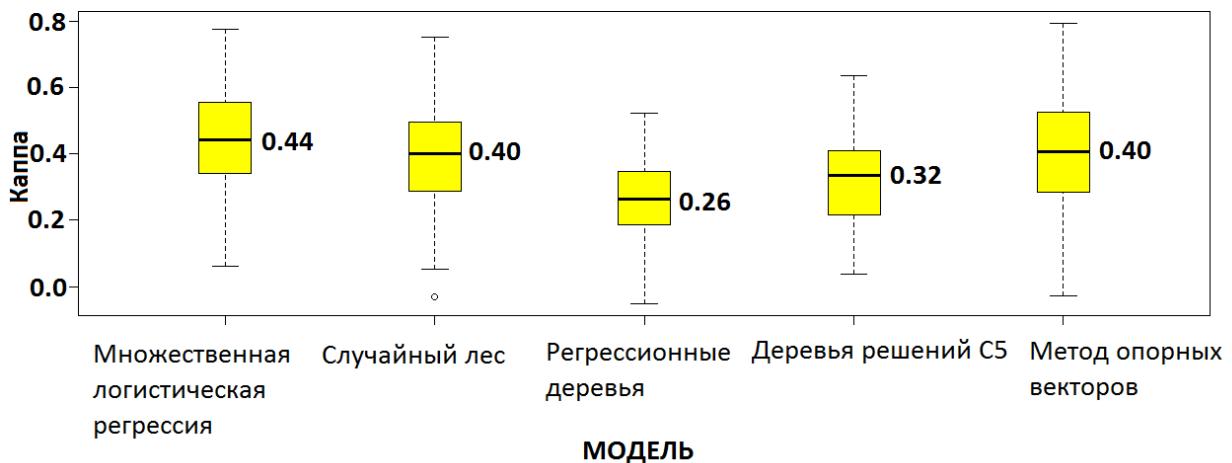


Рис.4.11 Медианы, доверительные интервалы и средние для каппы самовоспроизводимости моделей

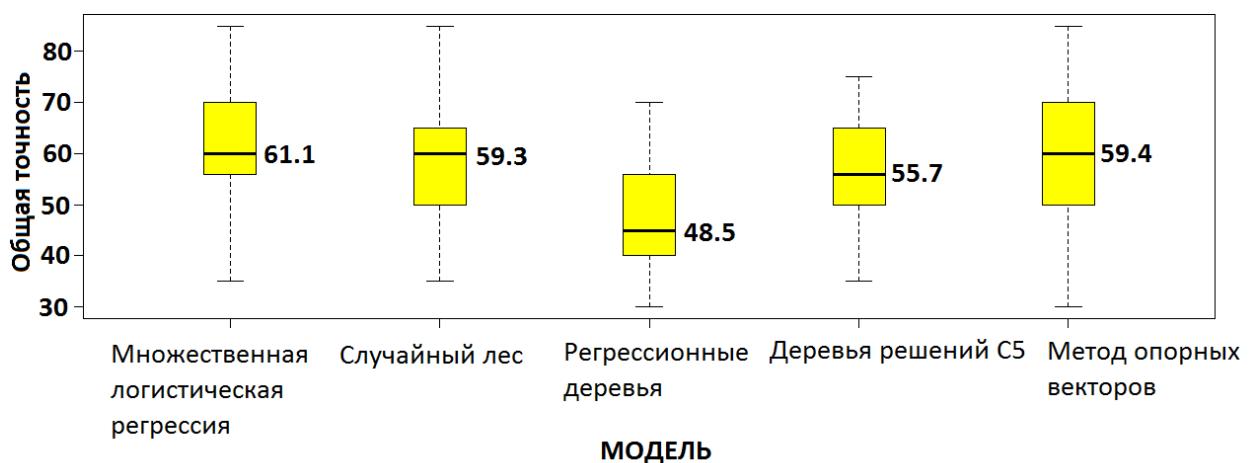


Рис.4.12 Медианы, доверительные интервалы и средние для общей точности моделей

Распределение коэффициента каппа и общей точности, а также средние значения для этих характеристик моделирования показали, что большие каппа самовоспроизводимости получаются при моделировании на основе множественной логистической регрессии (см. приложение 1,2, рис. 4.10,4.11,4.12). Распределение каппа при моделировании на основе метода случайного леса и опорных векторов имеют схожие формы кривых и средние значения, получаемые при моделировании этими методами также схожи (см. приложение 1,2, рис. 4.10,4.11,4.12). Из построенных графиков видно, что наиболее низкие результаты для коэффициентов каппа и общей точности

моделей характерны для метода регрессионных деревьев (см. приложение 1,2, рис. 4.10,4.11,4.12). Также отмечено, что при добавлении ковариат 1984 г. в процесс моделирования происходит снижение коэффициента каппа при моделировании на основе множественной логистической регрессии и значение этого коэффициента становится схожими с каппа самовоспроизводимости моделей метода случайного леса и опорных векторов (рис. 4.10.Б).

4.7.4 Анализ таблиц точности производителя и пользователя

Анализ таблиц точности производителя и пользователя при моделировании каждым из пяти методов, используя созданную нами базу данных и отобранные нами ковариаты для обучения моделей, показал следующие результаты (рис. 4.13, рис.4.14).

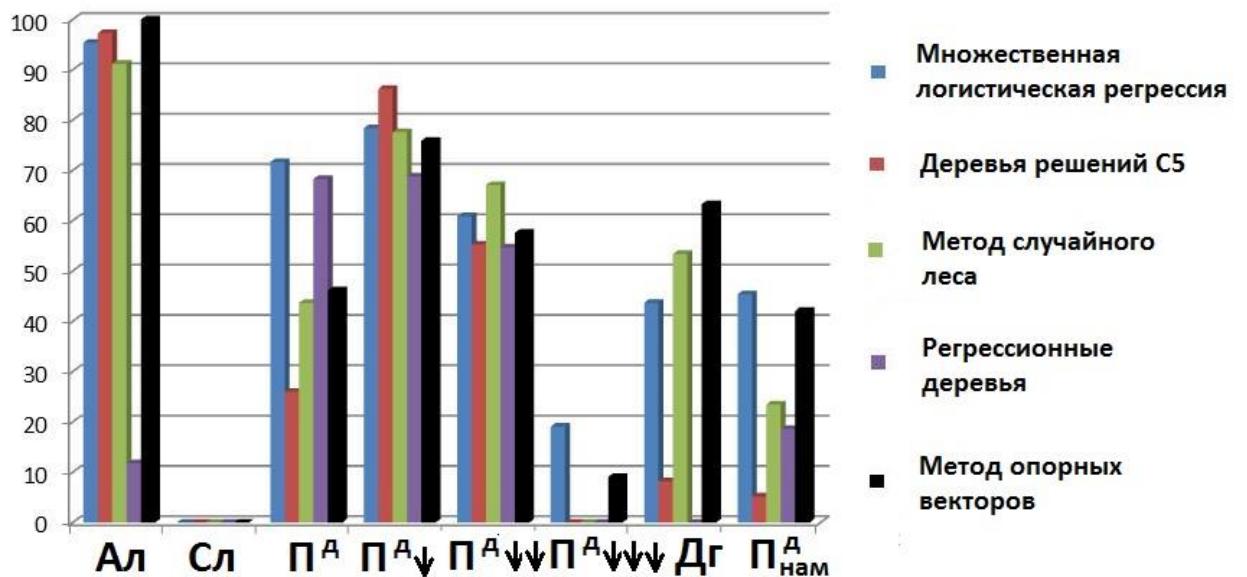


Рис.4.13 Точность производителя алгоритмов для каждой условной почвенной группы

Точность производителя показала, что дерново-подзолистые несмытые и сильносмытые почвы лучше всего выделялись алгоритмом множественной логистической регрессии (точность производителя составила 71,7 % и 19,1 %, соответственно), а дерново-подзолистые слабосмытые почвы - иерархически построенными моделями: деревьями решений C5 (точность производителя -

86,2 %), дерново-подзолистые среднесмытые почвы наилучшим образом предсказывает метод случайного леса (точность производителя 67,1 %), дерново-подзолистые намытые почвы лучше всего выделяются методами множественной логистической регрессией и опорных векторов (точности производителя 45,4 % и 42,1 %, соответственно), аллювиальные почвы отлично выделялись всеми методами (точность производителя более 90 %), кроме регрессионных деревьев, дерново-глеевые почвы лучшим образом предсказывает метод опорных векторов (точность производителя для этого алгоритма составляет 63,3 %) (рис. 4.13). В то же время светло-серые лесные почвы плохо предсказываются и отделяются от других почвенных выделов любым из этих методов (точность производителя равна 0 %).

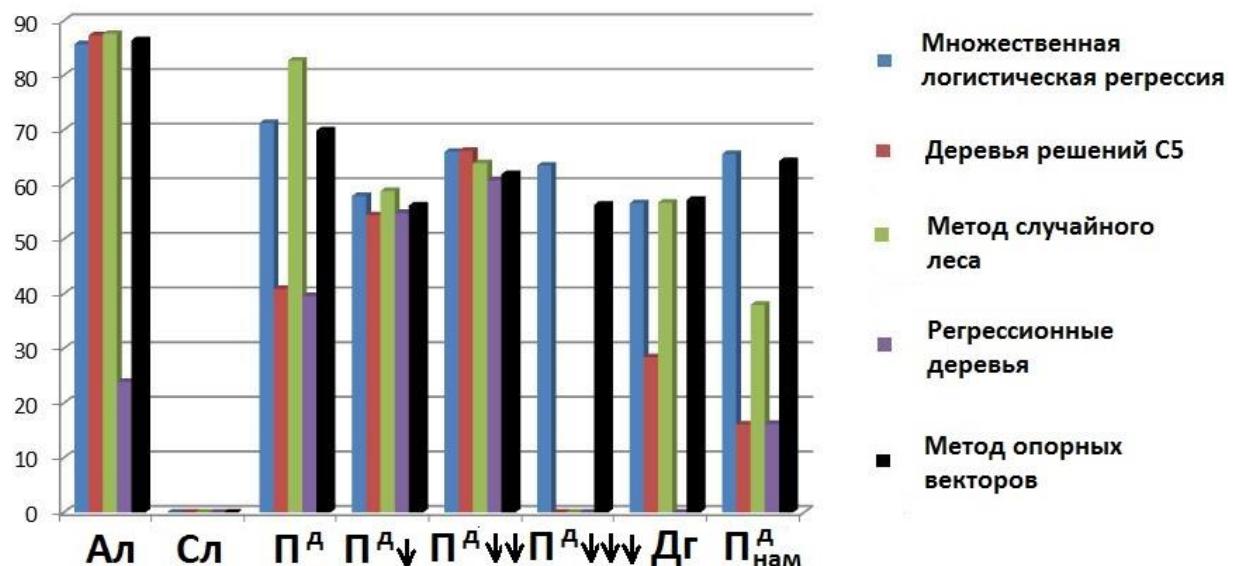


Рис.4.14 Точность пользователя алгоритмов для каждого условной почвенной группы

Точность пользователя показала, что дерново-подзолистые несмытые и намытые почвы лучше всего отделяются от других почвенных выделов алгоритмом множественной логистической регрессии (точность пользователя 49 % и 43,9 %, соответственно), все алгоритмы с более чем 50 %-ой вероятностью отделяют дерново-подзолистые слабосмытые, среднесмытые и аллювиальные почвы от других почвенных выделов (точность пользователя в этих случаях для всех моделей от 50,5 % до 68,5 %) (рис. 4.14). Дерново-

подзолистые сильносмытые почвы плохо отделяются от других почвенных выделов (лишь множественная логистическая регрессия и метод опорных векторов имеют точность пользователя больше нуля, а именно 18 и 12 % соответственно), точность пользователя для дерново-глеевых почв наибольшая в следующих алгоритмах: множественная логистическая регрессия и метод опорных векторов (по 29 %) (рис. 4.14).

Все модели с высокой степенью точности выделяют контура дерново-подзолистых слабосмытых и среднесмытых почв (табл. 4.3, 4.5, 4.7, 4.9, 4.11), так как дерново-подзолистые слабосмытые и среднесмытые почвы имеют наибольшее распространение на территории исследования и наибольшую долю точек в созданной нами базе данных, что увеличивает вероятность их правильного предсказания для того или иного пикселя растра. Немного хуже моделированию пространственного распределения подвергаются: аллювиальные почвы (точность производителя для таких моделей, как множественная логистическая регрессия, деревья решений, метод случайного леса и метод опорных векторов составила около 50 %) (табл. 4.3, 4.5, 4.7, 4.9, 4.11). Это связано с тем, что эти почвы преимущественно приурочены к таким элементам рельефа, как поймы рек и днища крупных оврагов, которые достаточно легко отделяются всеми использованными нами алгоритмами, кроме деревьев регрессии. Еще хуже предсказываются дерново-подзолистые намытые и дерново-глеевые почвы, что, по-нашему мнению, связано с низким пространственным разрешением топографической карты, которая легла в основу моделирования почвенного покрова территории (табл. 4.3, 4.5, 4.7, 4.9, 4.11). Хуже всего, в нашем случае, предсказываются дерново-подзолистые сильносмытые и светло-серые лесные почвы, так как их распространение на территории исследования достаточно невелико.

4.7.5 Анализ «устойчивых» пикселей на карте 2016 г. при моделировании

В ходе исследования были выделены пиксели и посчитана площадь «устойчивых» участков, которые классифицируются одинаковым образом в результате моделирования всеми методами. Сравнивалось пять карт со средним каппа. В результате была построена карта «устойчивых» участков исследуемой территории, пространственное распределение почвенного покрова на которых одинаково при моделировании на основе используемых алгоритмов (рис.4.15).

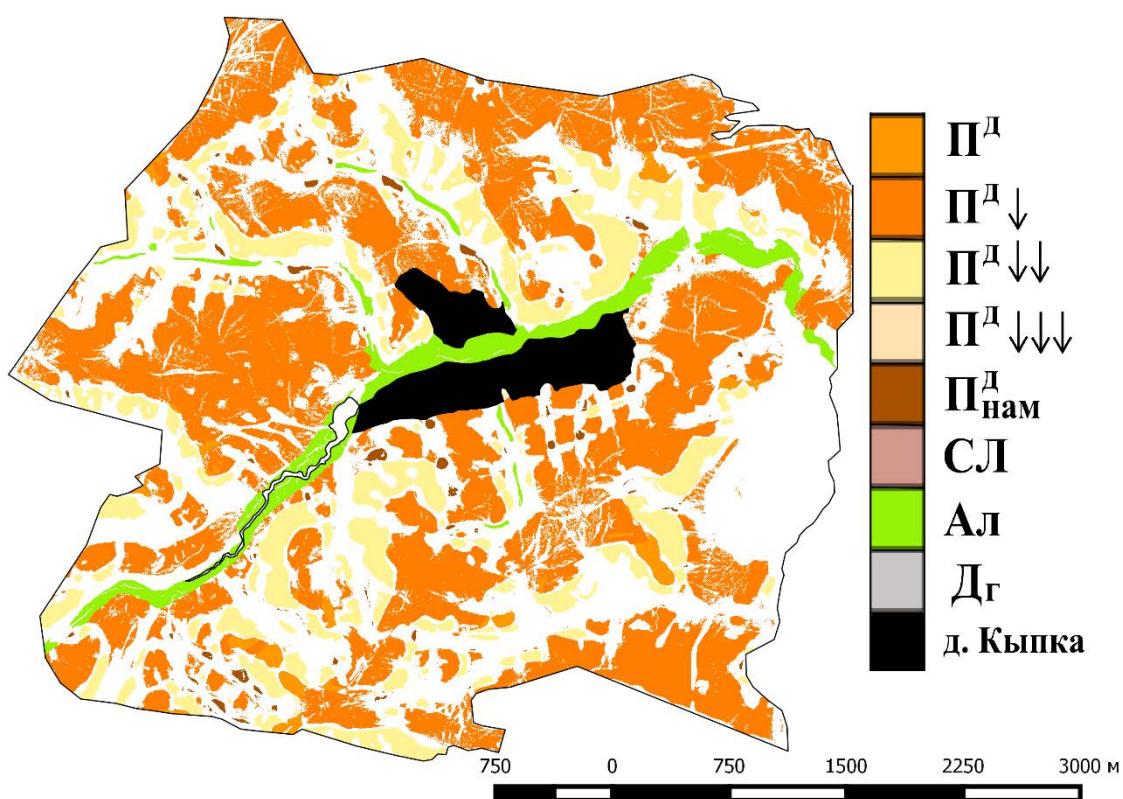


Рис.4.15 Карта распределения «устойчивых» участков при картографировании исследуемой территории

На построенной карте разными цветами отмечены пиксели, которые при моделировании на основе каждой из пяти моделей, были отнесены к одному и тому же классу объектов. Белым цветом выделены территории, которые различными методами классификации были отнесены к разным классам. Отмечено, что 54,82 % или 10,53 км² площади исследуемой территории

классифицируются всеми методами единообразно, из которых 4,09 % (или 0,79 км²) приходится на аллювиальные почвы, 1,51 % (или 0,29 км²) - на дерново-подзолистые несмытые почвы, 37,73 % (или 7,24 км²) - на дерново-подзолистые слабосмытые почвы, 10,94 % (или 2,1 км²) - на дерново-подзолистые среднесмытые почвы, 0,55 % (0,1 км²) - на дерново-подзолистые намытые почвы.

В процессе моделирования пространственного распределения почвенного покрова алгоритмы регрессионных деревьев и деревьев решений C5 не выделяли на исследуемой территории такие классы объектов как: светло-серые лесные, дерново-глеевые и дерново-подзолистые почвы. Три других алгоритма (множественная логистическая регрессия, метод случайного леса и метод опорных векторов) разделяли эту территорию на все восемь отобранных изначально условных почвенных групп. Поэтому были выделены пиксели и посчитана площадь «устойчивых» участков, которые классифицируются одинаковым образом в результате моделирования тремя этими методами. В результате была построена карта «устойчивых» участков исследуемой территории, пространственное распределение почвенного покрова на которых одинаково при моделировании на основе трех алгоритмов (рис.4.16).

Было выявлено, что 66,21 % или 12,71 км² площади исследуемой территории классифицируются множественной логистической регрессией, методом случайного леса и методом опорных векторов одинаково, из которых 4,09 % (или 0,79 км²) приходится на аллювиальные почвы, 0,69 % (или 0,13 км²) - на дерново-глеевые почвы, 0,21 % (или 0,04 км²) – на светло-серые лесные, 1,58 % (или 0,30 км²) - на дерново-подзолистые несмытые почвы, 43,93 % (или 8,43 км²) - на дерново-подзолистые слабосмытые почвы, 14,58 % (или 2,8 км²) - на дерново-подзолистые среднесмытые почвы, 0,01 % (или 0,002 км²) - на дерново-подзолистые сильносмытые почвы, 1,12 % (0,22 км²) - на дерново-подзолистые намытые почвы.

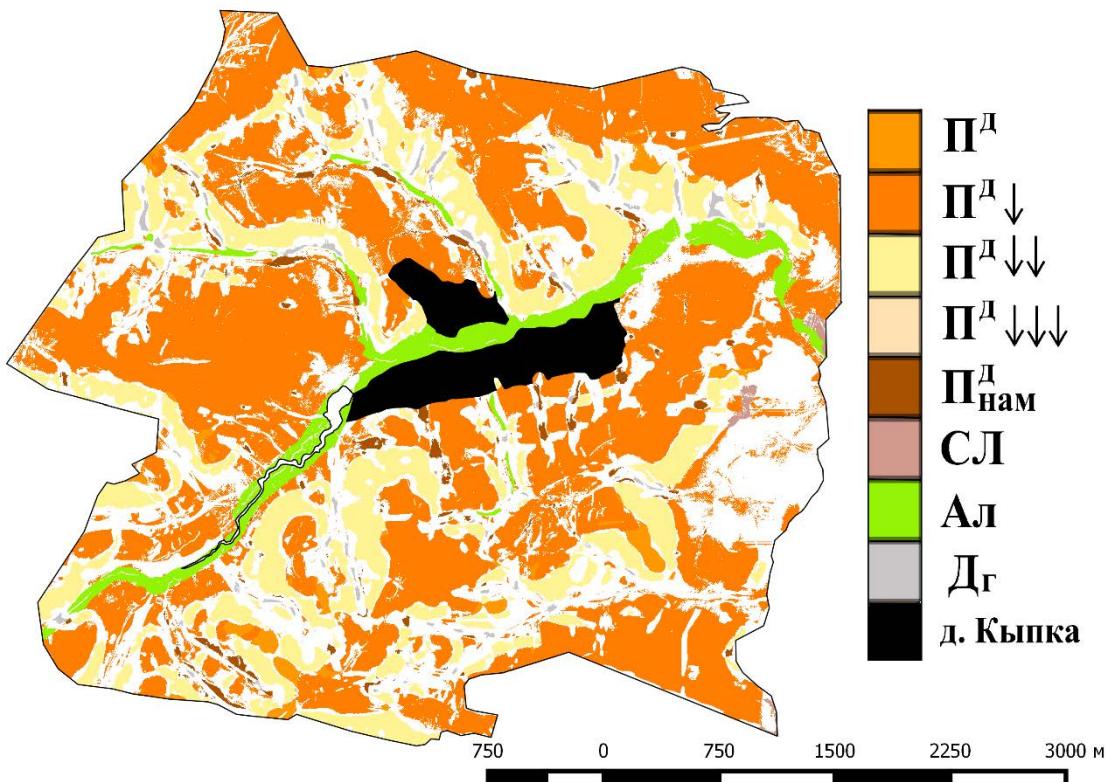


Рис.4.16 Карта распределения «устойчивых» участков при картографировании исследуемой территории (множественная логистическая регрессия, метод случайного леса и метод опорных векторов)

Была составлена сравнительная таблица, в которой для каждой условной почвенной группы были отражены занимаемые ими площади территории на карте 2016 г., построенной традиционным методом (рис.3.6), и площади, полученные по карте распределения «устойчивых» участков при картографировании исследуемой территории пятью методами и тремя методами (рис.3.6, 4.15, 4.16, табл.4.12).

Доля одинаково классифицируемых пикселей при моделировании на основе 3 методов (множественная логистическая регрессия, метод случайного леса и метод опорных векторов) незначительно увеличилась (примерно на 12 %) по сравнению с моделированием на основе 5 методов. Однако сравнение результатов моделирования на основе множественной логистической регрессии, метода случайного леса и метода опорных векторов показало, что эти три алгоритма однозначно выделяют некоторые участки, которые

принадлежат дерново-глеевым, серым лесным и дерново-подзолистым среднесмытым почвам.

Таблица 4.12
Доли «устойчивых» участков при моделировании для каждой условной почвенной группы

	Карта 2016 г. (рис.3.6)	Карта «устойчивых» участков для 5 методов (рис.4.15)	Карта «устойчивых» участков для 3 методов (рис.4.16)		
Почвенные выделы	Площадь, км ²	Площадь, км ²	Доля «устойчи- вых» участков, %	Площадь, км ²	Доля «устой- чивых» участков, %
A	0,96	0,79	82,3	0,79	82,3
Π ^д	0,95	0,29	30,5	0,30	31,6
Π ^д ↓	9,3	7,24	77,8	8,43	90,6
Π ^д ↓↓	1,7	2,1	123,5*	2,8	165*
Π ^д ↓↓↓	0,6	0	0	0,002	0,33
Π ^д _{нам}	2,95	0,1	3,4	0,22	7,46
СЛ	1,0	0	0	0,04	4,0
Дг	1,2	0	0	0,13	10,8

*В ходе моделирования произошла переклассификация некоторых пикселей и ошибочное их отнесение к классу дерново-подзолистых среднесмытых почв.

Больший процент «устойчивых» пикселей дерново-подзолистым слабосмытым почвам, что, по-видимому, связано с наибольшим распространением этих почв на территории исследования (табл.4.12). Также большая доля правильно определенных участков среди аллювиальных почв (82,3 %), среди дерново-подзолистых несмытых почв однозначно моделированием определяются лишь около трети всех территорий, занятых этими почвами по карте 2016 г. Классификация на основе всех алгоритмов приводит к увеличению площадей, занимаемых участками однозначно классифицированными как дерново-подзолистые среднесмытые почвы, по сравнению с картой 2016 г. Это связано с тем, что выборка, созданная нами для обучения моделей, способствует отнесению пикселей к этому классу.

Наименее устойчивыми (то есть наибольшая неопределенность при классификации используемыми моделями) оказались территории, которые располагались под лесной растительностью (рис.4.17, области под номерами 1,2,3), а также территории с депрессивными элементами рельефа (балки, овраги, ложбины) (рис.4.17, области под номерами 4,5,6). Наибольшая определенность при классификации наблюдалась на выровненных участках рельефа и территориях с углами наклона до 3^0 (рис.4.17).

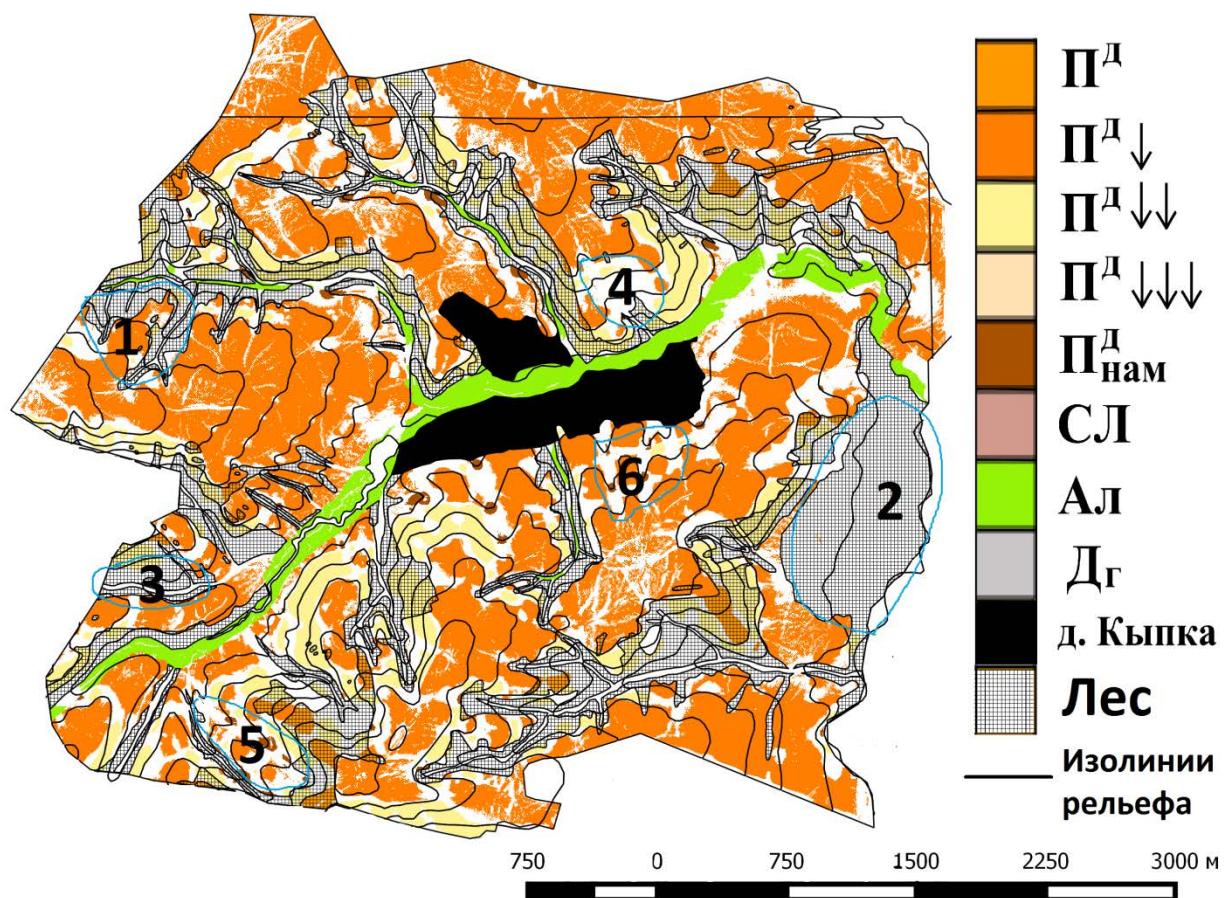


Рис.4.17 Карта распределения «устойчивых» участков при картографировании исследуемой территории с выделенными областями

Причиной такого разделения территории на «устойчивые» участки и «неустойчивые» участки, по нашему мнению, являются особенности созданной нами выборки, в которой большая часть точек обследования располагалась на выровненных участках рельефа, а точек на крутых склонах, в депрессиях рельефа и под лесными массивами было меньше. Поэтому

разные методы машинного моделирования по-разному разделяли эти территории на классы.

В ходе исследования было показано, что метод регрессионных деревьев путал практически все пиксели, отнесенные тремя методами к классу дерново-глеевых почв, с дерново-подзолистыми намытыми почвами. Алгоритм деревьев решений C5 40 % этих пикселей относил к аллювиальным почвам, а 60 % - к дерново-подзолистым намытым почвам.

4.7.6 Анализ положения «устойчивых» пикселей по карте 1984 г.

Были совмещены 2 карты: карта «устойчивых» пикселей и карта 1984 г. (рис. 4.18).

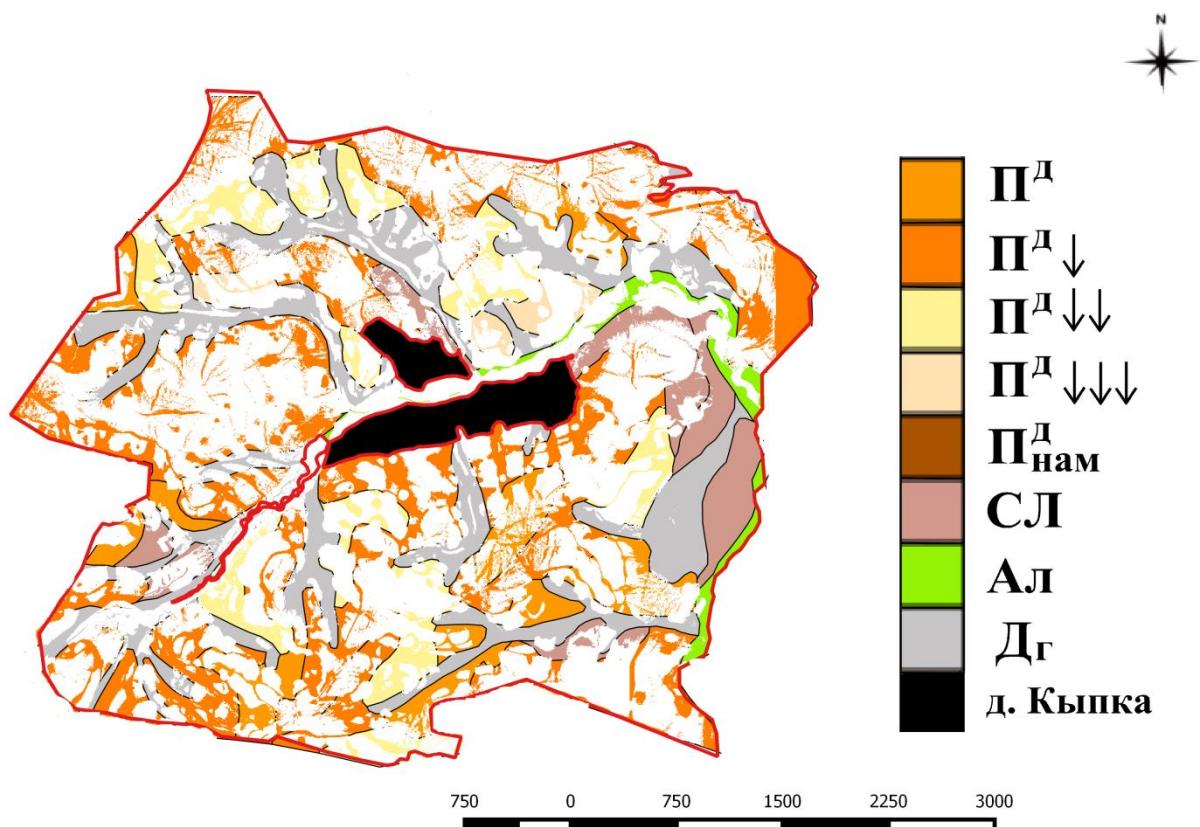


Рис.4.18 Карта распределения «устойчивых» участков при наложении на почвенную карту 1984 г. Белым цветом показаны «устойчивые» области.

На этой карте неустойчивые участки, как и на нашей карте, попадают в переувлажненные места и балки. Также неустойчивые участки соответствуют областям распространения серых лесных почв по карте 1984 г.

Карта 1984 г., которая создана для применения в процессе сельскохозяйственной деятельности, соответствует целям и задачам, поставленным перед ней разработчиками, так как области неустойчивости для этой карты в основном приурочены к балочным понижениям, лесным массивам и переувлажненным территориям, в то время как устойчивые пиксели распространены преимущественно на полях, на которых и ведется сельскохозяйственная деятельность (рис.4.18).

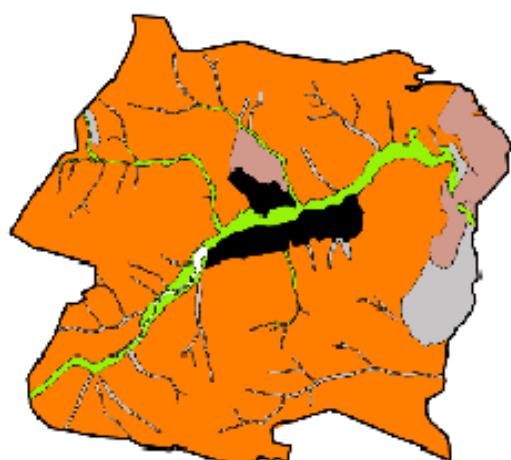
ГЛАВА 5. СРАВНЕНИЕ ЦИФРОВЫХ КАРТ С ПОЧВЕННОЙ КАРТОЙ 2016 г. И ПРОВЕРКА ПО СЛУЧАЙНОЙ НЕЗАВИСИМОЙ ВЫБОРКЕ

5.1 Сравнение карт почвенных типов, построенных различными методами, с картой 2016 г.

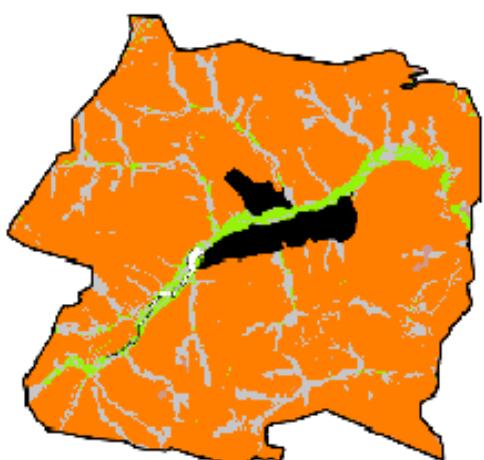
Сравнение карты почвенных типов со средней картой устойчивости модели с почвенной картой типов, полученной нами на основе традиционных методов, показало, что предсказание на основе алгоритма множественной логистической регрессии хорошо коррелирует с нашими представлениями о пространственном распределении почвенных типов на изучаемой территории: ареалы распространения дерново-подзолистых и аллювиальных почв полностью совпадают, ареалы распространения дерново-глеевых почв также совпадают с небольшимиискажениями. Контура светло-серых лесных почв на карте, построенной методом множественной логистической регрессии и созданной нами карте, абсолютно не совпадают (рис. 5.1).

Сравнение карт почвенных типов, построенных традиционными методами и на основе моделирования деревьями решений C5, показало, что предсказание пространственного распределения почвенных типов дерново-подзолистых и аллювиальных почв совпадает с нашими представлениями о распределении этих почв на изучаемой территории. Однако, контура светло-серых лесных и дерново-глеевых почв не выделяются на основе моделирования деревьями решений C5 (рис. 5.1).

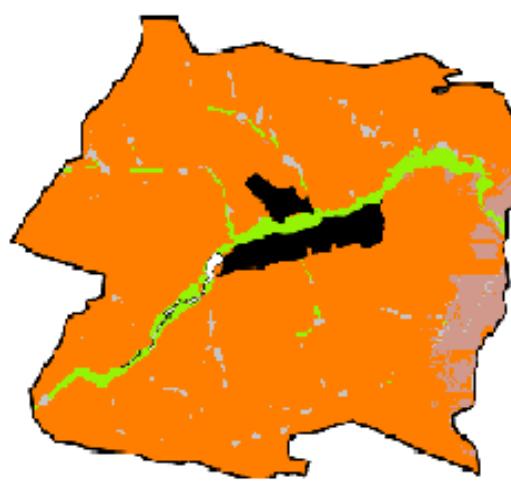
Сравнение карт почвенных типов, построенных традиционными методами и путем моделирования методом случайного леса, показало, что предсказание пространственного распределения почвенных типов дерново-подзолистых и аллювиальных почв совпадает с нашими представлениями о распределении этих почв на изучаемой территории. В некоторой степени схожи пространственные локации ареалов светло-серых лесных почв на созданной традиционными методами карте и карте, полученной путем моделирования методом случайного леса. Однако, контура дерново-глеевых почв плохо разделяются этим методом (рис. 5.1).



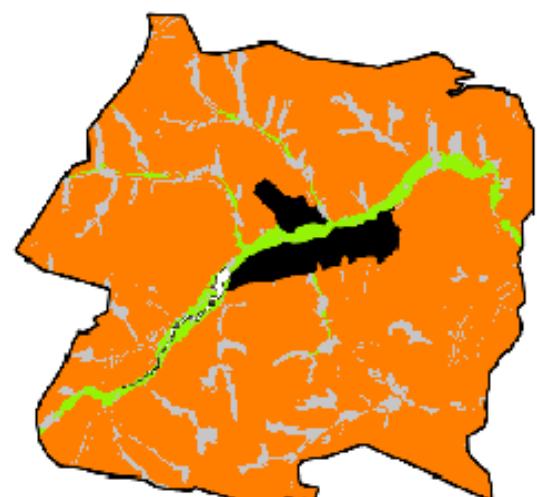
Традиционный метод с элементами
ГИС-анализа



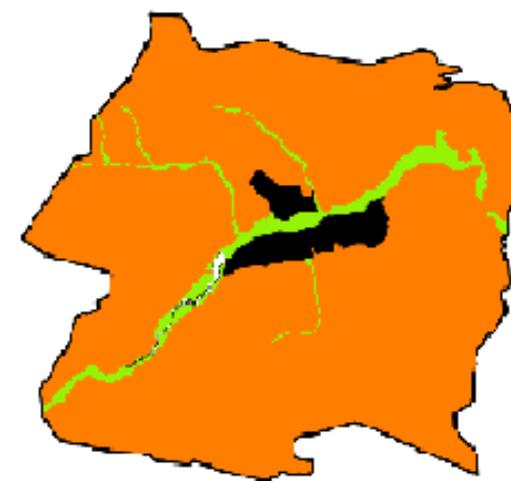
Множественная логистическая
регрессия



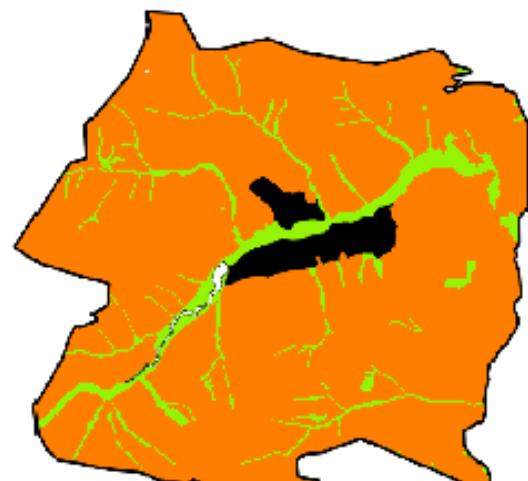
Метод случайного
леса



Метод опорных
векторов



Дерево решений С5



Регрессионные
деревья

Рис 5.1 Карты почвенных типов, построенных различными методами

Сравнение карты почвенных типов, построенной традиционными методами и карты, созданной на основе модели регрессионных деревьев со средней каппой устойчивости показало, что площади, занимаемые дерново-подзолистыми почвами по двум картам схожи. Также согласно карте, построенной на основе регрессионных деревьев, контура аллювиальных почв на ней хорошо коррелируют с контурами дерново-глеевых почв по построенной нами карте. Дерново-глеевые и светло-серые лесные почвы методом регрессионных деревьев не выделяются (рис. 5.1).

Сравнение карты почвенных типов, созданной на основе метода опорных векторов со средней каппой устойчивости, и карты, построенной нами, показало, что контура площадей, занимаемых дерново-подзолистыми и аллювиальными почвами по двум картам практически одинаковы. В тоже время выделы дерново-глеевых почв частично совпадают с построенной нами картой. Метод опорных векторов не позволяет выделять светло-серые лесные почвы (рис. 5.1).

Исследование показало, что пространственное распределение дерново-подзолистых и аллювиальных почв всеми алгоритмами в регионе исследования практически совпадает с нашими представлениями о распространении этого типа почв, отраженными в почвенной карте составленной традиционными методами. Методами опорных векторов и множественной логистической регрессии достаточно правильно выделяются дерново-глеевые почвы, которые располагаются в балочных понижениях. В то же время светло-серые лесные почвы всеми методами отделяются достаточно плохо (рис. 5.1).

5.2 Анализ схожести цифровых почвенных карт с картой 2016 г.

В результате сравнения по 200 случайным точкам почвенных карт (рис. 2.14), полученных на основе различных моделей, с почвенной картой 2016 г. были выявлено, что точность почвенных карт, созданных на основе цифровых

моделей, коэффициент каппа самовоспроизводимости которых характеризуется средними величинами для этих моделей, выше чем, точность карт, построенных на основе этих же видов моделей с максимальным значением коэффициента каппа (кроме метода деревьев решений C5 и множественной логистической регрессии, для которой коэффициенты каппа можно считать примерно одинаковыми). Это происходит из-за слишком большой настройки на исходные данные в случае, когда каппа модели максимальна, то есть получаемая модель оказывается слишком привязанной к конкретным данным (табл. 5.1).

Таблица 5.1
Сравнение почвенных карт с картой 2016 г.

Почвенные карты	Каппа, характеризующая самовоспроизводимость цифровой модели	Каппа	Общая точность, %
Карта 1984 г.	-	0,26	47
Множественная логистическая регрессия	средняя $k= 0,44$	0,50	66
	максимальная $k= 0,77$	0,51	67
Деревья решений C5	средняя $k= 0,32$	0,48	65
	максимальная $k= 0,64$	0,61	73
Метод случайного леса	средняя $k= 0,40$	0,46	65
	максимальная $k= 0,75$	0,44	62
Регрессионные деревья	средняя $k= 0,26$	0,57	70
	максимальная $k= 0,52$	0,38	56
Метод опорных векторов	средняя $k= 0,40$	0,47	65
	максимальная $k= 0,79$	0,45	65

Наибольшая схожесть построенной нами почвенной карты отмечается с почвенной картой, смоделированной на основе деревьев решений C5 с максимальной каппой: коэффициент каппа сравнения таких карт составил 0,61, что является существенным совпадением, а общая точность составила 73 % (табл. 5.1).

Худшие результаты сравнения получились с картой, построенной на основе модели регрессионных деревьев с максимальной каппой: каппа – 0,38 (удовлетворительное совпадение) и общая точность – 56%.

Исследование показало, что разброс каппа при проведении сравнения для карт, построенных на основе моделей с древовидной структурой и нашей картой, больше, чем при сравнении нашей карты с другими моделями. Например, карта, построенная на основе модели деревьев решений C5 со средней каппой самовоспроизводимости, при сравнении дает коэффициент каппа, равный 0,48 (хорошее совпадение), в то время как карта, построенная на основе этой же модели с максимальной каппой самовоспроизводимости, при сравнении дает коэффициент каппа, равный 0,61 (существенное совпадение) (табл. 5.1). Аналогичная ситуация наблюдается для моделей регрессионных деревьев. Происходит это из-за того, что правила, которые составляются этими моделями при обучении и на основе которых потом проходит классификация растра, имеют каждый раз хаотичный характер. В итоге каждый раз получаются сильно отличающиеся друг от друга карты, что можно наблюдать в главе 4 (рис 4.6, 4.8).

Наблюдается небольшое отличие капп при сравнении карты 2016 г. и карт, построенных на основе моделей с максимальной и средней каппой самовоспроизводимости, для следующих моделей: множественная логистическая регрессия, метод случайного леса и опорных векторов. То есть предсказание пространственного распределения условных почвенных групп всеми этими моделями на основе созданной выборки для исследуемой территории будет примерно одинаковым для каждой модели, и мало зависит от коэффициента каппа самовоспроизводимости модели, которая используется при предсказании.

Сравнение с картой 1984 г. показало удовлетворительное совпадение с построенной нами картой, каппа составил 0,26, а общая точность 47% (табл. 5.1).

Таким образом, среднее значение коэффициента каппа при сравнении карт, построенных путем моделирования, с картой 2016 г., построенной традиционными методами с элементами ГИС-анализа, составило 0,49, а средняя общая точность – 65%.

Тем самым, моделирование на основе цифровых почвенных методов показывает в целом хорошее совпадение с почвенной картой 2016 г., что дает возможность говорить о том, что карты, полученные путем цифрового моделирования, в целом соответствуют карте 2016 г.

5.3 Проверка карт по случайной независимой выборке

Проверка полученных карт по независимой случайной выборке (рис. 2.15) дает возможность судить о качестве модели и уровне точности, с которым можно экстраполировать результаты моделирования на всю территорию. Наилучшие результаты сравнения показал традиционный метод почвенного обследования с элементами ГИС-анализа, каппа составила 0,76 (существенное совпадение) и общая точность – 82% (табл.3.6, 5.2).

Таблица 5.2

Проверка построенных карт по случайной независимой выборке из 27 точек

Почвенные карты	Каппа, характеризующая самовоспроизводимость цифровой модели	Каппа	Общая точность, %
Традиционный метод почвенного обследования с элементами ГИС-анализа	-	0,76	82
Карта 1984 г.	-	0,35	49
Множественная логистическая регрессия	средняя $k= 0,44$	0,40	52
	максимальная $k= 0,77$	0,47	60
Деревья решений C5	средняя $k= 0,32$	0,46	60
	максимальная $k= 0,64$	0,47	60
Метод случайного леса	средняя $k= 0,40$	0,46	60
	максимальная $k= 0,75$	0,41	56
Регрессионные деревья	средняя $k= 0,26$	0,29	45
	максимальная $k= 0,52$	0,39	52
Метод опорных векторов	средняя $k= 0,40$	0,51	63
	максимальная $k= 0,79$	0,38	52

В тоже время, из полученных нами результатов сравнения следует, что наиболее качественной и точной моделью для экстраполяции результатов на

всю территорию ключевого участка является модель, основанная на методе опорных векторов, со средней каппой самовоспроизводимости модели. Каппа при сравнении карты, построенной на основе этой модели, и случайной независимой выборки показала хорошее совпадение и составила 0,51, а общая точность при сравнении этой карты с 27 точками проверки составила 63%. (табл. 5.2). Визуальная интерпретация результатов моделирования этим методом показывает схожесть границ почвенных выделов на карте, построенной этой моделью и традиционной почвенной карте 2016 г.

Сходные результаты сравнения по случайной независимой выборке дали четыре модели, а именно: две модели деревьев решений C5 с максимальной и средней кappой самовоспроизводимости, множественная логистическая регрессия с максимальной кappой и метод случайного леса со средней кappой (табл. 5.2). Каппа сравнения карт, построенных этими методами, и почвами случайной независимой выборки показала хорошее совпадение.

Худшие результаты сравнения показала модель на основе регрессионных деревьев со средней кappой. При сравнении почв случайной независимой почвенной выборки с почвами, отраженными на карте, построенной на основе этой модели, общая точность составила 45%, а коэффициент каппа оказался равным 0,29 (табл. 5.2). Обе карты смоделированные на основе регрессионных деревьев показали самые низкие результаты сравнения.

Для двух видов моделей: метод случайного леса и метод опорных векторов, - модели со средней кappой показали лучшие результаты, чем модели с максимальной кappой. В то же время, для регрессионных деревьев и множественной логистической регрессии наблюдается обратная картина. Проверка точности экстраполяции моделей на основе деревьев решений C5 по случайной независимой выборке дало практически одинаковые результаты со средней и максимальной кappой.

Проверка карты 1984 г. по случайной независимой выборке, показала, что эта карта имеет удовлетворительное совпадение ($k=0,35$, а общая точность составила 49%) (табл. 5.2).

Таким образом, самые точные результаты проверки по случайной независимой выборке из 27 точек показала карта 2016 г., составленная на основе полевых обследований традиционными методами почвенного картографирования с использованием элементов ГИС-анализа. Карта 1984 г. показала удовлетворительное совпадение с выборкой из 27 точек, причем средняя каппа сравнения по всем цифровым моделям оказалась выше, чем для карты 1984 г. (карта 1984 г. – 0,35, среднее по моделям – 0,42). А лучшей моделью при внешней проверке точности экстраполяции оказалась модель, основанная на методе опорных векторов со средней каппой самовоспроизводимости для этого вида моделирования.

Была проанализирована таблица сопряженности, в которой для каждой точки случайной независимой выборки были представлены почвенные классы полученные в результате моделирования (см. приложение 4). Получилось, что 14 из 27-ми точек выборки все методы цифровой почвенной картографии относили к одному и тому условному поченному выделу (рис. 5.2). В этих местах наши представления о почвенном покрове сошлись с предсказанными на основе моделирования почвенными классами. Десять из этих точек располагались на выровненных участках рельефа с углами наклона до 3^0 под пашней. Десять точек все модели отнесли к дерново-подзолистым слабосмытым почвам, две точки были отнесены к аллювиальным почвам, а две другие точки – к дерново-подзолистым среднесмытым почвам. Точки, которые разные модели относили к разным условным почвенным группам, располагались преимущественно под лесной растительностью, а также на более крутых склонах от 3^0 .

Анализ таблицы сопряженности (см. приложение 4) также показал, что все модели дерново-подзолистые несмытые, сильносмытые и намытые почвы нашей случайной независимой почвенной выборки ошибочно относят к

дерново-подзолистым среднесмытым почвам. 2 точки светло-серых лесных почв нашей выборки все модели также относят к дерново-подзолистым почвам (см. приложение 4).

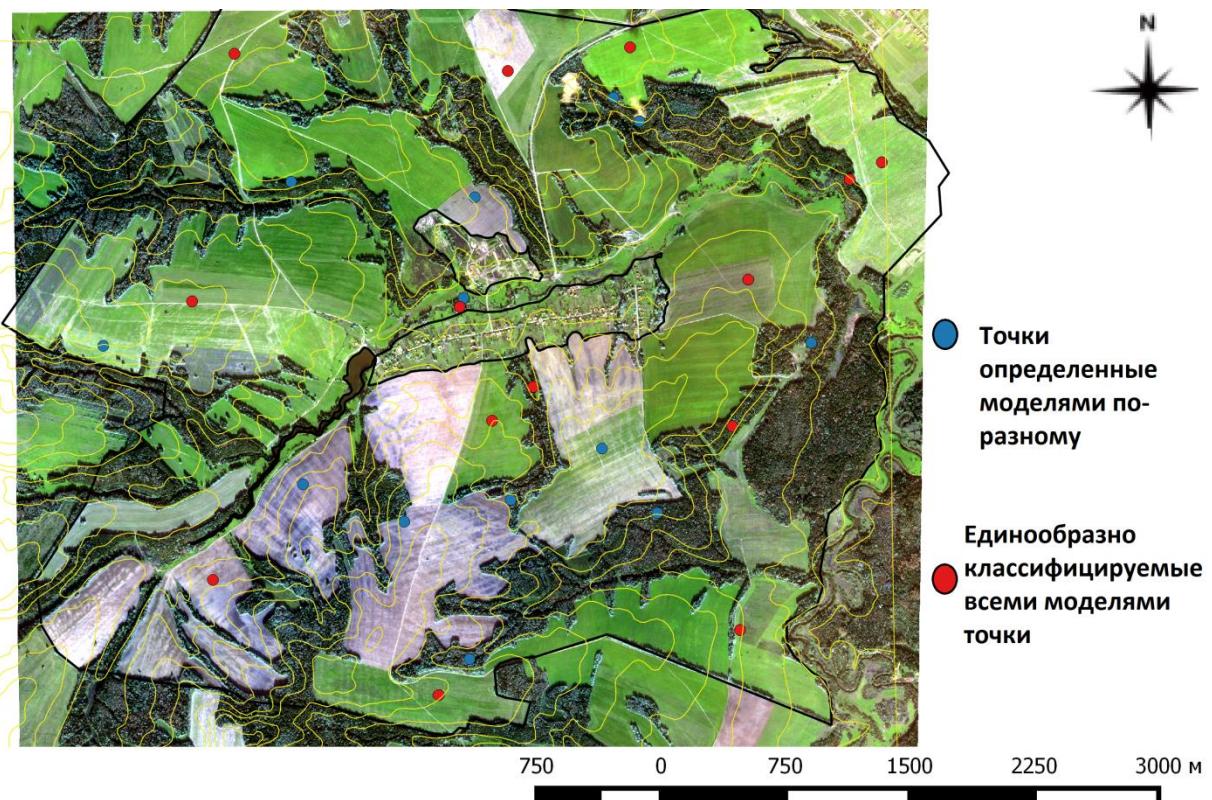


Рис. 5.2 Распределение единообразно классифицируемых всеми моделями точек

Все модели правильно выявляют территории, занятые дерново-подзолистыми слабосмытыми, среднесмытыми и аллювиальным почвам. Территории с крутыми склонами и территории, занятые лесной растительностью разные модели относят к разным условным почвенным группам. Такие результаты связаны с особенностями почвенного покрова территории.

Проверка построенных карт по случайной независимой выборке показала, что наибольшей общей точностью (82%) обладает почвенная карта, построенная традиционными методами. Цифровые карты только приближаются к ней по общей точности, показывая значения от 45-63% в зависимости от метода.

ГЛАВА 6. ПОСТРОЕНИЕ КАРТОГРАММЫ СОДЕРЖАНИЯ ОРГАНИЧЕСКОГО УГЛЕРОДА В ВЕРХНЕМ ГУМУСОВОМ СЛОЕ ПОЛЕЙ, ПОДСЧЕТ ЗАПАСОВ ГУМУСА И ОЦЕНКА ТОЧНОСТИ ПРЕДСКАЗАНИЯ

6.1 Отбор ковариат для пространственного моделирования содержания органического углерода в верхнем гумусовом слое

Для построения картограммы содержания органического углерода в верхнем гумусовом слое нами были отобраны точки обследования, находящиеся на полях исследуемого участка (рис. 2.16). Анализ коэффициентов корреляции между полученными данными по содержанию органического углерода и ковариатами окружающей среды показал, что коэффициенты корреляции не превышают значения 0,43 (табл. 6.1). Это, по-видимому, связано с тем, что топографическая карта, которая легла в основу расчетов морфометрических характеристик рельефа, имеет масштаб 1:50 000 и отражает элементы мезорельефа исследуемой территории, а микрорельеф не отражает. Содержание органического углерода в верхнем гумусовом слое зависит от микрорельефа местности и интенсивности сельскохозяйственного использования территории (Кузяков, Кузякова, 1997).

Значимая корреляция наблюдается со следующими морфометрическими характеристиками: крутизна склона, топографический индекс влажности, индекс всхолмленности рельефа, водосборные склоны, русловая сеть, абсолютные высоты рельефа и водосборная площадь (табл.6.1).

Поэтому в процессе пространственного моделирования распределения органического углерода были использованы только значимые ковариаты, описанные выше.

Таблица 6.1

Коэффициенты корреляции (r) между содержанием органического углерода (%) и ковариатами ($N=101$)

Ковариаты	r
Длина склона	0,11
Абсолютная высота	-0,18*
Текстура	-0,02
Продольная кривизна	-0,04
Максимальный вес	-0,14
Водосборная площадь	0,43
Долины	-0,01
Относительная позиция склона	0,06
Русловая сеть	-0,19
Кривизна в профиле	0,06
Фактор длины склона	-0,17
Основная кривизна	-0,03
Тангенциальная кривизна	-0,16
Топографический индекс позиции	-0,07
Отмывка рельефа	0,03
Индекс силы потока	0,01
Максимальная кривизна	-0,14
Превышение	-0,18
Крутизна склонов водосбора	-0,15
Индекс всхолмленности рельефа	-0,21
Индекс сходимости	-0,11
Минимальная кривизна	0,10
Кривизна в плане	-0,12
Крутизна склона	-0,21
Топографический индекс влажности	0,29
NDVI	0,16
Вектор всхолмленности рельефа	-0,12
Выпуклость поверхности	0,06
Аспект	-0,10
Индекс баланса массы	-0,03
Модифицированная водосборная площадь	0,11
Поперечная кривизна	-0,17
Индекс классификации рельефа на равнинах	0,02
Глубина долины	-0,18
Водосборные склоны	-0,21

*Примечание: жирным выделены строки со значимыми коэффициентами корреляции

6.2 Построение картограмм содержания органического углерода в верхнем гумусовом слое полей, оценка точности предсказания

В результате анализов среднее содержание органического углерода в верхних гумусовых слоях составило 0,94 %, стандартное отклонение – 0,47%, минимальное значение содержания органического углерода составило – 0,25%, а максимальное – 3,18% (см. приложение 5). Наибольшие значения содержания органического углерода наблюдались в гумусовых горизонтах светло-серых лесных почв, а также в локальных понижениях. Наименьшие значения среди всех верхних горизонтов наблюдались в дерново-подзолистых сильносмытых почвах. Близкие к нулю значения содержания были характерны для элювиальных, элювиально-иллювиальных и переходных горизонтов.

Построение картограмм содержания органического углерода проводилось методами ординарного и регрессионного кригинга. Для описания пространственного распределения содержания органического углерода в верхнем гумусовом слое полей были построены семивариограммы (рис. 6.1, табл. 6.2). Одна из них показывала изменение дисперсии разности значений содержания органического углерода в двух точках от расстояния для метода ординарного кригинга (рис.6.1,А, табл. 6.2), а другая отображала зависимость дисперсии разности значений остатков при проведении регрессионного кригинга (рис.6.1,Б, табл. 6.2).

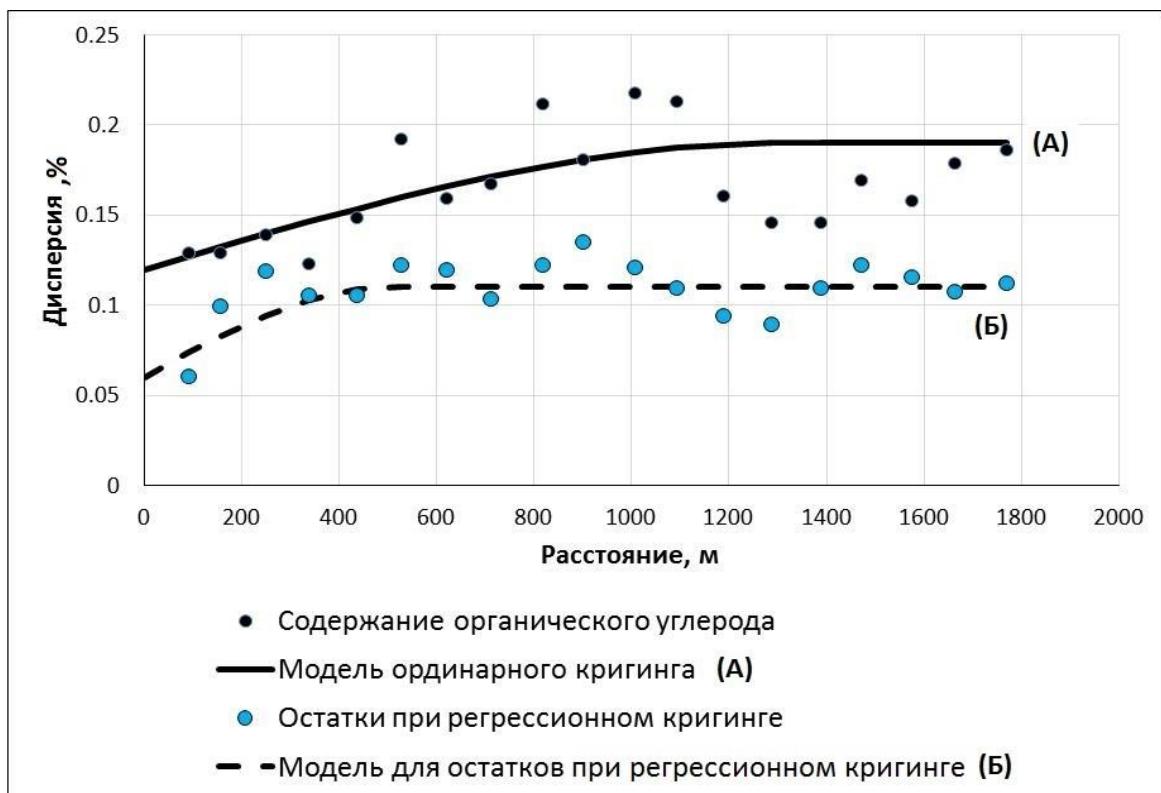


Рис. 6.1 Модели семивариограмм для содержания органического углерода при ординарном кригинге (А) и для остатков при регрессионном кригинге (Б)

Таблица 6.2

Характеристики семивариограмм, используемых при ординарном и регрессионном кригинге

Методы интерполяции	Модель ординарного кригинга	Модель для остатков при регрессионном кригинге
Тип семивариограммы	Сферическая	Сферическая
«Наггет»-эффект, %	0,12	0,06
Порог, %	0,071	0,05
Ранг, м	1309,75	500

Полученные семивариограммы показали, что регрессионный кригинг предпочтительнее перед ординарным при моделировании пространственного распределения органического углерода в верхних гумусовых слоях полей, так как кривая, описывающая дисперсию признака в случае семивариограммы для остатков при регрессионном кригинге, располагалась ниже кривой, описывающей дисперсию признака в случае семивариограммы для ординарного кригинга (рис.6.1).

На основе полученных семивариограмм (рис.6.1, табл. 6.2) проводилась интерполяция ординарным (рис. 6.2, А) и регрессионным кригингом (рис. 6.2, Б), и было получено две картограммы содержания органического углерода в верхнем гумусовом слое полей исследуемого участка.

При наложении на нее топографической карты содержание органического углерода на водораздельных поверхностях, в ложбинах, на расположенных частях склонов, плавно перетекающих в речные долины, было предсказано методом регрессионного кригинга более 1,2%, что несомненно является верным, так как на исследуемой территории идут активные процессы водной эрозии, которые приводят к смыву верхнего гумусосодержащего горизонта и намыву его в депрессии рельефа.

Согласно картограмме, построенной методом регрессионного кригинга, наименьшие значение содержания органического углерода характерны для территорий, занятых умеренными и крутыми склонами, например в северо-западной части ключевого участка (рис.6.2, Б). В то же время, картограмма, полученная на основе ординарного кригинга, в общем виде отражает закономерности распределения органического углерода на полях исследуемой территории. Согласно этой картограмме наименьшее содержание органического углерода в верхних гумусовых слоях также как и в случае картограммы, построенной методом регрессионного кригинга, характерно для северо-западной части участка, в центральной части территории содержание органического углерода составляет от 0,9% до 1,2% (рис.6.2, А). Поля юго-восточной части ключевого участка характеризуются самым большим содержанием органического углерода, что связано с рельефом этой части территории, который представляет из себя расположенную часть склона плавно переходящую к реке Убыть.

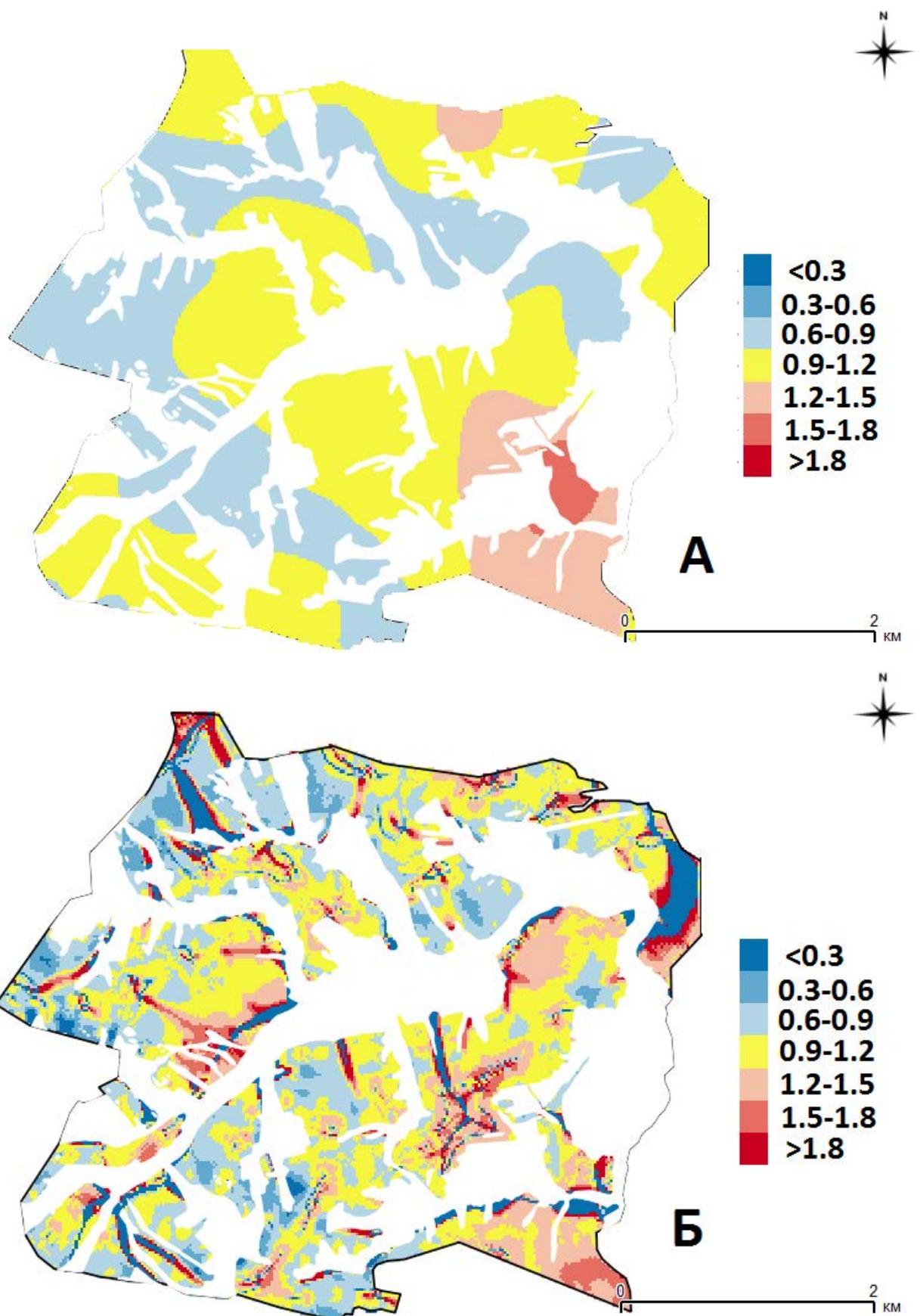


Рис. 6.2 Картограммы содержания органического углерода в верхнем гумусовом горизонте сельскохозяйственных полей, построенные методом ординарного кригинга (А) и методом регрессионного кригинга (Б)

Визуальное сравнение полученных картограмм показало, что картограмма, построенная методом регрессионного кrigинга, точнее отражает распределение органического углерода в пахотных слоях полей.

В ходе работы были построены картограммы, отражающие неопределенность предсказания ординарным кrigингом (рис. 6.3, А) и регрессионным кrigингом (рис. 6.3, Б). Были посчитаны средние ошибки предсказания при каждом методе интерполяции. Средняя ошибка предсказания при ординарном кrigинге составила 0,41%, а при регрессионном кrigинге – 0,32%. Полученные ошибки предсказания достаточно большие, но в то же время ошибка предсказания при регрессионном кrigинге меньше, чем при ординарном кrigинге.

Для ординарного кrigинга области с меньшей неопределенностью предсказания характерны для территорий, где имеется большое скопление точек обследования на достаточно близком расстоянии друг от друга (рис. 6.3, А). Для регрессионного кrigинга же области с наименьшей ошибкой предсказания характерны в окрестностях точек обследования в радиусе до 200 м (рис. 6.3, Б).

Визуальное сравнение и интерпретация картограмм содержания органического углерода в верхнем гумусовом слое, построенных методами регрессионного и ординарного кrigинга, показало, что на картограмме, построенной регрессионным кrigингом, выделяется большее количество элементарных контуров. Этот фактор в совокупности с меньшей неопределенностью предсказания дает основание сделать вывод о том, что регрессионный кrigинг является более предпочтительным и точным методом интерполяции, чем ординарный кrigинг. Поэтому построение картограммы запасов гумуса полей ключевого участка проводилось на основе картограммы, созданной путем регрессионного кrigинга (рис. 6.2, Б).

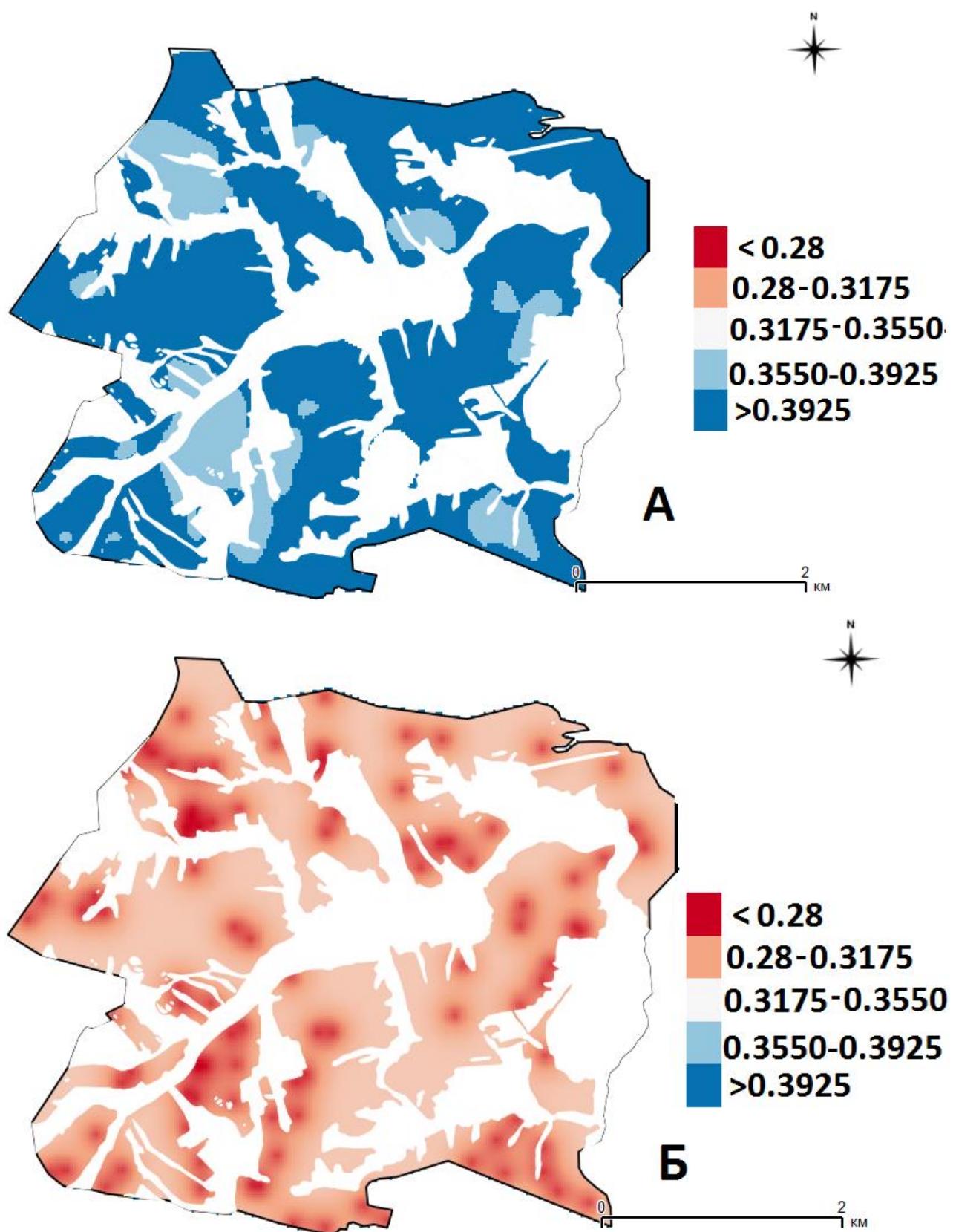


Рис. 6.3 Картограммы неопределенности предсказания ординарным кригингом (А) и регрессионным кригингом (Б)

6.3 Построение картограммы запасов гумуса

В ходе работы были рассчитаны запасы гумуса и построена картограмма содержания гумуса на полях ключевого участка (рис. 6.4) и картограмма ошибок предсказания (рис. 6.5).

Запасы гумуса в слое 0-20 см на сельскохозяйственных полях исследуемой территории преимущественно варьируют в пределах 20-60 т/га. По полученной карте отмечается увеличение количества гумуса в верхнем слое 0-20 см ближе к краям полей, а также в ложбинах, что связано с эрозионными процессами, происходящими в регионе исследования, которые приводят к смыву почвенного материала с высоких позиций рельефа и намыву этого материала в депрессионные позиции рельефа. Также отмечается низкие запасы гумуса на территориях с крутыми углами наклона поверхности, что также свидетельствует об эрозионных процессах, происходящих на этой территории.

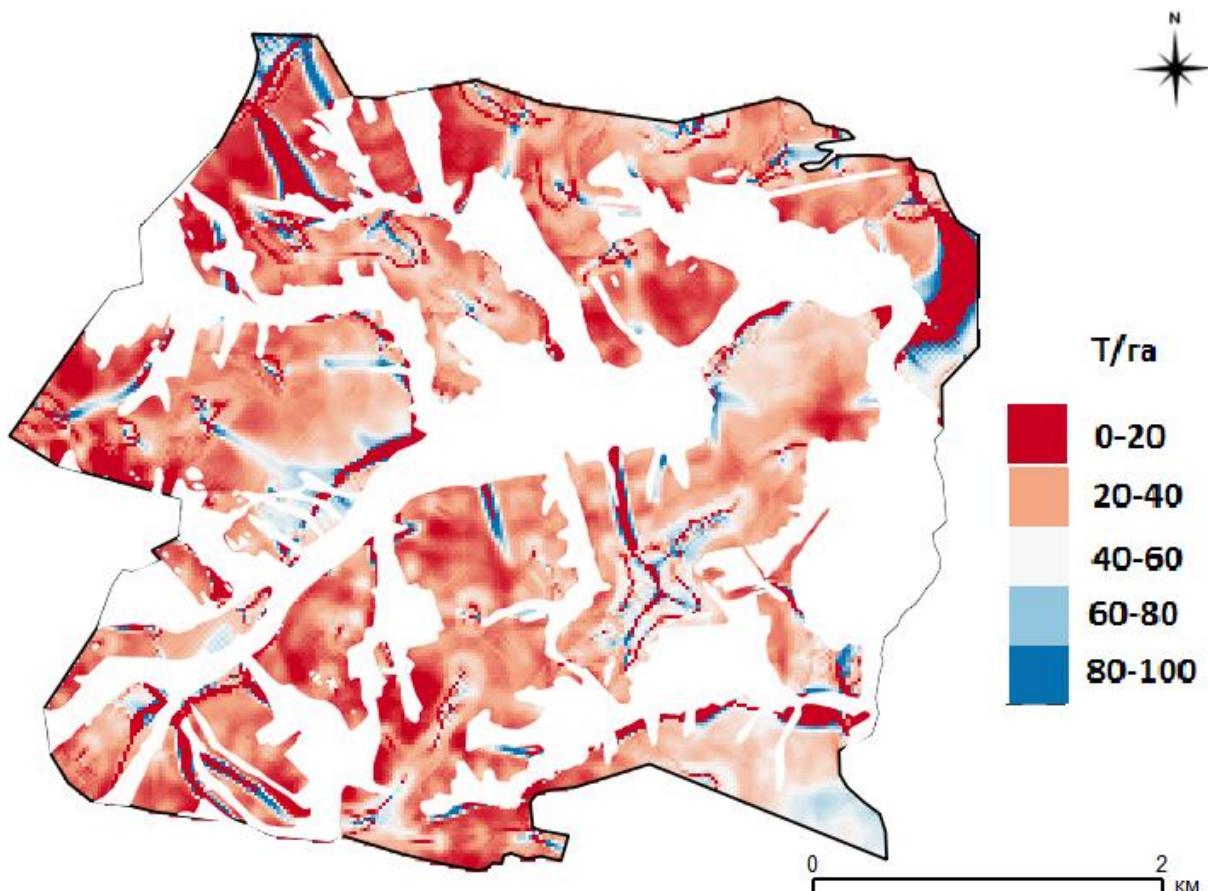


Рис. 6.4 Картограмма запасов гумуса полей в т/га в слое 0-20 см

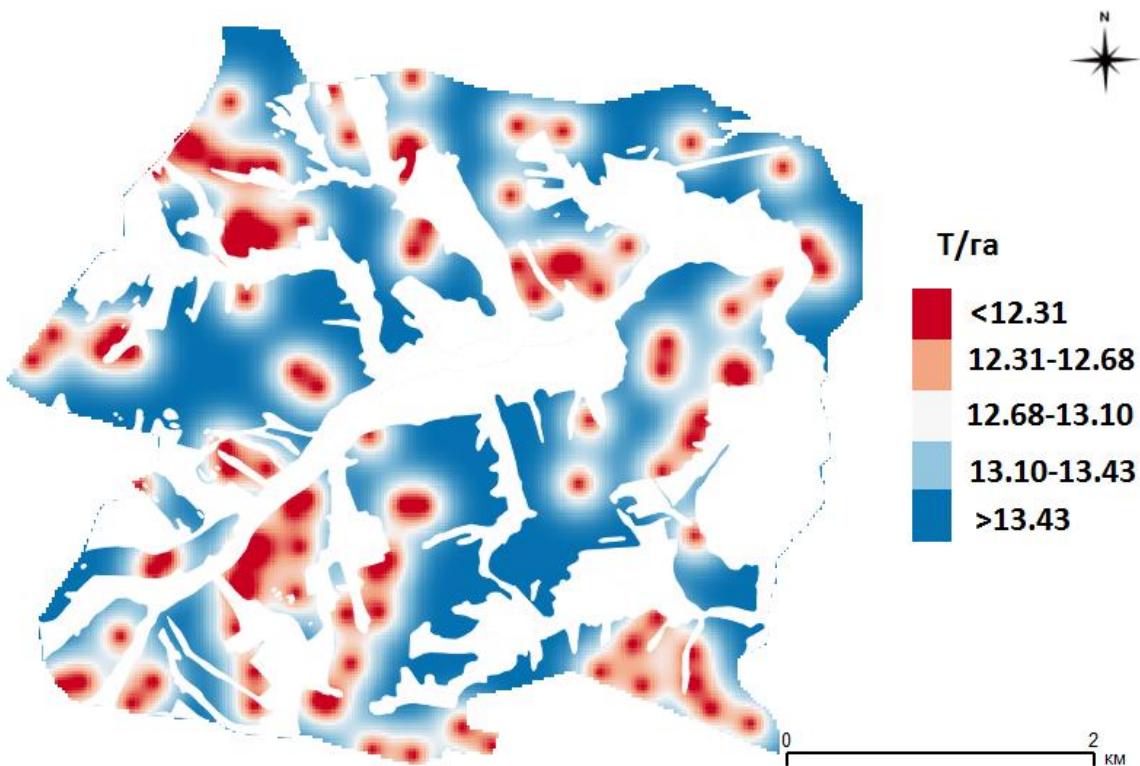


Рис. 6.5 Картограмма ошибок предсказания для картограммы запасов гумуса полей

Отмечается, что ошибка предсказания для территории исследования варьируется от 12 до 13,43 т/га. Согласно рисунку 7.5 ясно выделяются области с меньшей ошибкой предсказания, которые соответствуют окрестностям точек обследования.

Полученная картограмма запасов гумуса была наложена на карту организации пашни хозяйства ООО «Чура» (рис.2.4). Была составлена таблица, в которой были отражены средние запасы гумуса по полям и рабочим участкам (табл. 6.3).

Таблица 6.3

Средние запасы гумуса по полям и рабочим участкам

	Средние запасы гумуса по всему полю, т/га	Участок с наибольшими средними запасами гумуса, т/га	Участок с наименьшими средними запасами гумуса, т/га
Поле №1	45,4	Участок №4 (67,1)	Участок №7 (32,3)
Поле №2	36,7	Участок №6 (54,3)	Участок №3 (27,0)
Поле №3	46,0	Участок №5 (69,8)	Участок №1 (36,6)
Поле №5	37,5	Участок №7 (41,2)	Участок №5 (33,6)
Поле №6	44,9	Участок №6 (54,0)	Участок №4 (38,7)
Поле №7	43,8	Участок №2 (46,3)	Участок №3 (41,2)

Среднее значение запасов гумуса по всем полям исследуемой территории составило около 40,3 т/га. Запасы гумуса на исследуемых территориях можно считать очень низкими согласно показателям гумусного состояния почв, предложенным Орловым Д.С. и Гришиной Л.А (Гришина, Орлов, 1978). Наибольшие средние запасы гумуса согласно полученной картограмме характерны для поля №3 и составляют 46 т/га, а наименьшие – для поля №2, средние запасы гумуса которого составили 36,7 т/га (табл. 6.3, рис.2.4). Наибольшими запасами гумуса характеризуется рабочий участок №5 на поле №3 (рис.2.4). Средние запасы гумуса на этом участке составляют 69,8 т/га (табл. 6.3). Самыми низкими средними значениями запасов гумуса (27 т/га) среди всех рабочих участков данной территории обладает участок №3 поля №2 (табл. 6.3, рис.2.4). Общие запасы гумуса на всех полях были оценены нами как 922 т.

Таблица 6.4

Средние запасы гумуса по условным почвенным группам

Условный почвенный выдел	Средние запасы гумуса, т/га
Π^d	37,1
$\Pi^d \downarrow$	42,6
$\Pi^d \downarrow\downarrow$	32,8
$\Pi^d \downarrow\downarrow\downarrow$	28,2
$\Pi_{нам}^d$	38,2
СЛ	40,8

Полученная картограмма запасов гумуса была наложена на почвенную карту 8 условных почвенных групп, созданную традиционным методом обследования с элементами ГИС-анализа (рис.3.6). В итоге была составлена таблица, в которой были отражены средние запасы гумуса по 6 условным почвенным группам (все дерново-подзолистые почвы разной степени эродированности и светло-серые лесные почвы) (табл.6.4, рис. 6.6)

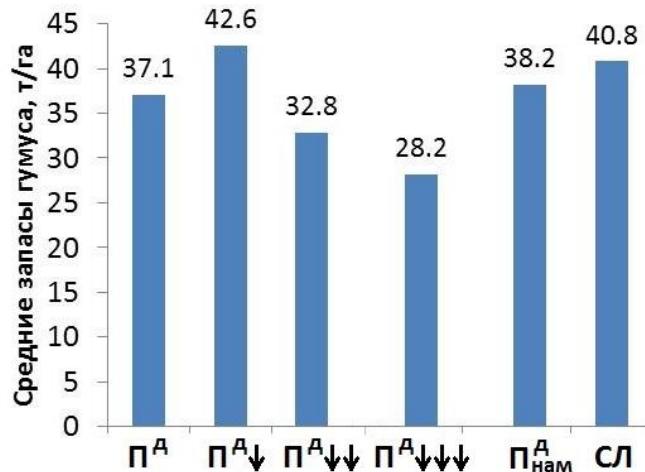


Рис. 6.6 Средние запасы гумуса по условным почвенным группам

В итоге, наибольшие средние запасы гумуса характерны для дерново-подзолистых слабосмытых почв (42,6 т/га), что связано с наибольшей площадью данной условной почвенной группы, занимаемой на исследуемых полях (табл.6.4, рис. 6.6). 37,1 т/га – средние запасы гумуса для дерново-подзолистых несмытых почв полей. Анализ таблицы и рисунка показал, что чем выше степень эродированности почвы, тем меньше средние запасы гумуса

для этой условной почвенной группы (табл.6.4, рис. 6.6). Самое низкое значение запасов гумуса среди почв характерно для дерново-подзолистых сильносмытых почв, так как они активней всего подвергаются водной эрозии, что приводит, в свою очередь, к активному смыву верхнего плодородного слоя почвы. Стоит отметить, что средние запасы гумуса для дерново-подзолистых намытых и светло-серых лесных почв полей в целом сравнимы с запасами гумуса для дерново-подзолистых слабосмытых почв (табл.6.4, рис. 6.6). Это связано с приуроченностью этих условных почвенных групп к депрессиям рельефа, в связи с чем из-за намыва для этих почв характерен увеличенный гумусовый слой, что способствует увеличению средних запасов гумуса для этих почв.

Таким образом, большая часть территории исследования характеризуется очень низкими запасами гумуса согласно показателям гумусного состояния почв, предложенным Орловым Д.С. и Гришиной Л.А (Гришина, Орлов, 1978). Это прежде всего связано с активными эрозионными процессами, протекающими на территории исследования. Наибольшие запасы гумуса характерны для поля №3, которое располагается на относительно выровненных участках рельефа исследуемой территории. Наибольшие запасы гумуса характерны для дерново-подзолистых слабосмытых почв, которые являются самыми распространенными на ключевом участке и больше других условных почвенных групп вовлечены в сельскохозяйственное производство.

ВЫВОДЫ

1. Почвенная карта, построенная на исследуемую территорию, включает 41 разновидность почв, которые относятся к четырем различным типам почв. Сравнение с картой 1984 г. показало, что границы основных почвенных типов на данной территории не претерпели существенных изменений за последние 30 лет. Использование современных ГИС-систем в совокупности со снимком высокого разрешения и качественной топографической основой позволило произвести уточнение контуров полей, оврагов и лесов.

2. Моделирование пространственного распределения почвенных выделов пятью методами автоматизированной классификации с обучением множественная логистическая регрессия, деревья решений C5, метод случайного леса, регрессионные деревья, метод опорных векторов на территорию исследуемого участка показало, что наилучшим образом выделяются дерново-подзолистые слабосмытые и среднесмытые почвы. На втором месте по значениям точности производителя находятся дерново-подзолистые несмытые и намытые, а также дерново-глеевые и аллювиальные почвы. Хуже всего методами автоматической классификации выделяются дерново-подзолистые сильносмытые и светло-серые лесные.

3. Проверка моделей на самовоспроизводимость (кросс-валидация) показала среднюю общую точность моделирования от 48 до 61%, а средняя каппа варьирует в диапазоне от 0,26 до 0,44. Уровень единобразия получаемых картографических материалов на основе моделей множественной логистической регрессии, характеризующихся различными коэффициентами каппа самовоспроизводимости этих моделей, самый высокий по сравнению с другими моделями.

4. Был выявлен минимальный набор ковариат, включающий характер землепользования, крутизну склона, различные виды кривизн, фактор длины склона и другие морфометрические характеристики, позволяющий использовать предложенные модели. Таким образом, было подтверждено, что

рельеф оказывает решающее значение на формирование почвенного покрова территории.

5. Сравнение цифровых карт между собой показало, что около 55 % территории классифицируются всеми методами единообразно. Наименьший процент неопределенности при классификации наблюдается для участков сельскохозяйственных полей, расположенных, в основном, на выровненных территориях с углами наклона поверхности до 3 градусов. Наибольший процент неопределенности при классификации выявлен для территорий под лесной растительностью и для территорий балок и оврагов.

6. Проверка построенных карт по случайной независимой выборке показала, что наибольшей общей точностью (82%) обладает почвенная карта, построенная традиционными методами. Цифровые карты только приближаются к ней по общей точности, показывая значения от 45-63% в зависимости от метода.

7. С экспертной точки зрения и по количественным характеристикам цифровые почвенные карты, построенные на основе моделей деревьев решений с максимальной каппой самовоспроизводимости модели и регрессионных деревьев со средней каппой самовоспроизводимости, ближе соответствуют карте 2016 г., по сравнению с другими цифровыми картами.

8. Содержание органического углерода в верхних 20 см сельскохозяйственных угодий на исследуемой территории колеблется от 0,4% до 3,7% (среднее - 1,0%); в среднем по полям запасы гумуса составили 40,3 т/га. Исследование показало, что построение картограммы содержания органического углерода регрессионным кригингом снижает ошибку предсказания запасов гумуса по сравнению с ординарным кригингом в два раза.

СПИСОК ЛИТЕРАТУРЫ

1. Андроников В.Л. Аэрокосмические методы изучения почв. – М.: Колос, 1979. 280 с.
2. Белоусова Н.И., Мешалкина Ю.Л. Пример гармонизации данных о бореальных почвах России // В сб.: Цифровая почвенная картография: теоретические и экспериментальные исследования. - М.: Почв. ин-т им. В.В. Докучаева, 2012. С. 108–115.
3. Васенев И.И., Васенева Э.Г. Перспективы использования цифровых почвенных карт и специализированных систем поддержки принятия решений для агроэкологической оптимизации земледелия в условиях Центрально-Черноземного региона России // В сб.: Цифровая почвенная картография: теоретические и экспериментальные исследования. - М.: Почв. ин-т им. В.В. Докучаева, 2012. С. 155–1.
4. Владыкина Н.И. Изменение климатических условий в северной зоне Удмуртии / Аграрная наука Евро-Северо-Востока. 2012. № 1. С. 4-8.
5. Вологжанина Т.В. Серые лесные почвы зоны широколиственных лесов Русской равнины: Автореферат дисс. докт.с.-х. наук. -М., 1984. 51 с.
6. Гедымин А. В. Использование изображения рельефа горизонталями при создании почвенных карт крупного масштаба. Методическое пособие.- М.: географический факультет МГУ, 1990. 24 с.
7. Геостатистика в почвоведении и экологии (интерактивный курс): Учебно-практическое пособие / Ю. Л. Мешалкина, И. И. Васенев, И. Ф. Кузякова, В. А. Романенков. — РГАУ-МСХА имени К.А. Тимирязева Москва, 2010. — С. 95.
8. Гришина Л.А., Орлов Д.С. Система показателей гумусного состояния почв // Проблемы почвоведения. - М., 1978. С. 42-47.
9. Дзюин Г.П., Дзюин А.Г. Типизация агроландшафтов для разработки адаптивно-ландшафтных систем земледелия в Удмуртской Республике / Достижения науки и техники АПК, 2013. № 6. С. 38-40.

10. Дмитриев Е.А. Теоретические и методологические проблемы почвоведения.- М.: ГЕОС, 2001. 374 с.
11. Добровольский Г.В., Урусевская И.С. География почв. 2-е изд., перераб. и доп. М: Колос, 2004. 458 с.
12. Добровольский Г. В., Урусевская И. С., Алябина И. О. Карта почвенно-географического районирования России масштаба 1:15 000 000 // Доклады по экологическому почвоведению. 2008. Т. 8. № 2. С. 1–18.
13. Докучаев П.М. Возможность использования данных дистанционного зондирования для отслеживания динамики экологической ситуации районов нефтедобычи (на примере Ханты-Мансийского Автономного Округа)// Материалы Международного молодежного научного форума «ЛОМОНОСОВ-2013».— М.: МАКС Пресс, 2013.
14. Докучаев П.М. Динамика экологической ситуации районов нефтедобычи в ханты-мансиjsком автономном округе (по данным дистанционного зондирования)// Материалы по изучению русских почв. 2014. № 8 (35). С. 135-139.
15. Докучаев П.М. Особенности почвенного покрова сельскохозяйственных угодий Глазовского района Удмуртской Республики. Материалы Международного молодежного научного форума «ЛОМОНОСОВ-2015» / — М.: МАКС Пресс, 2015а.
16. Докучаев П. М. Оценка динамики экологической ситуации районов нефтедобычи на основании данных дистанционного зондирования на примере Октябрьского района Ханты-Мансийского АО// Экология и промышленность России, 2016. №7. Т.20. С. 30-35. DOI:10.18412/1816-0395-2016-7-30-35.
17. Докучаев П.М. Почвы сельскохозяйственных угодий Глазовского района Удмуртской Республики// В кн.: Материалы межд. научной конф. XVIII Докучаевские молодежные чтения «Деградация и продовольственная безопасность России»/ СПб.: Издательский дом С.-Петербургского государственного университета. 2015б. С. 276-277.

18. Докучаев П.М., Мешалкина Ю.Л. Определение содержания органического углерода с помощью портативного спектро-радиометра// Вестник Оренбургского государственного университета, 2016а. №10. С. 58-62.
19. Докучаев П.М., Мешалкина Ю.Л. Сравнение результатов почвенного картографирования: с использованием ГИС-технологий и на основе автоматизированной классификации с обучением, на примере ключевого участка Чура Глазовского района Удмуртской Республики// В сб.: Почвоведение-продовольственной и экологической безопасности страны: тезисы докладов VII съезда Общества почвоведов им. В. В. Докучаева (Белгород, 15-22 августа 2016 г.). Часть II.– Москва-Белгород: Издательский дом «Белгород». 2016б. С. 299-300.
20. Ефимова Т.П., Ложкина Н.П., Тычинин В.А., Баранов В.И. Растительность// Природа Удмуртии. Ижевск, 1972. С. 145-201.
21. Жуйкова И.А., Пахомов М.М., Прокашев А.М. Динамика и становление современного растительного и почвенного покрова Вятско-Камского региона в голоцене //Учение о зонах природы на переломе тысячелетий. Н. Новгород, 2000. С. 109-123.
22. Кашин А.А. Исследование ландшафтной организации территории Удмуртии как фактора хозяйственного освоения и расселения населения. Дис. ... канд. геогр. наук: 25.00.23. Ижевск, 2015. – 162 с.
23. Качинский Н.А. Механический и микроагрегатный состав почвы, методы его изучения. –М.: Изд-во АН СССР, 1958. – 191 с.
24. Кириллова Н.П., Силёва Т.М., Ульянова Т.Ю., Савин И.Ю. Метод "совпадений" и его применение для построения цифровой крупномасштабной почвенной карты // Почвоведение, 2014. № 10. С. 1193-1203.
25. Кирьянова Е.Ю., Савин И.Ю. О возможностях оценки контрастности почвенного покрова Саратовского Поволжья по спутниковым данным Landsat// В сб.: Цифровая почвенная картография: теоретические и

- экспериментальные исследования. - М.: Почв. ин-т им. В.В. Докучаева, 2012. С. 189–208.
26. Классификация и диагностика почв России. Смоленск: Ойкумена, 2004. 342 с.
 27. Классификация и диагностика почв СССР. — М.: Колос, 1977. - 223 с.
 28. Ковриго В.П. Почвы Удмуртской Республики. Ижевск: РИО Ижевская ГСХА, 2004. 490 с.
 29. Козлов Д.Н. Цифровой анализ ландшафта в крупномасштабном картографировании структур почвенного покрова. Автореф...дисс..канд. геогр. н.-М., 2009. 27 с.
 30. Козлов Д.Н., Сорокина Н.П. Традиции и инновации в крупномасштабной почвенной картографии// В сб.: Цифровая почвенная картография: теоретические и экспериментальные исследования. - М.: Почв. ин-т им. В.В. Докучаева, 2012. С. 35–57.
 31. Козловский Ф.И. Почвенный индивидуум и методы его определения// Закономерности пространственного варьирования свойств почв и информационно-статистические методы изучения. – М.: Наука, 1970. С. 42-59.
 32. Колесков В. С. Почвы Юкаменской МТС УАССР// Записки Удмурт. научно-исслед. ин-та соц. культуры. Ижевск: Удмуртгиз, 1939. Вып. 8. С. 40-132.
 33. Кононова М.М. Органическое вещество почвы, его природа, свойства и методы изучения. М.: Изд-во АН СССР, 1963. 314 с.
 34. Королюк Т.В. Интерпретация космических изображений в системе методов цифровой почвенной картографии // В сб.: Цифровая почвенная картография: теоретические и экспериментальные исследования. - М.: Почв. ин-т им. В.В. Докучаева, 2012. С. 124–139.
 35. Кравцова В.И. Космические методы исследования почв.- М., 2005. 190 с.
 36. Красильников В.П. Почвенная номенклатура и корреляция. Петрозаводск, 1999. 435 с.

37. Кренке А.Н. Коррекция почвенных карт на основе данных дистанционного зондирования и цифровой модели рельефа // В сб.: Цифровая почвенная картография: теоретические и экспериментальные исследования. - М.: Почв. ин-т им. В.В. Докучаева, 2012. С. 284–301.
38. Кузякова И.Ф., Кузяков Я.В. Влияние микрорельефа на пространственное варьирование содержание гумуса в дерново-подзолистой почве в условиях длительного полевого опыта // Почвоведение. 1997 №7. С. 823-830.
39. Куст Г.С., Брызжев А.В., Розов С.Ю. Опыт применения метода эталонирования космических снимков для дешифрирования почвенного покрова сельскохозяйственных полей в Краснодарском крае// Доклады по экологическому почвоведению. 2010. Т. 13. № 1. С. 50-103.
40. Кутузова Н.Д., Куст Г.С., Розов С.Ю., Стома Г.В. Влияние пространственной неоднородности почвенных свойств на рост и урожайность сои// Почвоведение. М.: Наука. 2015. №3. С. 355-367.
41. Методика составления крупномасштабных почвенных карт с применением материалов аэрофотосъемки. М.: Изд-во АН СССР, 1962. - 114 с.
42. Мешалкина Ю.Л. Что такое «цифровая почвенная картография»? (обзор) // В сб.: Цифровая почвенная картография: теоретические и экспериментальные исследования. М.: Почв. ин-т им. В.В. Докучаева, 2012. С. 9–18.
43. Мищенко, Н.В, Трифонова Т.А. Оценка состояния почвенно-растительного покрова методами геоинформационного анализа – Владимир : ВладимирПолиграф, 2007. – 92 с.
44. Мясникова А.С., Васенев В.И. Построение 3D-модели загрязнения атмосферного воздуха г. Москвы и рекомендации по выбору газоустойчивого растительного ассортимента для озеленения участков г. Москвы// Вестник Российской университета дружбы народов. Серия: Агрономия и животноводство, 2015. № 3. С. 18-22.

45. Ногина Н.А. Вятско-Камская провинция дерново-подзолистых высокогумусированных почв и дерново-подзолистых почв со вторым гумусовым горизонтом //Почвенно-географическое районирование СССР. М.: Изд-во АН СССР, 1962. С. 84-87.
46. Общесоюзная инструкция по почвенным обследованиям и составлению крупномасштабных почвенных карт землепользований. М.: Колос, 1973. -96 с.
47. Пермяков Ф.И. Почвы Удмуртии, повышение их плодородия. Ижевск. 1972. 223 с.
48. Подсосова Т. К. Рельеф// Природа Удмуртии. Ижевск: Удмуртия, 1972. С. 37-64.
49. Поздняков А. И. Полевая электрофизика почв. М.: МАИК Наука/Интерperiодика, 2001. С. 188.
50. Почвенная карта мира.1:10000000. Гл. ред. В.А. Ковда. М.: ГУГК, 1975
51. Почвенная карта мира: Для высш. учеб. заведений. 1:15000000/ Глазовская, Фридланд М.: ГУГК, 1982.
52. Почвенная карта Удмуртской АССР. Масштаб 1:200000. Отв. ред. Р.К. Сигнаевский, Б.П. Теплых. М.: ГУГК, 1990.
53. Программа Почвенной карты СССР масштаба 1:2500000. М.: Почвенный ин-т им. В.В. Докучаева ВАСХНИЛ, 1972. 158 с.
54. Прокашев А.М. Почвы со сложным органопрофилем Вятско-Камского Предуралья. Автореф. дисс. д-ра. биол. н. - Киров, 2003.
55. Пузаченко М.Ю. Мультифункциональный ландшафтный анализ юго-запада Валдайской возвышенности. Автореф. дисс. канд. геогр. н. - М., 2009. 24 с.
56. Пузаченко Ю.Г., Федяева М.В., Козлов Д.Н., Пузаченко М.Ю. Методологические основания отображения элементарных геосистемных процессов // В сб.: Современные естественные и антропогенные процессы в почвах и геосистемах. – М.: Почв. ин-т им. В.В. Докучаева, 2006. С. 13-52.

57. Родина Л.Е., Базилевич Н.И. Динамика органического вещества и биологический круговорот в основных типах растительности. – Л.: Наука, 1965. 253 с.
58. Рожков В.А. Становление почвенной информатики // Почвоведение. 2002. № 7. С. 858—866.
59. Рухович Д.И., Вильчевская Е.В., Калинина Н.В., Королева П.В. Цифровая тематическая картография как смена доступных первоисточников и способ их использования// В сб.: Цифровая почвенная картография: теоретические и экспериментальные исследования. - М.: Почв. ин-т им. В.В. Докучаева, 2012. С. 58–86.
60. Рысин И.И. Почвы и земельные ресурсы // География Удмуртии: природные условия и ресурсы: учеб. пособие / под ред. И.И. Рысина. Ижевск: Изд. дом «Удмуртский университет», 2009. Ч. 1. 256 с. С. 182–204.
61. Савин И.Ю. Компьютерная имитация картографирования почв // В сб.: Цифровая почвенная картография: теоретические и экспериментальные исследования. - М.: Почв. ин-т им. В.В. Докучаева, 2012. С. 26–34.
62. Савин И.Ю., Танов Э.Р., Харзинов С. Использование вегетационного индекса NDVI для оценки качества почв пашни (на примере Баксанского района Кабардино-Балкарии)// Бюллетень Почвенного института имени В.В.Докучаева. М.: изд-во Почв. ин-т им. В.В. Докучаева, 2015. № 77. С. 51-65.
63. Самсонова В.П. Пространственная изменчивость почвенных свойств на примере дерново-подзолистых почв, М., 2008, 160 с.
64. Самсонова В. П., Благовещенский Ю. Н., Мешалкина Ю. Л. Пример использования Байесовского кригинга для выявления неоднородности распределения органического углерода на сельхозугодье // Почвоведение, 2017. № 3.
65. Самсонова В.П., Мешалкина Ю.Л. Исследование неопределенностей, возникающих при детальном цифровом почвенном картографировании

- агросерых почв Брянского ополья // В сб.: Цифровая почвенная картография: теоретические и экспериментальные исследования. - М.: Почв. ин-т им. В.В. Докучаева, 2012. С. 116–123.
66. Самсонова В. П., Мешалкина Ю. Л. Количественный метод сравнения почвенных карт и картограмм // Вестник Московского университета. Серия 17: Почвоведение. 2011. № 3. С. 3–5.
67. Сидорова В.А. Геостатистический анализ пространственной неоднородности сельскохозяйственных полей для целей точного земледелия. Дисс. канд. биол. н. Петрозаводск, 2011. 147 с.
68. Сидорова В.А., Соломатова Е.А. Пространственная вариабельность лесной подстилки в ельниках Восточной Фенноскандии.//Экологические функции почв Восточной Фенноскандии. - Петрозаводск: Карельский научный центр РАН, 2000. С.133-141.
69. Симакова М.С., Савин И.Ю. Использование материалов аэро- и космической съемки в картографировании почв: пути развития, состояние, задачи // Почвоведение, 1998. №11. С.1339-1347.
70. Смирнова Л.Г., Ченdev Ю.Г., Новых Л.Л., Украинский П.А., Новых И.Е. Микрозональные особенности профильного и пространственного распределения содержания гумуса в почвах склона // Научные ведомости. Серия естественные науки. 2011. № 15. С. 160-167.
71. Сорокина Н.П. Методология составления крупномасштабных агроэкологически ориентированных почвенных карт. М.: Изд-во Почвенный ин-т им. В.В. Докучаева РАСХН, 2006. - 161 с.
72. Сорокина Н.П., Козлов Д.Н. Методы цифровой почвенной картографии в задачах агроэкологической оценки земель // В сб.: Цифровая почвенная картография: теоретические и экспериментальные исследования. - М.: Почв. ин-т им. В.В. Докучаева, 2012. С. 140–154.
73. Тайчинов С.Н. Региональные особенности почв Южного Урала и принципы их качественной оценки// Докл. башкирских почвоведов к X

международному почвенному конгрессу. Уфа: БФАН СССР, 1973. Т.1.С. 63-71.

74. Трифонова Т.А., Мищенко Н.В. Оценка почвенно-продукционного потенциала речных бассейнов с использованием данных дистанционного зондирования // География и природные ресурсы. 2016. № 3. С. 17–25.
75. Тюлин В.В., Россохина М.В. Почвы со вторым гумусовым горизонтом Чепецко-Кильмезского водораздела //Почвоведение, 1967. № 7. С. 28-37.
76. Ульянова Т. Ю., Зборищук Ю. Н. Практические занятия по курсу Картография почв. Москва, 2005. С. 120.
77. Флоринский И. В. Гипотеза Докучаева как основа цифрового прогнозного почвенного картографирования (к 125-летию публикации)// Почвоведение, 2012, № 4, с. 500–506.
78. Фридланад В.М. Структуры почвенного покрова. – М.: Мысль, 1984. 239 с.
79. Шишов Л.Л., Рожков В.А., Столбовой В.С. Информационная база классификация почв // Почвоведение. 1985. № 9. С. 9-20.
80. Шоба С.А., Алябина И.О., Колесникова В.М., Молchanov Э.Н., Рожков В.А., Столбовой В.С., Урусевская И.С., Шеремет Б.В., Конюшков Д.Е. Почвенные ресурсы России. Почвенно-географическая БД. М.: ГЕОС, 2010. 128 с.
81. Юницкий В.П. Почвенные комплексы речных долин правобережья Малмыжского у. Вятской губ. и так называемые Малмыжские черноземы (по данным исследования 1927 г.)// Почвенные исследования в бывшей Вятской губернии (1926-1929 гг.), г.Киров, 1937. С. 3-42.
82. AAFC, Agriculture and Agri-Food Canada. The Canadian Soil Information System (CANSIS) and the National Soil Database (NSDB). [Электронный ресурс]. <http://sis.agr.gc.ca/cansis/index.html>. Редакция от: 04.08.2013.
83. Adhikari K., Bou Kheir R., Greve M. B., Bocher P. K., Malone B. P., Minasny B., McBratney A. B., Greve, M. H. High-resolution 3-D mapping of soil texture

- in Denmark// Soil Science Society of America Journal. 2012. Vol. 77. Issue 3. P. 860–876.
84. Advances in Digital Terrain Analysis / Zhou, Qiming; Lees, Brian; Tang, Guo-an (Eds.). 2008. XIV. 462 p.
85. Aksoy E., Panagos P., Montanarella L. Spatial prediction of soil organic carbon of Crete by using geostatistics. In: Minasny B., Malone B. P., McBratney A. B. (Eds.). Digital soil assessments and beyond. - London: CRC Press, 2012. P. 149–159.
86. Allison V.J., Yermakov Z., Miller R.M., Jastrow J.D., Matamala R. Assessing soil microbial community composition across landscapes: do surface soils reveal patterns?// Soil Sci. Soc. Am. J. 2007. Vol. 71. P. 730–734.
87. Aly Z., Bonn F.J., Magagi R. Analysis of the backscattering coefficient of saltaffected soils using modeling and RADARSAT-1 SAR data// Geosci. Remote Sens. IEEE Trans. 2007. Vol. 45. Issue 2. P. 332–341.
88. Anderson-Cook C.M., Alley M.M., Roygard J.K.F., Khosla R., Noble R.B., Doolittle J.A. Differentiating soil types using electromagnetic conductivity and crop yield maps// Soil Science Society of America Journal. 2002. Vol. 66. P. 1570 – 1652.
89. Arrouays D., Daroussin J., Kicin L. and Hassika P. Improving topsoil carbon storage prediction using a digital elevation model in temperate forest soils of France// Soil Science. 1998. Vol. 163. P. 103–108.
90. ASRIS. Australian Soil Resource Information System, Australian Department of Agriculture. [Электронный ресурс]. http://www.asris.csiro.au/index_ie.html. Редакция от: 30.06.2014
91. Awiti A.O., Walsh M.G., Shepherd K.D., Kinyamari J. Soil condition classification using infrared spectroscopy: a proposition for assessment of soil condition along a tropical forest-cropland chronosequence// Geoderma 2008. Vol. 143. P. 73–84.

92. Ballabio C., Comolli R. Mapping heavy metal content in soils with multi-kernel SVR and LiDAR derived data. In: Digital Soil Mapping - New York: Springer Science + Business Media B.V., 2010. P. 205-216.
93. Ballabio C., Fava F., Rosenmund A. A plant ecology approach to digital soil mapping, improving the prediction of soil organic carbon content in alpine grasslands// Geoderma. 2012. Vol. 187–188. P. 102–116.
94. Barthold F.K., Wiesmeier M., Breuer L., Frede H.-G., Wu J., Blank F.B. Land use and climate control the spatial distribution of soil types in the grasslands of Inner Mongolia// Journal of Arid Environments. 2013. Vol. 88. P. 194-205.
95. Behrens T., Foster H., Scholten T., Steinrucken U., Spies E.D., Goldschmitt M. Digital soil mapping using artificial neural networks// Journal of Plant Nutrition and Soil Science. 2005. Vol. 168. P. 21-33.
96. Behrens T., Zhu A.X., Schmidt K., Scholten T. Multi-scale digital terrain analysis and feature selection for digital soil mapping// Geoderma. 2010. Vol. 155. Issue 3-4. P. 175-185.
97. Bishop T., McBratney A.B. A comparison of prediction methods for the creation of field-extent soil property maps// Geoderma. 2001. Vol. 103. P. 149 – 160.
98. Bishop T., Minasny B. Digital soil-terrain modeling: the predictive potential and uncertainty. In: Grunwald S., Eds., Environmental Soil-Landscape Modeling—Geographic Information Technologies and Pedometrics. - CRC Press, 2006. P. 185–213.
99. Bockheim J.G., McLeod M. Soil distribution in the McMurdo dry valleys, Antarctica// Geoderma. 2008. Vol. 144. P. 43–49.
100. Boegh E., Soegaard H., Thomsen A. Evaluating evapotranspiration rates and surface conditions using Landsat TM to estimate atmospheric resistance and surface resistance// Remote Sensing of Environment. 2002. Vol. 79. P. 329 – 343.

101. Boettinger J.L. Environmental covariates for digital soil mapping in the Western USA. In: Digital Soil Mapping - New York: Springer Science + Business Media B.V., 2010. P. 17-27.
102. Boruvka L., Penizek V. A test of an artificial neural network allocation procedure using the Czech soil survey of agricultural land data. In: Lagacherie P.; McBratney A. B., Voltz M., eds. Digital soil mapping: An introductory perspective. – Amsterdam: Elsevier, 2007. P. 415-424.
103. Bou Kheir R., Greve M. H., Bocher P. K., Greve M. B., Larsen R. Predictive mapping of soil organic carbon in wet cultivated lands using classification-tree based models: the case study of Denmark// Journal of Environmental Management. 2010. Vol. 91. Issue 5. P. 50–60.
104. Breiman L. Random forests// Mach. Learn. 2001. Vol. 45. P. 5-32.
105. Breiman L., Friedman J. H., Olshen R. A., Stone, C. J. Classification and Regression Trees. Wadsworth. 1984.
106. Brest C.L., Goward S.N. Deriving surface albedo measurements from narrow-band satellite data// International Journal of Remote Sensing. 1987. Vol. 8. P. 351 – 367.
107. Brungard C.W., Boettinger J.L. Conditioned latin hypercube sampling: optimal sample size for Digital Soil Mapping of arid rangelands in Utah, USA. In: Digital Soil Mapping - New York: Springer Science + Business Media B.V., 2010. P. 67-75.
108. Brungard C.W., Boettinger J.L., Duniway M.C., Wills S.A., Edwards Jr. T.C. Machine learning for predicting soil classes in three semi-arid landscapes// Geoderma. 2015. Vol. 239-240. P. 68–83.
109. Buchanan S., Triantafyllis J. Mapping water table depth using geophysical and environmental variables// Ground Water. 2009. Vol. 47. P. 80-96.
110. Bui E.N. Soil survey as a knowledge system// Geoderma. 2003. Vol. 120. P. 17-26.

111. Bui E.N., Moran C.J. Disaggregation of polygons of surficial geology and soil maps using spatial modelling and legacy data// *Geoderma*. 2001. Vol. 103. P. 79 – 94.
112. Burrough P.A., Bouma J., Yates S.R. The state of the art in pedometrics // *Geoderma*. 1994. Vol. 62. P. 311-326.
113. Burrough P.A., MacMillan R.A. and Van Deursen W. Fuzzy classification methods for determining land suitability from soil-profile observation and topography// *Soil Science*. 1992. Vol. 43. Issue 2. P. 193–210.
114. Carré F., Jeannée N., Casalegno S., Lemarchand O., Reuter H.I. and Montanarella L. Mapping the CN ratio of the forest litters in Europe-Lessons for Global Digital Soil Mapping. В сб.: *Digital Soil Mapping* - New York: Springer Science + Business Media B.V., 2010. P. 217-225.
115. Castrignano A., Buttafuoco D., Comolli R. Using digital elevation model to improve soil pH prediction in an alpine doline// *Pedosphere*. 2011. Vol. 21. Issue 2. P. 259–270.
116. Chambers J. M. Linear models. Chapter 4 of *Statistical Models* in S eds J. M. Chambers and T. J. Hastie, Wadsworth & Brooks/Cole. 1992.
117. Chang C.C., Lin C.J. LIBSVM: a library for support vector machines. [Электронный ресурс] <https://www.csie.ntu.edu.tw/~cjlin/libsvm/> Редакция от: 01.12.2016.
118. Chen F., Kissel D.E., West L.T., Adkins W., Rickman D., Luval J.C. Mapping soil organic carbon concentrations for multiple fields with image similarity analysis// *Soil Sci. Soc. Am. J.* 2008. Vol. 72. P. 186–193.
119. Congalton R. G. A review of assessing the accuracy of classifications of remotely sensed data// *Remote Sensing of the Environment*. 1991. Vol. 37. P. 35-46.
120. Connolly J., Holden N.M., Ward S.M. Mapping peatland in Ireland using a rule-based methodology and digital data// *Soil Sci. Soc. Am. J.* 2007. Vol. 71. P. 492–499.

121. Crowley J.K. Mapping playa evaporite minerals with AVIRIS data: a 1st report from Death Valley, California// *Remote Sensing of the Environment*. 1993. Vol. 44. P. 337 – 356.
122. De Gruijter J.J., Walvoort D.J.J., van Gaans P.F.M. Continuous soil maps a fuzzy set approach to bridge the gap between aggregation levels of process and distribution models// *Geoderma*. 1997. Vol. 77. P. 169-195.
123. Diak G.R., Anderson M.D., Bland W.L., Norman J.M., Mecikalski J.M., Aune R.M. Agricultural management decision aids driven by real-time satellite data// *Bulletin of the American Meteorological Society*. 1998. Vol. 79. P. 1345 – 1355.
124. Diaz-Uriarte R., de Andres S.A. Gene selection and classification of microarray data using random forest// *BMC Bioinform*. 2006. Vol. 7. P. 1-13.
125. Digital soil mapping – as a support to production of functional maps/ Dobos E., Carré F., Hengl T., Reuter H.I., Tóth G. Luxemburg: European Communities, 2006. 68 pp.
126. Dijkshoorn J.A. SOTER database for Southern Africa (SOTERSAF). ISRIC — World Soil Information, Wageningen. 2003.
127. Dijkshoorn J.A., Huting J.R.M., Tempel P. Update of the 1:5 million Soil and Terrain Database for Latin America and the Caribbean (SOTERLAC; version 2.0). ISRIC — World Soil Information, Wageningen. 2005.
128. Dobos E., Bialkó T., Micheli E. and Kobza J. Legacy soil data harmonization and database development. B c6.: *Digital Soil Mapping* - New York: Springer Science + Business Media B.V., 2010. P. 309-323.
129. Donatelli M., Stockle C., Constantini E.A., Nelson R. SOILR: a model to estimate soil moisture and temperature regimes. 2002.
http://www.inea.it/isci/mdon/research/bottom_model_soil.htm.
130. Du C., Linker R., Shaviv A. Identification of agricultural soils using mid-infrared photoacoustic spectroscopy// *Geoderma*. 2008. Vol. 143. P. 85–90.
131. Farr T.G. The shuttle radar topography mission. *IEEE Aerospace Conference Proceedings*. 2000. 63 p.

132. Florinsky I.V. Combined analysis of digital terrain models and remotely sensed data in landscape investigations// *Prog. Phys. Geogr.* 1998. Vol. 22. Issue 1. P. 1-28.
133. Florinsky I.V. Digital Terrain Modeling in Soil Science and Geology. Elsevier / Academic Press, Amsterdam. 2011. 379 p.
134. Genovese G.P. Introduction to the MARS Crop Yield Forecasting System (MCYFS). Meeting on 4 and 5 October 2001. Luxembourg. Space Applications Institute. - Ispra, Italy: Joint Research Centre of the European Commission, 2001. 15 p.
135. Genxu W., Yuanshou L., Yibo W., Qingbo W. Effects of permafrost thawing on vegetation and soil carbon pool losses on the Qinghai — Tibet Plateau, China// *Geoderma*. 2008. Vol. 143. P. 143–152.
136. Gomez C., Viscarra Rossel R.A., McBratney A.B. Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: an Australian case study// *Geoderma*. 2008. Vol. 146. Issue 3–4. P. 403–411.
137. Gooley L., Huang J., Page D., Triantafilis J. Digital soil mapping of available water content using proximal and remotely sensed data// *Soil use and management*. 2014. Vol. 30. Issue 1. P. 139-151.
138. Grimm R., Behrens T., Märker M., Elsenbeer H. Soil organic carbon concentrations and stocks on Barro Colorado Island—digital soil mapping using Random Forest analysis// *Geoderma*. 2008. Vol. 146. P. 102–113.
139. Grunwald S. Multi-criteria characterization of recent digital soil mapping and modeling approaches// *Geoderma*. 2009. Vol. 152. P. 195–207.
140. Grunwald S., Thompson J.A., Boettinger J.L. Digital soil mapping and modeling at continental scales: finding solutions for global issues// *Soil Sci. Soc. Am. J.* 2011. Vol. 75. P. 1201-1213.
141. Hansen R.O., Racic L., Grauch V.J.S. Magnetic methods in near-surface geophysics. In: Butler, D.K. (ed.), *Near-Surface Geophysics: Investigations in Geophysics*. 2005. Vol. 13. P. 151–175.

142. Hartemink A. E., Krasilnikov P., Bockheim J. G. Soil maps of the world // *Geoderma*. 2013. Vol. 207. P. 256–267.
143. Häring T., Dietz E., Osenstetter S., Koschitzki T., Schröder B. Spatial disaggregation of complex soil map units: a decision-tree based approach in Bavarian forest soils// *Geoderma*. 2012. Vol. 185–186. P. 37–47.
144. Harmonized World Soil Database (version 1.0)/ FAO, IIASA, ISRIC, ISS-CAS, JRC. FAO, Rome, Italy and IIASA, Laxenburg, Austria. 2008.
145. Hengl T., Toormanian N., Reuter H.I., Malakouti M.J. Methods to interpolate soil categorical variables from profile observations: lessons from Iran// *Geoderma*. 2007. Vol. 140. P. 417–427.
146. Hewitt A.E. Predictive modelling in soil survey// *Soils and Fertilizers*. 1993. Vol. 56. P. 305– 314.
147. Hewitt A.E., Barringer J.R.F., Forrester G.J., McNeill S.J. Soilscapes Basis for Digital Soil Mapping in New Zealand. B c6.: Digital Soil Mapping - New York: Springer Science + Business Media B.V., 2010. P. 297-307.
148. Hijmans R.J., van Etten J. Raster: geographic data analysis and modeling. R Package Version 2.1-37 (<http://CRAN.R-project.org/package=raster>). 2013.
149. Hoosbeek M. R. Incorporating scale into spatio-temporal variability: applications to soil quality and yield data // *Geoderma*. 1998. Vol. 85. Issue 2-3. P. 113-131.
150. Howell D.W., Kim Y.G., Soil temperature modeling for soil survey. Poster. Environmental Systems Research Institute User Conference, San Diego, CA, USA. 2009.
151. Huang J., Lark R.M., Robinson D.A., Lebron I., Keith A.M., Rawlins B., Tye A., Kuras O., Raines M., Triantafilis J. Scope to predict soil properties at within-field scale from small samples using proximally sensed γ -ray spectrometer and EM induction data// *Geoderma*. 2014. Vol. 232-234. P. 69-80.

152. Huete A.R., Liu H.Q., Batchily K., van Leeuwen W. A comparison of vegetation indices global set of TM images for EOS-MODIS// Remote Sensing of Environment. 1997. Vol. 59. Issue 3. P. 440–451.
153. Isaaks E.H. and Srivastava R.M. An Introduction to Applied Geostatistics. – New York: Oksford. 1989. 561 p.
154. Johnson J. E., Shaw J. A., Lawrence R., Nugent P. W., Dobeck L. M., Spangler L. H. Long-wave infrared imaging of vegetation for detecting leaking CO₂ gas// Journal of Applied Remote Sensing. 2012. Vol. 6. P. 1-9.
155. Jordanovaa N., Jordanovaa D., Tsachevab T. Application of magnetometry for delineation of anthropogenic pollution in areas covered by various soil types// Geoderma. 2008. Vol. 144. P. 557–571.
156. Keitt T., Bivand R., Pebesma E., Rowlingson B. rgdal: Bindings for the Geospatial Data Abstraction Library. R package version 0.6–21. [Электронный ресурс]. <http://CRAN.R-project.org/package=rgdal>. 2009.
157. Kempen B., Brus D.J., Heuvelink G.B.M., Stoorvogel J.J. Updating the 1:50,000 Dutch soil map using legacy soil data: a multinomial logistic regression approach// Geoderma. 2009. Vol. 151. Issue 3–4. P. 311–326.
158. King T.V.V., Clark R.N., Ager C., Swayze G.A. Remote mineral mapping using AVIRIS data at Summitville, Colorado and the adjacent San Juan Mountains. Summitville Forum '95. Special Publication. Colorado. 1995.
159. Konyushkova M.V. Automatic interpretation of Quickbird Imagery for Digital Soil Mapping, North Caspian Region, Russia. B c6.: Digital Soil Mapping - New York: Springer Science + Business Media B.V., 2010. P. 103-111.
160. Kovačević M., Bajat B., Gajić B. Soil type classification and estimation of soil properties using support vector machines// Geoderma. 2010. Vol. 154. P. 340-347.
161. Lagacherie P., Gomez C., Bailly J.S., Baret F., Coulouma G. The use of hyperspectral imagery for digital soil mapping in Mediterranean Areas. B c6.: Digital Soil Mapping - New York: Springer Science + Business Media B.V., 2010. P. 93-102.

162. Lagacherie P., McBratney A.B. Chapter 1. Spatial soil information systems and spatial soil inference systems: perspectives for digital soil mapping: review article// *B c6.: Digital Soil Mapping. An introductory perspective.* Developments in soil science. 2007. Vol. 31. P. 137-150.
163. Lamorski K., Pachepsky Y., Slawiński C., Walczak R.T. Using support vector machines to develop pedotransfer functions for water retention of soils in Poland// *Soil Sci. Soc. Am. J.* 2008. Vol. 72. P. 1243–1247.
164. Le Bas C., King D., Jamagne M., Daroussin J. The European Soil Information System. European Soil Bureau Research, Luxembourg. 1998.
165. Lemercier B., Lacoste M., Loum M., Walter C. Extrapolation at regional scale of local soil knowledge using boosted classification trees: a two-step approach// *Geoderma*. 2012. Vol. 171. P. 75–84.
166. Levi M.R., Rasmussen C. Covariate selection with iterative principal component analysis for predicting physical soil properties// *Geoderma*. 2014. Vol. 219–220. P. 46–57.
167. Liaw A., Wiener M. Classification and regression by random forest// *R News*. 2002. Vol. 2. Issue 3. P. 18–22.
168. Liu J., Pattey E., Nolin M.C., Miller J.R., Ka O. Mapping within-field soil drainage using remote sensing, DEM and apparent soil electrical conductivity// *Geoderma*. 2008. Vol. 143. P. 261–272.
169. MacMillan R.A., Moon D.E., Coupé R.A. Automated predictive ecological mapping in a Forest Region of B.C., Canada, 2001–2005 // *Geoderma*. 2007. Vol. 140. Issue 4. P. 353-373.
170. Marchetti A., Piccini C., Francaviglia R., Santucci S. and Chiuchiarelli I. Estimating soil organic matter content by regression kriging. B c6.: *Digital Soil Mapping - New York: Springer Science + Business Media B.V.*, 2010. P. 241-253.
171. Marchetti A., Piccini C., Santucci S., Chiuchiarelli I., Francaviglia R. Simulation of soil types in Teramo province (Central Italy) with terrain parameters and remote sensing data// *Catena*. 2011. Vol. 85. P. 267-273.

172. Matt P.B., Johnson W.C. Thermoluminescence and new 14C age estimates for late Quaternary loesses in southwestern Nebraska// *Geomorphology*. 1996. Vol. 17. P. 115 – 128.
173. Maynard J.J., Johnson M.G. Scale-dependency of LiDAR derived terrain attributes in quantitative soil-landscape modeling: effects of grid resolution vs. neighborhood extent// *Geoderma*. 2014. Vol. 230–231. P. 29–40.
174. Mayr T., Rivas-Casado M., Bellamy P., Palmer R., Zawadzka J. and Corstanje R. Two methods for using legacy data in Digital Soil Mapping. B cб.: *Digital Soil Mapping* - New York: Springer Science + Business Media B.V., 2010. P. 191-202.
175. McBratney A.B., Mendonça Santos M.L., Minasny B. On digital soil mapping// *Geoderma*. 2003. Vol. 117. P. 3–52.
176. McBratney A.B., Minasny B., Cattle S., Vervoort R.W. From pedotransfer functions to soil inference systems// *Geoderma*. 2002. Vol. 109. P. 41-73.
177. McBratney A.B., Odeh I.O.A. Application of fuzzy sets in soil science: fuzzy logic, fuzzy measurements and fuzzy decisions// *Geoderma*. 1997. Vol. 77. P. 85–113.
178. McKenzie N.J., Ryan P.J., de Gruijter J.J. Spatial prediction of soil properties using environmental correlation// *Geoderma*. 1999. Vol. 89. Issue 1-2. P. 67-94.
179. Meirik E., Frazier B., Brown D., Roberts P., Rupp R. ASTER-based vegetation map to improve soil modeling in remote areas. . B cб.: *Digital Soil Mapping* - New York: Springer Science + Business Media B.V., 2010. P. 113-122.
180. Mendonça-Santos M.L., Dart R.O., Santos H.G., Coelho M.R., Berbara R.L.L., and Lumbreiras J.F. Digital soil mapping of topsoil organic carbon content of Rio de Janeiro State, Brazil. B cб.: *Digital Soil Mapping* - New York: Springer Science + Business Media B.V., 2010. P. 255-266.
181. Minasny B., McBratney A.B. A conditioned Latin hypercube method for sampling in the presence of ancillary information// *Computers & Geosciences* 2006. Vol. 32. Issue 9. P. 1378–1388.

182. Minasny B., McBratney A.B. Estimating the water retention shape parameter from sand and clay content// Soil Sci. Soc. Am. J. 2007a. Vol. 71. P. 1105–1110.
183. Minasny B., McBratney A.B. Incorporating taxonomic distance into spatial prediction and digital mapping of soil classes// Geoderma. 2007б. Vol. 142. P. 285–293.
184. Moonjun R., Farshad A., Shrestha D.P. and Vaiphasa C. Artificial Neural Network and Decision Tree in predictive soil mapping of Hoi Num Rin Sub-Watershed, Thailand. В сб.: Digital Soil Mapping - New York: Springer Science + Business Media B.V., 2010. P. 151-163.
185. Mulder V.L., de Bruin S., Schaepman M.E., Mayr T.R. The use of remote sensing in soil and terrain mapping—A review// Geoderma. 2011. Vol. 162. P. 1-19.
186. MultiSpec. A Freeware Multispectral Image Data Analysis System. [Электронный ресурс] / <https://engineering.purdue.edu/~biehl/MultiSpec/> - Редакция от: 02.08.2016.
187. Nauman T.W., Thompson J.A. Semi-automated disaggregation of conventional soil maps using knowledge driven data mining and classification trees//Geoderma. 2014. Vol. 213. P. 385–399.
188. Nauman T.W., Thompson J.A., Odgers N.P., Libohova Z. Fuzzy disaggregation of conventional soil maps using database knowledge extraction to produce soil property maps. В сб.: Digital Soil Assessments and Beyond: Proceedings of the Fifth Global Workshop on Digital Soil Mapping 2012, Sydney, Australia. – London: CRC Press, 2012. P. 203–207.
189. New M., Todd M., Hulme M., Jones P. Precipitation measurements and trends in the twentieth century// International Journal of Climatology. 2001. Vol. 21. P. 1922 – 1999.
190. Ninomiya Y., Fu B., Cudahy T.J. Detecting lithology with Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER)

- multispectral thermal infrared ‘radiance-at-sensor’ data// *Remote Sens. Environ.* 2005. Vol. 99. Issue 1–2. P. 127–139.
191. Odeh I. O.A., Leenaars J., Hartemink A., Amapu I. The challenges of collating legacy data for digital mapping of Nigerian soils. In: Minasny B., Malone B., McBratney A.B., Eds., *Digital Soil Assessments and Beyond*. – London: Taylor & Francis Group, 2012. P. 453–458.
192. Odgers N.P., Sun W., McBratney A.B., Minasny B., Clifford D. Disaggregating and harmonising soil map units through resampled classification trees// *Geoderma*. 2014. Vol. 214–215. P. 91–100.
193. Padarian J., Perez-Quezada J., Seguel S. Modeling the distribution of organic carbon in the soils of Chile. In: Minasny B., Malone B. P., McBratney A. B., Eds., *Digital soil assessments and beyond*. - London: CRC Press, 2012. P. 329–333.
194. Pahlavan Rad M. R., Toomanian N., Khormali F., Brungard C.W., Komaki C. B., Bogaert P. Updating soil survey maps using random forest and conditioned Latin hypercube sampling in the loess derived soils of northern Iran// *Geoderma*. 2014. Vol. 232–234. Issue 1. P. 97–106.
195. Park S., McSweeney K., Lowery B. Identification of the spatial distribution of soils using a process-based terrain characterization// *Geoderma*. 2001. Vol. 103. P. 249–272.
196. Pásztor L., Laborczi A., Takács K., Szabó J. Variations for the Implementation of SCORPAN’s “S”. In: Zhang G.-L., Brus D., Liu F., Song X.-D., Lagacherie P., Eds., *Digital Soil Mapping Across Paradigms, Scales and Boundaries*. Springer-Verlag, 2016. P.331-342.
197. Pebesma E.J. Multivariable geostatistics in S: the gstat package. *Computers & Geosciences*. 2004. Vol. 30. P. 683-691.
198. Pebesma E.J., Bivand R.S. Classes and methods for spatial data: the sp package. 2005.

199. Peng Y., Xiong X., Adhikari K., Knadell 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. P. 1-22.
200. Prasad A.M., Iverson L.R., Liaw A. Newer classification and regression tree techniques: bagging and random forests for ecological prediction// Ecosystems. 2006. Vol. 9. P. 181-199.
201. QGIS. Свободная географическая информационная система с открытым кодом. [Электронный ресурс]. / <http://www.qgis.org/ru/site/> - Редакция от: 27.02.2017.
202. Qi F., Zhu A.-X., Harrower M., Burt J.E. Fuzzy soil mapping based on prototype category theory// Geoderma. 2006. Vol. 136. P. 774–787.
203. Quinlan R. C4.5: Programs for Machine Learning. Morgan Kaufmann Publishers. [Электронный ресурс] <http://www.rulequest.com/see5-unix.html>. 1993.
204. R Core Team. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. Vienna. Austria 3-900051-07-0 [Электронный ресурс] <http://www.R-project.org/> . Редакция от: 06.03.2017.
205. Reeves, J.B. Near- versus mid-infrared diffuse reflectance spectroscopy for soil analysis emphasizing carbon and laboratory versus on-site analysis: Where are we and what needs to be done?// Geoderma. 2010. Vol. 158. Issue 1-2. P. 3-14.
206. Robinson D.A., Abdu H., Lebron I., Jones S.B. Imaging of hill-slope soil moisture wetting patterns in a semi-arid oak savanna catchment using time-lapse electromagnetic induction// J. Hydrol. 2012. Vol. 416. P. 39–49.
207. Roecker S.M. and Thompson J.A. Scale effects on terrain attribute calculation and their use as environmental covariates for Digital Soil Mapping. В сб.: Digital Soil Mapping - New York: Springer Science + Business Media B.V., 2010. P. 55-66.
208. SAGA. System for automated geoscientific analyses. [Электронный ресурс] / <http://saga-gis.org/en/index.html> - Редакция от: 28.02.2017.

209. Samsonova V. P., Pozdnyakov A. I., Meshalkina J. L. A study of disturbed soil cover using soil electrical resistivity and topographic data // International Journal of Ecology & Development. 2007. Vol. 8. P. 39–51.
210. Savin I.Yu, Prudnikova E.Yu, Vasilevva N.A., Veretelnikova I.V., Bairamov A.N. The color of soils as a basis for proximal sensing of their composition// Бюллетень Почвенного института имени В.В.Докучаева. М.: изд-во Почв. ин-т им. В.В. Докучаева. № 86. С. 46-52.
211. Sidorova V. A., Krasilnikov P. V. Soil-geographic interpretation of spatial variability in the chemical and physical properties of topsoil horizons in the steppe zone // Eurasian Soil Science. 2007. Vol. 40. P. 1042–1051.
212. Simin C., Rongqun Z., Liming L., Zhou De. A method of salt-affected soil information extraction based on a support vector machine with texture features// Mathematical and Computer Modelling. 2010. Vol. 51. P. 1319–1325.
213. Schmidt K., Behrens T., Scholten T. Instance selection and classification tree analysis for large spatial datasets in digital soil mapping// Geoderma. 2008. Vol. 146. Issue 1–2. P. 138–146.
214. Schmidt J., Hewitt A. Fuzzy land element classification from DTMs based on geometry and terrain position// Geoderma. 2004. Vol. 121. Issue 3–4. P. 243–256.
215. Scull P., Franklin J., Chadwick O.A. The application of decision tree analysis to soil type prediction in a desert landscape// Ecological Modelling. 2005. Vol. 181. P. 1–15.
216. Shary P.A., Sharay L.S., Mitusov A.V. Fundamental quantitative methods of and surface analysis // Geoderma. 2002. Vol. 107. P. 1-43.
217. Sinha A.K. Stratigraphic mapping of sedimentary formations in southern Ontario by ground electromagnetic methods// Geophysics. 1990. Vol. 55. P. 1148 – 1157.

218. Sinowski W., Auerswald K. Using relief parameters in a discriminant analysis to stratify geological areas with different spatial variability of soil properties// Geoderma. 1999. Vol. 89. P. 113–128.
219. Skidmore A.K., Varekamp C., Wilson L., Knowles E., Delaney J. Remote sensing of soils in a eucalypt forest environment// International Journal of Remote Sensing. 1997. Vol. 18. P. 39 – 56.
220. Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. American Web Soil Survey and Geographic (SSURGO). [Электронный ресурс]. <http://soildatamart.nrcs.usda.gov>. - Редакция от: 27.02.2017.
221. Stevens A., van Wesemael B., Bartholomeus H., Rossillon D., Tychon B., Ben-Dor E. Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils// Geoderma. 2008. Vol. 144. P. 395–404.
222. Stolbovoi V., McCallum I. Land Resources of Russia. 2002.
223. Stoorvogel J.J., Kempen B., Heuvelink G.B.M., de Bruin S. Implementation and evaluation of existing knowledge for digital soil mapping in Senegal// Geoderma. 2009. Vol. 149. P. 161–170.
224. Subburayalu S., Jenhani I., Slater B.K. Disaggregation of component soil series using possibilistic decision trees from an Ohio County soil survey map// Geoderma. 2014. Vol. 213. P. 334–345.
225. Sulaeman Y., Minasny B., McBratney A. B., Sarwani M., Sutandi A. Harmonizing legacy soil data for digital soil mapping in Indonesia// Geoderma. 2013. Vol. 192. P. 77-85.
226. Susskind J., Piraino P., Rokke L., Iredell L., Mehta A. Characteristics of the TOVS pathfinder path. A data set// Bulletin of the American Meteorology Society. 1997. Vol. 78. P. 1449 – 1472.
227. Taghizadeh-Mehrjardi R., Nabiollahi K., Minasny B., Triantafilis J. Comparing data mining classifiers to predict spatial distribution of USDA-family soil groups in Baneh region, Iran. Geoderma. 2015. Vol. 253-254. P. 67–77.

228. Therneau T.M., Atkinson, Beth, port R., Ripley B. rpart: Recursive Partitioning. (R package version 3.1-50). [Электронный ресурс] <http://CRAN.R-project.org/package=rpart>. Редакция от: 29.06.2015.
229. Van Niekerk H.S., Gutzmer J., Beukes N.J., Phillips D., Kiviets G.B. An $^{40}\text{Ar}/^{39}\text{Ar}$ age of supergene K – Mn oxyhydroxides in a post-Gondwana soil profile on the Highveld of South Africa// South African Journal of Science. 1999. Vol. 95. P. 450 – 454.
230. Vasques G.M., Grunwald S., Sickman J.O. Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra// Geoderma. 2008. Vol. 146. P. 14–25.
231. Venables W., Ripley B. Modern applied statistics with S, Statistics and Computing, 4th Edition. Springer. New York. 2002.
232. Viscarra Rossel R.A, Cattle S.R., Ortega A.S., Fouad Y. In situ measurements of soil colour, mineral composition and clay content by VIS–NIR spectroscopy// Geoderma. 2009. Vol. 150. P. 253–266.
233. Viscarra Rossel R.A., McKenzie N.J. and Grundy M.J. Using proximal soil sensors for Digital Soil Mapping. В сб.: Digital Soil Mapping - New York: Springer Science + Business Media B.V., 2010. P. 79-92.
234. Vitharana U.W.A., Van Meirvenne M., Simpson D., Cockx L., De Baerdemaeker J. Key soil and topographic properties to delineate potential management classes for precision agriculture in the European loess area// Geoderma. 2008. Vol. 143. P. 206–215.
235. Wang X., Xie H., Guan H., Zhou X. Different responses of MODIS-derived NDVI to root-zone soil moisture in semi-arid and humid regions// J. Hydrol. 2007. Vol. 340. P. 12–24.
236. Webster R. Soil science and geostatistics. In Krasilnikov P., Carré F. & Montanarella L., Eds, Geostatistics and Soil Geography, European Communities. 2008. Luxemburg. P. 1–11.
237. Webster R., Oliver M.A. Geostatistics for Environmental Scientists. Wiley, Chichester. 2007.

238. Wiesmeier M., Barthold F., Blank B., Kögel-Knabner I. Digital mapping of soil organic matter stocks using random forest modeling in a semi-arid steppe ecosystem// *Plant Soil*. 2011. Vol. 340. Issue 1. P. 7–24.
239. Woolard J.W., Colby J.D. Spatial characterization, resolution, and volumetric change of coastal dunes using airborne LIDAR: Cape Hatteras, North Carolina// *Geomorphology*. 2002. Vol. 48. Issue 1–3. P. 269–287.
240. Yigini Y., Panagos P. Reference area method for mapping soil organic carbon content at regional scale// *Procedia. Earth and Planetary Science*. 2014. Vol. 10. P. 330-338.
241. Zhang C., Fay D., McGrath D., Grennan E., Carton O.T. Statistical analysis of geochemical variables in soils of Ireland// *Geoderma*. 20086. Vol. 146. P. 378–390.
242. Zhao Y.C. and Shi X.Z. Spatial prediction and uncertainty assessment of soil organic carbon in Hebei Province, China. B c6.: *Digital Soil Mapping - New York*: Springer Science + Business Media B.V., 2010. P. 227-239.
243. Zhu A.X. A similarity model for representing soil spatial information// *Geoderma*. 1997. Vol. 77. P. 217-242.

ПРИЛОЖЕНИЕ № 1

Сводная таблица, отражающая каппу самовоспроизводимости моделей при использовании в качестве ковариат почвенной карты 1984 г. (+) и без нее (-).

Номер итерации	MNLR -	MNLR +	C5 -	C5 +	RF -	RF +	SVM -	SVM +	RT -	RT +
1	0,43	0,60	0,21	0,21	0,31	0,31	0,46	0,53	0,18	0,30
2	0,71	0,38	0,45	0,45	0,56	0,56	0,45	0,40	0,28	0,41
3	0,51	0,52	0,51	0,51	0,65	0,65	0,45	0,45	0,19	0,23
4	0,48	0,46	0,24	0,24	0,31	0,39	0,40	0,28	0,21	0,21
5	0,18	0,26	0,38	0,49	0,49	0,49	0,49	0,26	0,22	0,18
6	0,77	0,30	0,16	0,16	0,37	0,37	0,32	0,34	0,24	0,36
7	0,43	0,49	0,38	0,34	0,47	0,41	0,51	0,47	0,24	0,32
8	0,36	0,34	0,52	0,44	0,50	0,40	0,43	0,47	0,18	0,50
9	0,65	0,27	0,40	0,29	0,48	0,40	0,41	0,28	0,13	0,46
10	0,41	0,73	0,35	0,35	0,58	0,58	0,59	0,66	0,43	0,23
11	0,56	0,41	0,13	0,13	0,30	0,30	0,48	0,48	0,51	0,17
12	0,48	0,42	0,26	0,26	0,25	0,18	0,20	0,35	0,13	0,01
13	0,37	0,07	0,31	0,31	0,26	0,26	-0,02	0,24	0,32	0,45

Продолжение приложения №1

Номер итерации	MNLR -	MNLR +	C5 -	C5 +	RF -	RF +	SVM -	SVM +	RT -	RT +
14	0,23	0,41	0,54	0,54	0,67	0,67	0,47	0,42	0,30	0,10
15	0,59	0,44	0,51	0,51	0,36	0,37	0,36	0,29	0,35	0,30
16	0,43	0,36	0,34	0,25	0,44	0,44	0,38	0,36	0,00	0,51
17	0,41	0,65	0,52	0,52	0,64	0,64	0,63	0,44	0,22	0,39
18	0,50	0,49	0,40	0,33	0,34	0,34	0,55	0,55	0,30	0,14
19	0,09	0,24	0,29	0,37	0,41	0,33	0,28	0,29	0,44	0,21
20	0,14	0,29	0,54	0,54	0,41	0,26	0,47	0,35	0,06	0,28
21	0,72	0,28	0,04	0,04	0,47	0,44	0,13	0,46	0,29	0,19
22	0,47	0,41	0,41	0,41	0,50	0,49	0,49	0,49	0,26	0,48
23	0,57	0,37	0,36	0,41	0,38	0,38	0,55	0,52	0,19	0,57
24	0,39	0,19	0,37	0,37	0,39	0,31	0,24	0,24	0,18	0,55
25	0,43	0,46	0,36	0,36	0,46	0,46	0,52	0,66	0,24	0,15
26	0,06	0,42	0,16	0,16	0,23	0,11	0,32	0,23	0,44	0,35
27	0,54	0,30	0,42	0,42	0,28	0,41	0,43	0,44	0,34	0,48
28	0,54	0,22	0,25	0,16	0,09	0,09	0,35	0,21	0,26	0,21

Продолжение приложения №1

Номер итерации	MNLR -	MNLR +	C5 -	C5 +	RF -	RF +	SVM -	SVM +	RT -	RT +
29	0,47	0,27	0,15	0,15	0,39	0,39	0,35	0,24	0,37	0,23
30	0,38	0,54	0,27	0,27	0,45	0,52	0,53	0,52	0,35	0,29
31	0,38	0,53	0,43	0,45	0,47	0,49	0,68	0,61	0,32	0,43
32	0,41	0,39	0,31	0,31	0,58	0,58	0,43	0,46	0,15	0,15
33	0,62	0,54	0,28	0,28	0,55	0,48	0,36	0,30	0,23	0,11
34	0,38	0,38	0,44	0,44	0,56	0,55	0,63	0,40	0,20	0,23
35	0,18	0,45	0,34	0,34	0,46	0,52	0,66	0,45	0,29	0,10
36	0,31	0,29	0,17	0,23	0,29	0,18	0,31	0,25	0,12	0,20
37	0,16	0,50	0,35	0,35	0,38	0,44	0,55	0,63	0,14	0,25
38	0,12	0,17	0,22	0,22	0,29	0,29	0,29	0,23	0,21	0,21
39	0,50	0,64	0,28	0,28	0,48	0,54	0,56	0,50	0,28	0,27
40	0,54	0,37	0,14	0,14	0,38	0,38	0,24	0,34	0,37	0,30
41	0,61	0,13	0,25	0,32	0,28	0,29	0,25	0,20	0,46	0,10
42	0,26	0,44	0,25	0,55	0,26	0,32	0,56	0,56	0,40	0,13

Продолжение приложения №1

Номер итерации	MNLR -	MNLR +	C5 -	C5 +	RF -	RF +	SVM -	SVM +	RT -	RT +
43	0,55	0,14	0,25	0,25	0,24	0,16	0,04	0,06	0,16	0,17
44	0,37	0,53	0,29	0,29	0,46	0,52	0,47	0,46	0,30	0,14
45	0,61	0,38	0,39	0,23	0,29	0,29	0,50	0,44	0,09	0,50
46	0,16	0,14	0,13	0,13	0,13	0,08	0,13	0,14	0,27	0,24
47	0,47	0,35	0,47	0,47	0,49	0,49	0,41	0,42	0,30	0,25
48	0,38	0,45	0,46	0,46	0,45	0,45	0,66	0,49	0,12	0,11
49	0,31	0,19	0,14	0,21	0,53	0,45	0,59	0,39	0,28	0,25
50	0,45	0,36	0,22	0,22	0,34	0,40	0,41	0,34	0,45	0,32
51	0,57	0,28	0,23	0,23	0,38	0,38	0,21	0,13	0,20	0,26
52	0,65	0,46	0,42	0,42	0,45	0,46	0,57	0,17	0,46	0,14
53	0,25	0,26	0,46	0,46	0,34	0,33	0,37	0,38	0,52	0,12
54	0,40	0,46	0,33	0,25	0,10	0,10	0,37	0,35	0,40	0,53
55	0,39	0,43	0,27	0,27	0,41	0,41	0,19	0,43	0,26	0,28
56	0,25	0,38	0,20	0,20	0,32	0,25	0,26	0,32	0,40	0,34

Продолжение приложения №1

Номер итерации	MNLR -	MNLR +	C5 -	C5 +	RF -	RF +	SVM -	SVM +	RT -	RT +
57	0,53	0,12	0,11	0,11	-0,03	0,03	0,04	-0,01	0,26	0,02
58	0,60	0,04	0,25	0,25	0,40	0,40	0,12	-0,06	0,36	0,19
59	0,60	0,34	0,35	0,35	0,75	0,66	0,22	0,26	0,03	0,32
60	0,58	0,44	0,48	0,42	0,43	0,43	0,47	0,35	0,16	0,34
61	0,56	0,31	0,30	0,30	0,50	0,41	0,40	0,42	0,19	0,20
62	0,35	0,33	0,06	0,06	0,21	0,22	0,34	0,26	0,28	0,09
63	0,75	0,58	0,44	0,44	0,48	0,48	0,37	0,44	0,20	0,17
64	0,41	0,42	0,33	0,33	0,53	0,53	0,28	0,44	0,25	0,48
65	0,66	0,44	0,64	0,57	0,64	0,64	0,57	0,43	0,42	0,30
66	0,18	0,13	0,28	0,28	0,25	0,19	0,34	0,19	0,36	0,41
67	0,74	0,15	0,09	0,09	0,05	0,08	0,05	0,16	0,25	0,35
68	0,31	0,58	0,35	0,35	0,50	0,49	0,63	0,58	0,51	0,25
69	0,09	0,48	0,37	0,37	0,49	0,49	0,36	0,48	0,44	0,04
70	0,23	0,65	0,41	0,41	0,49	0,41	0,42	0,58	0,28	0,29

Продолжение приложения № 1

Номер итерации	MNLR -	MNLR +	C5 -	C5 +	RF -	RF +	SVM -	SVM +	RT -	RT +
71	0,48	0,10	0,43	0,46	0,48	0,48	0,20	0,20	0,29	-0,01
72	0,13	0,43	0,30	0,19	0,35	0,27	0,36	0,36	0,30	0,11
73	0,35	0,21	0,16	0,34	0,32	0,47	0,25	0,20	0,28	0,09
74	0,72	0,67	0,37	0,37	0,49	0,44	0,58	0,58	0,04	0,17
75	0,71	0,43	0,18	0,18	0,14	0,14	0,55	0,62	0,23	0,23
76	0,29	0,16	0,20	0,14	0,32	0,32	0,39	0,20	0,08	0,00
77	0,55	0,38	0,29	0,28	0,28	0,28	0,39	0,33	0,34	0,48
78	0,66	0,62	0,34	0,34	0,48	0,48	0,47	0,54	0,18	0,14
79	0,47	0,41	0,20	0,20	0,38	0,44	0,47	0,41	0,13	0,08
80	0,55	0,36	0,29	0,29	0,34	0,34	0,48	0,40	0,19	0,10
81	0,30	0,41	0,18	0,21	0,16	0,16	0,35	0,38	0,12	0,41
82	0,36	0,28	0,39	0,39	0,33	0,38	0,31	0,32	0,21	0,33
83	0,35	0,39	0,07	0,07	0,22	0,22	0,23	0,22	0,20	-0,19
84	0,35	0,32	0,38	0,53	0,63	0,51	0,56	0,50	0,26	0,11

Продолжение приложения № 1

Номер итерации	MNLR -	MNLR +	C5 -	C5 +	RF -	RF +	SVM -	SVM +	RT -	RT +
85	0,43	0,33	0,20	0,20	0,27	0,27	0,31	0,31	0,36	0,04
86	0,33	0,41	0,26	0,01	0,53	0,46	0,61	0,49	0,35	0,43
87	0,50	0,02	0,12	0,12	0,19	0,10	0,11	0,18	0,27	0,47
88	0,45	0,34	0,28	0,28	0,34	0,34	0,41	0,38	0,25	0,00
89	0,69	0,66	0,49	0,49	0,65	0,72	0,72	0,59	0,14	0,23
90	0,47	0,55	0,35	0,35	0,53	0,60	0,79	0,80	0,20	0,27
91	0,46	0,46	0,20	0,30	0,31	0,31	0,36	0,32	0,30	0,51
92	0,51	0,53	0,60	0,60	0,61	0,58	0,75	0,59	0,39	0,17
93	0,25	0,08	0,16	0,16	0,25	0,25	-0,03	0,22	0,37	0,40
94	0,54	0,31	0,42	0,58	0,51	0,34	0,41	0,46	0,14	0,25
95	0,33	0,32	0,42	0,42	0,18	0,34	0,28	0,36	0,33	0,34
96	0,63	0,31	0,38	0,38	0,51	0,49	0,27	0,28	0,33	0,10
97	0,43	0,65	0,54	0,54	0,25	0,32	0,56	0,53	0,44	0,23
98	0,49	0,36	0,35	0,35	0,40	0,40	0,26	0,34	0,13	0,31

Приложение приложения № 1

Номер итерации	MNLR -	MNLR +	C5 -	C5 +	RF -	RF +	SVM -	SVM +	RT -	RT +
99	0,32	0,49	0,35	0,45	0,69	0,69	0,54	0,50	0,44	0,18
100	0,60	0,27	0,50	0,50	0,52	0,44	0,25	0,06	-0,05	0,23
Среднее	0,44	0,38	0,32	0,32	0,40	0,39	0,40	0,38	0,26	0,25
Медиана	0,44	0,38	0,33	0,33	0,40	0,40	0,40	0,38	0,26	0,23
Стандартное отклонение	0,17	0,15	0,13	0,14	0,15	0,15	0,17	0,15	0,12	0,15

*Примечания: MNLR- множественная логистическая регрессия, C5 – дерево решений C5, RF – метод случайного леса, SVM – метод опорных векторов, RT – регрессионные деревья. Жирным цветом выделены минимальные, максимальные и средние значения, а также среднее, медиана и стандартное отклонение.

ПРИЛОЖЕНИЕ № 2

Сводная таблица, отражающая общую точность моделей при использовании в качестве ковариат почвенной карты

1984 г. (+) и без нее (-).

Номер итерации	MNLR -	MNLR +	C5 -	C5 +	RF -	RF +	SVM -	SVM +	RT -	RT +
1	56	70	45	45	50	50	60	65	35	50
2	80	60	65	65	75	75	65	60	50	60
3	65	65	65	65	75	75	60	60	40	45
4	70	60	40	40	50	56	56	45	35	40
5	45	56	65	70	70	70	70	56	45	45
6	85	56	50	50	60	60	56	60	50	56
7	60	65	65	60	70	65	70	65	45	60
8	60	60	75	70	70	65	70	70	35	65
9	75	50	60	50	65	60	60	50	50	65
10	56	80	56	56	70	70	70	75	60	40
11	70	65	56	56	60	60	70	70	65	45
12	60	56	45	45	45	40	40	50	60	35
13	50	35	65	65	56	56	35	56	50	60
14	50	60	70	70	80	80	65	60	56	35
15	70	60	70	70	60	60	56	50	56	56
16	60	56	56	50	60	60	56	56	30	70
17	60	75	70	70	75	75	75	65	40	56
18	65	60	56	50	50	50	65	65	45	30
19	35	40	45	50	56	50	45	45	65	56
20	35	50	70	70	60	50	65	56	40	50
21	80	60	56	56	70	70	56	70	56	50

Продолжение приложения № 2

Номер итерации	MNLR -	MNLR +	C5 -	C5 +	RF -	RF +	SVM -	SVM +	RT -	RT +
22	65	60	60	60	65	65	65	65	45	65
23	70	56	56	56	56	56	70	65	40	75
24	56	40	56	56	56	50	45	45	45	65
25	65	60	56	56	60	60	65	75	45	45
26	30	60	56	56	56	50	60	50	65	60
27	65	50	60	60	50	60	60	60	56	65
28	65	45	50	45	45	45	60	56	50	45
29	60	45	35	35	56	56	50	40	56	35
30	60	65	50	50	60	65	65	65	56	56
31	60	70	65	65	65	65	80	75	56	60
32	56	56	50	50	70	70	60	60	40	40
33	70	65	45	45	65	60	50	45	45	45
34	56	60	65	65	70	70	75	60	40	40
35	40	60	56	56	60	65	75	60	45	35
36	50	50	50	50	56	50	56	50	50	40
37	40	65	60	60	56	60	70	75	35	45
38	35	35	40	40	45	45	45	40	40	40
39	60	75	56	56	65	70	70	65	45	50
40	70	50	40	40	56	56	45	50	56	50
41	70	25	40	45	45	45	40	35	60	40
42	50	60	50	70	50	56	70	70	56	40
43	65	35	45	45	45	40	30	30	40	45
44	50	70	56	56	65	70	65	65	56	35
45	70	56	60	45	50	50	65	60	30	70

Продолжение приложения № 2

Номер итерации	MNLR -	MNLR +	C5 -	C5 +	RF -	RF +	SVM -	SVM +	RT -	RT +
46	50	60	60	60	56	50	56	60	45	40
47	60	56	65	65	65	65	60	60	50	45
48	60	65	70	70	65	65	80	70	35	40
49	45	40	40	45	65	60	70	56	45	40
50	60	56	56	56	56	60	60	56	65	56
51	70	50	50	50	60	60	50	40	40	50
52	75	65	65	65	65	65	75	50	65	45
53	45	50	70	70	60	60	60	60	70	35
54	60	65	60	56	45	45	60	56	56	75
55	56	60	50	50	60	60	45	60	45	45
56	45	56	45	45	50	45	45	50	65	50
57	70	45	45	45	35	40	40	40	50	35
58	75	40	56	56	60	60	45	30	56	40
59	70	50	60	60	85	80	45	45	35	56
60	70	65	70	65	70	70	70	60	40	56
61	70	50	56	56	65	60	60	60	45	35
62	60	50	45	45	50	50	60	50	45	40
63	85	75	70	70	70	70	65	70	40	30
64	60	60	65	65	70	70	56	60	45	60
65	75	60	75	70	75	75	70	60	56	50
66	45	35	50	50	50	45	56	40	56	56
67	85	40	45	45	45	45	45	45	45	60
68	65	70	56	56	65	65	75	70	70	50
69	45	60	56	56	65	65	56	60	70	35

Продолжение приложения № 2

Номер итерации	MNLR -	MNLR +	C5 -	C5 +	RF -	RF +	SVM -	SVM +	RT -	RT +
70	45	75	60	60	65	60	60	70	45	50
71	60	50	70	70	70	70	56	56	60	20
72	35	56	50	40	50	45	50	50	50	40
73	56	45	45	60	56	65	50	45	50	30
74	80	80	65	65	65	65	75	75	40	56
75	80	56	40	40	35	35	65	70	40	40
76	50	50	60	56	60	60	65	56	40	40
77	70	56	50	50	50	50	56	50	60	65
78	80	70	50	50	60	60	60	65	40	45
79	65	56	45	45	56	60	60	56	35	35
80	70	50	50	50	50	50	60	56	45	45
81	56	60	50	56	40	40	60	60	35	56
82	56	45	60	60	50	56	50	50	45	56
83	60	56	35	35	45	45	45	45	45	20
84	60	50	56	70	75	65	70	65	50	30
85	60	50	50	50	50	50	50	50	60	35
86	56	56	45	25	65	60	70	60	56	60
87	70	25	35	35	40	35	35	40	50	65
88	70	50	45	45	50	50	56	56	45	35
89	80	75	65	65	75	80	80	70	30	50
90	60	65	50	50	65	70	85	85	45	40
91	60	60	45	50	50	50	56	50	50	65
92	60	70	75	75	75	75	85	75	56	40
93	56	35	45	45	50	50	30	45	65	56

Продолжение приложения № 2

Номер итерации	MNLR -	MNLR +	C5 -	C5 +	RF -	RF +	SVM -	SVM +	RT -	RT +
94	65	50	60	70	65	56	60	60	35	40
95	56	50	60	60	45	56	50	56	50	56
96	75	50	60	60	65	65	45	45	50	45
97	65	75	70	70	50	56	70	70	60	40
98	65	60	65	65	65	65	56	60	35	56
99	56	65	60	65	80	80	70	65	60	35
100	70	50	70	70	65	60	50	40	30	56
Номер итерации	MNLR -	MNLR +	C5 -	C5 +	RF -	RF +	SVM -	SVM +	RT -	RT +
Среднее	61,13	56,16	55,68	55,63	59,28	58,9	59,35	57,14	48,51	47,81
Медиана	60,00	56,00	56,00	56,00	60,00	60,00	60,00	60,00	60,00	56,00
Стандартное отклонение	11,92	11,25	9,94	10,40	10,37	10,46	11,69	10,85	11,92	11,25

*Примечания: MNLR- множественная логистическая регрессия, C5 – дерево решений C5, RF – метод случайного леса, SVM – метод опорных векторов, RT – регрессионные деревья. Жирным цветом выделены минимальные, максимальные и средние значения, а также среднее, медиана и стандартное отклонение.

ПРИЛОЖЕНИЕ № 3

Сводная таблица, отражающая каппу самовоспроизводимости моделей множественной логистической регрессии при использовании в качестве ковариат почвенной карты 1984 г. (+) и без нее (-).

5 м (-)	20 м (-)	50 м (-)	5 м (+)	20 м (+)	50 м (+)
0,77	0,72	0,52	0,73	0,53	0,58
0,75	0,62	0,52	0,67	0,53	0,57
0,74	0,62	0,47	0,66	0,53	0,56
0,72	0,60	0,46	0,65	0,51	0,54
0,72	0,56	0,44	0,65	0,51	0,51
0,71	0,52	0,44	0,65	0,50	0,50
0,71	0,51	0,43	0,64	0,50	0,50
0,69	0,51	0,43	0,62	0,48	0,48
0,66	0,49	0,41	0,60	0,47	0,47
0,66	0,49	0,40	0,58	0,46	0,46
0,65	0,49	0,39	0,58	0,45	0,43
0,65	0,46	0,39	0,55	0,43	0,43
0,63	0,45	0,38	0,54	0,42	0,42
0,62	0,44	0,38	0,54	0,41	0,41
0,61	0,44	0,38	0,53	0,41	0,40
0,61	0,42	0,37	0,53	0,41	0,38
0,60	0,42	0,37	0,53	0,41	0,38
0,60	0,41	0,35	0,52	0,41	0,37
0,60	0,40	0,34	0,50	0,41	0,36
0,59	0,40	0,33	0,49	0,40	0,36
0,58	0,39	0,32	0,49	0,40	0,36
0,57	0,39	0,31	0,49	0,39	0,36
0,57	0,39	0,30	0,48	0,39	0,36
0,56	0,38	0,30	0,46	0,38	0,35
0,56	0,38	0,30	0,46	0,38	0,35
0,55	0,38	0,29	0,46	0,38	0,35
0,55	0,38	0,29	0,46	0,37	0,35
0,55	0,38	0,29	0,46	0,37	0,34
0,54	0,37	0,28	0,45	0,36	0,34
0,54	0,37	0,28	0,45	0,36	0,33
0,54	0,37	0,28	0,44	0,34	0,33
0,54	0,37	0,27	0,44	0,34	0,32
0,53	0,36	0,27	0,44	0,34	0,32
0,51	0,35	0,27	0,44	0,34	0,32
0,51	0,35	0,26	0,43	0,34	0,31
0,50	0,35	0,26	0,43	0,33	0,31
0,50	0,34	0,26	0,43	0,33	0,31
0,50	0,34	0,25	0,42	0,32	0,30

Продолжение приложения № 3

5 м (-)	20 м (-)	50 м (-)	5 м (+)	20 м (+)	50 м (+)
0,49	0,34	0,25	0,42	0,32	0,30
0,48	0,34	0,24	0,42	0,32	0,29
0,48	0,34	0,24	0,41	0,31	0,29
0,48	0,34	0,23	0,41	0,31	0,29
0,47	0,34	0,23	0,41	0,31	0,28
0,47	0,33	0,23	0,41	0,31	0,28
0,47	0,33	0,21	0,41	0,30	0,27
0,47	0,32	0,21	0,41	0,30	0,27
0,47	0,32	0,21	0,39	0,30	0,26
0,46	0,32	0,21	0,39	0,29	0,26
0,45	0,32	0,21	0,38	0,29	0,26
0,45	0,31	0,21	0,38	0,29	0,25
0,43	0,31	0,20	0,38	0,29	0,24
0,43	0,31	0,20	0,38	0,29	0,24
0,43	0,31	0,19	0,38	0,28	0,24
0,43	0,30	0,19	0,37	0,28	0,24
0,43	0,30	0,19	0,37	0,28	0,24
0,41	0,30	0,19	0,36	0,27	0,24
0,41	0,30	0,18	0,36	0,27	0,23
0,41	0,29	0,18	0,36	0,26	0,23
0,41	0,29	0,18	0,35	0,26	0,23
0,40	0,28	0,17	0,34	0,26	0,22
0,39	0,28	0,17	0,34	0,25	0,21
0,39	0,27	0,17	0,34	0,25	0,21
0,38	0,27	0,17	0,33	0,25	0,21
0,38	0,26	0,16	0,33	0,25	0,21
0,38	0,26	0,16	0,32	0,24	0,20
0,38	0,26	0,15	0,32	0,24	0,20
0,37	0,25	0,15	0,31	0,24	0,19
0,37	0,25	0,14	0,31	0,23	0,19
0,36	0,25	0,14	0,31	0,23	0,19
0,36	0,24	0,14	0,30	0,23	0,18
0,35	0,24	0,14	0,30	0,22	0,18
0,35	0,24	0,14	0,29	0,21	0,17
0,35	0,22	0,12	0,29	0,21	0,17
0,35	0,22	0,12	0,28	0,20	0,17
0,33	0,22	0,12	0,28	0,20	0,17
0,33	0,21	0,11	0,28	0,20	0,17
0,32	0,21	0,11	0,27	0,19	0,17
0,31	0,21	0,11	0,27	0,19	0,17
0,31	0,20	0,11	0,27	0,18	0,16

Продолжение приложения № 3

5 м (-)	20 м (-)	50 м (-)	5 м (+)	20 м (+)	50 м (+)
0,31	0,20	0,10	0,26	0,18	0,15
0,30	0,19	0,10	0,26	0,17	0,14
0,29	0,19	0,09	0,24	0,17	0,13
0,26	0,19	0,09	0,22	0,17	0,13
0,25	0,19	0,08	0,21	0,17	0,13
0,25	0,18	0,07	0,19	0,16	0,13
0,25	0,17	0,06	0,19	0,16	0,12
0,23	0,17	0,06	0,17	0,12	0,12
0,23	0,17	0,06	0,16	0,11	0,12
0,18	0,17	0,04	0,15	0,10	0,11
0,18	0,16	0,03	0,14	0,10	0,07
0,18	0,16	0,03	0,14	0,07	0,07
0,16	0,16	0,03	0,13	0,07	0,06
0,16	0,15	0,03	0,13	0,06	0,03
0,14	0,15	0,01	0,12	0,04	0,03
0,13	0,14	0,00	0,10	0,02	0,03
0,12	0,13	-0,01	0,08	0,01	0,01
0,09	0,06	-0,02	0,07	0,00	-0,01
0,09	0,05	-0,05	0,04	-0,02	-0,03
0,06	0,04	-0,05	0,02	-0,05	-0,07

ПРИЛОЖЕНИЕ № 4

Таблица сравнения 27-ми точек случайной независимой выборке

Обследование 2016 г.	Карта 2016 г.	Карта 1984 г.	MNLR		C5		RF		SVM		RT	
			средняя $k=0,44$	максим альная $k=0,77$	средняя $k=0,32$	максим альная $k=0,64$	средняя $k=0,40$	максим альная $k=0,75$	средняя $k=0,40$	максим альная $k=0,79$	средняя $k=0,26$	максим альная $k=0,52$
Π ^д	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓↓
Дг	Дг	Π ^д	Дг	Дг	Π ^д	Дг	Π ^д ↓	Π ^д ↓	Дг	Π ^д нам	Ал	Ал
Π ^д ↓	Π ^д ↓	Π ^д ↓↓	Π ^д ↓	Π ^д ↓								
Π ^д ↓	Π ^д ↓	Π ^д ↓↓	Π ^д ↓↓	Π ^д ↓↓	Π ^д ↓	Π ^д ↓	Π ^д ↓↓	Π ^д ↓↓	Π ^д ↓	Π ^д ↓↓	Π ^д ↓	Π ^д ↓
Π ^д нам	Π ^д нам	Дг	Π ^д ↓↓	Π ^д ↓↓								
Π ^д ↓↓	Π ^д ↓↓↓	Π ^д ↓	Π ^д ↓↓	Π ^д ↓↓	Π ^д ↓↓	Π ^д ↓↓	Π ^д ↓↓	Π ^д ↓↓↓	Π ^д ↓↓	Π ^д нам	Π ^д нам	Π ^д ↓↓
Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓↓
Π ^д ↓↓↓	Π ^д ↓↓↓	Π ^д ↓↓↓	Π ^д ↓↓	Π ^д ↓↓								
Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓	Π ^д ↓
Π ^д	Π ^д	Π ^д ↓	Π ^д ↓↓	Π ^д ↓↓								

Продолжение приложения № 4

П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓↓	П ^д ↓	П ^д ↓
П ^д _{нам}	П ^д _{нам}	П ^д	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓↓	П ^д ↓↓	П ^д ↓↓
П ^д ↓	П ^д ↓	П ^д	П ^д _{нам}	П ^д ↓	П ^д ↓	П ^д _{нам}	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д _{нам}
СЛ	СЛ	СЛ	П ^д _{нам}	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓
П ^д ↓↓	П ^д ↓↓	ДГ	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д _{нам}	П ^д ↓↓
П ^д ↓	СЛ	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д _{нам}	П ^д ↓
П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓
Ал	Ал	ДГ	Ал	Ал	Ал	Ал	Ал	Ал	Ал	Ал	Ал	Ал
П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓	П ^д ↓
П ^д	Ал	Ал	Ал	Ал	Ал	Ал	Ал	Ал	Ал	Ал	Ал	Ал
Ал	Ал	Ал	Ал	Ал	Ал	Ал	Ал	Ал	Ал	Ал	Ал	Ал
П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓	П ^д ↓↓
СЛ	СЛ	СЛ	П ^д	П ^д	П ^д	П ^д	П ^д	СЛ	СЛ	П ^д ↓↓	СЛ	П ^д

Продолжение приложения № 4

$\Pi^d \downarrow$	$\Pi^d \downarrow$	$\Pi^d \downarrow$	$\Pi^d \downarrow$	$\Pi^d \downarrow$	$\Pi^d \downarrow$	$\Pi^d \downarrow$	$\Pi^d \downarrow$	$\Pi^d \downarrow$	$\Pi^d \downarrow$	$\Pi^d \downarrow$	$\Pi_{\text{нам}}^d$	$\Pi^d \downarrow$	
$\Pi^d \downarrow \downarrow \downarrow$	$\Pi^d \downarrow \downarrow \downarrow$	$\Delta \Gamma$	$\Pi^d \downarrow \downarrow$										
$\Delta \Gamma$	$\Pi_{\text{нам}}^d$	$\Delta \Gamma$	$\Pi^d \downarrow \downarrow$										
$\Pi^d \downarrow$	$\Pi^d \downarrow$	Π^d	$\Pi^d \downarrow$										

*Примечания: MNLR - множественная логистическая регрессия, C5 – дерево решений C5, RF – метод случайного леса, SVM – метод опорных векторов, RT – регрессионные деревья, k – коэффициент каппа самовоспроизводимости модели.
Обозначения почв приведены в главе 3.

ПРИЛОЖЕНИЕ №5

Содержание органического углерода и физической глины в верхнем
гумусовом слое 0-20 см на сельскохозяйственных полях исследуемого
участка

Номер точки	Содержание органического углерода, %	Содержание физической глины, %
Объем выборки	101	166
Среднее	0,94	41
Ст. отклонение	0,47	5
Минимум	0,25	27
Нижний дециль	0,43	37
Нижний квартиль	0,65	38
Медиана	0,86	40
Верхний квартиль	1,21	37
Верхний дециль	1,43	47
Максимум	3,18	61