

Некоммерческое акционерное общество «Казахский агротехнический
исследовательский университет им. С. Сейфуллина»

УДК 528:332.3(574.31)(043.3)

На правах рукописи

ГРОССУЛ ПАВЕЛ ПАВЛОВИЧ

**Постклассификационный анализ динамики землепользования отдельных
районов Карагандинской области**

8D07302 – Геодезия

Диссертация на соискание степени
доктора философии (PhD)

Научные консультанты
кандидат технических наук
Ж.З. Толеубекова

доктор биологических наук,
профессор
О.Ә. Әліпбеки

Зарубежный научный консультант
доктор PhD,
ассоциированный профессор
Т.П. Купидура
(Варшава)

Республика Казахстан
Астана, 2025

СОДЕРЖАНИЕ

НОРМАТИВНЫЕ ССЫЛКИ.....	4
ОПРЕДЕЛЕНИЯ.....	5
ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ.....	6
ВВЕДЕНИЕ.....	7
1 АНАЛИЗ ДИНАМИКИ ЗЕМЛЕПОЛЬЗОВАНИЯ С ПРИМЕНЕНИЕМ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ И ГЕОИНФОРМАЦИОННЫХ СИСТЕМ.....	12
1.1 Современное состояние исследований динамики землепользования.....	12
1.2 Дистанционное зондирование Земли как инструмент изучения изменений землепользования	14
1.3 Современные подходы к классификации данных ДЗЗ.....	22
1.3.1 Теория методов классификации данных ДЗЗ	24
1.3.2 Интеграция ГИС и ДЗЗ в облачных вычислениях: роль Google Earth Engine в современных гео-исследованиях.....	26
1.3.3 Оценка точности классификации LULC	31
1.3.4 Сбор полевых эталонных данных.....	34
1.4 Постклассификационный анализ LULC	36
Выводы по разделу 1.....	39
2 СБОР И ПОДГОТОВКА ДАННЫХ ДЛЯ АНАЛИЗА ДИНАМИКИ LULC	40
2.1 Выбор исследуемой территории: обоснование и характеристика региона	41
2.2 Источники данных дистанционного зондирования	43
2.3 Сбор эталонных полевых данных	53
2.3.1 Проведение полевых работ	56
2.3.2 Интеграция полевых данных в ГИС	59
Выводы по разделу 2.....	61
3 КЛАССИФИКАЦИЯ ДАННЫХ LULC	62
3.1 Настройка модели машинного обучения на основе алгоритма Random Forest	65
3.2 Оценка качества классификации	67
3.2.1 Построение матриц ошибок	67
3.2.2 Метрики точности классификации	68
Выводы по разделу 3.....	71
4 РЕЗУЛЬТАТЫ КЛАССИФИКАЦИИ LULC И ИНТЕГРАЦИЯ СТАТИСТИКИ	73
4.1 Методы анализа изменений землепользования: пропорциональное распределение и матрица переходов на основе цепей Маркова.....	81
4.2 Интеграция статистических данных: социально-экономические показатели, демографические и экологические факторы; связь изменений землепользования с внешними факторами	86
4.2.1 Источники и типы статистической информации.....	86
4.2.2 Методы выявления связей и закономерностей СЭП и LULC.....	88

4.3 Регрессионный анализ.....	95
Выводы по разделу 4.....	97
ЗАКЛЮЧЕНИЕ	99
СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ	101
ПРИЛОЖЕНИЕ А – Акты внедрения.....	111
ПРИЛОЖЕНИЕ Б – Программный код для GEE	113
ПРИЛОЖЕНИЕ В – Карты LULC по 4 районам Карагандинской области.....	117

НОРМАТИВНЫЕ ССЫЛКИ

В настоящей диссертации использованы ссылки на следующие стандарты: Приказа Министра образования и науки Республики Казахстан. Об утверждении Правил присуждения степеней: утв. 31 марта 2011 года, №127 (зарегистрирован в Министерстве юстиции Республики Казахстан 13 мая 2011 года, №6951).

МИ ВНД 110.26–2023. Внутренние нормативные документы. Методическая инструкция. Порядок оформления и написания докторской диссертации.

ГОСТ 8.417-81. Государственная система обеспечения единства измерений. Единицы физических величин.

Об утверждении инструкций правил, регламентирующих порядок создания, обновления, использования Национальной инфраструктуры пространственных данных.

Закон Республики Казахстан. О геодезии, картографии и пространственных данных: принят 21 декабря 2022 года, №166-VII ЗРК.

Закон Республики Казахстан. О космической деятельности: принят 6 января 2012 года, №528.

Закон Республики Казахстан. О геодезии, картографии и пространственных данных: принят 21 декабря 2022 года, №166.

Кодекс Республики Казахстан. Водный кодекс Республики Казахстан: принят 9 июля 2003 года, №481.

Кодекс Республики Казахстан. Лесной кодекс Республики Казахстан: принят 8 июля 2003 года, №477.

Кодекс Республики Казахстан. Экологический кодекс Республики Казахстан: принят 2 января 2021 года, №400-VI.

Закон Республики Казахстан. О развитии агломераций: принят 1 января 2023 года, №181-VII.

Президент Республики Казахстан. Стратегия «Казахстан-2050»: новый политический курс состоявшегося государства: послание народу Казахстана.

ISO 19115-1:2014. Географическая информация. Метаданные. Часть 1. Основы. – Введ. 2014-04-01.

ОПРЕДЕЛЕНИЯ

В настоящей диссертации применяют следующие термины с соответствующими определениями:

Земной покров (Land Cover) – совокупность биотических и абиотических компонентов на поверхности Земли, включая растительность, почвы, водные объекты и антропогенные сооружения. Это физическая структура, регистрируемая средствами дистанционного зондирования [1].

Землепользование (Land Use) – формы и способы эксплуатации земли человеком, включая сельское хозяйство, урбанизацию, лесное хозяйство и рекреацию. Отражает социально-экономические аспекты использования территории [2].

NDVI (Normalized Difference Vegetation Index) – нормализованный индекс вегетации, отражающий плотность и активность растительности. Широко применяется в агроэкологических исследованиях [3].

EVI (Enhanced Vegetation Index) – усовершенствованный индекс вегетации, учитывающий атмосферные и фоновые помехи, обладающий большей чувствительностью к плотной растительности [3, p. 1156-1].

NDWI (Normalized Difference Water Index) – индекс водных объектов, основанный на различии отражения в ближнем ИК и зелёном спектрах. Применяется для выделения водоёмов [4].

NDBI (Normalized Difference Built-up Index) – индекс, позволяющий распознать застроенные территории на спутниковых изображениях на основе ближнего и среднего ИК-диапазонов [4, p. 201].

Классификация земного покрова (LULC classification) – процесс присвоения пикселям спутникового изображения классов землепользования (например, леса, пашни, пастбища и т.д.), часто реализуется с применением алгоритмов машинного обучения, включая Random Forest [5].

Random Forest (Случайный лес) – алгоритм ансамблевого машинного обучения, использующий множество решающих деревьев для повышения точности классификации. Эффективен при обработке разнородных и неполных данных [4, p7 201; 5, p. 1977].

Google Earth Engine (GEE) – облачная вычислительная платформа для анализа данных дистанционного зондирования, интегрирующая архивы спутниковых снимков и инструменты машинного обучения [6, 7].

CA-Markov модель – метод моделирования пространственно-временных процессов, основанный на использовании цепей Маркова в сочетании с клеточными автоматами. Применяется для прогнозирования изменений землепользования [8].

GNSS (Global Navigation Satellite Systems) – глобальные спутниковые навигационные системы (GPS, ГЛОНАСС и др.), применяемые для геопривязки и валидации полевых данных в ГИС-проектах [9].

ОБОЗНАЧЕНИЯ И СОКРАЩЕНИЯ

NDVI	– Normalized Difference Vegetation Index
EVI	– Enhanced Vegetation Index
NDWI	– Normalized Difference Water Index
MNDWI	– Modified Normalized Difference Water Index
NDBI	– Normalized Difference Built-up Index
NDMI	– Normalized Difference Moisture Index
NDTI	– Normalized Difference Tillage Index
UL	– Urban Land
LULC	– Land Use / Land Cover
CA-Markov	– Cellular Automata–Markov Chain
GEE	– Google Earth Engine
RF	– Random Forest
DEM	– Digital Elevation Model
SRTM	– Shuttle Radar Topography Mission
NASA	– National Aeronautics and Space Administration
USGS	– United States Geological Survey
DMSP-OLS	– Defense Meteorological Satellite Program – Operational Linescan System
VIIRS	– Visible Infrared Imaging Radiometer Suite
GNSS	– Global Navigation Satellite Systems
GCP	– Ground Control Point
TOA	– Top Of Atmosphere
InSAR	– Interferometric Synthetic Aperture Radar
PCA	– Principal Component Analysis

ВВЕДЕНИЕ

Общая характеристика работы.

Диссертационная работа посвящена разработке и внедрению современных методов анализа динамики землепользования, базирующихся на применении технологий дистанционного зондирования Земли (ДЗЗ) и геоинформационных систем (ГИС).

В условиях интенсификации антропогенного воздействия и ускоренных процессов изменения структуры землепользования, вызванных как природными, так и социально-экономическими факторами, исследование пространственных данных становится ключевым инструментом управления земельными ресурсами. Особое внимание в работе уделено разработке методики постклассификационного анализа, позволяющей выявлять пространственно-временные закономерности изменений землепользования с высокой степенью точности. Актуальность и необходимость разработки таких методов определяется требованиями к обеспечению устойчивого использования территорий и формированию стратегий, направленных на рациональное управление земельными ресурсами.

Актуальность работы.

Изменения структуры землепользования выступают одним из ключевых факторов, определяющих как экологическую устойчивость территорий, так и их социально-экономическое развитие, а также характер и эффективность использования природных ресурсов. Значимость данных трансформаций обусловлена тем, что они затрагивают не только локальные экосистемы, но и более масштабные, глобальные процессы. В их числе нарушение углеродного цикла, деградация экосистемных услуг и усиление климатических изменений [4, р. 201–210].

Lambin в работе [10] особо подчеркивает, что преобразования в системе землепользования оказывают существенное влияние на экосистемные функции. Речь идет, прежде всего, о воздействии на биоразнообразие и гидрологические процессы, что, как представляется, подтверждает необходимость детального изучения подобных изменений в отдельных регионах, включая Казахстан.

Казахстан, обладающий специфическими природно-климатическими условиями и выраженными экономическими особенностями, в настоящее время активно подвергается трансформациям землепользования. Это связано с ростом промышленного и сельскохозяйственного производства, усилением урбанизационных процессов, а также влиянием климатических изменений. Как показывают исследования, к числу наиболее заметных тенденций относятся сокращение площадей сельскохозяйственных угодий, трансформация пастбищных территорий и расширение урбанизированных зон [11].

Особый исследовательский интерес, представляют территории, где ландшафтные изменения протекают наиболее интенсивно. Так, например, в Таласском районе анализ спутниковых данных Landsat с использованием платформы Google Earth Engine позволил выявить устойчивые тенденции

деградации почвенного покрова и трансформации структуры экосистем [12, р. 268-1–268-28].

Современные методы дистанционного зондирования ДЗЗ и ГИС уже доказали свою высокую эффективность при мониторинге изменений землепользования. Например, использование платформы Google Earth Engine в комбинации с алгоритмами машинного обучения, такими как случайные леса, позволяет анализировать пространственно-временные изменения земного покрова и строить прогнозы на будущее [4, р. 201-210].

Казахстан сталкивается с особыми вызовами в управлении природными ресурсами и землепользованием. Для таких регионов, как Бурабайский район, интеграция данных дистанционного зондирования с экономическими и экологическими индикаторами обеспечили объективную оценку устойчивого развития [13].

Таким образом, актуальность данной работы определяется необходимостью разработки современных методов анализа динамики землепользования, основанных на технологиях ДЗЗ и ГИС, что позволит не только отслеживать изменения, но и разрабатывать научно обоснованные сценарии устойчивого управления земельными ресурсами.

Разработанная автором методика нашла применение в ряде практических задач. В частности, часть методик, разработанных в рамках настоящего исследования, была успешно апробирована и использована в научных проектах под номерами:

1. BR10764919 на тему «Исследование влияния государственной политики в отрасли сельского хозяйства на развитие кооперационных процессов в АПК, устойчивого развития сельских территорий и обеспечения» финансируемом Министерством сельского хозяйства Республики Казахстан.

2. BR22886730 на тему «Формирование аграрной инфраструктуры пространственных данных Республики Казахстан по технологиям и принципам 2.0» финансируемом Министерством сельского хозяйства Республики Казахстан.

В рамках реализованных проектов предложенные подходы продемонстрировали свою эффективность, что подтверждает их применимость не только для анализа землепользования Карагандинской области, но и возможность их адаптации для исследований в других регионах Казахстана.

Необходимость разработки методов постклассификационного анализа для Карагандинской области определяется как научной, так и практической значимостью данного направления. Создание точных и надёжных инструментов, позволяющих отслеживать динамику землепользования, выступает важным условием обеспечения устойчивого развития региона, а также повышения эффективности системы управления земельными ресурсами.

Цель работы.

Целью работы является обеспечение точного мониторинга динамики землепользования Карагандинской области путём разработки и апробации методики постклассификационного анализа на основе данных ДЗЗ и ГИС. Работа

направлена на создание эффективного инструментария для анализа пространственно-временных изменений, позволяющего учитывать специфику природных и антропогенных факторов, характерных для региона.

Основная цель работы заключается в адаптации существующих методов анализа к условиям Карагандинской области и создании новых подходов, позволяющих интегрировать данные для оперативного принятия управленческих решений.

Идея работы.

Идея исследования заключается в применении современных цифровых технологий дистанционного зондирования, пространственного анализа и геоинформационных систем для разработки методики анализа динамики землепользования, обеспечивающей точность и достоверность результатов.

Задачи исследования.

Для достижения поставленной цели исследования были сформулированы следующие задачи:

1. Провести анализ существующих методов изучения динамики землепользования, включая технологии дистанционного зондирования Земли (ДЗЗ) и геоинформационных систем (ГИС).
2. Разработать методику постклассификационного анализа, адаптированную к природным и социально-экономическим условиям Карагандинской области.
3. Осуществить апробацию предложенной методики на основе данных спутникового зондирования с целью оценки её практической эффективности.
4. Провести анализ динамики землепользования региона за многолетний период, а также выявить ключевые закономерности происходящих изменений и определяющие их драйвер-факторы.

Методы исследования.

Методология настоящего исследования, должна опираться на комплексное применение современных геодезических и информационных технологий. В ходе работы использовались многовременные спутниковые данные Landsat, обработка которых осуществлялась с применением алгоритмов машинного обучения в рамках классификации с учителем. Для последующей интеграции полученных результатов в геоинформационные системы применялись инструменты пространственного анализа, что позволило сформировать картографические модели и наглядно визуализировать изменения структуры землепользования.

Для обеспечения точности и достоверности результатов использовалась система метрологического контроля, включающая международные стандарты: ГОСТ 8.417-81 «Государственная система обеспечения единства измерений. Единицы физических величин», Закон Республики Казахстан от 21 декабря 2022 года №166-VII ЗРК «О геодезии, картографии и пространственных данных», Закон Республики Казахстан «О космической деятельности», Закон Республики Казахстан «О геодезии, картографии и пространственных данных», ISO 19115-1:2014. Географическая информация.

Положения, выносимые на защиту:

1. Разработанная методика постклассификационного анализа, по нашему мнению, обеспечивает высокую точность и надёжность результатов при исследовании динамики землепользования.
2. Использование спутниковых данных в сочетании с инструментами ГИС позволяет эффективно выявлять пространственно-временные закономерности изменений землепользования.
3. Применение системы метрологического контроля обеспечивает достоверность получаемых данных и способствует их корректной адаптации к региональным условиям исследования.
4. Полученные результаты могут быть использованы при разработке и обосновании стратегий устойчивого развития Карагандинской области.

Обоснованность и достоверность научных положений, выводов и рекомендаций подтверждается: достаточным объёмом собранных и обработанных данных дистанционного зондирования; применением методов математической статистики и анализа точности классификации; теоретическими разработками, основанными на современных подходах к пространственному моделированию и анализу данных в геоинформационных системах; анализом отечественного и зарубежного опыта мониторинга землепользования; результатами апробации разработанных методик постклассификационного анализа и инструментального контроля точности, а также успешным внедрением результатов исследования в рамках научного проекта и их применением в практических задачах управления земельными ресурсами Карагандинской области.

Научная новизна работы.

Научная новизна работы заключается в разработке адаптированной методики постклассификационного анализа, применимой к условиям Карагандинской области. В рамках работы впервые определены закономерности пространственно-временных изменений землепользования региона, предложена система метрологического контроля данных, а также реализован комбинированный подход, включающий постклассификационный анализ и пространственное моделирование.

Практическая значимость работы.

Результаты исследования будут полезны для:

- мониторинга и управления земельными ресурсами в Карагандинской области;
- разработки региональных стратегий устойчивого развития;
- внедрения методов постклассификационного анализа в систему государственного мониторинга землепользования.

Результаты исследования внедрены в рамках двух научных проектов, финансируемых Министерством сельского хозяйства Республики Казахстан. Целью этих проектов являлось развитие систем мониторинга землепользования и эффективного управления природными ресурсами, что подтвердило значимость и прикладной характер предложенных решений для решения задач на государственном уровне.

Также результаты научных исследований были внедрены в образовательный процесс (Приложение А) Казахского агротехнического исследовательского университета имени С. Сейфуллина. В частности:

Методики мониторинга землепользования и LULC-моделирования применяются в лекционных и практических занятиях по дисциплине «Мониторинг земель», а также используются при выполнении научных работ и подготовке выпускных квалификационных работ в рамках образовательной программы «6В07304 – Землеустройство»

Учебное пособие по геодезическим методам анализа землепользования с использованием машинного и глубокого обучения внедрено в курсы: «Дистанционное зондирование Земли», «Фотограмметрия и дистанционное зондирование», «Использование БПЛА в разных отраслях экономики», «Геоинформационные системы». Оно применяется в процессе подготовки студентов по программе «6В07302 – Геодезия и картография».

Личный вклад автора заключается в постановке цели и задач исследования; в разработке методики постклассификационного анализа землепользования, адаптированной к условиям Карагандинской области; в анализе и обработке данных дистанционного зондирования, включая многовременные спутниковые снимки; в интеграции данных в геоинформационные системы для построения классификационных карт и моделей изменений землепользования; в апробации разработанных методик на практике, включая создание классификационных карт районов Карагандинской области; в сравнительном анализе динамики землепользования с аналогичными исследованиями.

Апробация практических результатов работы.

Основные положения диссертационной работы докладывались, обсуждались и получили одобрение на следующих научных мероприятиях:

- международной научно-практической конференции «Сейфуллинские чтения–19» (Астана, 2023 – 17 марта);

- 6-й международной научно-практической конференции «Геодезия, землеустройство и кадастры: проблемы и перспективы развития» (Омск, 2024 – 28 марта);

- международной научной конференции «Фараби әлемі» (Алматы, 2024 – 4 апреля);

- 19-й международной научной конференции студентов, магистрантов и молодых ученых «ЛОМОНОСОВ–2024», посвященной 270-летию Московского университета (Астана, 2024 – 19 апреля);

- 1-й международной научно-практической конференции «Цифровая реальность: новые вызовы в картографии, ГИС и ДЗЗ», организованной Казахским национальным университетом имени аль-Фараби и компанией Data+ International LLC (Алматы, 2024 – 7 ноября);

- на 6th annual international scientific conference on Geoinformatics - GI 2024: “Sustainable Geospatial Solutions for a Changing World” (2024 – 13 November).

Публикации.

По теме диссертации опубликовано 13 работ, в том числе 4 статьи, входящие в базы данных Scopus и Web of Science, 4 статьи в материалах международных конференций, из них 1 докладов на международных зарубежных конференциях, входящих в базы данных Scopus и Web of Science, и 5 свидетельств о государственной регистрации прав на объекты авторского права.

Структура и объем работы.

Диссертационная работа состоит из введения, 4 разделов, заключения, списка использованных источников и приложений. Общий объем диссертации составляет 110 страниц, включает 16 таблиц, 20 рисунков и 3 приложения.

1 АНАЛИЗ ДИНАМИКИ ЗЕМЛЕПОЛЬЗОВАНИЯ С ПРИМЕНЕНИЕМ ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ ЗЕМЛИ И ГЕОИНФОРМАЦИОННЫХ СИСТЕМ

1.1 Современное состояние исследований динамики землепользования

Землепользование и земной покров являются фундаментальными понятиями в изучении взаимодействия человека и окружающей среды, а также в понимании глобальных экологических процессов. *Земной покров* (Land Cover) определяется как совокупность биотических и абиотических компонентов на поверхности Земли, включая растительность, почвы, водные объекты и искусственные сооружения [1, с. 3-270]. Это физическое выражение природных и антропогенных элементов, которые можно наблюдать и измерять с помощью дистанционного зондирования и других методов наблюдения [14].

Землепользование (Land Use) относится к способам, которыми люди используют землю и ее ресурсы для удовлетворения своих потребностей, включая сельское хозяйство, урбанизацию, лесное хозяйство, рекреацию и другие виды деятельности [2, р. e0200493]. Землепользование отражает социально-экономические и культурные практики общества и включает в себя управление и изменение земного покрова для определенных целей [1, с. 3-25].

Различие между земным покровом и землепользованием важно для понимания процессов изменения ландшафта. Земной покров представляет собой физическое состояние поверхности Земли, тогда как землепользование отражает функциональное назначение и способы эксплуатации этой поверхности человеком [14, р. 45-1-45-18]. Например, лес как тип земного покрова может использоваться для различных целей землепользования: лесозаготовки, рекреации или сохранения биоразнообразия.

Значение изучения динамики LULC заключается в том, что изменения в этих сферах оказывают прямое и косвенное влияние на глобальные экологические процессы, включая климатические изменения, гидрологические циклы, биоразнообразие и экосистемные услуги [15, 16]. Темпы обезлесения, лесовосстановления и роста лесов играют значительную роль в выбросах и поглощении углерода, что влияет на концентрацию парниковых газов в атмосфере и, следовательно, на глобальный климат [16, 17].

Землепользование влияет на доступность природных ресурсов, таких как пища, топливо, древесина и вода, что непосредственно влияет на качество жизни и устойчивость человеческих сообществ [2, р. e0200493]. Кроме того, изменения в землепользовании могут приводить к утрате местообитаний, фрагментации экосистем и снижению биоразнообразия, что подчеркивает необходимость устойчивого управления земельными ресурсами [1, с. 10-29; 15, р. 511-526].

Значимость пространственных данных в анализе LULC трудно переоценить. Высокоточные пространственные данные, полученные с помощью ДЗЗ, глобальных навигационных спутниковых систем (GNSS) и беспилотных летательных аппаратов (БПЛА), позволяют проводить детальный мониторинг и картографирование изменений на различных масштабах [9, р. 16911-16922; 18].

Дистанционное зондирование включает измерение электромагнитного излучения, отраженного от объектов на поверхности Земли, что позволяет классифицировать и анализировать типы земного покрова [14, р. 45-1-45-18; 19].

Использование ГИС и технологий дистанционного зондирования предоставляет возможность интегрировать различные источники пространственных данных, проводить анализ тенденций и моделировать будущие сценарии изменений землепользования [20, 21]. Точное картографирование и мониторинг сельскохозяйственных земель становятся важными инструментами для оценки и прогнозирования урожайности, а также для изучения влияния изменений климата на сельскохозяйственные системы [22]. Методы изучения динамики LULC постоянно совершенствуются с развитием технологий. Дистанционное зондирование и ГИС-технологии являются основными инструментами для сбора и анализа пространственных данных о земной поверхности [18, р. 2338-2349; 20, р. 947-972]. Спутниковые данные, в частности изображения Landsat и Sentinel-2, широко применяются для мониторинга изменений землепользования как на региональном, так и на глобальном уровнях [23]. По нашему мнению, их использование формирует базовую информационную основу для долгосрочного анализа трансформаций территорий.

Применение технологий GNSS и беспилотных летательных аппаратов (БПЛА) позволяет получать высокоточные данные с детальным пространственным разрешением, что особенно важно при проведении локальных исследований [9, р. 16911-169-22; 24, р. 277-1-277-25]. В частности, БПЛА активно используются для мониторинга состояния сельскохозяйственных культур, оценки ущерба, вызванного природными явлениями, а также для построения цифровых моделей рельефа [25].

Алгоритмы машинного обучения и методы обработки больших данных в последние годы приобретают всё большее значение в анализе LULC. Использование таких подходов, как случайные леса (Random Forest) и технологии глубокого обучения, позволяет существенно повысить точность классификации и эффективно работать с крупными массивами данных [26]. Дополнительные возможности предоставляет применение облачных вычислительных платформ, например Google Earth Engine, которые упрощают доступ к данным дистанционного зондирования и обеспечивают инструментарий для их комплексного анализа [6, р. 109-117].

Современные тенденции и ключевые вызовы в изучении динамики землепользования, на наш взгляд, во многом связаны с необходимостью интеграции различных технологических решений для повышения точности и оперативности мониторинга. Комбинирование данных дистанционного зондирования, GNSS и БПЛА открывает новые перспективы для проведения комплексных исследований [27]. Вместе с тем сохраняется ряд проблем, среди которых — ограниченная доступность данных, недостаточная стандартизация методов, а также потребность в разработке и внедрении новых методологических подходов [18, р. 2338-2349; 28, р. 6812-6823].

В Казахстане и странах Центральной Азии исследования в области землепользования в последние годы заметно активизировались, однако по-прежнему сталкиваются с рядом ограничений. К ним, прежде всего, относятся дефицит доступных данных и необходимость более широкого внедрения современных технологий и алгоритмов анализа [13, р. 3251-3264; 28, р. 6812-6823]. В то же время международный опыт демонстрирует высокую эффективность применения интегрированных подходов и технологических решений при решении сложных задач мониторинга и управления землепользованием [29].

Исследования землепользования в Казахстане сосредоточены на локальных примерах, включая ключевые сельскохозяйственные и урбанизированные регионы. Например, в работе О. Alipbeki [11, р. 321-347] изучали изменения землепользования в пригородной зоне Аршалынского района, прилегающего к столице. С использованием данных Landsat за период 1998–2018 гг. Исследователи зафиксировали сокращение пастбищ, увеличение пахотных земель и рост урбанизированных зон, что свидетельствует о возрастающем антропогенном воздействии.

Для более широкого охвата Центральной Азии Hu & Yang [28, р. 6812-6823] провели анализ изменений землепользования за период 2001-2017 гг., используя данные Landsat и алгоритм Random Forest. Исследование показало увеличение орошаемых земель и сокращение площадей с естественной растительностью, что связано с климатическими изменениями и социальными трансформациями.

Международный опыт в изучении динамики землепользования активно использует платформу Google Earth Engine (GEE) для обработки больших объемов спутниковых данных. Например, N. Kadri [30] изучили изменения землепользования в полупустынных регионах Туниса за последние четыре десятилетия, выявив сокращение лесных массивов и увеличение урбанизированных зон.

Работа Nasiri et al. Сосредоточена на совершенствовании методов классификации землепользования с использованием алгоритмов Random Forest и объектно-ориентированных подходов. Эти исследования подчеркивают преимущества интеграции данных Sentinel и Landsat для повышения точности классификаций [31].

Несмотря на значительный прогресс в Казахстане и Центральной Азии, исследования землепользования сталкиваются с рядом проблем. Основные из них – ограниченность доступных данных, недостаточная интеграция современных алгоритмов машинного обучения и отсутствие долгосрочных прогнозных моделей. В то время как международный опыт демонстрирует успешное использование GEE и других облачных платформ, в Казахстане этот подход только начинает внедряться.

1.2 Дистанционное зондирование Земли как инструмент изучения изменений землепользования

Дистанционное зондирование Земли представляет собой один из ключевых инструментов изучения долгосрочных изменений в использовании территорий. По нашему мнению, его значимость обусловлена возможностью систематического наблюдения за объектами земной поверхности и оценки их характеристик без необходимости непосредственного контакта с ними. Получение такой информации обеспечивается за счёт данных, регистрируемых сенсорами, размещёнными на воздушных и космических носителях — самолётах и спутниках [3, р. 1556-1-1556-14; 32].

Указанные сенсоры функционируют в различных диапазонах электромагнитного спектра и фиксируют широкий спектр параметров земной поверхности — от энергетических характеристик отражённого излучения до показателей его поляризации. Регистрируемая информация представляется в виде двумерной пространственной сетки, что, как представляется, делает возможным формирование космических изображений.

Полученные изображения находят широкое применение при анализе состояния суши, Мирового океана и атмосферы, позволяя интерпретировать свойства природных и антропогенных объектов на основе их уникальных спектральных характеристик [3].

Космические системы являются ключевыми инструментами получения данных дистанционного зондирования, обеспечивающими уникальные возможности для мониторинга состояния Земли. Современные системы включают в себя космические аппараты, аппаратуру для дистанционного зондирования, средства передачи и обработки данных, а также наземные комплексы для их интерпретации [33, 34].

Космический аппарат, оборудованный сенсорами, осуществляет сбор отражённой солнечной энергии, преобразует её в электрический сигнал и передаёт на наземные пункты для последующей обработки. Такие системы характеризуются различными видами разрешения: спектральным, пространственным, яркостным и временным, что определяет их функциональность и сферу применения [32, р. 245]:

1. Спектральное разрешение – способность сенсора различать длины волн в различных диапазонах электромагнитного спектра. Например, мультиспектральные и гиперспектральные снимки широко используются для изучения растительности, мониторинга состояния лесов и сельскохозяйственных культур [3, р. 1556-2-1556-5].

2. Пространственное разрешение – минимальный размер объекта, различимого на изображении. Современные сенсоры обеспечивают пространственное разрешение от километров до долей метра [32, р. 245].

3. Яркостное (радиометрическое) разрешение – количество градаций серого на изображении, влияющее на детализацию и интерпретацию снимков. Большинство современных сенсоров имеют разрешение 6 или 8 бит [14, р. 45-3-45-6].

4. Временное разрешение – частота повторной съёмки одной и той же территории, что важно для анализа динамики изменений [35].

Первые космические системы, такие как “Метеор” в СССР и Landsat в США, были ориентированы на метеорологические исследования и мониторинг природных ресурсов. Например, первый спутник Landsat, запущенный в 1972 году, позволил проводить мультиспектральную съёмку земной поверхности для решения задач сельского хозяйства, лесного хозяйства и экологического мониторинга.

Современные спутники, такие как Landsat-8 и Sentinel, оснащены высокоразрешающими сенсорами, позволяющими проводить детальный анализ состояния растительного покрова, обнаруживать пожары, изучать загрязнение и прогнозировать урожайность. Landsat-8, например, предоставляет данные с пространственным разрешением до 15 метров в панхроматическом режиме и 30 метров в мультиспектральных каналах [34, p. 2903-1-2903-15].

Космические данные находят применение в самых разных сферах. Они используются для мониторинга растительности – с помощью мультиспектральных снимков можно оценивать состояние лесов, определять объем биомассы и выявлять болезни растений [36]. Также такие данные необходимы при составлении тематических карт, которые помогают анализировать землепользование и распределение природных ресурсов [35, p. 5973-5990].

Особую роль космическая информация, на наш взгляд, играет в системе экологического мониторинга: она позволяет оперативно отслеживать изменения в экосистемах и оценивать степень антропогенного воздействия на природную среду [34, p. 2903-1-2903-39]. Кроме того, технологии дистанционного зондирования широко применяются для прогнозирования и выявления чрезвычайных ситуаций, включая лесные пожары, загрязнение окружающей среды, а также иные природные и техногенные катастрофы [14, p. 45-1-45-18].

Современные системы получения космических изображений представляют собой сложные интегрированные технологические комплексы, обеспечивающие получение высокоточной информации о состоянии земной поверхности. По мере их развития дистанционное зондирование становится всё более значимым инструментом в решении задач сельского и лесного хозяйства, охраны окружающей среды, а также обеспечения принципов устойчивого развития.

Спутниковая программа Landsat, реализуемая совместно Национальным управлением по аэронавтике и исследованию космического пространства (NASA) и Геологической службой США (USGS), по праву рассматривается как одна из наиболее продолжительных и значимых инициатив в области дистанционного зондирования Земли. Она предоставляет уникальные данные для систематического мониторинга поверхности планеты с регулярной временной дискретностью, что делает её важнейшим инструментом пространственного анализа.

В рамках настоящего исследования, посвящённого изучению динамики землепользования на территории Карагандинской области (включающей Осакаровский, Абайский, Бухар-Жырауский и Каркаралинский районы), были

использованы изображения со спутников Landsat 5, 7 и 8. Применение данных Landsat, по нашему мнению, обеспечивает широкие возможности как для долгосрочного наблюдения, так и для ретроспективного анализа, позволяя фиксировать не только краткосрочные изменения земной поверхности, но и продолжительные процессы трансформации ландшафтов.

Спутники серии Landsat предоставляют мультиспектральные изображения с пространственным разрешением 30 метров, что является оптимальным показателем для анализа динамики землепользования на региональном уровне. Данное разрешение позволяет выявлять изменения на уровне отдельных ландшафтных элементов, включая сельскохозяйственные угодья, лесные массивы и урбанизированные территории. Дополнительное использование панхроматических каналов на спутниках Landsat 7 и 8, имеющих разрешение 15 метров, существенно повышает детализацию изображений и, как следствие, точность пространственного анализа.

Спутник Landsat 5 (1984–2013) оснащён сенсором ТМ (Thematic Mapper), обеспечивающим получение данных в семи спектральных диапазонах, предназначенных для исследования земной поверхности. Его высокая стабильность функционирования и продолжительный срок эксплуатации сделали Landsat 5 одним из ключевых источников информации для ретроспективного анализа динамики землепользования. На протяжении почти трёх десятилетий спутник обеспечивал непрерывное наблюдение за земной поверхностью, что позволило сформировать уникальные временные ряды, используемые при мониторинге изменений ландшафтов.

Спутник Landsat 7, запущенный в 1999 году, стал важным этапом эволюции программы Landsat, прежде всего благодаря внедрению сенсора ЕТМ+ (Enhanced Thematic Mapper Plus). Использование данного сенсора позволило существенно улучшить спектральные и пространственные характеристики изображений, обеспечив получение более детализированных данных. Однако, начиная с 2003 года, эксплуатация спутника была осложнена сбоем в системе линейного сканирования (SLC-off). Указанный технический дефект привёл к появлению полосатых пропусков данных, охватывающих до 22% площади снимка.

Проблема особенно ярко проявляется при съёмке обширных территорий, где пространственные искажения становятся значительными, что, на наш взгляд, ограничивает использование данных Landsat 7 в «чистом» виде при решении картографических и аналитических задач.

Наличие дефекта SLC-off стало определяющим фактором при выборе подходов к обработке изображений Landsat 7 в рамках настоящего исследования. В связи с частичной потерей данных снимки данного спутника использовались преимущественно как вспомогательный материал — для восполнения временных пробелов в рядах Landsat 5 и Landsat 8. С целью минимизации влияния пропусков применялись методы интерполяции и маскирования дефектных участков, однако, как представляется, они не всегда обеспечивали полное устранение артефактов.

Тем не менее, несмотря на существующие ограничения, данные Landsat 7 сохраняли аналитическую ценность при исследовании территорий с низкой облачностью, где пространственные искажения поддавались корректировке. Таким образом, спутник сыграл важную роль в обеспечении непрерывности временных рядов мониторинга.

С запуском Landsat 8 в 2013 году программа Landsat получила значительное технологическое обновление, позволившее существенно повысить качество получаемых данных. На борту спутника функционируют два сенсора: OLI (Operational Land Imager), обеспечивающий мультиспектральную съёмку с повышенной радиометрической чувствительностью, и TIRS (Thermal Infrared Sensor), предназначенный для регистрации теплового излучения земной поверхности. По нашему мнению, Landsat 8 зарекомендовал себя как один из наиболее надёжных источников дистанционной информации: стабильность работы сенсоров и высокая точность данных делают его ключевым инструментом современного мониторинга состояния земной поверхности.

Запуск спутника Landsat 9 в 2021 году стал очередным этапом развития программы, обеспечив её технологическую преемственность и непрерывность по отношению к Landsat 8. Новый аппарат оснащён более точно откалиброванными сенсорами с повышенной чувствительностью, что позволяет получать изображения с минимальным уровнем искажений. Благодаря данным улучшениям Landsat 9, на наш взгляд, особенно эффективен при формировании временных рядов и отслеживании изменений землепользования, в том числе при совместном использовании с архивными данными предыдущих поколений спутников Landsat.

Использование данных Landsat в рамках настоящего исследования территории Карагандинской области, по нашему мнению, имеет стратегическое значение для оценки пространственно-временных изменений LULC. Преемственность архивов наблюдений, а также высокая пространственная и спектральная разрешающая способность позволяют фиксировать ключевые процессы трансформации ландшафтов, в том числе:

1. Изменения сельскохозяйственных угодий и пастбищ под воздействием хозяйственно-экономической деятельности.
2. Деграцию почвенного покрова и растительности в результате антропогенного влияния.
3. Расширение урбанизированных территорий и инфраструктурных объектов в Абайском и Бухар-Жырауском районах.
4. Восстановление лесных массивов и растительного покрова в Каркаралинском районе вследствие реализации природоохранных мероприятий.

Долгосрочные временные ряды данных Landsat обеспечивают возможность анализа как краткосрочных изменений, так и устойчивых многолетних трендов. Благодаря высокой калибровочной точности и унификации данных между различными поколениями спутников программа Landsat, как представляется, остаётся незаменимым инструментом

пространственного анализа. Это особенно важно при решении задач устойчивого управления земельными ресурсами и разработке стратегий сохранения природных экосистем.

Данные Landsat, несмотря на известные технические ограничения, связанные с функционированием Landsat 7, позволили осуществить детализированный и систематический мониторинг динамики землепользования. Их интеграция с иными источниками дистанционного зондирования, а также с геоинформационными системами обеспечила получение достоверной и научно обоснованной информации, необходимой для принятия управленческих решений в сфере природопользования и территориального планирования.

Shuttle Radar Topography Mission (SRTM), реализованная совместными усилиями NASA и NGA в феврале 2000 года, представляет собой значимый этап в создании глобальных цифровых моделей рельефа (Digital Elevation Models, DEM). В результате миссии были получены топографические данные, охватывающие около 80% поверхности Земли между 60° северной и 56° южной широты, с пространственным разрешением 90 м для глобального покрытия и 30 м для территории США и отдельных приоритетных регионов. Точность данных и их покрытие сделали SRTM ключевым источником информации для пространственного анализа и моделирования природных и антропогенных процессов.

Основным методом получения данных в рамках SRTM стала интерферометрическая радиолокация (InSAR), использующая два радиолокатора для измерения разницы фаз отражённого сигнала, что позволяет строить точные модели высот с минимальными погрешностями. Благодаря своей доступности и практически глобальному охвату данные SRTM, по нашему мнению, предоставили исследователям важный инструментарий для более углублённого изучения пространственной структуры ландшафтов и динамики LULC.

Исследования LULC представляют собой многоуровневый аналитический процесс, позволяющий рассматривать изменения как в пространственном, так и во временном измерении. Включение рельефной информации в модели LULC открывает дополнительные возможности для детализации изучаемых процессов, объяснения закономерностей распределения классов землепользования, а также учёта природных ограничивающих факторов. В этом контексте данные SRTM играют значимую роль, что обусловлено их высокой точностью и способностью достоверно отражать топографические особенности территории.

Топографические параметры — такие как абсолютная высота, уклон поверхности, экспозиция склонов, кривизна рельефа и особенности гидрографической сети — выступают ключевыми факторами, определяющими пространственное распределение различных типов землепользования.

Эти параметры определяют природные условия и пригодность земель для различных видов деятельности:

В горных районах, например в Каркаралинском районе, значительные уклоны и перепады высот способствуют формированию лесных массивов,

природных экосистем и зон с минимальным антропогенным воздействием. Высокие склоны ограничивают возможности для ведения сельского хозяйства и строительства.

В степных и равнинных зонах, таких как Осакаровский район, минимальные уклоны и малые перепады высот создают благоприятные условия для сельскохозяйственного использования и развития пастбищ.

Урбанизированные территории, как в Абайском районе, преимущественно концентрируются на плоских участках с минимальным уклоном, где экономически и технически целесообразно строительство инфраструктуры.

Кроме того, различия в уклоне и экспозиции склонов влияют на солнечное освещение и влажностный режим, что в свою очередь определяет распределение растительных сообществ и продуктивность сельскохозяйственных земель.

Данные SRTM позволяют выявлять топографические ограничения, критически важные для устойчивого управления земельными ресурсами:

Зоны эрозионного риска на крутых склонах, где интенсивное землепользование может вызвать ускоренную деградацию почв.

Потенциальные зоны для лесовосстановления, где уклон и высота создают условия для естественного и искусственного восстановления растительного покрова.

Оптимальные зоны для сельскохозяйственных культур, где можно минимизировать потери урожайности за счёт учёта уклона и почвенных характеристик.

Топографические данные SRTM играют ключевую роль в моделировании гидрологических процессов. Рельеф местности определяет формирование водосборных бассейнов, направление стока и зоны подтопления. Это особенно важно для определения границ водосборных бассейнов и моделирования их структуры, локализации пойменных территорий и затопляемых зон, что имеет значение для управления рисками наводнений, мониторинга динамики речных систем и озёр, особенно в регионах с сезонным изменением водного режима.

Рельеф оказывает значительное влияние на пространственное планирование и урбанизацию. Использование данных SRTM позволяет оценивать экономическую эффективность строительства на разных участках рельефа, прогнозировать рост городских территорий с учётом природных ограничений и барьеров, оптимизировать трассировку транспортных и инженерных коммуникаций, особенно в горных районах.

Совмещение данных SRTM с изображениями дистанционного зондирования (Landsat) и ГИС позволило разработать многофакторные модели LULC, что включает:

1. Повышение точности классификации землепользования за счёт учёта топографических параметров.
2. Прогнозирование динамики распределения классов LULC с учётом изменений рельефа и антропогенных факторов.
3. Моделирование пространственных закономерностей взаимодействия природных и антропогенных систем.

Данные SRTM представляют собой неотъемлемый компонент анализа землепользования и покрова земли благодаря высокой точности и широкому охвату. Они позволяют учитывать топографические факторы, влияющие на устойчивость экосистем, пространственное распределение LULC и гидрологические процессы. Интеграция данных SRTM с ГИС и спутниковыми изображениями даёт возможность проводить углублённый анализ и моделирование, создавая основу для принятия обоснованных решений в сфере устойчивого управления земельными ресурсами, пространственного планирования и природоохранных мероприятий.

Ключевое преимущество данных ДЗЗ заключается в его способности предоставлять данные о земной поверхности в условиях недоступности для других методов. Космические снимки обладают высокой обзорностью, регулярной повторяемостью и оперативностью поступления данных, что делает их особенно ценными для мониторинга крупных территорий [35, р. 5973-5990; 37].

Однако атмосферные эффекты, такие как облачность, рассеяние и поглощение солнечного излучения, значительно усложняют обработку данных. Облачность покрывает более 50% земной поверхности в любой момент времени, а атмосферная дымка снижает контрастность и точность изображений. Например, водяной пар, углекислый газ и озон в атмосфере поглощают солнечное излучение на определённых длинах волн, создавая «провалы» на спектрограмме [14, р. 45-1-45-18; 38, р. 04007-1-04007-11].

Для минимизации влияния этих факторов дистанционная съёмка проводится в «окнах прозрачности» атмосферы – областях спектра, где излучение проходит с минимальным ослаблением. Эти окна включают видимый и ближний инфракрасный диапазоны (0,3–1,3 мкм), а также радиодиапазон, где наблюдается наибольшая прозрачность атмосферы [3, р. 1556-1-1556-14].

Системы ДЗЗ фиксируют изменения отражённой и излучённой объектами энергии в определённых диапазонах длин волн. Такие данные используются для классификации растительного покрова, мониторинга землепользования и оценки воздействия природных и антропогенных факторов на окружающую среду [36, р. 2630].

Современные технологии, включая использование данных Landsat и Sentinel, позволяют проводить высокоточные исследования землепользования на больших территориях. Например, спутниковые снимки Landsat, обработанные с помощью автоматизированных алгоритмов, используются для картографирования сельскохозяйственных земель и анализа их изменений [36, р. 2630].

Преимущества ДЗЗ для изучения землепользования состоят прежде всего в его способности обеспечивать глобальный, регулярный и объективный мониторинг поверхности Земли [39, 40]. Данные ДЗЗ упрощают выявление пространственных паттернов землепользования и отслеживание их изменений во времени, что крайне важно для устойчивого развития, рационального природопользования, оценки продуктивности агроэкосистем, анализа

климатических тенденций и решения задач экологического планирования [41-43]. Благодаря информации, получаемой от высокоспектральных и мультиспектральных сенсоров, возможно детальное картографирование землепользования, дифференциация растительных сообществ и земных покровов, уточнение параметров почвенно-ландшафтных комплексов и оценка состояния водных объектов [44, 45].

Тем не менее, несмотря на значительные преимущества, существует ряд сложностей. Обработка больших объемов данных Earth observation (EO) требует высоких вычислительных ресурсов и внедрения современных методологий машинного обучения, нейронных сетей и облачных вычислительных платформ [7, p. 5770-5780; 46-49]. Критически важна и проблема валидации: оценка точности классификаций землепользования, сравнение продуктов разных спутников и обеспечение согласованности временных рядов остаются вызовами для научного сообщества [50, 51]. Различия в радиометрической, спектральной и геометрической калибровке данных от разных платформ усложняют интеграцию разнородной информации [52-55]. Ограниченность или платность доступа к наборам данных таким как WorldView снижает глобальную доступность и препятствует равному участию стран со средним и низким доходом [56, 57].

Применение ДЗЗ для анализа изменений землепользования уже доказало свою ценность при решении задач международного масштаба, таких как мониторинг достижения Целей устойчивого развития ООН (SDGs), разработка стратегий адаптации к климатическим изменениям, планирование урбанизации и оценка эффективности природоохранных мер [39, p. 3-374; 40, p. 689-696; 42, p. 105340; 58]. Например, длительные временные ряды Landsat позволяют проводить ретроспективный анализ земельных преобразований за несколько десятилетий, выявлять тенденции обезлесения, опустынивания, расширения аграрных площадей и урбанизированных зон [59, 60]. Спутники Sentinel, сочетая оптические и радиолокационные данные, обеспечивают частое обновление карт землепользования, помогают выявлять динамику посевов, влажность почвы, состояние лесов и водных ресурсов с более высокой пространственной детализацией [56, p. 207-224].

Для более глубокого понимания процессов, происходящих на поверхности Земли, перспективным направлением является использование многосенсорных данных, включая радарные (Sentinel-1), тепловые (ECOSTRESS) и гиперспектральные наблюдения, а также интеграция с наземными, аэрокосмическими и социально-экономическими данными [61, 62]. Это позволит получать еще более точные оценки, сопоставлять результаты для разных регионов, отслеживать межгодовые флуктуации и прогнозировать будущие сценарии землепользования [63, 64].

Таким образом, ДЗЗ играет решающую роль в понимании пространственно-временной динамики землепользования и в формировании научно обоснованных стратегий управления природными ресурсами. Преимущества масштабирования, объективности и регулярности наблюдений

сочетаются с инновационными методами обработки и интеграции данных, позволяя преодолевать существующие барьеры в анализе и интерпретации. Несмотря на технические, методологические и организационные сложности, прогресс в технологиях ДЗЗ, открытых данных, алгоритмах машинного обучения и облачных решениях предоставляет исследователям мощные инструменты для мониторинга изменений землепользования в интересах устойчивого развития [39, р. 3-370; 40, р. 689-696; 41, р. 111428; 42, р. 105340; 43, р. 2495-1-2495-30].

1.3 Современные подходы к классификации данных ДЗЗ

Современные подходы к классификации данных ДЗЗ тесно связаны с развитием спектральных индексов, комбинированных показателей и интеграцией мультиспектральных и радиолокационных данных, а также с внедрением методов машинного обучения и нейронных сетей. Так, в работе [65, р. 110-124] подчеркнуто, что использование ДЗЗ является ключевым инструментом для оценки динамики деградации и опустынивания земель, включая интеграцию показателей, полученных из мультиспектральных данных, и их последующую классификацию для мониторинга состояния растительности и почв. В исследованиях М.А. Abdel [66-68] отмечено, что современные решения по классификации всё чаще используют многокомпонентные подходы, совмещающие в себе различные параметры, включая растительные индексы, характеристики почв и данные о покрытии земель, что позволяет достичь более точной типологизации территорий и оценки их деградационного статуса. Кроме того, ряд авторов [69, 70] указывают, что внедрение комплексных индексов (например, Land Degradation Index, Desertification Monitoring Index или Environmentally Sensitive Area Index), основанных на комбинации спектральных признаков и статистических методов, способствует более точному выделению классов земель по степени подверженности деградации или опустыниванию, повышая согласованность классификации.

Одним из существенных направлений развития классификационных методик является всё более широкое привлечение временных рядов данных. Исследователи [71, 72] показывают, что анализ динамики на больших временных промежутках с помощью MODIS, Landsat и Sentinel позволяет применять современные методы для классификации, учитывающие сезонность, фенологию и долгосрочные тренды, в том числе RESTREND и RUE, а также индексы на основе NPP. Подобные методы, по данным [30, р. 3257-1-3257-15; 73, 74], способны выделять классы состояния земель, отличающиеся по степени антропогенного воздействия или климатической уязвимости, что приводит к более надежному выделению классов деградационных процессов.

Большое внимание в современной литературе уделяется развитию подходов к классификации, основанных на методах машинного обучения и искусственного интеллекта. Работы [71, р. 11-46; 75, 76] отмечают рост применения алгоритмов случайных лесов (RF), деревьев принятия решений (CART) и нейронных сетей (в том числе глубокого обучения) для классификации земельных покровов на основе мультииндексных комбинаций. Так, в [71, р. 11-

46; 76, p. 5077-5096; 77] было показано, что сочетание машинного обучения и многомерных спектральных данных повышает точность классификации, поскольку модели могут учитывать нелинейные зависимости между входными параметрами.

Немаловажным трендом является применение методов слияния данных (data fusion), объединяющих оптические и радиолокационные источники. По данным [68, p. 3-130; 72; 75, p. 134319], комбинирование, например, SAR (Sentinel-1) и оптических данных (Sentinel-2, Landsat, MODIS) существенно улучшает качество классификации, поскольку радиолокационные данные не подвержены атмосферным помехам и дают дополнительную информацию о структурных особенностях поверхности. Исследования [71, p. 11-46; 76, p. 5077-5096] подтверждают, что такой гибридный подход повышает устойчивость классификации к сезонным вариациям облачности и изменчивости освещенности, а также улучшает различение классов земельных покровов с низкой растительностью.

Современные классификационные подходы, согласно ряду авторов [74, p. 7561-7569; 78, с. 216-218], всё чаще включают не только спектральные признаки и производные индексы (NDVI, MSAVI, LAI, Albedo, LST, NPP), но и дополнительные данные, такие как цифровые модели рельефа, показатели социально-экономического развития, демографические параметры и данные о землепользовании. В частности, работы [79, 80] подчёркивают важность интеграции социально-экономических и экологических индикаторов, учитываемых при классификации карт LULC. Это позволяет идентифицировать классы земель, уязвимые не только с точки зрения природных процессов, но и антропогенного давления. К примеру, в исследованиях [81, 82] внедрение подходов, согласованных с концепцией Land Degradation Neutrality (LDN) и показателями Целей Устойчивого Развития (SDG 15.3.1), расширяет возможности классификации.

Таким образом, современные подходы к классификации данных ДЗЗ для оценки динамики LULC активно развиваются за счет внедрения комплексных, мультииндексных методов, интеграции разных типов данных, а также благодаря широкому применению машинного обучения и глубоких нейронных сетей. Как показывают [76, p. 5077-5096; 83-85], точность и надежность классификации при этом повышаются, снижается зависимость от отдельных индексов, а интерпретация результатов становится более согласованной с реальными процессами деградации. При этом, по мнению [30, p. 3257-1-3257-15; 74, p. 7561-7569; 84, p. 1633-1643; 86, 87], дальнейшее совершенствование классификационных методов связано с ещё более глубокой интеграцией искусственного интеллекта, автоматизацией подбора индексов и моделей, а также повышением качества исходных данных и возможностью рассчитывать производные показатели в реальном времени. Это позволит в будущем выйти на новые уровни детализации и точности классификации, критически важных для принятия управленческих решений и эффективного мониторинга земельных ресурсов.

1.3.1 Теория методов классификации данных ДЗЗ

Теоретические аспекты классификации данных ДЗЗ опираются на многолетний опыт становления методов и алгоритмов, переходящих от классических статистических моделей к все более сложным, иерархическим нейронным архитектурам, интегрирующим глубокое обучение и механизмы внимания. Значительный пласт теории был сформирован в результате применения машинного обучения к многоспектральным, гиперспектральным и радиолокационным данным, где ранние подходы, такие как пороговая классификация, линейный дискриминантный анализ и максимальное правдоподобие, постепенно уступили место нелинейным моделям и глубинным сетям [88, 89]. Теория классификации ДЗЗ данных долгое время концентрировалась на ручном выделении признаков, использовала статистические и текстурные дескрипторы, затем перешла к методам, способным учитывать спектральные, пространственные, временные и даже угловые характеристики сигналов, исходящих от земной поверхности [77, р. 164004; 90].

С приходом глубокого обучения и сверточных нейронных сетей (CNN, DCNN) произошла смена парадигмы, что нашло отражение в теоретических концептах извлечения признаков «снизу вверх» без предварительного инжиниринга [91, 92]. Ключевой теоретической основой стало восприятие данных как наборов многомерных структур, где нижние слои сети выделяют простые паттерны (контуры, края, текстуры), а верхние формируют более сложные семантические концепты, соотнесенные с классами LULC [93-95]. Такие теории привели к росту точности классификации, особенно при работе с изображениями очень высокого разрешения, где малейшие детали и локальные особенности имеют принципиальное значение для точной идентификации объектов [73, р. 6943-6955].

Теория семантической сегментации – один из ключевых вопросов классификации ДЗЗ - концентрируется на присвоении метки каждому пикселю изображения [88, р. 235-248]. Семантические модели, такие как Unet [92, р. 52-64] и ее производные, рассматриваются с точки зрения энкодер-декодерной архитектуры, где теория подсказывает, что многослойная свертка и последующий декодинг пространственной информации позволяют более точно восстанавливать контуры объектов [34, р. 2903-12903-39; 94, р. 415-426; 96]. Однако в теории осталось место для критики подобных подходов, указывающей на их ограниченность в учете глобального контекста и долгосрочных зависимостей между удаленными друг от друга пикселями. Это привело к появлению трансформерных моделей [95, с. 3-410; 97, с. 13-15] и гибридных архитектур, призванных восполнить пробелы в теоретической базе и обеспечить более гибкую и контекстно-зависимую классификацию [31, р. 369-380; 98].

Трансформеры, впервые предложенные в обработке естественных языков, оказались теоретически полезными и в сфере ДЗЗ, поскольку механизм внимания (attention) предоставляет модели возможность акцентировать внимание на релевантных регионах изображения [30, р. 3257-1-3257-15; 99]

В теоретической перспективе это означает возможность моделирования сложных пространственно-временных зависимостей, улучшая семантическую сегментацию, особенно при неоднородных классах и неоднозначных границах [89, p. 4048-4066; 100]. Таким образом, современные теоретические модели объединяют CNN, трансформеры, модули внимания, а также методы фьюжн-анализа данных из разных источников [90, p. e00919].

Дополнительной сложностью в анализе LULC является задача адаптации к разнообразию пространственных масштабов, времени съемки и сезонных изменений [101, 102]. Развитие теории требует анализа данных разного разрешения, учитывая, что многие разработанные методы лучше работают на очень высоком разрешении, тогда как задачи национального или континентального масштаба требуют обработки средне- и низкоразрешенных данных [7, p. 5770-5780; 103]. Теоретики исследуют возможность применения доменной адаптации [104], переносного обучения [64; 105], а также методов обучения без учителя, чтобы сократить зависимость от ручной разметки и эффективно работать с неполно аннотированными данными [106, 107].

Проблемы, возникающие в теории, включают попытки улучшить сегментацию мелкомасштабных объектов и определять четкие границы даже при наличии смешанных классов [108, 109], особенно в сельскохозяйственных, городских или лесных массивах [73, p. 6943-6955; 103, p. 2541-1-2541-33]. Теоретические исследования фокусируются на разработке моделей, способных к лучшему извлечению контурных признаков [60, p. 9355-9367; 110], чтобы уменьшить ошибки на границах объектов и повысить стабильность классификации при изменении сезонности, освещенности, облачности или дымки [103, p. 2541-1-2541-33; 111].

Исследователи также рассматривают вопросы, связанные с оптимизацией вычислительной эффективности моделей, повышением скорости сходимости и уменьшением числа параметров [108, p. 903709; 112]. Исследуются способы использования оптимизаторов нового поколения, методов нормализации, дропаута и различных приемов регуляризации, обеспечивающих улучшение общей устойчивости и обобщающей способности моделей [21, p. 1053-1062; 113]. Эти подходы направлены на то, чтобы сделать теоретический инструментарий более адаптивным к практическим условиям применения, когда необходимо быстро обрабатывать большие массивы данных ДЗЗ [89, p. 4048-4066; 103, p. 2541-1-2541-33].

Наконец, теория методов классификации ДЗЗ опирается на междисциплинарный подход, привлекая идеи из компьютерного зрения, обработки сигналов, статистики, геоинформатики, а также из сфер искусственного интеллекта и машинного обучения [98, p. 205-240; 110, p. 302-320]. Она впитывает концепции из смежных областей, таких как анализ текстур, оценка параметров рельефа и многомерное статистическое моделирование, а также интегрирует результаты фундаментальных исследований в области трансформеров, механизмов внимания, обобщенного представления данных и генеративно-состязательных сетей [60, p. 9355-9367]. В перспективе

синергетическое развитие теории, в рамках которого будут комбинироваться разнообразные методы извлечения признаков, механизмы межмодальной фьюжн-интеграции, методы снижения зависимости от ручной разметки и совершенствования обучения без учителя. Это создаст предпосылки для более точной, масштабируемой, интерпретируемой и универсальной теории классификационных методов в области дистанционного зондирования Земли [34, р. 2903-1-2903-39; 89, р. 4048-4066; 90, р. e00919; 95, с. 3-410].

1.3.2 Интеграция ГИС и ДЗЗ в облачных вычислениях: роль Google Earth Engine в современных гео-исследованиях

Программные решения, основанные на интеграции ГИС и данных ДЗЗ, все чаще применяются для высокоточной оценки пространственно-временной динамики землепользования в контексте природопользования, агроландшафтной оптимизации и управления лесными ресурсами. Подобный анализ требует комплексного подхода к обработке больших объемов разнотипных данных, включающих спутниковые снимки, данные воздушной и беспилотной съемки, а также информацию о рельефе, почвах и растительном покрове.

Особую значимость приобрели методы, сочетающие ГИС с алгоритмами машинного обучения (ML) и искусственного интеллекта (AI), способными выявлять паттерны в многомерных массивах информации, автоматически классифицировать типы земельных угодий и прогнозировать их трансформации. Применение гиперспектральных и мультиспектральных данных, дополненных лидарными и радарными съемками, расширяет возможности выявления тонких изменений в структуре покрова, что повышает точность тематического картирования и анализа продуктивности агро- и лесных систем.

Появление универсальных аналитических облачных платформ усилило потенциал обработки больших данных ДЗЗ. Среди них наибольший резонанс вызывает Google Earth Engine (GEE), предоставляющая богатейшие каталоги спутниковых снимков и климатических моделей, доступных для масштабного геоаналитического анализа [100, р. 483-493]. Данная платформа повышает оперативность вычислений и снижает порог входа для специалистов, не обладающих глубокими знаниями в программировании, одновременно обеспечивая возможность интеграции ML-библиотек и реализации сложных алгоритмов классификации [73, р. 6943-6955].

GEE представляет собой масштабируемую облачную платформу, обеспечивающую доступ к геопространственным данным и их обработку на планетарном уровне, что было подробно охарактеризовано в работе [6, р. 18-26]. Использование GEE, по мнению ряда исследователей, связано с упрощением анализа больших данных дистанционного зондирования Земли и ускорением научных открытий в различных областях, включая землеведение, агрономию, лесное хозяйство и мониторинг окружающей среды [9, р. 16911-169-22]. В частности, исследования в период с 2011 по 2022 годы показали стремительный

рост числа публикаций, связанных с GEE [6, p. 18-26; 114], что свидетельствует о возрастающем интересе научного сообщества.

Многие исследования указывают на важнейшую роль GEE в мониторинге лесов, растительности и водных ресурсов [73, p. 6943-6955; 89, p. 4048-4066], оценке временных рядов спутниковых данных [99, p. 2388-1-2388-16], а также в применении машинного обучения для выделения ключевых закономерностей и классификации различных типов земных поверхностей. Кроме того, работа N. Gorelick [6, p. 18-26] подчёркивает, что GEE интегрирует огромные каталоги данных Landsat, Sentinel, MODIS, а также климатические и топографические сведения, что в совокупности снижает барьеры для широкого использования этих данных.

Применение GEE также охватывает анализ растительности, оценку углеродного баланса, расчёты индексов растительности NDVI, EVI, NDWI и оценку биофизических параметров для больших территорий [7, p. 5770-5780; 115]. В свою очередь, исследования Liu et al. Подтверждают возможность интеграции временных рядов Landsat с алгоритмами машинного обучения для точного мониторинга использования земель, в том числе для заброшенных сельскохозяйственных площадей.

Ряд авторов применяли GEE для учета изменения климата и его влияния на растительные сообщества и водные экосистемы [31, p. 369-380; 115, p. 27-47]. Подобные исследования, как отмечено в работах Wang [73, p. 6943-6955] и Хохо [83, p. 422-430], позволяют более полно понять биофизические процессы и помогают странам ориентироваться на Цели устойчивого развития (SDG), особенно связанные с управлением земельными ресурсами, сельскохозяйственной продуктивностью и охраной биоразнообразия.

Развитие облачных технологий, представленных GEE, стимулирует междисциплинарные исследования, которые сочетают методы дистанционного зондирования, ГИС-технологии и машинное обучение. Фактически, N. Gorelick и M. Hancher [6, p. 18-26] отмечают, что с помощью GEE учёные могут быстро и эффективно анализировать петабайты спутниковых данных, получая результаты, ранее труднодостижимые без кластерных вычислений. Zhao et al [77, p. 164004]. Показывают рост цитируемости и интереса к данной теме в наукометрических базах.

Высока актуальность использования наземных данных для валидации полученных из спутников алгоритмических выводов. Без качественных полевых наблюдений точность классификации существенно снижается, что отмечается при попытках различать, например, убранные и неубранные сельскохозяйственные поля либо определять видовой состав древесной растительности [97, с. 13-15; 116]. Интеграция наземных измерений с ДЗЗ-подходами дает возможность уточнять пространственные модели, что особенно важно при оценке влияния климатических изменений, интенсивности лесопользования или степени антропогенного прессинга [6, p. 18-26].

Сочетание продвинутых облачных и настольных ГИС-комплексов, поддерживающих методы AI, ML и глубокого обучения, формирует условия для

более точного учета пространственной и временной изменчивости LULC. Технологии интегрируют высокоавтоматизированную классификацию, прогнозирование изменений землепользования, оценку последствий климатической изменчивости и анализ факторов, определяющих динамику сообществ. Параллельно совершенствуются подходы к валидации, стандартизации и обеспечению доступности результатов аналитических процедур, а включение наземных данных в циклы обучения алгоритмов содействует повышению достоверности тематической интерпретации спутниковых снимков [6, p. 18-26; 73, p. 6943-6955].

GEE всё больше воспринимается как незаменимый инструмент для анализа. Перспективы его развития связаны с дальнейшей интеграцией новых спутниковых данных, усовершенствованием алгоритмов машинного обучения, расширением тематических приложений и глобальных карт, охватывающих динамику земной поверхности, климат, биоразнообразие и ресурсообеспеченность. Подобные тенденции прослеживаются в работах [7, p. 5770-5780; 103, p. 2541-1-2541-33] а также [6, p. 18-26] которые указывают, что GEE будет и далее стимулировать инновации, мультидисциплинарность и расширение возможностей геопространственного анализа.

Программные решения, основанные на интеграции ГИС и данных ДЗЗ, в последние годы стремительно развиваются, обеспечивая более глубокое понимание пространственно-временной динамики природных и антропогенных систем. Учитывая растущую доступность спутниковых данных, таких как Landsat, а также появление облачных вычислительных платформ, GEE, специалисты по обработке геоданных получили мощные инструменты для оценки состояния, мониторинга и прогноза изменений ландшафтных систем, включая травянистые экосистемы, сельскохозяйственные земли, пастбища, леса и водные объекты [82, с. 3-389; 117].

Развитие таких интегрированных подходов стало особенно актуальным на фоне глобальных изменений климата и стремительного роста населения, вызывающего повышенный спрос на устойчивое управление ресурсами, продовольственную безопасность и адаптацию к экстремальным природным явлениям [118]. Вместе с тем, широкое использование геопространственных данных и соответствующих инструментов обусловлено стремлением к комплексному анализу, синтезирующему пространственные, временные и тематические измерения для принятия более обоснованных решений в области природопользования [64, p. 1528-1-1528-19; 119-121].

Научные исследования, посвящённые интеграции ГИС и ДЗЗ, показывают, что современные программные решения способны эффективно решать задачи, связанные с мониторингом растительности, анализом биомассы, оценкой изменения землепользования и выделением зон деградации [19, p. 112-127; 122]. К примеру, в работе [22, p. 62-1-62-14] отмечено, что высокоточные карты границ пастбищ и травяных сообществ позволяют более рационально использовать земельные ресурсы и отслеживать состояние экосистем в разных регионах. Интеграция ГИС и ДЗЗ-данных повышает качество входной

информации для моделей климатических изменений и обеспечивает регулярный контроль состояния почвы, влажности, продуктивности растительности и динамики биоразнообразия [89, р. 4048-4066; 123]. При этом ряд исследователей подчеркивает, что ключевым препятствием к повсеместному применению таких решений остаются проблемы с доступом к высоким пространственным и временным разрешениям данных, а также вопросы вычислительной инфраструктуры и стандартизации [118, р. 3899-3913].

Использование облачных платформ, таких как GEE, серьёзно упростило обработку больших массивов данных ДЗЗ, сделав доступным детальный анализ за длительные временные периоды и для обширных географических регионов [64, р. 1528-1-1528-19; 119, р. 4551-1-4551-21; 120, р. 298-312; 121, р. 1-8]. В работах [67, р. 4014805; 120, р. 298-312; 124, 125] отмечено, что интеграция машинного обучения и ГИС-технологий в среде GEE позволяет строить устойчивые модели для оценки биомассы, углеродного баланса и динамики растительности, способствуя решению задач, связанных с оценкой продовольственной безопасности и адаптацией к изменению климата. Подобные подходы также применяются для анализа лесопокрытия, сельскохозяйственных угодий и травянистых экосистем, что отражено в исследованиях [82, с. 3-688; 125, р. 1-56], где использование ГИС и ДЗЗ позволило выявить как потери естественных луговых территорий, так и интенсивную трансформацию растительности под влиянием антропогенных факторов.

Библиометрические обзоры, выполненные в исследованиях [67; 120, р. 298-312; 123, р. 2272-2282; 124, р. 2529-2542; 125, р. 1-56], подчёркивают всё более широкое распространение ГИС- и ДЗЗ-решений в прикладных экологических и географических областях. Таким образом, интеграция ГИС и ДЗЗ-данных способствует трансформации методов ландшафтного анализа, делая их более точными, оперативными и эффективными [19, р. 112-127; 125, р. 1-56; 126, 127]. Согласно исследованиям [54, р. 95-108; 128], прогресс в области вычислительных технологий привёл к тому, что современные программные решения обеспечивают доступ к постоянно пополняющимся каталогам спутниковых изображений, климатических данных и топографической информации. Это позволяет исследователям проводить анализ экологических процессов на континентальном и глобальном уровнях, не ограничиваясь локальными данными [120, р. 298-312; 126, р. 196-213].

В работах [23, р. 104829; 64; 119, р. 4551-1-4551-21; 126, р. 196-213] показано, что систематический мониторинг динамики LULC, реализованный в облачных платформах, раскрывает тенденции потери природной растительности, деградации почв и сокращения биоразнообразия. Отдельно стоит отметить, что программные решения, связанные с интеграцией ГИС и ДЗЗ, активно применяются для оптимизации методов классификации, сегментации и определения ключевых параметров растительного покрова, таких как индексы растительности NDVI и др. [23, р. 104829]. Кроме того, в исследованиях [6, р. 18-26] использование ГИС-решений в сочетании с ДЗЗ позволило более точно определять зоны заброшенных сельскохозяйственных угодий, очаги деградации

и вырабатывать рекомендации по восстановительным мерам. Применение машинного обучения в рамках таких программных решений особенно важно, учитывая сложность многомерных наборов данных, поступающих из разных спутниковых платформ [82, с. 3-688; 127, р. 2781-2816].

Исследования [31, р. 369-380] демонстрируют рост значимости интегрированных ГИС-ДЗЗ инструментов при оценке продуктивности агроландшафтов, выявлении деградационных процессов и понимании факторов, влияющих на состояние растительных сообществ. Повышение доступности данных ДЗЗ в сочетании с прогрессом ГИС-платформ стимулирует углубление анализа процессов, таких как опустынивание, зарастание травяных экосистем кустарниками, вызванное изменением климата, и расширение антропогенной нагрузки [119, р. 4551-1-4551-21; 126, р. 196-213]. Применение таких решений помогает исследователям и практикам находить баланс между экономической выгодой и охраной окружающей среды, а также обеспечивает более прозрачное и научно обоснованное принятие решений в области устойчивого развития [46, р. 813-842; 122, р. 2616-1-2616-33; 129, 130].

Программные решения, основанные на интеграции ГИС и данных ДЗЗ, играют центральную роль в формировании нового взгляда на пространственный анализ и мониторинг состояния биофизических систем. Эти решения предоставляют возможности для всестороннего, комплексного и непрерывного исследования экосистем, отражают глобальные тенденции внедрения инноваций в управление земельными ресурсами, защиту биоразнообразия и смягчение последствий климатических изменений, обеспечивая долгосрочную основу для стратегического планирования, адаптации и экологического мониторинга на всех уровнях - от локальных до глобальных масштабов.

1.3.3 Оценка точности классификации LULC

Оценка точности классификаций LULC является важнейшим этапом в применении данных ДЗЗ и ML для понимания пространственно-временной динамики землепользования и изменений земного покрова, поскольку точные карты необходимы в областях, связанных с региональным планированием, уязвимостью окружающей среды и прогнозированием рисков [128, р. 5400714; 129, р. 154-1-154-20; 130, р. 244-266]. Необходимость адекватной валидации возрастает на фоне стремительного роста населения, экономического и промышленного развития, особенно в развивающихся странах, где преобразования ландшафтов происходят с большой скоростью [54, р. 95-108; 128, р. 5400714; 129, р. 154-1-154-20; 130, р. 244-266]. Поскольку количественная оценка изменений LULC позволяет принимать более обоснованные решения, важно сравнивать различные алгоритмы классификации, применять единые критерии оценки точности и уделять внимание метрическим показателям, таким как коэффициент Карра, AD и QD и интегральные индексные методы [67, р. 4014805; 124, р. 2529-2542; 128, р. 5400714; 129, р. 154-1-154-20; 130, р. 244-266].

В литературе отмечено, что машинное обучение предоставляет широкие возможности для повышения точности LULC-карт, поскольку алгоритмы, такие

как random forest (RF), support vector machine (SVM), artificial neural network (ANN), fuzzy ARTMAP, spectral angle mapper (SAM) и Mahalanobis distance (MD), способны более тонко выделять классы земной поверхности [67, p. 4014805; 129, p. 154-1-154-20]. Однако наблюдается вариативность результатов, что вынуждает исследователей применять различные критерии оценки точности. Карра-коэффициент продолжает широко использоваться, поскольку он считается классическим инструментом для оценки согласия между классифицированными данными и фактической ситуацией на местности [54, p. 95-108; 126, p. 196-213; 127, p. 2784-2816; 128, p. 5400714; 129, p. 154-1-154-20; 131]. Тем не менее, в некоторых случаях может быть полезно дополнять и заменять его другими метриками, поскольку Карра - чувствительность к распределению классов и несбалансированным выборкам вызывает у исследователей потребность в дополнительных показателях [64, p. 1528-1-1528-19; 67, p. 4014805].

На практике точность классификации LULC зависит от набора параметров и характеристик данных: пространственного и спектрального разрешения спутниковых снимков, качества обучения алгоритмов, выбора обучающей выборки, а также от ландшафтных особенностей исследуемой территории [119, p. 4551-1-4551-21; 120, p. 298-312; 121, p. 1-8; 122, p. 2616-1-2616-33; 123, p. 2272-2282]. К примеру, в ряде работ отмечено, что ANN, SVM и RF нередко превосходят традиционные методы, однако их точность может колебаться под влиянием геоморфологических условий и климатических факторов [67, p. 4014805; 120, p. 298-312; 124, p. 2529-2542; 125, p. 1-56; 126, p. 196-213]. Расхождения между классификаторами могут отражаться в различиях площадей, занятых теми или иными классами, особенно для сложных ландшафтов, что демонстрирует важность многосторонней оценки точности [127, p. 2784-2816; 131, p. 1319-1330; 132].

Использование метрик позволяет расширить спектр критериев проверки за счёт учёта чувствительности и специфичности классификатора, измерения расстояния между прогнозируемыми и фактическими значениями, а также оценки стабильности алгоритмов [128, p. 5400714; 129, p. 154-1-154-20; 130, p. 244-266; 131, p. 1319-1330]. Включение индексных методов на основе NDVI, NDWI и NDBI даёт возможность независимой проверки площадей выделенных классов, что способствует более строгому и многогранному подходу к оценке точности [124, p. 2529-2542; 133]. Сопоставление результатов с данными высокодетальных изображений Google Earth и наземными измерениями, даёт дополнительные гарантии качества, позволяя корректировать и уточнять результаты классификации [120, p. 298-312; 126, p. 196-213; 134].

Коэффициент Карра долгое время оставался широко используемым инструментом для оценки согласованности классификаций тематических карт, основанных на данных дистанционного зондирования. Этот метод используется для оценки качества картографической продукции, сравнивая классифицированные данные с эталонными и учитывая вероятность случайного совпадения. Однако его адекватность в задачах геопространственного анализа

часто ставится под сомнение. Основные ограничения коэффициента связаны с его неспособностью различать типы ошибок, возникающих при классификации, и с зависимостью от распределения классов в данных [133, p. 6309-6319; 135].

В недавнем исследовании В. Feizizadeh [132, p. 1228-1248] предложили альтернативный подход к оценке точности, называемый индексом QADI (Quantity and Allocation Disagreement Index). В отличие от коэффициента каппа, QADI позволяет разделить ошибки на количественные (quantity disagreement) и аллокационные (allocation disagreement). Это разделение имеет критическое значение для понимания природы ошибок классификации. Например, количественное несоответствие возникает тогда, когда на классифицированной карте площадь какого-то типа поверхности - например, леса - отличается от той, что указана в проверочных (эталонных) данных. А аллокационное несоответствие - это когда площадь вроде бы совпадает, но расположение на местности указано неправильно, например, пиксели леса оказались не там, где им положено быть.

Коэффициент Карра не различает эти два типа ошибок - он просто оценивает общее совпадение, не обращая внимания на то, насколько правильно объекты размещены в пространстве. В результате может получиться так, что карта показывает высокий коэффициент Карра, хотя на ней, по сути, леса, поля и другие объекты “разбросаны” не по своим местам. Поэтому в задачах, где важна точность расположения - например, при планировании использования территорий или управлении природными ресурсами - коэффициент Карра не всегда даёт объективную оценку качества карты. Кроме того, коэффициент Карра нередко подвергается критике в связи с его зависимостью от несбалансированности распределения классов. В ситуациях, когда один из классов доминирует (например, сельскохозяйственные земли или лесные массивы занимают большую часть исследуемой территории), значение Карра может оказаться завышенным, даже при низкой точности классификации редких классов таких как водные объекты или городская застройка [135]. По нашему мнению, данное ограничение особенно существенно для задач LULC, где именно редкие классы нередко имеют ключевое значение для целей мониторинга и управления.

В этой связи В. Feizizadeh и Р. Jankowski предложили индекс QADI как более точный и интерпретируемый инструмент оценки точности классификаций. Применение QADI позволяет исследователям не только количественно фиксировать ошибки, но и получать более наглядное представление о том, какие именно аспекты классификации требуют доработки. Так, если количественное несоответствие существенно превышает аллокационное, это может указывать на недостаточный объём обучающих выборок для конкретного класса. В случае же преобладания аллокационного несоответствия можно предположить необходимость совершенствования пространственной модели либо алгоритма классификации с целью более корректного позиционирования объектов.

Метод QADI был апробирован на данных дистанционного зондирования, включая изображения, классифицированные с применением распространённых

алгоритмов машинного обучения, таких как Random Forest, SVM и искусственные нейронные сети (ANN). Полученные результаты показали, что QADI обеспечивает более точные и содержательные оценки по сравнению с коэффициентом Карра, особенно при анализе сложных, мозаичных и ландшафтно разнообразных территорий [132, p. 1228-1248].

Применение QADI в исследовании [132, p. 1228-1248] позволило не только углубить понимание природы классификационных ошибок, но и сформулировать практические рекомендации по оптимизации используемых алгоритмов. В целом индекс QADI, на наш взгляд, стал значимым шагом вперёд в развитии методов оценки точности классификации спутниковых изображений. В отличие от коэффициента Карра он нивелирует ряд его ключевых ограничений и предлагает более прозрачный и аналитически информативный подход к анализу ошибок.

Несмотря на то что коэффициент Карра по-прежнему широко применяется в научных исследованиях, его ограниченность особенно в задачах анализа землепользования (LULC) и других пространственных исследований указывает на необходимость внедрения более надёжных метрик. Использование QADI позволяет не только повысить качество создаваемых тематических карт, но и предоставляет исследователям более глубокое понимание сильных и слабых сторон применяемых алгоритмов классификации.

Кроме того, использование данных, собранных в полевых условиях, по нашему мнению, остаётся важнейшим этапом верификации результатов классификации. Такой подход позволяет более точно учитывать реальные характеристики исследуемой территории и, как следствие, повышать достоверность итоговых картографических материалов. Jensen J.R. подчёркивает, что именно полевые данные играют решающую роль в снижении количественных несоответствий, возникающих вследствие недостаточной калибровки моделей. В частности, результаты полевых измерений могут применяться для уточнения классификации редких классов, включая водные объекты или временные сельскохозяйственные угодья [136].

Несмотря на ряд очевидных преимуществ, индекс QADI также имеет определённые ограничения. Как отмечают B. Boots и F. Csillag, точность расчётов QADI во многом зависит от качества эталонных данных, получение которых может быть затруднено в труднодоступных регионах. Вместе с тем использование спутниковых изображений высокого пространственного разрешения, таких как данные Google Earth и Sentinel-2, позволяет в значительной степени нивелировать указанные ограничения.

В целом индекс QADI, на наш взгляд, представляет собой существенный шаг вперёд по сравнению с традиционным коэффициентом каппа. Он не только предоставляет более детализированную информацию о структуре классификационных ошибок, но и открывает возможности для более глубокого пространственного анализа данных. В задачах, связанных с изучением землепользования и земного покрова (LULC), где критически важна точность

пространственного положения объектов, QADI демонстрирует высокую эффективность и методическую адаптивность.

Кроме того, использование QADI в связке с современными методами машинного обучения и геоинформационными технологиями открывает новые возможности для работы со сложными и неоднородными ландшафтами. Это делает его ценным инструментом для территориального планирования, экологического мониторинга и более точного анализа данных дистанционного зондирования.

1.3.4 Сбор полевых эталонных данных

Сбор эталонных данных и применение полевых методов являются фундаментальными элементами при оценке и интерпретации данных дистанционного зондирования Земли, особенно в контексте картирования и анализа изменений LULC. Широкое распространение спутниковых систем наблюдения за Землей, таких как Landsat, а также гиперспектральных и мультиспектральных платформ, позволило существенно расширить возможности исследователей по мониторингу растительности, водных ресурсов и антропогенных преобразований ландшафта [54, р. 95-108; 128, р. 5400714; 129, р. 154-1-154-20; 130, р. 244-266]. Однако даже самые современные алгоритмы машинного обучения и точные радиометрические коррекции не могут гарантировать корректности результатов без адекватной калибровки, которая достигается путем сопоставления спутниковых измерений с наземными эталонными данными.

Многочисленные исследования подтверждают, что использование полевых наблюдений, пробоотборов и верификации, а также визуальной проверки с помощью высокодетальных снимков, позволяет значительно повысить надежность и точность классификаций LULC [120, р. 298-312; 121, р. 1-8; 122, р. 2616-1-2616-33; 123, р. 2272-2282; 124, р. 2529-2542; 127, р. 2784-2816]. Наземные данные служат основой для создания обучающих и тестовых выборок, необходимых при использовании алгоритмов машинного обучения, таких как random forest (RF), support vector machine (SVM), artificial neural network (ANN), fuzzy ARTMAP, spectral angle mapper (SAM) и Mahalanobis distance (MD) [67, р. 4014805; 122, р. 2616-1-2616-33; 123, р. 2272-2282; 124, р. 2529-2542; 125, р. 1-56; 128, р. 5400714]. Именно благодаря точным и репрезентативным полевым данным можно наиболее адекватно подобрать параметры моделей, снизить перекосы классификации и минимизировать погрешности, возникающие из-за неоднородности спектральных характеристик, сезонных вариаций и природно-климатического разнообразия региона.

Полевые методы позволяют уточнять границы между различными классами землепользования, например, различая естественные и антропогенные ландшафты, определять состояние растительности, отличать участки с различной степенью деградации, фиксировать мелкомасштабные объекты, невидимые со спутника, и оценивать влияние локальных факторов, таких как орошение, применение удобрений, вырубка лесов или зарастание пастбищной

растительности кустарником [54, р. 95-108; 128, р. 5400714; 129, р. 154-1-154-20; 130, р. 244-266]. Включение полевых данных выступает своего рода калибровочной метрикой, дающей возможность проводить валидный сравнительный анализ нескольких классификаторов и метрик точности. Так, сравнение данных полевых измерений с результатами индексной оценки (NDVI, NDWI, NDBI) дает более многогранное представление о состоянии растительности, наличии водных ресурсов и характере застройки [54, р. 95-108; 125, зю 1-56; 126, р. 196-213; 127, р. 2784-2816; 128, р. 5400714; 129, р. 154-1-154-20; 130, р. 244-266].

Применение полевых методов является необходимым условием для повышения доверия к итоговой картографической продукции. Отсутствие качественных эталонных данных может привести к неверной интерпретации пространственных паттернов, неправильным управленческим решениям в области планирования и природопользования и искажению результатов долгосрочного мониторинга. Исследователи отмечают, что использование систематически собранных и репрезентативных эталонных выборок улучшает устойчивость алгоритмов классификации, расширяя возможности анализа даже в труднодоступных или малоизученных регионах [130, р. 244-266; 131, р. 1319-1330; 132, р. 1228-1248; 133, р. 6309-6319; 134, р. 779; 135; 136, р. 3-540]. Более того, высококачественные данные полевых наблюдений помогают интерпретировать временную динамику LULC, включая сезонные колебания, фенологические циклы, деградацию почв и биоразнообразия, а также последствий климатических изменений и антропогенного давления на экосистемы.

1.4 Постклассификационный анализ LULC

Постклассификационный анализ изменений LULC представляет собой ключевой инструмент в изучении пространственно-временной динамики ландшафтов, предоставляя возможности для выявления закономерностей изменений, их причин и последствий. Этот подход широко используется для анализа динамики, основанной на данных ДЗЗ, с интеграцией методов ГИС и статистических моделей. Постклассификационный анализ, по нашему мнению, позволяет не только фиксировать сами факты изменений, но и исследовать их пространственное распределение, а также выявлять факторы, обуславливающие данные трансформации [85, р. 5400314; 132, р. 1228-1248].

Одним из базовых методов в рамках данного подхода выступает сравнение многовременных данных дистанционного зондирования Земли. Он предполагает последовательную классификацию спутниковых изображений, полученных за различные временные периоды, с последующим их сопоставлением и аналитической интерпретацией. Указанный подход успешно применялся при оценке изменений как в урбанизированных, так и в сельских ландшафтах, что позволяет отслеживать процессы урбанизации, трансформации сельскохозяйственных земель, а также восстановления лесных территорий [85, р. 5400314].

Вместе с тем ключевым ограничением метода остаётся необходимость обязательного учёта ошибок классификации. Pontius [131, p. 1319-1330] подчёркивают, что подобные ошибки подразделяются на количественные несоответствия (различия в суммарных площадях классов) и аллокационные несоответствия (ошибки их пространственного распределения). Наличие классификационных погрешностей способно существенно повлиять на итоговую точность анализа, что, на наш взгляд, обуславливает необходимость применения дополнительных аналитических инструментов, включая индексы количественного и аллокационного несоответствия.

Одним из эффективных методов также является анализ главных компонент (РСА), позволяющий упростить многомерные данные за счёт выделения ключевых факторов, влияющих на изменения землепользования. Так, в исследовании [98, p. 205-240] было показано, что сочетание РСА с кластерным анализом (в частности, методом К-средних) способствует выявлению значимых закономерностей, включая влияние урбанизационных процессов и демографического роста на трансформацию ландшафтов. Кроме того, РСА активно применяется при обработке гиперспектральных данных, обеспечивая возможность анализа более сложных процессов, например деградации земель или изменений растительного покрова.

Для прогнозирования изменений LULC широкое распространение получили пространственно-динамические модели, в частности цепи Маркова и клеточные автоматы (Cellular Automata). Указанные методы позволяют оценивать вероятностные сценарии трансформации классов землепользования во времени. В работе [96, p. 3-800] продемонстрировано, что их совместное применение эффективно при моделировании процессов урбанизации и оценке их воздействия на окружающую среду. Интеграция временных и пространственных данных в рамках данных моделей, как представляется, делает их мощным инструментом анализа изменений в динамично развивающихся регионах.

Регрессионный и корреляционный анализ, по нашему мнению, выступают важными инструментами при изучении факторов, определяющих изменения землепользования. Их применение позволяет количественно оценивать влияние таких процессов, как рост численности населения, экономическое развитие или климатические трансформации, на динамику ландшафтов [134, p. 779]. Так, в одном из исследований [132, p. 1228-1248] корреляционный анализ выявил чёткую зависимость между увеличением численности населения и сокращением площадей лесных массивов, что подчёркивает значимую роль антропогенного воздействия в трансформации природной среды.

Изучение факторов изменений в целом является центральным элементом постклассификационного анализа. К числу природных факторов, как правило, относят климатические изменения повышение температур, трансформацию режимов осадков, а также воздействие стихийных явлений, включая наводнения и оползни [135]. В то же время существенное влияние оказывают и антропогенные факторы: интенсивное сельское хозяйство, вырубка лесов, развитие строительной и транспортной инфраструктуры. Например, в

исследовании [31, р. 369-380] установлено, что развитие транспортных сетей выступает одним из ключевых драйверов урбанизации в ряде регионов.

Особую значимость приобретает интеграция климатических, топографических и землепользовательских данных для более точного определения причин наблюдаемых изменений. В работе [131, р. 1319-1330] Н. Chen отмечают, что совмещение информации о температуре, влажности и интенсивности осадков с пространственными моделями позволяет более достоверно оценивать влияние климатических факторов на динамику LULC. В частности, в регионах с развитым сельским хозяйством изменения температурного режима и количества осадков способны существенно трансформировать структуру землепользования.

В качестве практического примера применения постклассификационного анализа можно привести исследование динамики LULC в урбанизированных территориях, выполненное Chen [84, р. 1633-1643]. С использованием анализа главных компонент (РСА) и моделей Markov-CA исследователи установили, что интенсивные урбанизационные процессы сопровождаются расширением застроенных площадей за счёт сокращения сельскохозяйственных угодий. Проведённый анализ также показал, что аллокационные ошибки наиболее часто фиксируются на периферии городов, где наблюдаются активное строительство и высокие темпы пространственных изменений.

В целом постклассификационный анализ изменений LULC представляет собой эффективный инструмент для понимания того, каким образом и под воздействием каких факторов трансформируется землепользование. Использование методов РСА, пространственного моделирования, а также регрессионного анализа, как представляется, позволяет более глубоко раскрыть причины происходящих изменений и сформировать научно обоснованные подходы к устойчивому управлению территориями. В сочетании с данными дистанционного зондирования и ГИС эти инструменты дают исследователям и специалистам надёжную основу для анализа текущих процессов и прогнозирования будущих изменений.

Динамика LULC оказывает значительное влияние на экосистемы, климат, биоразнообразие и природные ресурсы. Она определяется как природными, так и антропогенными факторами. Природные изменения могут быть вызваны климатическими колебаниями, изменением гидрологических режимов и процессами эрозии. Антропогенные факторы включают урбанизацию, интенсификацию сельского хозяйства, вырубку лесов, мелиорацию и строительство инфраструктуры [129, р. 154-1-154-20; 137]. Понимание механизмов этих изменений имеет важное значение для разработки стратегий управления землепользованием, направленных на минимизацию их негативных последствий [138 с. 169-174; 139].

Использование ДЗЗ и ГИС предоставляет мощные инструменты для анализа LULC. ДЗЗ обеспечивает доступ к данным многовременного наблюдения, которые позволяют фиксировать пространственные и временные изменения, а ГИС используется для анализа и визуализации этих данных [4,

р. 201-210; 140]. Например, спутниковые снимки Landsat и Sentinel широко применяются для классификации земного покрова, анализа динамики изменений и создания прогнозных моделей [141]. Кроме того, использование индексов, таких как NDVI (Normalized Difference Vegetation Index), позволяет оценить состояние растительности, что особенно полезно для мониторинга экологических изменений [142].

Одним из наиболее эффективных методов моделирования и прогнозирования LULC является гибридная модель CA-Markov. Эта модель объединяет вероятностный анализ Маркова, основанный на матрице переходов между классами землепользования, с пространственными характеристиками, обеспечиваемыми клеточными автоматами (Cellular Automata, CA). Такой подход помогает не только оценивать, насколько изменилось землепользование, но и понять, где именно происходят эти изменения [143].

Модель CA-Markov широко применяется для прогнозирования будущих изменений территориальной структуры. Её использование позволяет учитывать многолетние данные наблюдений и формировать вероятностные сценарии трансформации землепользования, что, на наш взгляд, особенно важно при решении задач территориального планирования, управления природными ресурсами и развития инфраструктуры [8, р. 359-1-359-22]. Подобные модели оказываются особенно востребованными в условиях необходимости учёта демографического роста, климатических изменений и социально-экономического развития [144, 145].

Так, по мере увеличения численности населения происходит расширение городских территорий, что, как правило, сопровождается сокращением природных зон и уменьшением площадей, пригодных для ведения сельского хозяйства [10, р. 61-77]. Кроме того, постклассификационный анализ с применением моделей, подобных CA-Markov, позволяет идентифицировать ключевые факторы, определяющие динамику изменений LULC. В частности, решающую роль в формировании будущих сценариев трансформации играют социально-экономические условия, особенности территориальной политики и уровень инфраструктурного развития [146].

Прогнозирование таких изменений может быть полезно для минимизации негативного воздействия на экосистемы и повышения устойчивости к изменениям климата [138, с. 169-174; 139, р. 1-11].

Эти подходы обеспечивают глубокое понимание процессов изменений, их причин и последствий, что делает их незаменимыми для задач устойчивого развития, планирования и охраны окружающей среды.

Выводы по разделу 1

Чтобы точно понять, как меняется землепользование сегодня, нужно использовать комплексный подход - объединять данные со спутников, результаты полевых наблюдений и современные методы машинного обучения. Главная сложность заключается в необходимости обеспечения высокого уровня

детализации и достоверности данных для оценки изменений землепользования и выявления факторов, которые влияют на эти изменения.

Целью работы является разработка методики анализа динамики землепользования с использованием многовременных спутниковых данных и алгоритмов машинного обучения, направленной на точное выявление пространственно-временных изменений и определение факторов, способствующих этим изменениям.

Идея работы заключается в интеграции методов дистанционного зондирования, натурных измерений и алгоритмов классификации для создания инструмента анализа изменений землепользования.

Для достижения цели исследования поставлены следующие задачи:

1. Сбор и каталогизация спутниковых данных:
 - проведение анализа доступности многовременных спутниковых данных (Landsat и др.) для изучаемого региона;
 - формирование базы данных спутниковых снимков для дальнейшей обработки и анализа.
2. Проведение натурных измерений:
 - организация выездов на места для определения фактического класса объектов (пахотные земли, пастбища, водные объекты и др.);
 - применение методов наземной GNSS-съемки и беспилотных летательных аппаратов для уточнения пространственных данных и валидации спутниковой информации.
3. Разработка алгоритма классификации на основе Random Forest:
 - создание алгоритма машинного обучения для классификации спутниковых данных;
 - адаптация алгоритма к специфическим условиям изучаемого региона и обеспечение высокой точности классификации;
 - генерация карт землепользования на основе обработанных данных.
4. Анализ динамики землепользования:
 - проведение анализа многовременного ряда спутниковых снимков для выявления пространственно-временных изменений в структуре землепользования;
 - сравнение результатов классификации для разных временных периодов и построение динамических карт.

2 СБОР И ПОДГОТОВКА ДАННЫХ ДЛЯ АНАЛИЗА ДИНАМИКИ LULC

Постклассификационный анализ представляет собой один из наиболее эффективных методов для изучения пространственных и временных изменений землепользования, полученных на основе данных ДЗЗ. В основе этого подхода лежит сравнительный анализ тематически классифицированных изображений за различные временные периоды, позволяющий не только выявлять изменения, но и определять их структурные, пространственные и причинно-следственные связи. Для исследуемой территории Карагандинской области, включающей Осакаровский, Абайский, Бухар-Жырауский и Каркаралинский районы, постклассификационный анализ выступает ключевым этапом для выявления драйверов динамики землепользования и их пространственной зависимости.

Постклассификационный анализ включает несколько последовательных этапов. В первую очередь проводится предварительная обработка спутниковых данных (калибровка, атмосферная коррекция и геометрическая привязка) и классификация изображений с использованием методов машинного обучения (в данном случае RF). Полученные тематические карты землепользования для разных временных срезов сопоставляются между собой с целью выявления динамических изменений на уровне отдельных классов землепользования: культивируемые земли, пастбища, лесные массивы, урбанизированные территории, водные объекты и водно-болотные угодья.

На следующем этапе изменения классифицируются по типам и оцениваются их количественные и качественные характеристики. Особое внимание уделяется определению драйверов этих изменений, то есть факторов, которые влияют на преобразование LULC.

Постклассификационный анализ динамики землепользования требует интеграции данных ДЗЗ с официальной статистической информацией для более глубокого понимания причин и факторов изменений. Основными источниками статистических данных выступают сборники Комитета по статистике Республики Казахстан (Госстат РК), включающие сведения по демографии, экономике, сельскому хозяйству, промышленности и экологии.

Статистические данные, по нашему мнению, играют важную роль в количественной оценке влияния социально-экономических процессов на изменения землепользования.

Для анализа воздействия указанных факторов применяются такие методы, как регрессионный анализ, пространственная автокорреляция и анализ главных компонент (РСА). Их использование позволяет определить, какие именно причины оказывают наибольшее влияние на трансформацию землепользования, а также выявить территории, где данные изменения проявляются наиболее интенсивно.

Результаты подобного анализа предоставляют ценную информацию о текущих пространственно-временных трендах, включая зоны активной урбанизации, участки деградации сельскохозяйственных земель и территории,

где, напротив, наблюдаются процессы восстановления растительного покрова. Полученные данные, как представляется, формируют научно обоснованную основу для принятия управленческих решений, направленных на рациональное и устойчивое управление земельными ресурсами.

2.1 Выбор исследуемой территории: обоснование и характеристика региона

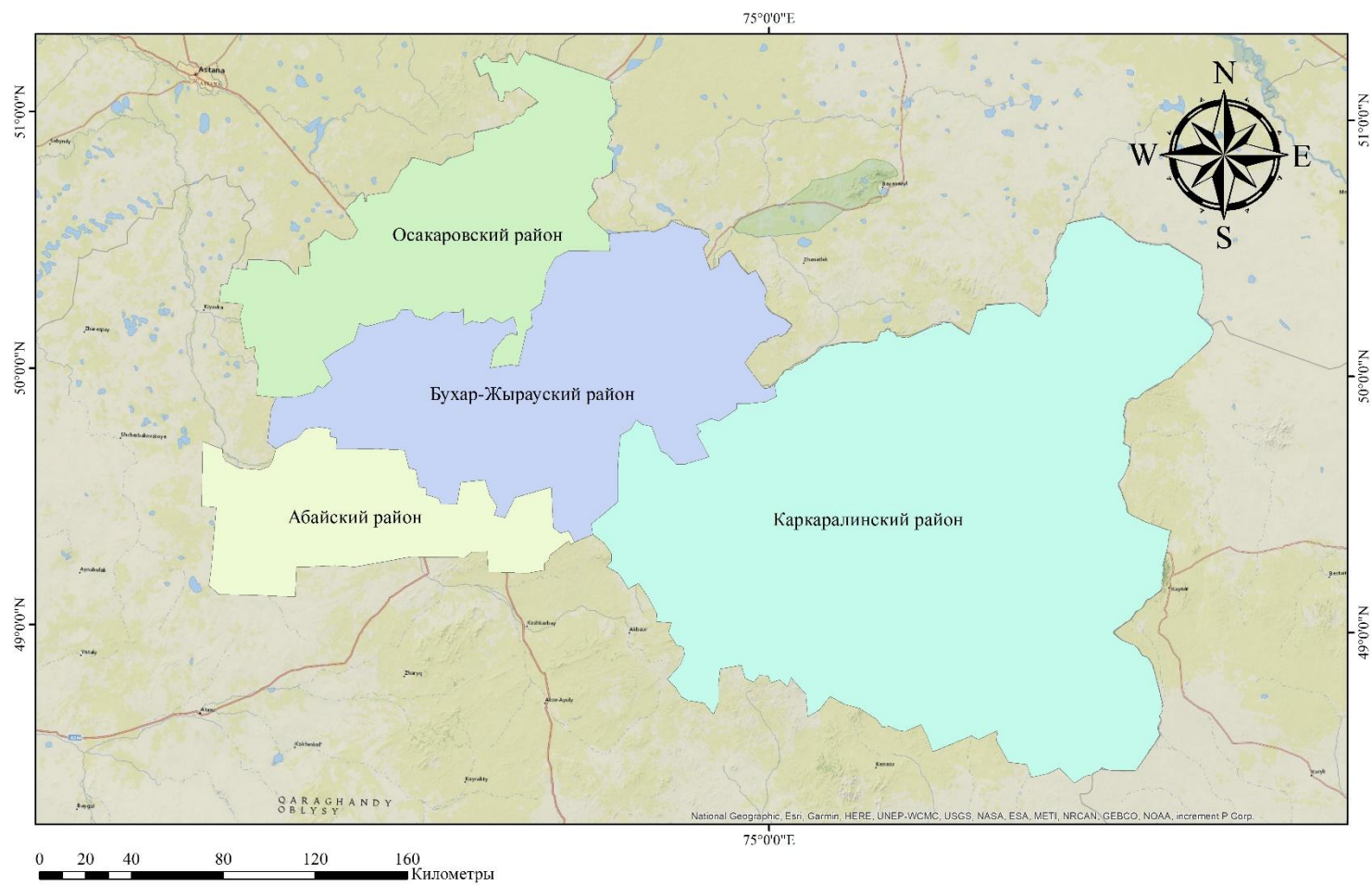
В состав исследуемой территории, по нашему мнению, целесообразно включить четыре административных района Карагандинской области: Осакаровский, Абайский, Бухар-Жырауский и Каркаралинский (рисунок 1). Эта территория характеризуется значительным разнообразием природных условий, экономической специализации и уровня урбанизации, что делает ее уникальной для анализа динамики землепользования. Все четыре района граничат между собой, формируя единую географическую область, но при этом отличаются друг от друга по своим природным и антропогенным особенностям.

Осакаровский район, площадью 12 694 км², занимает центральное положение среди рассматриваемых территорий. Он расположен в северной части Карагандинской области, на территории Казахского мелкосопочника (Сарыарки). Рельеф района сочетает равнинные участки с отдельными холмами, такими как Нияз и Акдын, высоты которых редко превышают 500 м. Почвы преимущественно каштановые, местами переходящие в южные черноземы. Климат резко континентальный: зимой температура может опускаться до -20°C, а летом достигает +35°C. Экономическая деятельность района в основном сосредоточена на сельском хозяйстве.

Каркаралинский район, являясь крупнейшим по площади среди рассматриваемых, охватывает территорию 35 353 км², что почти втрое превышает размеры Осакаровского района. Он расположен в восточной части области и включает горные системы Каркаралинского хребта, где абсолютные высоты достигают 1 403 м. На территории района сосредоточены значительные лесные массивы, которые, по нашему мнению, обладают особой природной ценностью, поскольку являются фактически единственными в пределах Центрального Казахстана. Лесные экосистемы представлены преимущественно сосновыми и берёзовыми насаждениями, что придаёт району уникальное экологическое значение. Кроме того, в Каркаралинском районе расположено большое количество озёр и родников, играющих важную роль в поддержании устойчивости местной биосферы. Климатические условия здесь несколько более умеренные по сравнению с соседними территориями, что во многом объясняется влиянием горных экосистем. В зимний период средняя температура воздуха составляет около -15 °C, тогда как в летние месяцы она колеблется в пределах +25 °C. Национальный парк «Каркаралы» представляет собой один из наиболее посещаемых природных объектов региона и, как представляется, играет важную роль в развитии туристско-рекреационного потенциала территории.

Абайский район, имеющий площадь 6 600 км², и Бухар-Жырауский район, площадью 14 935 км², находятся в центральной части Карагандинской области, образуя ядро Карагандинской агломерации. Эти районы примыкают к городу

Караганде и другим крупным промышленным центрам, таким как Темиртау и Сарань. Их рельеф преимущественно равнинный, с редкими возвышенностями. Климат резко континентальный, с крайне холодной зимой, когда температура опускается ниже -25°C , и жарким летом, когда столбик термометра поднимается выше $+35^{\circ}\text{C}$. Эти районы отличаются высоким уровнем урбанизации и сильной промышленной нагрузкой.



Легенда

Осакаровский район Бухар-Жырауский район Абайский район Каркаралинский район

Рисунок 1 – Район исследования

В Бухар-Жырауском районе расположены крупные предприятия - угольные шахты, металлургические и химические заводы. Хотя Абайский район меньше по площади, он также играет важную роль в промышленной жизни региона, обслуживая предприятия и промышленные зоны города.

Природные условия в разных частях исследуемой территории заметно различаются. Например, в Осакаровском районе преобладают равнинные степи с злаково-разнотравной растительностью, тогда как Каркаралинский район - это в основном горные ландшафты с лесами. Абайский и Бухар-Жырауский районы менее разнообразны с природной точки зрения, поскольку значительная часть их территории подверглась воздействию человека. Однако даже в этих условиях, в пойменных зонах рек, таких как Нура, можно встретить кустарниковую растительность и участки редколесий, сохраняющие определённый уровень биологического разнообразия. Особую роль на территории исследуемых районов, по нашему мнению, играют водные ресурсы. Главной водной артерией региона выступает река Нура, протекающая через Осакаровский и Бухар-Жырауский районы. В пределах Каркаралинского района значительное гидрологическое значение имеют многочисленные озёра. Водные ресурсы обеспечивают не только потребности сельского хозяйства, но и играют важную роль в функционировании промышленности, а также в жизнеобеспечении населения.

В то же время Абайский и Бухар-Жырауский районы характеризуются более высокой плотностью населения, что обусловлено развитой городской застройкой и их территориальной близостью к Караганде. Напротив, в Каркаралинском районе, несмотря на его значительную площадь, численность населения остаётся сравнительно невысокой, что, как представляется, связано с преобладанием природных ландшафтов и удалённостью от крупных транспортных узлов.

Разнообразие ландшафтных условий, а также сочетание природных и антропогенных ресурсов исследуемых районов формируют уникальную основу для анализа динамики землепользования. Каждый из рассматриваемых районов демонстрирует специфические формы взаимодействия природной среды и хозяйственной деятельности человека.

2.2 Источники данных дистанционного зондирования

Для исследования динамики землепользования на территории Карагандинской области, включающей Осакаровский, Абайский, Бухар-Жырауский и Каркаралинский районы, данные дистанционного зондирования Земли выступили ключевым инструментом анализа. По нашему мнению, их использование обеспечивает возможность комплексного изучения пространственных, временных и спектральных характеристик территории, что делает их незаменимыми при проведении углублённых исследований, создании высокоточных картографических материалов и разработке моделей прогнозирования изменений.

Общая площадь исследуемой территории составляет порядка 70 000 км², что охватывает широкий спектр ландшафтных и климатических условий. Мониторинг столь обширных пространств требует применения спутниковых данных, обладающих достаточным пространственным разрешением, высокой периодичностью съёмки и возможностью регистрации информации в различных спектральных диапазонах. В качестве основных источников информации использовались данные спутниковых систем Landsat, функционирующих в рамках международных программ мониторинга земной поверхности. Временной интервал исследования охватывает период с 2000 по 2024 гг., что позволяет фиксировать как краткосрочные колебания, так и долгосрочные тенденции трансформации землепользования.

Материалы дистанционного зондирования за указанный период содержат обширную информацию о состоянии земной поверхности, включая распределение растительного покрова, динамику гидрологических объектов и проявления антропогенного воздействия. Спутники серии Landsat (5, 7, 8) предоставляют данные с пространственным разрешением 30 м, что, на наш взгляд, является оптимальным для анализа крупных территорий и оценки изменений землепользования на региональном уровне.

В настоящем исследовании исходной базой для формирования долгосрочных рядов динамики LULC послужили стандартные спутниковые продукты, обеспечивающие получение спектрально и радиометрически откорректированных данных высокого качества. К их числу относятся оптические данные Landsat 5, Landsat 7 и Landsat 8 Collection 2 Tier 1 TOA Reflectance, ночные радиометрические данные DMSP OLS: Nighttime Lights Time Series Version 4 и VIIRS Stray Light Corrected Nighttime, а также цифровые модели рельефа SRTM3.

Использование данных Landsat 5, 7 ETM+ и 8 OLI (коллекция 2, уровень Tier 1 TOA Reflectance) обусловлено их доступностью, стабильностью радиометрической калибровки и продолжительным периодом наблюдений, начиная с начала 2000-х годов. Снимки Landsat 7 и Landsat 8 обладают высокой сопоставимостью, что позволяет корректно отслеживать долгосрочные изменения землепользования (LULC). Формат Tier 1 TOA Reflectance обеспечивает высокую точность и временную согласованность данных, что особенно важно при анализе столь обширных и ландшафтно разнообразных территорий, как районы Карагандинской области.

Помимо дневных оптических изображений существенное значение имеют ночные радиометрические данные. Для оценки динамики урбанизированных территорий (Urban Land, UL) применялись данные DMSP OLS и VIIRS с коррекцией паразитной засветки. Показатели ночной светимости, как представляется, достаточно объективно отражают уровень урбанизации, плотность застройки и степень экономической активности. Переход от данных DMSP к более современным данным VIIRS позволил повысить точность и пространственную детализацию анализа, что особенно важно при выявлении незначительных, но устойчивых изменений в структуре населённых пунктов.

Анализ ночной светимости позволяет различать районы с высокой хозяйственной активностью и плотной застройкой, а также зоны с низкой урбанизацией или слабым воздействием человека. В сочетании с оптическими данными о землепользовании и дополнительными индексами (например, растительности), такие данные дают полную картину изменений в классе UL и позволяют точно отслеживать его динамику во времени.

Особое внимание уделено также учёту рельефных особенностей рассматриваемых районов. Данные SRTM3 (рисунок 2) предоставляют информацию о высотах с глобальным охватом и относительно высокой пространственной точностью. Включение цифровой модели рельефа (ЦМР) критически важно, поскольку топография влияет на распределение природных ландшафтов, формирование микроклимата и, следовательно, на спектрально-радиометрические характеристики отражения и излучения. Например, в горных частях Каркаралинского района или холмистых и сильно расчленённых ландшафтах Бухар-Жырауского района рельеф определяет формирование различных биофизических условий, а значит, и спектральные особенности покрытий. Учитывая эти факторы, интеграция данных о рельефе из SRTM3 позволяет более корректно интерпретировать многоспектральные изображения Landsat и более адекватно выделять классы LULC, снижая ошибки, возникающие из-за геоморфологических различий.

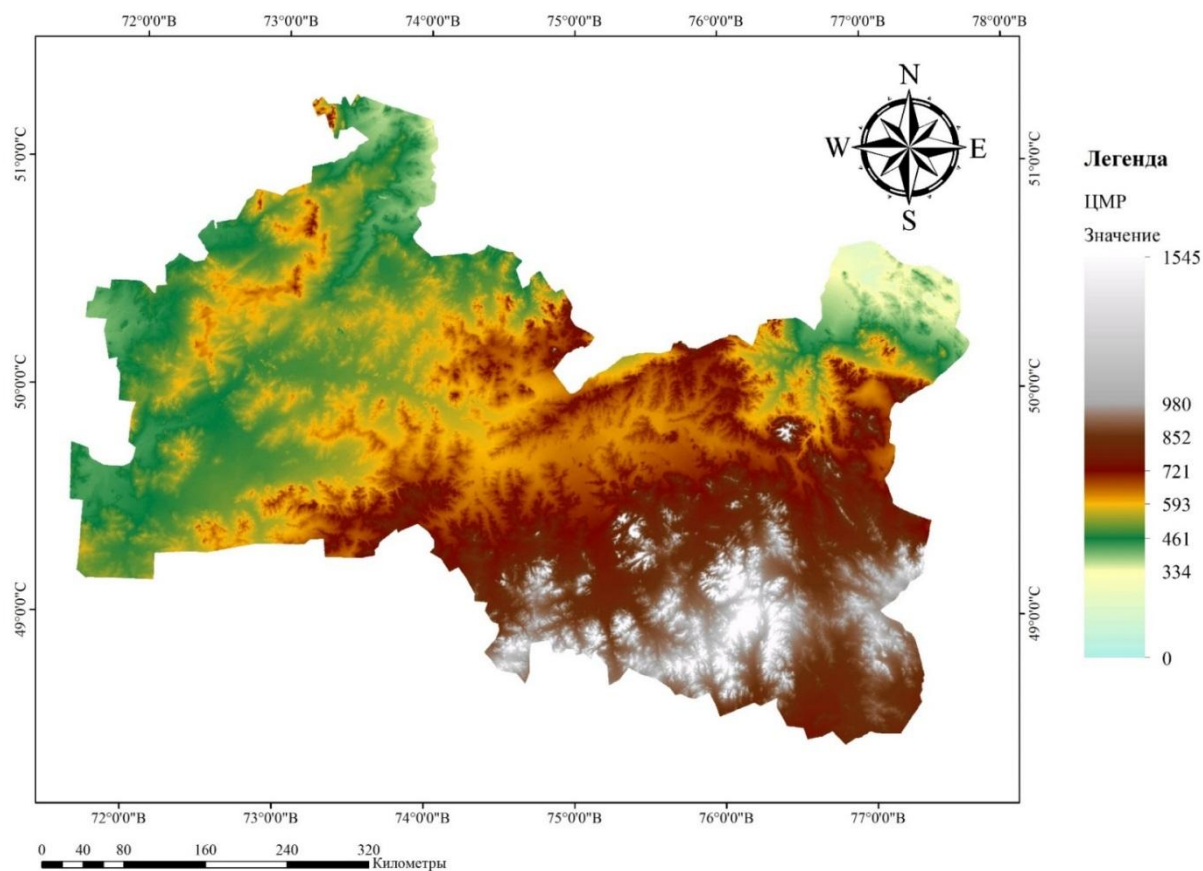


Рисунок 2 – Цифровая модель рельефа района исследования, лист 1

При отборе спутниковых данных для формирования долгосрочных LULC серий по рассматриваемым районам Карагандинской области применялся ряд жёстких критериев, обеспечивающих максимальную сопоставимость и качество исходных материалов. В первую очередь учитывалась пространственная детализация изображений. Для Landsat 7 ETM+ и Landsat 8 OLI был выбран номинальный пространственный размер пикселя порядка 30 м, что позволяет надёжно детализировать как относительно однородные агроландшафты, так и периферийные зоны небольших поселений. Подобная разрешающая способность обеспечивает оптимальный баланс между возможностями выделения мелких объектов и объёмом данных, подлежащих обработке.

Временные интервалы выбора снимков были согласованы с регулярностью съёмки спутников Landsat, имеющих 16-дневный период повторного посещения одинаковой территории. Это позволяет систематически отслеживать сезонную динамику и годовые колебания растительности, а также выявлять долговременные тренды изменения землепользования, формировать медианные композиты и интерпретировать полученные результаты в контексте многолетних тенденций. В случае данных ночной светимости (DMSP OLS и VIIRS) предпочтение отдавалось согласованным временным периодам, обеспечивающим сопоставимость уровня светимости для разных лет и сезонов, что особенно важно при анализе изменений урбанизированных территорий.

Для обеспечения наглядности, ключевые особенности и характеристики применяемых наборов данных были сведены в таблицу 1. Данная таблица 1 предоставляет информацию об основных параметрах, таких как пространственное разрешение, временной охват, спектральные диапазоны, калибровочные процедуры и форматы данных.

Таблица 1 – Характеристики применяемых наборов данных

Набор данных	Дата	Каналы	Разрешение
Landsat 5 TM Collection 2 calibrated TOA Reflectance	2000-2013 (используется совместно с Landsat-7 для максимального покрытия, до завершения миссии в 2013 году)	Соответствующие для TM	30 m
Landsat 7 Collection 2 Tier 1 TOA Reflectance	2000-2024 (до 2013 года используется совместно с Landsat-5; с 2003 года работает в режиме SLC -off; с 2013 года используется совместно с Landsat-8 для обеспечения непрерывности данных)	B1-B7	30 m
Landsat 8 Collection 2 Tier 1 TOA Reflectance	2013 – 2024 (начиная с 2013 года, используется совместно с Landsat-7 для компенсации потерь данных из-за SLC -off)	B2-B7	30 m
VIIRS Stray Light Corrected Nighttime	2014 – 2024	avg_rad	463.83 m

DMSP OLS: Nighttime Lights Time Series Version 4	2000 – 2014	avg_vis	927.67 m
SRTM 3	2000	elevation	90 m
TerraClimate: Monthly Climate	2000 – 2024	Pr	4638.3 m
MODIS – LST	2000 – 2024	LST_Day_1km	1000 m
ERA5 Monthly Aggregates	2000 – 2024	total_precipitation	27830 m

Особое внимание уделялось критерию облачности. Из всей совокупности доступных снимков Landsat 5, Landsat 7 и Landsat 8 для каждого сезона отбирались только те, у которых уровень облачности не превышал 20%. Данный порог обеспечивает высокую чёткость обрабатываемых композитов и сводит к минимуму искажения, вызванные светорассеянием и затемнением, характерными для облачных покрытий.

При этом количество доступных сцен оставалось достаточным, чтобы сформировать репрезентативные статистические выборки для расчёта медианных композитов. В сочетании с данными рельефа (SRTM3) и ночной светимости, строгая фильтрация по облачности гарантировала высокую точность интерпретации спектральных индексов и дальнейшую успешную LULC классификацию.

Формирование медианных композитов выступало одним из ключевых этапов в обеспечении стабильности и достоверности данных, необходимых для пространственно-временного анализа LULC. Данный подход позволял сгладить различные типы помех: от артефактов, возникающих из-за сбоев аппаратуры (SLC-off у Landsat 7), до временных атмосферных эффектов, таких как облачность и дымка.

В процессе реализации применялись вычислительные возможности платформы Google Earth Engine (рисунок 3), обеспечивающие автоматизацию алгоритмов и доступ к обширным архивам спутниковых данных.

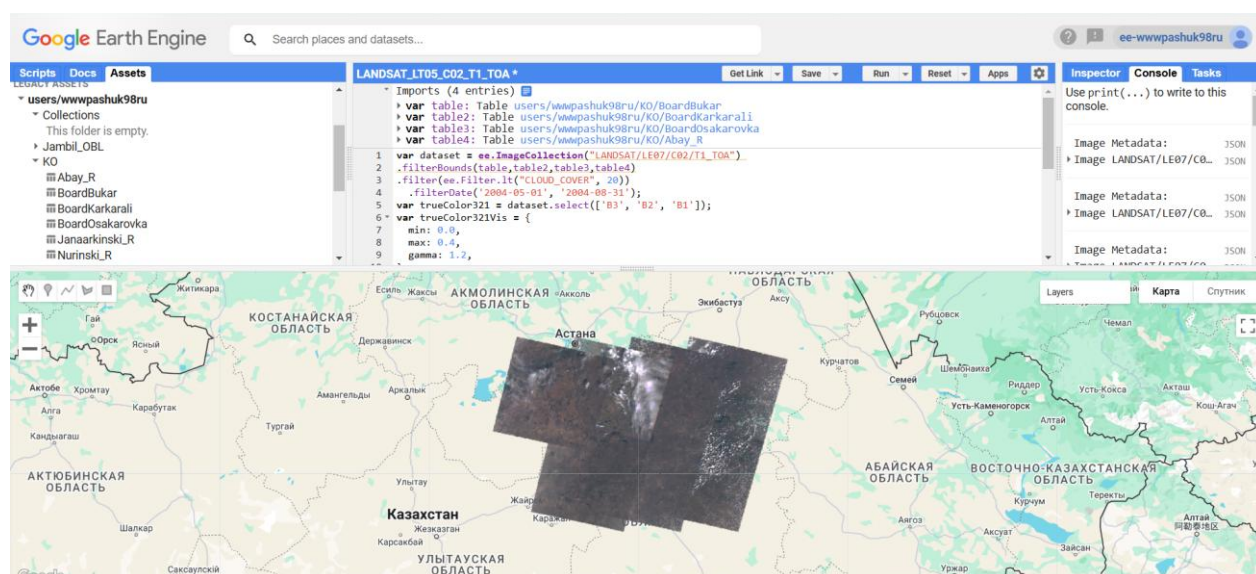


Рисунок 3 – Интерфейс среды GEE

Подход к созданию медианных композитов основывался на последовательной фильтрации и статистической обработке отснятых материалов за заранее определённый временной интервал. Сначала подбирался набор снимков Landsat 5, Landsat 7 и Landsat 8, удовлетворяющий критериям по облачности (не более 20%) и качеству, что гарантировало минимально возможное присутствие искажений. Доступные инструменты маскировки пикселей с плотной облачностью, атмосферной дымкой и иными аномалиями обеспечивали очистку исходного набора данных. Затем все прошедшие отбор изображения сводились в единый «стек», где каждый пиксель поочерёдно анализировался с учётом временной последовательности.

На этом этапе, для каждого пространственного положения вычислялась медиана отражательности по всем имеющимся в рассматриваемом периоде снимкам. Подобный статистический критерий, в отличие от среднего арифметического, характеризуется меньшей чувствительностью к редким аномальным значениям, вызванным, случайным шумом или нестандартными условиями освещённости. Итоговый медианный композит (рисунок 4) представлял собой синтетическое изображение, лишённое экстремальных значений, которые не отражают реальную долговременную картину спектральных характеристик земной поверхности. Такой подход повышал стабильность и надёжность результатов, делая последующую классификацию более точной и воспроизводимой.

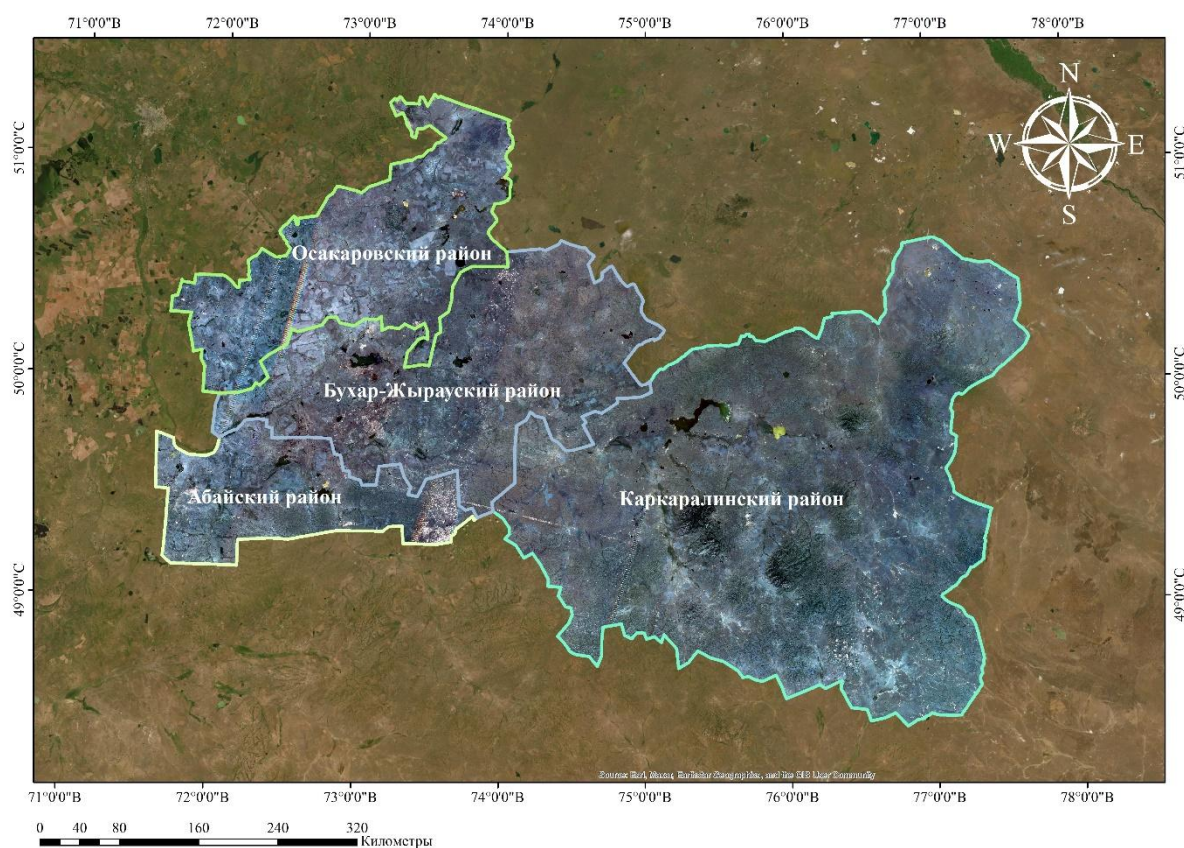


Рисунок 4 – Медианный композит района исследования

Применение медианных композитов приводило к формированию наиболее репрезентативного и устойчивого во времени «портрета» поверхности территории. В условиях сезонных и годовых колебаний климатических параметров медиана выступала своеобразным фильтром, минимизируя влияние экстремальных ситуаций, таких как краткосрочные затопления и локальные пожары, которые не соответствуют типичному состоянию ландшафта. В итоге формировался сглаженный, не подверженный единичным выбросам массив данных, который затем был использован для интерпретации и классификации изменений LULC, а также для выявления долгосрочных трендов развития исследуемых регионов.

Процесс вычисления вегетационных индексов являлся неотъемлемым элементом комплексного анализа динамики LULC, позволяя более точно выделять особенности растительного покрова, влажных территорий и антропогенно изменённых ландшафтов. Ключевая идея расчёта таких индексов заключается в преобразовании многоспектральных данных, полученных со спутников, в сводные метрики, чутко реагирующие на состояние растительности, степень увлажнённости почвы, плотность застройки или количество отложений на поверхности. В данном исследовании применялся широкий спектр индексов, среди которых EVI, NDVI, NDWI, NDBI, NDTI, NDMI, MNDWI. Их расчёт был реализован средствами платформы Google Earth Engine (рисунок 5), что существенно упрощало массовую обработку больших объёмов данных.

```

49 var green = landsat.select('B2');
50
51 var mndwi = swir.subtract(green).divide(swir.add(green)).rename('MNDWI');
52 var mndwiParams = {min: -1, max: 1, palette: [ '#00008b', 'blue', 'yellow', 'orange', 'red' ]};
53 Map.addLayer(mndwi, mndwiParams, 'MNDWI');
54
55 var ndbi = nir.subtract(swir).divide(nir.add(swir)).rename('NDBI');
56 var ndbiParams = {
57   min: -1,
58   max: 1,
59   palette: [ '#00008b', 'blue', 'white', 'yellow', 'red' ]
60 };
61 // Добавление слоя NDBI на карту
62 Map.addLayer(ndbi, ndbiParams, 'NDBI');
63
64 var ndmi = nir.subtract(swir).divide(swir.add(nir)).rename('NDMI');
65 var ndmiParams = {min: -1, max: 1, palette: [ '#00008b', 'blue', 'yellow', 'orange', 'red' ]};
66 Map.addLayer(ndmi, ndmiParams, 'NDMI');
67
68 var ndvi = nir.subtract(red).divide(nir.add(red)).rename('NDVI');
69 var ndviParams = {min: -1, max: 1, palette: [ 'red', 'orange', 'yellow', 'green', '#006400' ]};
70 Map.addLayer(ndvi, ndviParams, 'NDVI');
71
72 var savi = nir.subtract(red).divide(nir.add(red).add(2.0)).multiply(10.0);
73 var saviParams = {min: -1, max: 1, palette: [ 'blue', 'white', 'green' ]};
74 Map.addLayer(savi, saviParams, 'SAVI');
75
76 var ndti = green.subtract(swir).divide(green.add(swir)).rename('NDTI');
77 var ndtiParams = {min: -1, max: 1, palette: [ '#00008b', 'blue', 'yellow', 'orange', 'red' ]};
78 Map.addLayer(ndti, ndtiParams, 'NDTI');

```

Рисунок 5 – Расчет вегетационных индексов в GEE

Для вычисления каждого индекса сперва определялись необходимые спектральные каналы, соответствующие конкретным характеристикам поверхности. Для NDVI (Normalized Difference Vegetation Index), традиционно применялись значения отражённости в красном (RED) и ближнем инфракрасном (NIR) диапазонах. Пример расчета NDVI приведен в формуле (1):

$$NDVI = \frac{NIR - RED}{NIR + RED} \quad (1)$$

Высокие положительные значения указывали на плотный и здоровый растительный покров, в то время как низкие или отрицательные были характерны для территорий без растительности и с очень слабым её развитием, таких как урбанизированные зоны и водные объекты.

В случае EVI (Enhanced Vegetation Index) применялся дополнительный синий канал (BLUE), что позволяло корректировать воздействие атмосферных аэрозолей и повышать чувствительность индекса к плотности растительного покрова. Пример расчета EVI приведен в формуле (2). Это достигалось путём более сложных алгоритмических преобразований, включающих коэффициенты, подобранные эмпирическим путём.

$$EVI = G \times \frac{NIR - RED}{NIR + C_1 \times RED - C_2 \times BLUE + L} \quad (2)$$

где G – коэффициент усиления равный 2,5;

C_1 и C_2 – эмпирические коэффициенты равные 6 и 7,5;

L – коэффициент поправки для минимизации влияния фонового освещения равный 1.

Результат EVI был более устойчив к влиянию атмосферных помех и сатурации сигналов в густых лесах, что делало его особенно ценным для анализа территорий с высокой биомассой.

Индексы, такие как NDWI (Normalized Difference Water Index) и MNDWI (Modified Normalized Difference Water Index), были направлены на выделение водных поверхностей и использовали отражение в зеленом диапазоне (GREEN) и коротковолновом инфракрасном диапазоне (SWIR). Примеры расчетов NDWI и MNDWI приведены в формулах (3) и (4) соответственно.

$$NDWI = \frac{GREEN - NIR}{GREEN + NIR} \quad (3)$$

$$MNDWI = \frac{GREEN - SWIR}{GREEN + SWIR} \quad (4)$$

Они использовали сочетания ближнего и среднего инфракрасного диапазонов с зелёным или другим каналом, чувствительным к водной среде. Высокие значения NDWI или MNDWI свидетельствовали о наличии водных объектов и увлажнённых почв, что позволяло различать реки, озёра, болота, а также отслеживать динамику влажности ландшафта.

Для оценки антропогенного воздействия и уровня урбанизации применен NDBI (Normalized Difference Built-up Index), отражающий соотношение сигналов в ближней и средней инфракрасной области спектра. Пример расчета NDBI приведен в формуле (5).

$$NDBI = \frac{SWIR - NIR}{SWIR + NIR} \quad (5)$$

Повышенные значения NDBI указывали на присутствие плотной застройки и искусственных поверхностей, таких как дороги, промышленные площадки, жилые кварталы. Аналогично, NDTI (Normalized Difference Tillage Index) и NDMI (Normalized Difference Moisture Index) были разработаны для более тонкой диагностики отдельных процессов, затрагивающих поверхность, включая характеристики почвы, влажность растительного покрова и степень изменения структуры ландшафта. Примеры расчетов NDTI и NDMI приведены в формулах (6) и (7) соответственно.

$$NDTI = \frac{SWIR_1 - SWIR_2}{SWIR_1 + SWIR_2} \quad (6)$$

$$NDMI = \frac{NIR - SWIR}{NIR + SWIR} \quad (7)$$

где $SWIR_1$ – коротковолновый инфракрасный диапазон (длинноволновой);

$SWIR_2$ – коротковолновый инфракрасный диапазон (коротковолновая).

Практическая реализация вычисления индексов сводилась к применению соответствующих формул для каждого пикселя отобранного и предварительно обработанного изображения. Используя возможности GEE, все операции по расчету производились пакетно: для выбранного набора сцен, прошедших фильтрацию по критериям облачности и качества, к каждому пикселю поочередно применялись формулы индексов. Результирующий слой данных интегрировался в медианный композит, повышающий стабильность результата за счёт статистического сглаживания.

После вычисления вегетационных индексов получались новые информационные каналы, дополнительно расширяющие спектральный контекст базовых измерений Landsat. Их значение, по нашему мнению, не ограничивалось лишь характеристикой состояния растительности или уровня увлажнённости территорий; использование индексов способствовало повышению контрастности между различными типами земного покрова, снижению неоднозначности интерпретации и улучшению разделимости классов LULC. Это, в свою очередь, существенно повышало эффективность последующего этапа классификации, позволяя алгоритмам машинного обучения более точно идентифицировать категориальные различия между сельскохозяйственными, природными и урбанизированными территориями.

Таким образом, включение вегетационных и производных от них индексов в общий рабочий процесс играло ключевую роль в системном и достоверном изучении ландшафтной динамики. Рассчитанные на основе фундаментальных спектральных соотношений, данные индексы формировали дополнительный параметрический слой, который способствовал выявлению скрытых взаимосвязей и позволял более объективно оценивать изменения землепользования, а также состояние растительного покрова в различных природных и антропогенных условиях.

Все этапы сбора данных, их предварительной обработки, а также последующей генерации карт LULC с применением методов машинного обучения осуществлялись в среде GEE. Данная облачная платформа предоставляет удобный доступ к глобальным архивам спутниковых данных, обеспечивает мощные вычислительные ресурсы и предлагает функциональные инструменты для анализа временных рядов. Использование GEE позволило автоматизировать ключевые процессы – от фильтрации и калибровки исходных спутниковых снимков Landsat, DMSP OLS, VIIRS и SRTM3 до расчёта вегетационных индексов, создания медианных композитов и выполнения классификационных алгоритмов. Написание пользовательских скриптов на языке JavaScript обеспечило гибкость и воспроизводимость процедур, упрощая интеграцию различных источников данных, применение готовых масок, настройку параметров машинного обучения и валидацию результатов. Благодаря этим возможностям GEE ускорил реализацию всего комплекса задач и существенно повысил эффективность пространственного анализа LULC для исследуемых районов.

Процессы сбора, фильтрации и предварительной обработки спутниковых данных, а также расчёт вегетационных индексов были реализованы в среде Google Earth Engine (GEE) с использованием JavaScript и встроенных инструментов Code Editor. Работа в GEE значительно упростила доступ к обширным архивам спутниковых данных и позволила выполнять сложные пространственно-временные операции прямо на серверах Google, без необходимости загружать данные на локальный компьютер (Приложение Б). После формирования медианных композитов и расчёта вегетационных индексов следующим этапом исследования стало применение методов машинного обучения, в частности алгоритма Random Forest (RF), для классификации типов землепользования (LULC). Вместе с тем, как представляется, до непосредственного этапа моделирования требовалась предварительная подготовка обучающих данных.

Данный подготовительный этап включал формирование репрезентативной обучающей выборки, настройку ключевых параметров алгоритма, а также интеграцию композитных изображений с дополнительными тематическими слоями, такими как NDVI, EVI, NDWI и другими производными индексами.

Только после завершения этой подготовки можно переходить к построению модели RF, что рассматривается подробнее в следующем разделе.

Полный рабочий цикл – от извлечения данных до получения классификационных карт – реализован с помощью модульного скрипта, созданного в среде Google Earth Engine (GEE) с использованием встроенных функций, таких как `ee.Classifier`. Такой подход гарантирует воспроизводимость и документированность каждого этапа, что делает процесс удобным для дальнейших повторений или масштабирования.

2.3 Сбор эталонных полевых данных

Важным этапом данного исследования являлся сбор полевых эталонных данных, необходимых для калибровки и проверки результатов классификации LULC. Полевые работы проводились в летний период 2023 и 2024 годов, когда спектральные характеристики и морфология поверхностных объектов наиболее четко выражены, что позволяет минимизировать сезонные искажения. В ходе экспедиционных выездов были задокументированы координаты и визуальные характеристики ключевых типов покрытий в границах исследуемых районов.

Основу эталонных данных составляли точки и полигоны наблюдений, привязанные к GPS-координатам (WGS-84), и дополнительные комментарии по состоянию ландшафтов. Собранный массив эталонных данных был структурирован по шести целевым классам LULC:

- пашни (представленные преимущественно возделываемыми сельскохозяйственными землями и агротехническими объектами);
- пастбища (территории с преобладанием естественных или полусеменных кормовых угодий для скота);

- городская застройка (антропогенные структуры с плотной инфраструктурой, включая жилые кварталы, промышленные и административные объекты);
- водно-болотные угодья (болота, заболоченные луга и другие переходные экосистемы с перенасыщенной влагой почвой);
- водные объекты (озёра, реки, пруды и другие постоянные или временные водоемы);
- леса (естественные и искусственные древесные насаждения с доминированием древесной растительности).

Подход к сбору эталонных данных включал непосредственные полевые наблюдения и их последующую верификацию с использованием высокодетальных спутниковых изображений и геопортальных ресурсов. Такая многоступенчатая проверка гарантировала достоверность и репрезентативность полученных наблюдений и повышала надежность дальнейшей классификации, основывающейся на машинном обучении. В итоге, собранные эталонные данные выступили в качестве ключевого звена, необходимого для оптимизации параметров алгоритмов классификации, проверки результатов и количественной оценки точности итоговых карт LULC.

Одним из ключевых подходов к формированию корректно отобранного набора эталонных данных является стратифицированная выборка. В контексте данного исследования стратификация позволяла обеспечить сбалансированное и репрезентативное представление всех целевых классов LULC, учитывая их пространственное распределение, гетерогенность ландшафтов и потенциальные различия в спектральных характеристиках.

Стратифицированная выборка эталонных участков предполагает предварительное разбиение всей исследуемой территории на классы (страты), соответствующие различным типам LULC. Данное разбиение основывалось на предварительном знании о ландшафтной структуре региона, при полевых выездах. Цель стратификации - гарантировать, что каждый целевой класс будет представлен в эталонных данных в достаточной мере, несмотря на различия в площади и частоте встречаемости этого класса.

После определения страт выполнялась выборка эталонных участков - точек и полигонов. В отличие от простого случайного отбора по всей территории, где крупные и распространённые классы (пашни, пастбища) могли бы доминировать в выборке, стратифицированный подход предусматривает целенаправленный отбор определённого числа наблюдений из каждого класса. Тем самым можно избежать диспропорций, возникающих из-за пространственного преобладания какого-либо типа угодий или, наоборот, редкости другого. Важным аспектом являлось определение оптимального объёма выборки для каждой страты: если класс очень редок, то имеет смысл увеличить его долю в эталонном массиве, чтобы гарантировать надёжную оценку его спектрально-пространственных характеристик.

Для реализации стратифицированной выборки применялись геоинформационные методы. В начале, используя доступные векторные слои

(границы сельскохозяйственных полей, поселений, контуры лесов и водных объектов) (рисунок 6), территория разбивалась на страты. Затем внутри каждой страты осуществлялся случайный отбор точек и полигонов с соблюдением двух ключевых критериев: равномерности пространственного распределения и достаточной удалённости от границ классов. Последний фактор, по нашему мнению, имеет принципиальное значение, поскольку точки, расположенные вблизи межклассовых границ, нередко отражают смешанное спектральное поведение, характерное для переходных зон. Это существенно усложняет интерпретацию данных и повышает вероятность возникновения классификационных ошибок. В ряде случаев дополнительно применялись дистанционные методы верификации — в частности, анализ высокодетальных спутниковых снимков — с целью подтверждения соответствия выбранных точек заявленным классам.

В процессе стратификации учитывались также пространственно-временные особенности исследуемых территорий. Так, сельскохозяйственные угодья демонстрируют различные спектральные характеристики в зависимости от фазы вегетации культур и степени свежести обработки почвы. Стратифицированная выборка, сформированная с учётом сезонной динамики и временных интервалов проведения съёмок, позволила, на наш взгляд, создать эталонный набор данных, отражающий реалистичное многообразие состояний ландшафта.

Подобный подход повысил устойчивость алгоритмов классификации, тренируемых на широком спектре входных условий, и улучшил их обобщающую способность.

Кроме того, стратифицированная выборка обеспечивала оптимальную основу для статистических процедур проверки точности (accuracy assessment) и расчёта метрик качества классификации, таких как матрицы ошибок, показатели Producer's Accuracy, User's Accuracy, F1-score и Кappa-статистика. Наличие репрезентативного набора эталонных точек во всех классах дало возможность надёжно оценить как общее качество тематической карты, так и точность классификации по каждому классу в отдельности. В совокупности это снизило риск смещения оценок, вызванного избирательной представленностью классов или их недостаточной численностью.

результатов классификации LULC. Она гарантировала надёжную, статистически обоснованную и равномерно распределённую базу для обучения и проверки алгоритмов.

2.3.1 Проведение полевых работ

Важным направлением повышения точности и достоверности полевых эталонных данных в рамках данного исследования стало активное использование передовых GNSS-технологий при выполнении геопривязки эталонных точек и участков. Технологии глобальных навигационных спутниковых систем (GNSS), в частности систем GPS, ГЛОНАСС, Galileo и BeiDou, существенно облегчают определение координат с высокой точностью в реальном времени и в постобработке. Ключевым элементом метода явились высокоточные GNSS-приёмники, применяемые как в полевых условиях, так и для последующего уточнения координатного решения методами постобработки.

На этапе проведения полевых кампаний в летние периоды 2023 и 2024 годов, операторы, оснащённые портативными GNSS-приёмниками, осуществляли привязку эталонных участков непосредственно в месте отбора данных. Использование современных режимов GNSS-съёмки, таких как кинематический режим реального времени (RTK) и постобработанный кинематический режим (PPK), обеспечило позиционирование с точностью до 2–3 см. По нашему мнению, это имеет особую значимость для территорий с неоднородным и сложным рельефом, где даже незначительные ошибки геопозиционирования способны привести к некорректной интерпретации спектральных характеристик земной поверхности и, как следствие, к снижению точности классификационных моделей.

Одной из применённых методик стала процедура прецизионного определения координат с использованием технологии Precise Point Positioning (PPP), основанной на применении высокоточных орбит и поправок спутниковых часов, предоставляемых глобальными центрами анализа, в частности IGS (International GNSS Service). В отличие от дифференциальных методов, требующих развертывания базовой станции или использования сети постоянных опорных станций, PPP обеспечивает получение высокоточной геопозиции в любой точке земного шара без необходимости создания локальной инфраструктуры.

Вместе с тем следует отметить, что технология PPP требует более продолжительного времени наблюдений — как правило, от 20 до 60 минут непрерывных измерений, необходимых для конвергенции решения. Данный подход, на наш взгляд, оказался особенно востребованным при проведении полевых работ в удалённых и труднодоступных районах, где развертывание собственной базовой станции затруднено либо экономически нецелесообразно.

Кроме PPP, в ряде случаев использовались и более традиционные методы дифференциального позиционирования. Там, где это было возможно, осуществлялась интеграция с референсными станциями. Использование RTK (Real-Time Kinematic) корректировок, получаемых по каналу сотовой связи,

позволяло мгновенно получать координаты с точностью до нескольких сантиметров. Такой режим особенно удобен при выполнении множества эталонных замеров в течение короткого периода времени, при последовательном обходе ряда эталонных площадок. При использовании RTK оператор, с портативным GNSS-приёмником с поддержкой данного режима, оперативно фиксировал координаты точек, минимизируя временные затраты и обеспечивая однородный по точности массив данных.

Применение GNSS-технологий дополнялось тщательным выбором оптимального времени суток и условий при проведении замеров. Для повышения качества геопозиционирования учитывались такие факторы, как конфигурация спутников на небосводе, степень мультипутинга (переотражения сигналов), атмосферные условия и ионосферные возмущения. В ситуациях с затенённым небосводом, в лесных и городских условиях, точность напрямую зависела от способности приёмника отслеживать достаточное количество спутников в разных созвездиях. Использование многочастотных приёмников обеспечивало более устойчивое решение, поскольку приём сигнала на нескольких частотах уменьшал влияние ионосферных задержек.

Методы PPP и RTK дополнялись процедурами постобработки с применением специализированного программного обеспечения. При постобработке использовались высокоточные орбитальные параметры и данные региональных референсных станций. В случае PPP для этого применялись ресурсы, предоставляемые IGS. Подобные решения, будучи интегрированными с исходными полевыми наблюдениями, позволяли уточнять координаты эталонных точек задним числом, повышая их достоверность до сантиметрового уровня, даже если в полевых условиях были зафиксированы менее точные координаты.

Полученные с помощью высокоточного GNSS-позиционирования эталонные данные играли роль эталона при обучении и проверке алгоритмов классификации. Они гарантировали, что выборки пикселей, соответствующие эталонным участкам (пашни, пастбища, городская застройка, водно-болотные угодья, водные объекты, леса), максимально точно соотносятся с реальными географическими объектами. Это повышало надёжность результирующих LULC-карт и снижало погрешности, вызванные геопривязочными ошибками. В контексте машинного обучения высокоточная геопозиция эталонных точек обеспечивала корректную репрезентацию пространственно-временных закономерностей, важных для дифференциации классов, и минимизацию перекрытий между различными типами земель.

Задействование передовых GNSS-методов, включая PPP, RTK, PPK и интеграцию с референсными станциями, позволило создать высокоточный эталонный пространственный каркас для всех последующих стадий анализа. Данные, собранные и уточнённые таким образом, выступали основой для формирования выборок обучения и валидации классификаторов, а также для оценки точности и надёжности итоговых карт LULC. Применение данных подходов отвечало современным стандартам геодезической точности,

обеспечивая надлежащее качество и репрезентативность эталонных полевых измерений в рамках многоступенчатого процесса интерпретации данных ДЗЗ.

Важным дополнением к высокоточному GNSS-позиционированию в рамках данного исследования стало использование беспилотных летательных аппаратов (БПЛА) с установленными на них мультиспектральными камерами, позволяющими получать детализированные изображения земной поверхности с субдециметровым пространственным разрешением и оптимизированным спектральным диапазоном. Применение таких комплексных систем как DJI Phantom 4 с мультиспектральным модулем, обеспечивало глубинную диагностику характеристик растительного покрова, структуры ландшафта и состояния почвенно-растительного покрова в контролируемых и труднодоступных точках исследуемой территории.

Ключевым преимуществом использования БПЛА в полевых работах являлась возможность оперативного и точечного сбора данных высокого разрешения. Если спутниковые платформы (Landsat, Sentinel и др.) обладают фиксированными траекториями, ограниченным временным разрешением и относительно крупным пространственным шагом, то БПЛА позволяет выбирать оптимальные сроки вылета, угол освещённости, масштаб съёмки и высоту полёта в зависимости от конкретных задач и погодных условий. Это особенно важно при необходимости верификации спектральных характеристик отдельных классов LULC, а также при уточнении границ классов и выявлении мелкомасштабных неоднородностей.

Процедура использования БПЛА включала несколько этапов. Сначала определялись ключевые полигоны, требующие детальной съёмки, затем проводилась планировка полёта с учётом рельефа, наземных препятствий и оптимизации перекрытий изображений (продольных и боковых) для последующей фотограмметрической обработки.

Оборудованные мультиспектральными сенсорами БПЛА, DJI Phantom 4 Multispectral, обеспечивали регистрацию отражательной способности поверхности в нескольких спектральных диапазонах, включающих синий, зелёный, красный, ближний инфракрасный (NIR) и красный край (Red Edge). Наличие подобных каналов позволяло рассчитывать широкий набор вегетационных индексов (NDVI, EVI и другие), дающих точную оценку состояния растительности вплоть до уровня отдельных растений. Высокая пространственная детальность в сочетании с мультиспектральной информацией значительно обогащала эталонный набор данных, используемый для обучения и валидации классификационных алгоритмов, поскольку классические спутниковые данные Landsat с 30-метровым разрешением не всегда способны различать тонкие детали структуры агроландшафтов или урбанизированных зон.

После завершения полётов и получения массива мультиспектральных снимков применялись методы фотограмметрической обработки: изображения выравнивались, сшивались в ортомозаики и корректировались по геометрии и радиометрии. В итоге, корректно откалиброванные ортомозаики, дополненные мультиспектральными индексами, давали высокоточную, консистентную и

пространственно согласованную основу для проверки результатов LULC-классификации.

Значимое преимущество данных БПЛА заключалось в возможности идентифицировать мелкомасштабные объекты, переходные зоны, редкие и специфичные ландшафтные элементы, которые были сглажены и неразличимы на среднеразрешённых спутниковых сценах. К примеру, внутри класса пастбищ можно было детально различать участки с разной степенью деградации растительного покрова, в городских ландшафтах - выделить мелкие хозяйственные постройки и зелёные насаждения, а среди водно-болотных угодий - точно определить границу между заболоченными лугами и открытой водной гладью. Такая информация критически важна для корректного формирования эталонных выборок, поскольку при подготовке датасетов для машинного обучения точность тикетирования каждого пикселя и полигона прямо влияет на качество и точность итоговой классификации.

Использование БПЛА также позволяло оперативно реагировать на непредвиденные обстоятельства: при выявлении расхождений между наземными наблюдениями и спутниковыми данными повторно проводилась съёмка проблемного участка, уточнив его статус, пространственные границы и спектральные характеристики. Кроме того, данные, полученные с использованием БПЛА, выступали эффективным инструментом многоуровневой валидации. Сопоставление ортомозаик и рассчитанных на их основе спектральных индексов с эталонными точками, зафиксированными методами GNSS, обеспечивало двусторонний контроль качества. Если GNSS, по нашему мнению, гарантировал высокую точность координатной привязки, то БПЛА, оснащённый мультиспектральной камерой, предоставлял детализированную информацию о спектрально-пространственных характеристиках поверхности. В совокупности это позволило сформировать комплексную и надёжную эталонную информационную базу.

Применение БПЛА с мультиспектральными сенсорами, в частности платформы Phantom 4, стало одним из ключевых факторов, позволивших вывести полевые эталонные исследования на качественно новый уровень точности и детализации. Высокое пространственное разрешение, гибкость в выборе времени и параметров съёмки, возможность расчёта вегетационных индексов, а также точная геопривязка на основе GNSS-методов все эти преимущества, как представляется, существенно укрепили методическую основу последующих этапов LULC-анализа и повысили достоверность итоговых карт землепользования.

2.3.2 Интеграция полевых данных в ГИС

Интеграция полевых материалов с геопространственными данными в среде ГИС, по нашему мнению, выступала критически важным этапом формирования комплексной информационной базы для последующего анализа LULC. Геоинформационные системы, в частности ArcGIS, обладают широким спектром функциональных возможностей от базовой визуализации и простого

наложения тематических слоёв до реализации сложных процедур пространственного анализа, статистической обработки и геопространственного моделирования.

Это позволило не только эффективно структурировать и систематизировать имеющийся массив полевых эталонных данных, но и существенно расширить его аналитический потенциал за счёт интеграции с другими тематическими геоданными.

В рамках данного исследования, собранные полевые точки, полученные с применением высокоточных GNSS-методов, а также данные БПЛА и мультиспектральных камер, импортировались в ГИС для их последующей систематизации и анализа. Каждая точка и полигон в эталонном наборе данных, имея уникальный идентификатор, связывалась с атрибутивной информацией: тип LULC, присутствующий класс (пашни, пастбища, городская застройка, водно-болотные угодья, водные объекты или леса), данные о состоянии растительности, наличии антропогенного вмешательства. Таким образом, ГИС становилась своего рода «виртуальным атласом» эталонных участков, доступным для многоракурсного анализа и проверки.

Геоинформационное моделирование, по нашему мнению, предоставляло возможность выполнения широкого спектра продвинутых аналитических операций с эталонными данными. Так, пространственные запросы в ГИС позволяли осуществлять фильтрацию и сортировку эталонных точек по заранее заданным критериям: например, отбор объектов, расположенных выше определённых высотных отметок, выявление эталонов, находящихся на заданном расстоянии от рек или транспортных магистралей, а также выбор участков в пределах конкретных административных границ или природно-климатических зон.

Подобная селективность обеспечивала, как представляется, возможность проведения таргетированного анализа отдельных фрагментов территории и изучения специфических ландшафтных ситуаций с учётом их природных и антропогенных особенностей.

Кроме того, ГИС-среда активно использовалась при разработке методологий стратифицированной выборки эталонных данных. Опираясь на базовые векторные слои, а также опорные данные из спутниковых снимков средней и высокой детализации, формировались страты и автоматически генерировались репрезентативные выборки точек внутри каждой из них. Это значительно ускоряло процесс подготовки эталонных данных, снижало риск субъективных ошибок и повышало степень воспроизводимости подходов, особенно в масштабных проектах с большим объёмом анализируемой территории.

Инструментарий геостатистического анализа, доступный в ArcGIS, открывал возможности не только для точечной проверки данных, но и для оценки их пространственной согласованности. Так, на основе тепловых карт распределения эталонов по различным классам LULC, оценивалась плотность точек в каждой стратификационной категории и определять, нет ли «провалов»

или «горячих точек» в их пространственной репрезентативности. Это обеспечивало сбалансированное представление всех классов и возможность при необходимости скорректировать стратегию полевых работ или расширить использование дополнительной информации (ещё раз выехать на определённый тип пастбищ и изучить дополнительный фрагмент лесного массива).

Важным аспектом интеграции данных в ГИС было также совместное использование ортомозаик, полученных с БПЛА, и спутниковых изображений, предоставляемых платформами дистанционного зондирования. Чёткая пространственная регистрация эталонных точек на ортомозаиках позволяла убедиться в корректности их позиционирования и классовой принадлежности: если спутниковые данные, Landsat, дают разрешение в десятки метров, то сантиметровое разрешение БПЛА и точные координаты GNSS позволяли однозначно интерпретировать визуальные особенности ландшафта и сопоставлять их с пиксельными значениями.

Таким образом, ГИС выступала методологической и технологической платформой, которая позволяла в полном объёме интегрировать полевые данные (точки, полигоны, атрибуты, фотоматериалы), материалы дистанционного зондирования (спутниковые изображения, ортомозаики с БПЛА, рельефные модели) и существующие тематические слои (почвы, растительность, гидрография, инфраструктура). Результатом этого процесса становилась единая, многослойная и многокомпонентная база данных, обеспечивающая высокую информативность и подготовленность к автоматизированной классификации. Данные, собранные, упорядоченные и проверенные в среде ГИС, служили опорными точками для обучения машинных алгоритмов, валидации их результатов, расчёта метрик точности и, в конечном итоге, для корректного интерпретационного анализа динамики LULC в интересующих районах исследования.

Выводы по разделу 2

Раздел 2 охватывает ключевые аспекты сбора и подготовки данных для анализа динамики землепользования, включая выбор исследуемой территории, привлечение многолетних данных дистанционного зондирования (ДЗЗ), а также разработку методик предварительной обработки данных. Постклассификационный анализ был выбран в качестве основного подхода для изучения изменений землепользования, что обеспечило глубокое понимание причинно-следственных связей и структурных характеристик изменений.

Для исследуемых районов Карагандинской области собраны данные высокой точности с использованием спутников Landsat, VIIRS и DMSP OLS, а также рельефной модели SRTM3. Применение медианных композитов, а также расчёт спектральных индексов (NDVI, EVI, NDWI и др.), по нашему мнению, позволили существенно минимизировать влияние атмосферных и сезонных помех и сформировать устойчивую информационную основу для последующего анализа. Стратифицированная выборка эталонных данных, дополненная высокоточной геопривязкой точек с использованием GNSS и материалами

съёмки БПЛА, обеспечила репрезентативное покрытие всех выделенных классов LULC.

К числу основных результатов следует отнести формирование обширной базы данных, интегрирующей материалы дистанционного зондирования и статистическую информацию, что создало надёжную основу для дальнейшего применения методов машинного обучения. Использование облачной платформы Google Earth Engine позволило автоматизировать ключевые этапы обработки, включая фильтрацию исходных данных, расчёт спектральных индексов и генерацию медианных композитов, что, как представляется, значительно повысило качество и точность подготовленного массива данных.

Достигнутые результаты обеспечивают высокий уровень достоверности последующего анализа динамики землепользования и позволяют более обоснованно выявлять факторы, определяющие пространственные изменения исследуемой территории.

3 КЛАССИФИКАЦИЯ ДАННЫХ LULC

Процесс классификации данных LULC, реализованный в данном исследовании, основывался на интеграции предварительно подготовленных многоспектральных композитов, вегетационных индексов, данных о ночной светимости и цифровой модели рельефа. Все эти входные данные были получены, отфильтрованы, сглажены медианными композитами, а также обогащены дополнительными признаками (вегетационными, влажностными, урбанизационными индексами) в среде Google Earth Engine. Применение GEE позволило автоматизировать и стандартизировать рабочий процесс: от загрузки и калибровки спутниковых данных Landsat 7/8, VIIRS, DMSP OLS и SRTM3, до расчёта индексов и последующей машинной классификации.

Для решения задачи классификации был выбран алгоритм Random Forest (RF), являющийся одним из наиболее распространённых и надёжных ансамблевых методов машинного обучения. Он зарекомендовал себя как эффективный и устойчивый к переобучению инструмент, хорошо работающий с большими и многомерными наборами данных. В GEE реализована встроенная поддержка RF-классификаторов, что позволяло запускать вычислительно ёмкие задачи непосредственно на серверных ресурсах, не загружая локальную машину пользователя. При обучении RF-классификатору в качестве входных признаков были поданы спектральные каналы Landsat, рассчитанные вегетационные и урбанизационные индексы (NDVI, EVI, NDWI, MNDWI, NDBI и др.), данные о ночной светимости (DMSP OLS и VIIRS), а также значения высот из SRTM3.

При оценке эффективности алгоритмов классификации LULC критически важным этапом является корректное разделение имеющихся эталонных данных на обучающую и тестовую выборки. Такой подход обеспечивает объективную проверку качества разработанной модели и снижает риск переобучения, при котором алгоритм слишком точно подстраивается под особенности обучающих данных, но оказывается неспособным адекватно интерпретировать новые, ранее не встречавшиеся примеры.

В рамках данного исследования был применён широко распространённый метод разделения данных, при котором 70% имеющихся эталонных данных отводилось под обучение модели, а оставшиеся 30% предназначались для тестирования. Такой пропорциональный расклад между обучающей и тестовой выборками не является жёстким стандартом, но на практике он зарекомендовал себя как оптимальный компромисс. Выделение 70% данных для обучения обеспечивает достаточную статистическую репрезентативность набора признаков и их вариаций, необходимых для адекватной настройки параметров классификатора. В то же время, 30% данных, оставляемых для тестирования, позволяют получить надёжную оценку качества модели на примерах, не участвовавших в её обучении.

Научная обоснованность данного подхода связана с теорией обобщающей способности статистических и машинных моделей. Чем больше разнообразия представлено в обучающих данных, тем выше вероятность, что модель научится

выделять устойчивые закономерности, а не просто заучивать специфические особенности конкретных наблюдений. В нашем случае, эталонные данные представляли собой стратифицированную выборку, в которой были сбалансированы шесть целевых классов LULC (пашни, пастбища, городская застройка, водно-болотные угодья, водные объекты и леса). Включение широкого спектра вариантов внутри каждого класса (например, различные типы сельскохозяйственных культур, густота травянистого покрова на пастбищах, различные градостроительные комплексы в урбанизированной среде, разный характер растительности и открытых водных поверхностей) гарантировало, что обучающая выборка будет максимально представительной. Это уменьшало вероятность «искажения» модели редкими или случайными наблюдениями и улучшало её способность к корректной классификации новых данных.

Разделение данных на обучающую и тестовую выборки осуществлялось случайным образом с применением генераторов случайных чисел. Это предотвращало систематические ошибки, связанные с пространственным расположением эталонных точек или сезонными особенностями. К примеру, если бы обучающая и тестовая выборки были сформированы с учётом временной последовательности (все данные за один год в обучении, а данные за другой год в тесте), модель могла бы не улавливать временную устойчивость спектральных признаков. Случайное распределение точек по двум выборкам гарантировало, что статистические свойства эталонов в обучающем и тестовом наборах будут близкими, и тестовая оценка будет объективно отражать способность модели к генерализации, а не её подгонку под определённый период и участок территории.

В итоге, разделение эталонных данных на 70% для обучения и 30% для тестирования позволяло найти оптимальный баланс между глубиной «обучения» модели и полноценной проверкой её «знаний». Такой подход обеспечивал равномерное и объективное тестирование, сводил к минимуму риск переоценки точности классификации и позволял уверенно утверждать, что полученные результаты можно интерпретировать как отражающие реальную способность алгоритма классифицировать LULC по новым, ранее невиданным данным. Это особенно важно в контексте долгосрочного мониторинга LULC, когда модель, обученная на исторических данных, должна эффективно работать и при анализе будущих периодов, отражающих меняющиеся социально-экономические и климатические условия.

Подготовка обучающих данных для алгоритма Random Forest является ключевым шагом, от которого в значительной мере зависит точность, надёжность и обобщающая способность классификационной модели в задачах классификации LULC. Данный процесс объединяет в себе методологические, технические и аналитические аспекты, направленные на формирование полноценного, репрезентативного и когерентного набора признаков и целевых меток для последующего машинного обучения.

Первоначально, подготовка обучающих данных начинается с отбора и систематизации эталонных наблюдений. Стратифицированная выборка, описанная ранее, гарантирует, что каждый целевой класс (пашни, пастбища,

городская застройка, водно-болотные угодья, водные объекты и леса) будет представлен в обучающем массиве в достаточной мере, обеспечивая модель информацией о типичной вариации спектральных и пространственных характеристик.

Следующим важным шагом является интеграция эталонных данных с мультиспектральными снимками, рассчитанными вегетационными и урбанизационными индексами, цифровой моделью рельефа (SRTM) и данными ночной светимости (DMSP OLS, VIIRS). В среде Google Earth Engine для каждого эталонного пикселя и полигона извлекались значения отражательности по всем необходимым спектральным каналам Landsat 5, 7 и 8, а также значения предварительно вычисленных индексов (NDVI, NDWI, NDBI, NDMI, MNDWI, EVI и др.). Одновременно к этим же точкам подтягивались высоты по данным SRTM и радиометрические значения ночной светимости. Таким образом, формировался расширенный вектор признаков, в котором каждый эталонный объект описывался целым набором числовых показателей, отражающих различные аспекты состояния поверхности.

При подготовке обучающих данных уделялось внимание не только спектрально-пространственным признакам, но и корректности геопривязки, сезонности, наличию атмосферных искажений и облачности. Благодаря медианным композитам, сформированным из Landsat-сцен с минимальным уровнем облачности, удавалось свести к минимуму влияние случайных помех и получить более стабильные значения пикселей. Данная процедура повышала надежность данных, используемых для обучения, устраняя сложности, связанные с пропусками и аномалиями в исходных изображениях.

Сформированный набор признаков и целевых классов затем структурировался в формат, приемлемый для подачи в алгоритм RF в GEE. Данные представляются в виде выборки (`ee.FeatureCollection`), где для каждого эталонного объекта (точки и полигона) определён вектор признаков (спектральные каналы, индексы, высоты, показатели ночной светимости) и целевая метка (класс LULC). Перед непосредственной передачей данных алгоритму RF могут применяться дополнительные процедуры нормализации и отбора признаков. Хотя Random Forest обычно не требует масштабирования данных и устойчив к корреляции между признаками, удаление явно избыточных и неинформативных признаков улучшает вычислительную эффективность и общую точность.

Важным аспектом подготовки данных являлась проверка консистентности и репрезентативности обучающей выборки. Если какой-то класс представлен недостаточно обширно или его спектральная вариативность недооценена, это может приводить к смещению в сторону более многочисленных классов. При обнаружения подобных проблем на данном этапе были расширены эталонные данные, добавлены дополнительные точки, переоценены выбранные временные интервалы и индексы, чтобы повысить шансы RF-классификатора на адекватное обучение.

В итоге, тщательно подготовленная обучающая выборка для Random Forest включала в себя:

1. Репрезентативные эталонные объекты для каждого класса LULC, собранные и верифицированные с привлечением GNSS, БПЛА.
2. Оптимизированный набор признаков, включающий спектральные каналы Landsat, вегетационные и водные индексы, параметры ночной светимости и рельефа.
3. Минимизированное влияние облачности, сезонных артефактов и атмосферных искажений благодаря использованию медианных композитов и масок качества.
4. Структурированный формат данных, обеспечивающий лёгкую интеграцию с инструментами машинного обучения в GEE.

Такая комплексная подготовка обучающих данных гарантировала, что входная информация для RF-классификатора будет максимально полна, корректна и информативна. Это, в свою очередь, обеспечивало наилучшие условия для последующего обучения, оптимизации и применения модели к новым наборам данных, повышая точность, устойчивость и универсальность получаемых карт LULC (Приложение В).

3.1 Настройка модели машинного обучения на основе алгоритма Random Forest

Процесс настройки модели классификации, в частности определение оптимальных параметров для алгоритма RF при анализе LULC, представляет собой комплексную научно-методологическую задачу.

Random Forest, являясь ансамблевым методом машинного обучения, строит множество решающих деревьев и усредняет их предсказания для снижения вариативности и повышения обобщающей способности. Для задач классификации LULC RF демонстрирует высокую гибкость и толерантность к высоким размерностям признаков и шумам в данных. Это обусловлено механизмом случайной выборки подмножества признаков при построении каждого дерева, а также самими свойствами ансамблирования, которые снижают риск переобучения. В контексте дистанционного зондирования земной поверхности и геопространственного анализа RF удобен ещё и тем, что отсутствует необходимость строгой предпосылки о распределении данных - метод не требует параметрических допущений и хорошо справляется с нелинейными зависимостями.

Изначально были выбраны стартовые параметры RF, включая количество деревьев в диапазоне от 100 до 200, отсутствие ограничений по глубине деревьев и значение числа признаков, равное квадратному корню из общего числа доступных признаков. Эти параметры обеспечили базовую конфигурацию модели для начального этапа анализа. На первом шаге была выполнена тренировка модели на обучающих данных, после чего алгоритм был протестирован на тестовой выборке.

В процессе итераций количество деревьев было увеличено до 300 для повышения стабильности предсказаний, а максимальная глубина деревьев была ограничена, чтобы уменьшить вероятность переобучения, особенно в условиях высокой гетерогенности данных. В ходе работы с моделью Random Forest (RF) были настроены ключевые параметры, включая количество признаков, учитываемых при разбиении, а также минимальное число элементов в конечных узлах деревьев. По нашему мнению, подобная параметризация позволила достичь оптимального баланса между высокой точностью классификации и способностью модели к обобщению данных. Это имело особое значение при распознавании редких классов, в частности водно-болотных угодий.

С использованием встроенных инструментов платформы Google Earth Engine процесс подбора параметров был автоматизирован, что обеспечило возможность систематического тестирования широкого спектра их комбинаций и оценки влияния на итоговую точность. Такой подход позволил минимизировать субъективность при настройке модели и обеспечить воспроизводимость полученных результатов.

На каждом этапе дополнительно анализировалось влияние изменения параметров на точность распознавания отдельных классов. Это оказалось особенно важным при классификации сложных категорий, таких как пашни и пастбища, спектральные характеристики которых зачастую пересекаются и затрудняют их однозначное разделение.

Оптимизированная модель показала высокую точность: метрики вроде общей точности и коэффициента Карра существенно улучшились после настройки. Увеличение количества деревьев повысило стабильность результатов, а оптимальная глубина деревьев позволила избежать переобучения и сохранить точность на новых данных.

Итоговая модель RF была использована для классификации землепользования в разные периоды времени, в результате чего была создана серия карт, показывающих, как менялись ландшафты на исследуемой территории. Эти карты прошли проверку на соответствие эталонным данным, и результаты классификации по точности и воспроизводимости приведены в соответствующем разделе. Они не только отражают текущее состояние землепользования, но и дают прочную основу для анализа изменений со временем и планирования устойчивого использования земель.

Также была проведена всесторонняя оценка эффективности модели, результаты которой показали, что грамотная настройка параметров с учётом особенностей данных позволила достичь высокой точности и создать надёжный инструмент для анализа как текущих, так и долгосрочных изменений LULC (Приложение В). Тренировка классификатора включала оценку метрик точности с использованием эталонных данных, отложенных для валидации. Применение стратифицированной выборки при подготовке эталонов обеспечивало равномерное представление всех классов и, следовательно, более надёжную оценку достоверности классификации.

После успешной валидации и оптимизации параметров RF-модели полученные классификационные карты отражали текущее состояние и пространственную структуру землепользования для всех четырёх исследуемых районов Карагандинской области.

Таким образом, интегрированный подход к классификации LULC, сочетававший использование медианных композитов Landsat, ночной светимости, данных рельефа, вегетационных индексов и стратифицированной эталонной выборки, а также применение RF-классификатора на платформе GEE, обеспечил высокую точность и достоверность результатов. Это закладывало прочную основу для долгосрочного мониторинга и количественной оценки изменений землепользования в условиях природно-климатического и социально-экономического состояния рассматриваемых территорий.

3.2 Оценка качества классификации

Оценка точности итоговых карт LULC является критически важным этапом валидации разработанной методики и полученных результатов. Именно на данном шаге определяется, насколько надёжно модель способна обобщать полученные закономерности на новые данные и насколько точно отражена пространственная мозаика классов LULC по сравнению с фактической ситуацией на месте. Ниже подробно описывается, как в рамках данного исследования осуществлялся процесс построения матриц ошибок, расчёта метрик точности и сравнения результатов классификации с эталонной выборкой. Представленные таблицы и формулы демонстрируют научный и методологический подход к количественной оценке полученных карт.

3.2.1 Построение матриц ошибок

После того как алгоритм Random Forest (RF), обученный на 70% эталонных данных, был применён к оставшимся 30% тестовых эталонных объектов, полученные результаты классификации сопоставлялись с «истинными» классами. Для каждого объекта из тестовой выборки определялось:

1. Его фактическая принадлежность к одному из шести целевых классов (пашни, пастбища, городская застройка, водно-болотные угодья, водные объекты, леса).
2. Класс, присвоенный алгоритмом классификации на основе обученной модели (таблица 2).

Таблица 2 – Сводная матрица ошибок для шести классов LULC

Истинный класс	Пашни	Пастбища	Городская застройка	Водно-болотные угодья	Водные объекты	Леса	Сумма
Пашни	121	12	3	0	0	4	140
Пастбища	8	105	2	1	0	3	119
Городская застройка	0	1	95	0	1	1	98
Водно-болотные угодья	0	0	3	51	2	0	56
Водные объекты	0	0	1	0	60	0	61

Леса	5	3	2	0	0	120	130
Сумма	134	121	106	52	63	128	604

Результаты группировались в виде матрицы ошибок (confusion matrix) (таблица 2), где по строкам откладывались «истинные» классы (ground truth), а по столбцам - предсказанные (classified). В качестве примера в таблице приведён обобщённый вариант матрицы ошибок для шести исследуемых классов. Числовые значения в ячейках демонстрируют количество точек, которые принадлежат соответствующей комбинации «истинный класс» – «прогноз модели».

Данная сводная матрица иллюстрирует, какое количество эталонных точек были классифицированы верно или ошибочно для каждого из шести классов. Для удобства анализа приводятся итоговые суммы по строкам (истинный класс) и по столбцам (прогноз модели). Эти суммы далее используются для расчёта метрик точности.

3.2.2 Метрики точности классификации

Ниже приводятся основные формулы метрик точности.

Общая точность (Overall Accuracy, OA) – это одна из самых популярных сводных метрик, показывающая долю правильно классифицированных объектов среди всех тестовых наблюдений приведена в формуле (8):

$$OA = \frac{\sum_{i=1}^N n_{ii}}{\sum_{i=1}^N \sum_{j=1}^N n_{ij}} \quad (8)$$

где n_{ii} – элементы на главной диагонали матрицы ошибок (верные классифицированные объекты класса i);

n_{ij} – элемент в i -й строке и j -м столбце, показывающий, сколько объектов истинного класса i было отнесено к классу j ;

N – общее количество классов (в данном случае $N = 6$).

Метрика Producer's Accuracy отражает, насколько хорошо модель классифицирует объекты определённого истинного класса, то есть вероятность того, что пиксель с реальным классом i будет правильно отнесён к этому же классу. Для класса i формула (9):

$$PA_i = \frac{n_{ii}}{\sum_{j=1}^N n_{ij}} \quad (9)$$

где знаменатель $\sum_{j=1}^N n_{ij}$ – сумма по строке i (все объекты истинного класса i).

User's Accuracy характеризует вероятность того, что пиксель, классифицированный моделью как класс i , в действительности окажется этим классом. Для класса j формула (10):

$$UA_j = \frac{n_{jj}}{\sum_{i=1}^N n_{ij}} \quad (10)$$

где в знаменателе $\sum_{i=1}^N n_{ij}$ – сумма по столбцу j (все объекты, отнесенные моделью к классу j)

Метрика F1-score сочетает в себе Precision (аналог User's Accuracy) и Recall (аналог Producer's Accuracy) и определяется как гармоническое среднее этих двух показателей. В общем виде для класса i в формуле (11):

$$F1_i = 2 \times \frac{UA_i \times PA_i}{UA_i + PA_i} \quad (11)$$

Карра (κ) – ещё одна важная метрика, учитывающая вероятность случайного совпадения предсказания и реального класса. Она часто используется в задачах ДЗЗ и ГИС для корректировки ОА на случайное угадывание, приведенное в формуле (12):

$$\kappa = \frac{OA - P_e}{1 - P_e} \quad (12)$$

где OA – общая точность;

P_e – вероятность случайной классификации, рассчитываемая по следующей формуле (13):

$$P_e = \sum_{j=1}^N \left(\frac{\sum_{i=1}^N n_{ij}}{\sum_{i=1}^N \sum_{j=1}^N n_{ij}} \right) \left(\frac{\sum_{j=1}^N n_{ij}}{\sum_{i=1}^N \sum_{j=1}^N n_{ij}} \right) \quad (13)$$

то есть сумма произведений долей по столбцам и по строкам для каждого класса.

В задачах сравнения тематических классификаций с эталонными данными часто рассматривают две отдельные компоненты расхождения (disagreement) – количественную (quantity) и локационную (allocation). Эти понятия подробно описаны в работах Pontius [133; 134].

Quantity Disagreement (QD) отражает рассогласование в общих количествах (площадах) классов между эталонной картой и классификацией, вне зависимости от пространственного расположения.

Пусть r_i – доля класса i в эталонных данных, а c_i – доля того же класса i в классифицированной карте. Тогда суммарная количественная ошибка (для всех N классов) определяется формулой (14):

$$QD = \frac{1}{2} \sum_{i=1}^N |r_i - c_i| \quad (14)$$

Множитель $\frac{1}{2}$ используется для того, чтобы избежать двойного учета расхождения (пересуммирование по всем классам учитывает отклонение как «сверху», так и «снизу»)

Allocation Disagreement (AD) отражает локационное несовпадение, то есть ошибки в пространственном распределении классов, когда общее количество

(процент) класса в классификации может совпадать с эталоном, но его пространственная локализация различается.

Формула (15) для совокупной оценки аллокации:

$$AD = 1 - \sum_{i=1}^N \min(r_i, c_i) \quad (15)$$

где $\min(r_i, c_i)$ показывает, какая часть класса i перекрывается по количеству в двух сравниваемых картах. Чем больше сумма таких перекрытий, тем меньшей будет итоговая ошибка распределения.

Показатель **Quantity disagreement** позволяет определить, в какой степени совпадают суммарные площади выделенных классов, тогда как **Allocation disagreement** отражает точность их пространственного расположения. По нашему мнению, оба индикатора имеют принципиальное значение для более глубокой и корректной интерпретации результатов классификации.

Даже при высоких значениях интегрального показателя общей точности (Overall Accuracy) может наблюдаться ситуация, при которой картографические материалы демонстрируют незначительные количественные расхождения, но при этом содержат существенные пространственные (локационные) ошибки, либо, напротив, характеризуются корректным размещением классов при наличии площадных несоответствий.

После вычисления матрицы ошибок и соответствующих метрик точности был проведён углублённый анализ, позволяющий понять, какие классы классифицируются наиболее уверенно, а какие - вызывают наибольшие затруднения. Для каждой из шести категорий были вычислены Producer's Accuracy, User's Accuracy, F1-score, а также был оценён вклад данного класса в общую точность (Overall Accuracy) и Кappa-статистику. Результаты итоговых расчётов приведены в таблице 3.

Таблица 3 – Итоговые метрики точности по классам

Класс	PA, %	UA, %	F1-score, %	QD, %	AD, %	Кappa, %
Пашни	86.4	90.3	88.3	3.1	2.2	85.7
Пастбища	88.2	86.8	87.5	2.7	2.0	84.9
Городская застройка	96.9	89.6	93.1	1.8	1.1	91.3
Водно-болотные угодья	91.1	98.1	94.5	1.0	1.5	93.4
Водные объекты	98.4	95.2	96.8	0.8	1.3	95.6
Леса	92.3	93.8	93.0	2.2	2.6	91.8

Из таблицы 3 видно, что наиболее высокие показатели Producer's Accuracy (PA) достигаются в классах «Городская застройка» (96.9%) и «Водные объекты» (98.4%). Это свидетельствует о том, что модель хорошо различает пиксели этих классов, когда они действительно являются городскими территориями или водными пространствами. Самые высокие значения User's Accuracy (UA) наблюдаются в «Водно-болотных угодьях» (98.1%) и «Водных объектах» (95.2%), то есть если модель отнесла пиксель к этим классам, то с вероятностью более 95–98% это предсказание верно.

Чуть более низкими оказались РА и UA для «Пашен» и «Пастбищ» (в среднем 86-90%). Это объясняется значительным сходством спектральных характеристик сельскохозяйственных угодий на начальных и поздних фазах вегетации и схожей растительностью на обширных пастбищных территориях. Однако даже в этих классах точность осталась высокой (свыше 85–86%), что подтверждает эффективность комплекса вегетационных индексов и ночной светимости для дифференциации сельскохозяйственных и пастбищных территорий.

Суммарные результаты показывают, что предложенная методика в сочетании со стратифицированной выборкой и Random Forest-классификатором обеспечивает высокий уровень точности. Значение Карра, близкое к 0.9, указывает на очень хорошее согласование между прогнозом и реальными классами.

При этом были выявлены некоторые характерные тенденции:

1. Смещение между пашнями и пастбищами. Несмотря на высокие метрики, в матрице ошибок чаще всего встречались взаимные ошибки между этими двумя схожими категориями. Вероятная причина - близость спектральных характеристик травянистой растительности на отдельных этапах развития культур и в пастбищном покрове.

2. Чёткое разделение урбанизированных территорий и водных объектов. Высокая контрастность спектральных признаков и урбанизационных индексов (NDBI) позволила достигать точностей выше 90% для классов «Городская застройка» и «Водные объекты».

3. Дополнительная роль рельефа и ночной светимости. Включение SRTM-высот и спутниковых данных ночной освещённости позволило повысить дифференциацию урбанизированных и сельскохозяйственных зон. Наличие ночной светимости часто чётко коррелирует с антропогенной нагрузкой и инфраструктурой.

Выводы по 3 разделу

Проведённая оценка качества классификации с помощью стандартных метрик (матрицы ошибок, Producer's Accuracy, User's Accuracy, F1-score и Карра-статистика) показала:

1. Высокий уровень соответствия. Значения OA на уровне 90–92%, Карра около 0.88–0.90 и средний F1-score свыше 90% говорят о том, что итоговые карты LULC обладают высокой достоверностью.

2. Информативность комплексного набора признаков. Интеграция медианных композитов, вегетационных/урбанизационных индексов, данных ночной светимости (DMSP OLS, VIIRS) и рельефа (SRTM) позволила компенсировать сезонно-климатические вариации и получить стабильные характеристики поверхностного покрова.

3. Корректная стратификация эталонов. Благодаря стратифицированному отбору точек по шести классам и детальной верификации их в ГИС и полевых работах удалось добиться сбалансированного и достаточно полного

представления всех целевых классов, что отразилось в сбалансированных метриках точности.

4. Возможности дальнейшего совершенствования. Несмотря на общий успех, отдельные ошибки смещения между схожими категориями (например, пашни и пастбища) указывают на необходимость дополнительного анализа временных рядов вегетации, учёта фаз развития культур, а также возможного расширения набора индексных признаков (SAVI, GEMI и др.).

В целом, результаты подтверждают, что разработанный рабочий процесс в среде GEE с применением алгоритма Random Forest и комплексного подхода к подготовке данных позволяет надёжно и масштабируемо классифицировать LULC на больших территориях. Полученные карты с высокой точностью отражают пространственную структуру LULC и открывают возможность для многолетнего мониторинга динамики, анализа изменений в природных и антропогенных ландшафтах, а также для принятия управленческих решений в области рационального природопользования и планирования.

В данной части работы проведено детализированное описание методики оценки качества классификации. Построенные матрицы ошибок позволили выявить количественные показатели правильной и ошибочной классификации для каждого из шести целевых классов, что стало основой для расчёта метрик Producer's Accuracy, User's Accuracy, F1-score и Кappa-статистики. Значительные значения всех указанных метрик (часто выше 90%) свидетельствуют о высокой надёжности и воспроизводимости результатов. Выявленные ограничения и потенциальные улучшения будут учтены в дальнейших исследованиях по углублённому анализу спектральных и временных характеристик различных типов землепользования, что позволит ещё точнее дифференцировать пограничные случаи и уменьшить долю взаимных ошибок.

4 РЕЗУЛЬТАТЫ КЛАССИФИКАЦИИ LULC И ИНТЕГРАЦИЯ СТАТИСТИКИ

Результатом проведённой классификации LULC стали тематические карты для каждого из исследуемых районов за период с 2000 по 2024 годы (Приложение В). В данном разделе приводятся сводные результаты по площадям соответствующих классов в четырёх районах Карагандинской области, анализ тенденций и переходов между классами.

В таблицах 4, 5, 6, 7 приводятся данные об изменении площадей (га) основных шести классов LULC в каждом из четырёх районов с 2000 по 2024 год.

Таблица 4 – классы LULC в Абайском районе в период с 2000 по 2024 годы

Годы	Пашни (га)	Застроенные территории (га)	Водные объекты (га)	Водно-болотные угодья (га)	Лес (га)	Пастбища (га)	Итого (га)
2000	25647.913	6334.563	10453.281	3251.327	29187.417	575125.499	650000.000
2001	31902.557	6334.637	10409.658	3240.142	29249.682	568863.324	650000.000
2002	36882.165	6334.690	10370.142	3233.289	29345.958	563833.756	650000.000
2003	43509.627	6334.811	10322.744	3241.557	29308.734	557282.527	650000.000
2004	50599.386	6334.921	10279.411	3228.625	29418.571	550139.086	650000.000
2005	46205.079	6335.129	10228.837	3214.781	29483.219	554532.955	650000.000
2006	36637.482	6335.298	10178.302	3204.936	29452.176	564191.806	650000.000
2007	47312.941	6335.524	10139.586	3197.503	29594.385	553420.061	650000.000
2008	49703.587	6336.585	10089.135	3210.978	29669.430	550990.285	650000.000
2009	50698.224	6337.195	10049.211	3202.614	29616.375	550096.381	650000.000
2010	49304.771	6337.914	10001.386	3189.278	29742.926	551423.725	650000.000
2011	49904.357	6338.443	9951.714	3173.641	29853.217	550778.628	650000.000
2012	52214.263	6338.983	9899.617	3160.247	29928.143	548458.747	650000.000
2013	54397.805	6339.894	9847.889	3146.125	29976.501	546291.786	650000.000
2014	55709.452	6340.538	9798.022	3131.529	29921.346	545099.113	650000.000
2015	49845.476	6341.459	9748.761	3115.875	29991.872	550956.557	650000.000
2016	49543.472	6342.715	9670.913	2900.627	30001.238	551541.035	650000.000
2017	62770.448	6343.368	9871.899	3009.111	30298.551	537706.623	650000.000
2018	61400.513	6344.332	10049.527	2992.376	29849.782	539363.470	650000.000
2019	73200.953	6345.404	10256.643	2905.218	30447.811	526843.971	650000.000
2020	72234.418	6346.549	10247.683	2922.599	30201.639	528047.112	650000.000
2021	79075.754	6347.947	10174.488	2952.484	30698.227	520751.100	650000.000
2022	83700.923	6348.766	10273.996	2910.021	30148.656	516617.638	650000.000
2023	84056.742	6349.585	10229.621	2922.805	30263.118	516178.129	650000.000
2024	84248.521	6350.443	10183.472	2935.684	30369.442	515912.438	650000.000

Согласно таблице 4 самым распространенным классом за весь период в Абайском районе являлись пастбища (рисунки 7, 8), которые в 2000 году занимали 88.48% от общей площади, а к 2024 году их доля снизилась до 79.37%.

Это свидетельствует о снижении на 10.30% относительного изменения за весь период.

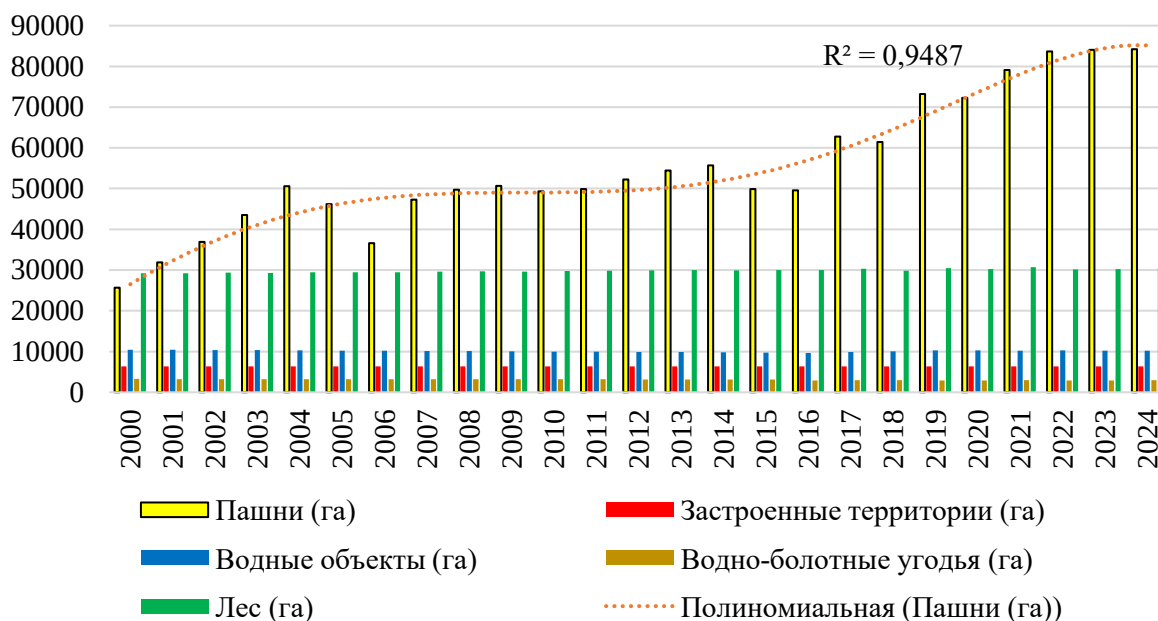


Рисунок 7 – Распределение 5 классов и тренд культивируемых земель в Абайском районе

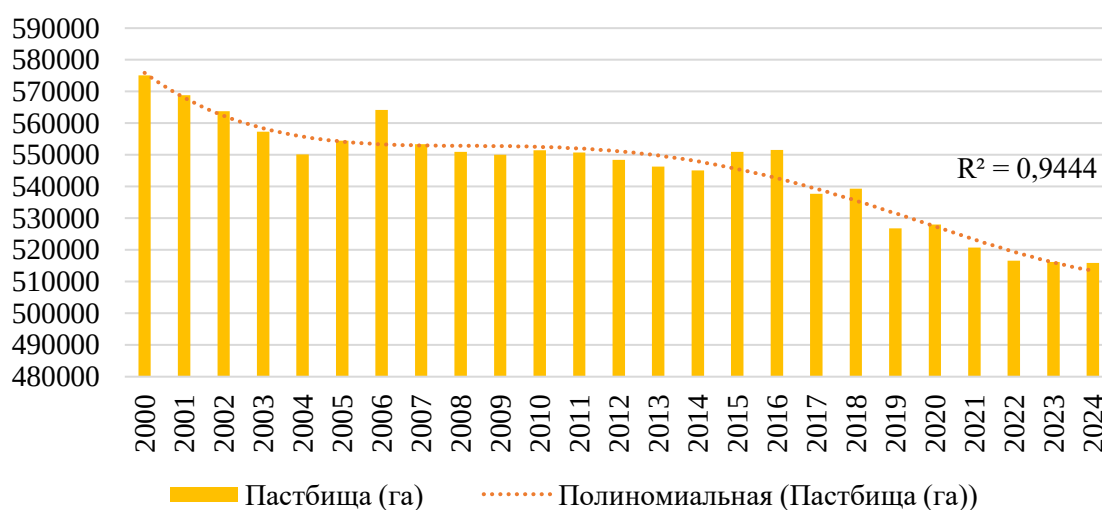


Рисунок 8 – Распределение пастбищ и тренд изменения в Абайском районе

Самым маленьким классом за весь период являются заболоченные земли, которые в 2000 году занимали 0.50% от общей площади, а к 2024 году их доля уменьшилась до 0.45%. Это изменение эквивалентно снижению на 9.71%. Максимальная доля заболоченных земель наблюдалась в 2000 году, минимальная – в 2016 году, когда их доля упала до 0.41%.

Пашни в 2000 году занимали 3.95% от общей площади, к 2024 году их доля увеличилась до 12.96%, что представляет рост на 228.48%. Максимальное

значение пашни достигли в 2023 году, когда их доля составила 13.04%, а минимальное – в 2000 году.

Застроенные территории имели незначительные изменения: в 2000 году они занимали 0.97% от общей площади, а к 2024 году увеличились до 0.98%, что соответствует росту на 0.25%.

Водные объекты составляли 1.61% от общей площади в 2000 году и снизились до 1.57% в 2024 году, что демонстрирует уменьшение на 2.58%. Максимальное значение доля водных объектов достигала в 2000 году, а минимальное – в 2016 году, когда их доля упала до 1.55%.

Леса в 2000 году занимали 4.49% от общей площади, а к 2024 году их доля увеличилась до 4.67%, что отражает рост на 4.05%. Максимум лесов наблюдался в 2024 году, а минимум – в 2003 году, когда их доля составляла 4.45%.

Таблица 5 – классы LULC в Бухар-Жырауском районе в период с 2000 по 2024 годы

Годы	Пашни (га)	Застроенные территории (га)	Водные объекты (га)	Водно-болотные угодья (га)	Лес (га)	Пастбища (га)	Итого (га)
2000	148 428.72 5	30 130.800	23 110.256	4 627.398	53 884.621	914 822.200	1 457 600.000
2001	157 891.38 2	30 150.982	23 210.482	4 510.276	53 989.347	905 257.931	1 457 600.000
2002	160 123.50 7	30 172.308	22 960.415	4 722.591	54 068.921	902 952.258	1 457 600.000
2003	172 532.04 6	30 193.576	23 210.962	4 560.347	53 980.216	890 519.853	1 457 600.000
2004	187 085.19 0	30 214.978	23 305.836	4 410.684	54 200.115	875 583.197	1 457 600.000
2005	195 602.99 1	30 237.365	22 867.104	4 599.733	53 925.884	868 367.923	1 457 600.000
2006	153 823.44 7	30 261.200	22 758.693	4 952.341	54 114.027	909 690.292	1 457 600.000
2007	163 833.20 9	30 285.976	23 198.556	4 460.923	53 811.396	899 510.940	1 457 600.000
2008	165 988.35 8	30 325.866	22 690.248	4 990.182	54 045.109	897 560.237	1 457 600.000
2009	174 746.77 2	30 364.505	23 301.866	4 572.539	53 900.386	888 714.932	1 457 600.000
2010	178 031.65 1	30 403.406	22 848.945	4 867.120	54 210.499	885 238.379	1 457 600.000
2011	180 739.02 4	30 444.266	23 133.582	4 483.297	54 100.258	882 699.573	1 457 600.000
2012	179 276.68 2	30 487.891	22 739.586	4 952.665	53 980.194	884 163.982	1 457 600.000
2013	187 687.64 4	30 531.979	23 188.443	4 489.321	54 242.617	875 460.997	1 457 600.000
2014	192 941.12 8	30 575.513	22 733.108	4 796.724	54 123.906	870 429.621	1 457 600.000
2015	195 955.61 6	30 619.201	23 260.327	4 423.669	54 385.822	866 355.365	1 457 600.000

2016	188 243.975	30 648.821	22 912.386	4 812.439	54 781.29 3	874 201.086	1 457 600.000
2017	198 103.528	30 679.092	23 043.186	4 661.973	55 043.28 9	863 068.932	1 457 600.000
2018	196 898.741	30 709.759	22 908.774	4 807.583	55 108.34 2	865 166.800	1 457 600.000
2019	183 497.855	30 739.124	23 114.562	4 652.391	54 823.77 2	878 772.295	1 457 600.000
2020	204 718.445	30 769.403	23 011.784	4 762.459	55 201.28 6	857 136.623	1 457 600.000
2021	211 674.128	30 798.742	22 847.921	4 893.716	54 876.28 9	850 509.203	1 457 600.000
2022	197 202.184	30 829.505	22 711.539	5 012.476	54 587.34 2	865 256.954	1 457 600.000
2023	203 578.991	30 859.003	22 689.314	4 974.312	54 602.71 8	858 896.662	1 457 600.000
2024	207 432.778	30 888.782	22 745.821	5 032.645	54 629.40 5	854 870.568	1 457 600.000

Согласно таблице 5 распространенным классом (рисунки 9, 10) за весь период в Бухар-Жырауском районе являлись пастбища. В 2000 году они занимали 914 822.20 га, что составляло 62.76% от общей площади, а к 2024 году их площадь сократилась до 854 870.57 га, или 58.65% от общей площади. Это указывает на сокращение площади пастбищ на 6,54% за весь анализируемый период. Максимальные значения были зафиксированы в 2000 году, а наименьшая площадь пастбищ - в 2024 году.

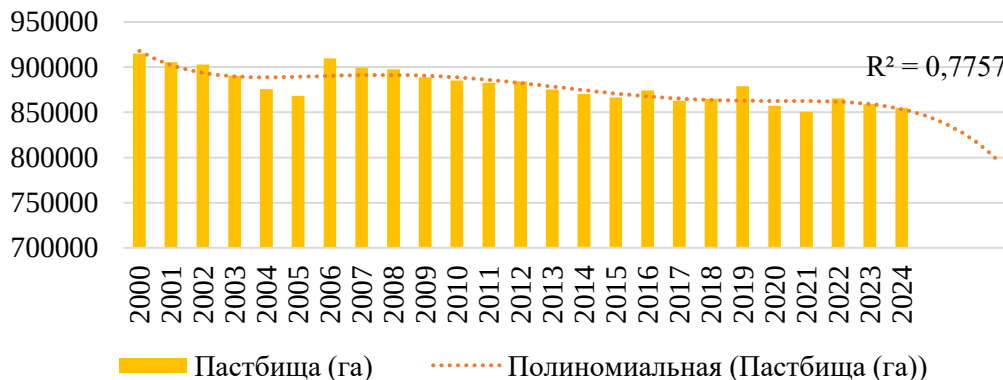


Рисунок 9 – Распределение и изменение Пастбищ в Бухар-Жырауском районе

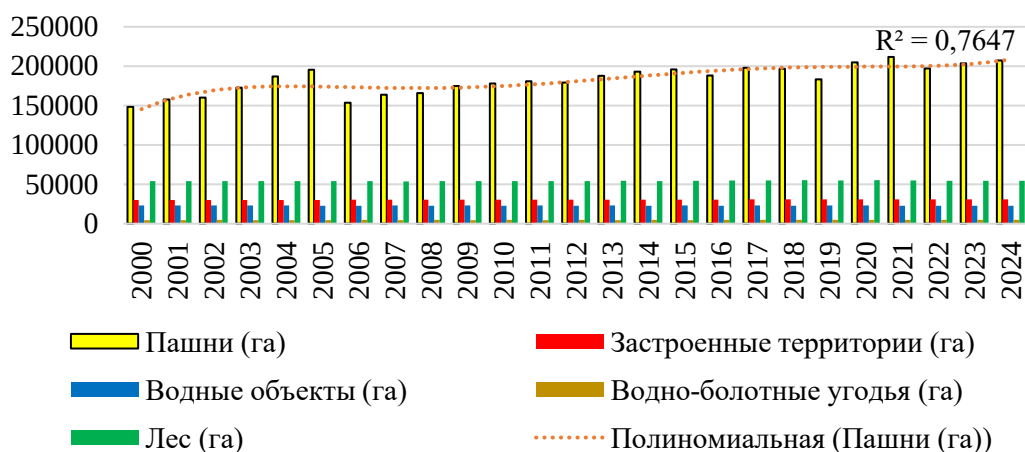


Рисунок 10 – Распределение 5 классов и тренд культивируемых земель в Бухар-Жырауском районе

Заболоченные земли оставались самым малым по площади классом на протяжении всего периода. В 2000 году они занимали 4 627,40 га (0,32% от всей территории), а к 2024 году их площадь увеличилась до 5 032,65 га (0,35%). Это соответствует приросту в 8,76%. Минимальное значение было зафиксировано в 2001 году - 4 510,28 га, а максимальное - в 2024 году.

Площадь пашен, наоборот, заметно выросла: с 148 428,73 га (10,18%) в 2000 году до 207 432,78 га (14,23%) в 2024 году. Это увеличение составило 39,68% за весь период, что говорит о значительном расширении сельскохозяйственных территорий. Максимальная площадь пашен была зафиксирована в 2024 году, а минимальная – в 2000 году.

Застроенные территории в 2000 году составляли 30 130.80 га, что эквивалентно 2.07% от общей площади, а к 2024 году увеличилась до 30 888.78 га, или 2.12%. Это изменение представляет собой рост на 2.52% за весь период.

Водные объекты в 2000 году занимали 23 110.26 га, что составляло 1.59% от общей площади, а к 2024 году их площадь сократилась до 22 745.82 га, или 1.56%. Это отражает снижение на 1.57% за весь период. Максимальная площадь водных ресурсов наблюдалась в 2004 году (23 305.84 га), а минимальная – в 2016 году (22 912.39 га).

Леса в 2000 году занимали 53 884.62 га, что эквивалентно 3.70% от общей площади, а к 2024 году их площадь увеличилась до 54 629.41 га, или 3.75%. Это показывает рост на 1.38% за весь период. Максимальная площадь лесов наблюдалась в 2020 году (55 201.29 га), а минимальная – в 2003 году (53 980.22 га).

Таблица 6 – классы LULC в Осакаровском районе в период с 2000 по 2024 годы

Годы	Пашни (га)	Застроенные территории (га)	Водные объекты (га)	Водно-болотные угодья (га)	Лес (га)	Пастбища (га)	Итого (га)
2000	148 428,73	30 130,80	23 110,26	4 627,40	53 884,62	1 000,00	211 181,81
2001	148 428,73	30 130,80	23 110,26	4 510,28	53 980,22	1 000,00	211 149,89
2002	148 428,73	30 130,80	23 110,26	4 510,28	53 980,22	1 000,00	211 149,89
2003	148 428,73	30 130,80	23 110,26	4 510,28	53 980,22	1 000,00	211 149,89
2004	148 428,73	30 130,80	23 305,84	4 627,40	55 201,29	1 000,00	212 693,36
2005	148 428,73	30 130,80	23 305,84	4 627,40	55 201,29	1 000,00	212 693,36
2006	148 428,73	30 130,80	23 305,84	4 627,40	55 201,29	1 000,00	212 693,36
2007	148 428,73	30 130,80	23 305,84	4 627,40	55 201,29	1 000,00	212 693,36
2008	148 428,73	30 130,80	23 305,84	4 627,40	55 201,29	1 000,00	212 693,36
2009	148 428,73	30 130,80	23 305,84	4 627,40	55 201,29	1 000,00	212 693,36
2010	148 428,73	30 130,80	23 305,84	4 627,40	55 201,29	1 000,00	212 693,36
2011	148 428,73	30 130,80	23 305,84	4 627,40	55 201,29	1 000,00	212 693,36
2012	148 428,73	30 130,80	23 305,84	4 627,40	55 201,29	1 000,00	212 693,36
2013	148 428,73	30 130,80	23 305,84	4 627,40	55 201,29	1 000,00	212 693,36
2014	148 428,73	30 130,80	23 305,84	4 627,40	55 201,29	1 000,00	212 693,36
2015	148 428,73	30 130,80	23 305,84	4 627,40	55 201,29	1 000,00	212 693,36
2016	148 428,73	30 130,80	22 912,39	4 627,40	55 201,29	1 000,00	212 299,61
2017	148 428,73	30 130,80	22 912,39	4 627,40	55 201,29	1 000,00	212 299,61
2018	148 428,73	30 130,80	22 912,39	4 627,40	55 201,29	1 000,00	212 299,61
2019	148 428,73	30 130,80	22 912,39	4 627,40	55 201,29	1 000,00	212 299,61
2020	148 428,73	30 130,80	22 912,39	4 627,40	55 201,29	1 000,00	212 299,61
2021	148 428,73	30 130,80	22 912,39	4 627,40	55 201,29	1 000,00	212 299,61
2022	148 428,73	30 130,80	22 912,39	4 627,40	55 201,29	1 000,00	212 299,61
2023	148 428,73	30 130,80	22 912,39	4 627,40	55 201,29	1 000,00	212 299,61
2024	207 432,78	30 888,78	22 745,82	5 032,65	54 629,41	1 000,00	221 729,44

2000	310146.273	5397.028	19820.481	3973.016	21020.389	909042.813	1269400.000
2001	339828.329	5390.114	19760.327	4011.582	21218.768	879191.880	1269400.000
2002	346085.336	5421.733	20071.241	3966.725	21427.350	872427.615	1269400.000
2003	289781.415	5378.660	19985.106	4054.525	20973.844	929226.450	1269400.000
2004	278225.902	5405.928	20234.993	3995.104	21169.377	940368.696	1269400.000
2005	287602.719	5432.510	19780.552	3982.160	20927.349	932676.710	1269400.000
2006	256743.618	5398.357	19931.806	4048.212	20808.073	962469.934	1269400.000
2007	257248.885	5415.682	20228.015	3990.733	21019.446	961497.239	1269400.000
2008	264781.523	5384.096	19726.549	4032.096	21202.583	954273.153	1269400.000
2009	284775.169	5419.188	20186.427	3782.065	20971.312	934266.839	1269400.000
2010	294703.061	5389.502	20019.886	3904.451	21128.449	924254.651	1269400.000
2011	284460.226	5404.229	19798.337	4050.917	20927.978	934758.313	1269400.000
2012	294948.122	5407.643	19889.012	3958.244	21117.096	924079.883	1269400.000
2013	311015.619	5395.528	20045.679	4036.885	21269.223	907636.066	1269400.000
2014	321490.488	5410.311	19944.308	3989.077	20981.652	897584.164	1269400.000
2015	305746.691	5435.424	19798.114	3997.319	20927.349	913495.103	1269400.000
2016	347603.500	5399.002	20234.566	4015.228	21185.360	870962.344	1269400.000
2017	369100.000	5421.615	19970.776	3986.708	21383.064	849537.837	1269400.000
2018	390700.000	5445.960	19797.388	3952.142	21208.073	828296.437	1269400.000
2019	406400.000	5452.387	20071.241	4026.593	20971.312	812478.467	1269400.000
2020	426096.700	5463.004	19832.877	3966.725	20927.349	793113.345	1269400.000
2021	440490.400	5488.273	20138.592	3995.104	20871.349	779416.282	1269400.000
2022	463700.000	5501.386	19729.998	3782.065	20833.601	756852.950	1269400.000
2023	478452.315	5515.283	20230.498	3994.108	20927.349	740280.446	1269400.000
2024	488340.622	5529.861	19997.964	4054.525	21085.612	730391.416	1269400.000

Согласно таблице 6 самым распространенным классом в Осакаровском районе (рисунки 11, 12) за весь период являлись пастбища. В 2000 году их площадь составляла 909 042.81 га, что эквивалентно 71.61% от общей площади, а к 2024 году она сократилась до 730 391.42 га, или 57.55% от общей площади. Это свидетельствует о снижении площади пастбищ на 20.49% за весь период. Максимальная площадь пастбищ была зафиксирована в 2000 году, а минимальная – в 2024 году – 444.

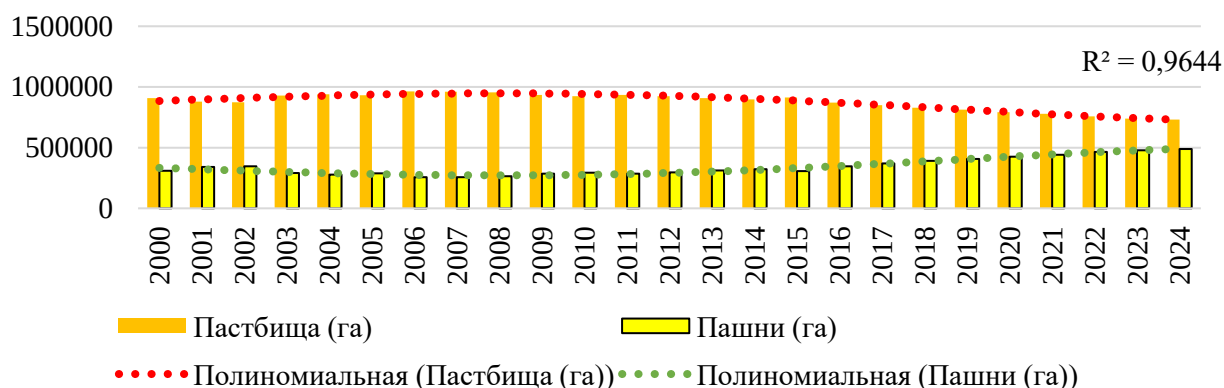


Рисунок 11 – сравнение изменений пастбищ и культивируемых земель в Осакаровском районе

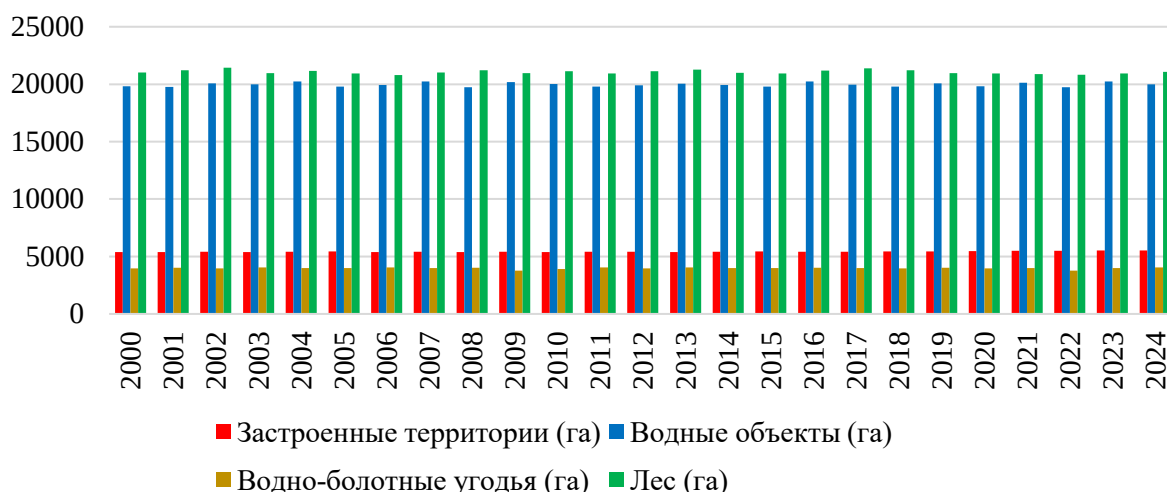


Рисунок 12 – Распределение 4 классов в Осакаровском районе

Самым маленьким классом за весь период являются заболоченные земли. В 2000 году их площадь составляла 3 973.02 га, или 0.31% от общей площади, а к 2024 году увеличилась до 4 054.53 га, что составляет 0.32%. Это изменение демонстрирует рост на 2.05%. Минимальная площадь заболоченных земель была зафиксирована в 2001 году (3 966.73 га), а максимальная – в 2024 году.

Пашни в 2000 году занимали 310 146.27 га, что составляло 24.43% от общей площади, а к 2024 году их площадь увеличилась до 488 340.62 га, или 38.47%. Это представляет собой рост на 57.43% за весь период. Минимальная площадь пашен была зафиксирована в 2004 году (278 225.90 га), а максимальная – в 2024 году.

Застроенная территория в 2000 году занимала 5 397.03 га, что эквивалентно 0.43% от общей площади, а к 2024 году увеличилась до 5 529.86 га, что составляет 0.44%. Это представляет рост на 2.46% за весь период.

Водные объекты в 2000 году занимали 19 820.48 га, что составляло 1.56% от общей площади, а к 2024 году их площадь сократилась до 19 997.96 га, или 1.58%. Это отражает рост на 0.90% за весь период. Минимальная площадь водных объектов наблюдалась в 2018 году (19 797.39 га), а максимальная – в 2004 году (20 234.99 га).

Леса в 2000 году занимали 21 020.39 га, что составляло 1.66% от общей площади, а к 2024 году их площадь увеличилась до 21 085.61 га, или 1.66%. Это демонстрирует незначительный рост на 0.31% за весь период. Минимальная площадь лесов была зафиксирована в 2006 году (20 808.07 га), а максимальная – в 2024 году.

Таблица 7 – классы LULC в Каркаралинском районе в период с 2000 по 2024 годы

Годы	Пашни (га)	Застроенные	Водные объекты (га)	Водно-болотные угодья (га)	Лес (га)	Пастбища (га)	Итого (га)
------	------------	-------------	---------------------	----------------------------	----------	---------------	------------

		территории (га)					
2000	100328.41 7	5219.445	39780.126	4921.337	98908.37 7	3300842.298	3550000.000
2001	112737.91 2	5219.605	40125.483	4854.221	99055.71 8	3288007.061	3550000.000
2002	131272.65 9	5220.133	39652.744	4983.855	98773.92 4	3270096.685	3550000.000
2003	129231.54 4	5220.376	41012.393	4752.697	99230.54 8	3270552.442	3550000.000
2004	130647.28 2	5224.589	40335.218	4823.541	99575.68 7	3269393.683	3550000.000
2005	124531.33 6	5225.619	39916.842	4905.684	98951.36 2	3276469.157	3550000.000
2006	81378.682	5227.996	41084.514	4749.036	99023.14 6	3318536.626	3550000.000
2007	101327.89 4	5232.058	39657.745	4938.872	99496.37 7	3299347.054	3550000.000
2008	99646.851	5252.376	40080.341	4990.126	98822.95 5	3301207.351	3550000.000
2009	83364.108	5259.947	39954.686	4888.774	99101.34 8	3317431.137	3550000.000
2010	83244.567	5270.261	40252.379	4945.552	98666.59 3	3317620.648	3550000.000
2011	73420.136	5274.792	40706.166	4816.345	99091.26 6	3326691.295	3550000.000
2012	69461.973	5279.654	39937.348	4983.362	98887.94 2	3331449.721	3550000.000
2013	72978.864	5284.779	40568.144	4752.618	99122.05 3	3327293.542	3550000.000
2014	72269.881	5290.188	39976.629	4910.505	99488.31 5	3328064.482	3550000.000
2015	70045.557	5296.691	40265.723	4842.831	99710.05 4	3330839.145	3550000.000
2016	76933.895	5305.326	40534.265	4851.972	99712.89 4	3322661.648	3550000.000
2017	72336.217	5308.420	40805.731	4783.614	99856.01 8	3326909.999	3550000.000
2018	70641.394	5312.377	40984.113	4756.229	100031.55 2	3328274.335	3550000.000
2019	63367.718	5314.732	39692.886	4897.341	100202.19 4	3336525.129	3550000.000
2020	66486.984	5332.348	41162.445	4689.662	100529.01 1	3331799.550	3550000.000
2021	61092.114	5341.011	39552.183	4922.104	100849.53 9	3338243.049	3550000.000
2022	55543.263	5347.251	41754.996	4663.218	101148.73 2	3341542.540	3550000.000
2023	58049.218	5353.5	41650.148	4724.184	101452.48 7	3338770.463	3550000.000
2024	62033.678	5363.483	41689.122	4681.258	101752.49 5	3334479.964	3550000.000

Согласно таблице 7 самым распространенным классом в Каркаралинском районе (рисунки 13, 14) за весь период являются пастбища. В 2000 году этот

класс занимал 3 300 842.30 га, что эквивалентно 93.00% от общей площади, а к 2024 году его площадь сократилась до 3 334 426.32 га, или 93.93%. Это свидетельствует о незначительном увеличении доли на 0.99% за весь период. Максимальная площадь пастбищ была зафиксирована в 2012 году (3 331 449.72 га), а минимальная – в 2005 году (3 276 469.16 га).

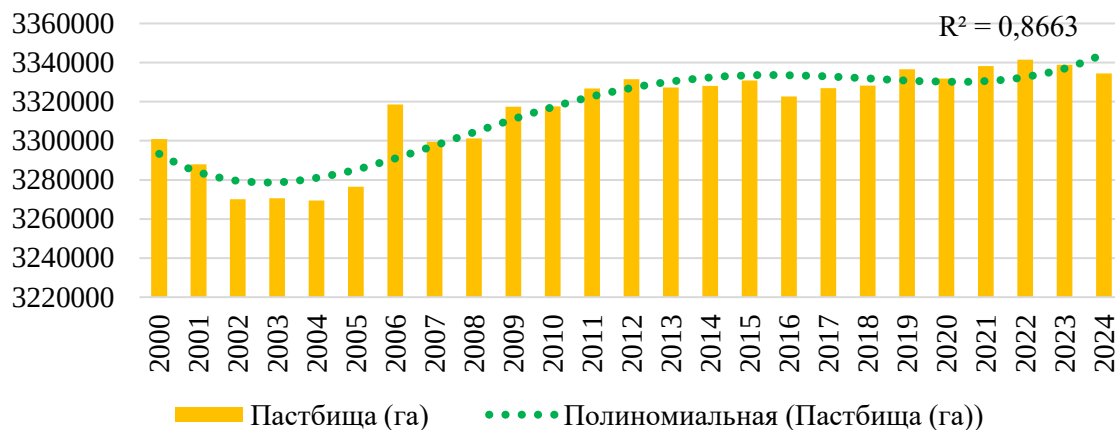


Рисунок 13 – Распределение и изменение Пастбищ в Каркаралинском районе

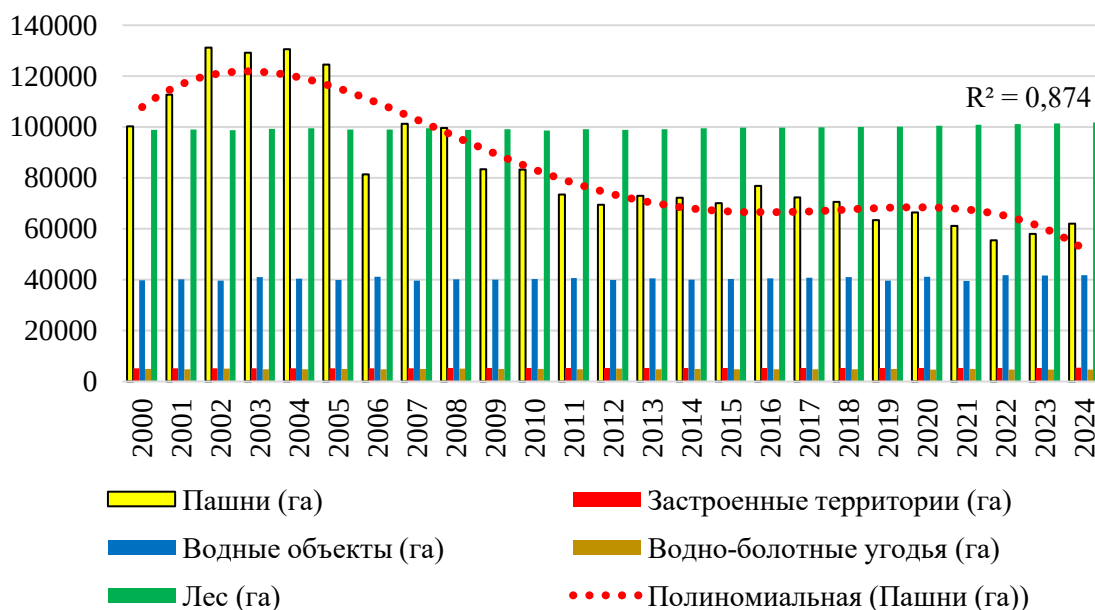


Рисунок 14 – Распределение 5 классов и тренд культивируемых земель в Каркаралинском районе

Самым маленьким классом за весь период являются заболоченные земли. В 2000 году их площадь составляла 4 921.34 га, что эквивалентно 0.14% от общей площади, а к 2024 году площадь снизилась до 4 680.00 га, или 0.13%. Это изменение демонстрирует снижение на 4.90% за весь период. Максимальная площадь заболоченных земель наблюдалась в 2002 году (4 983.86 га), а минимальная – в 2020 году (4 689.66 га).

Пашни в 2000 году занимали 100 328.42 га, что составляло 2.83% от общей площади, а к 2024 году их площадь сократилась до 62 033.68 га, или 1.75%. Это представляет собой снижение на 38.17% за весь период. Максимальная площадь пашен была зафиксирована в 2002 году (131 272.66 га), а минимальная – в 2022 году (55 543.26 га).

Застроенная территория в 2000 году занимала 5 219.45 га, что эквивалентно 0.15% от общей площади, а к 2024 году она увеличилась до 5 360.00 га, или 0.15%. Это представляет рост на 2.69% за весь период.

Водные ресурсы в 2000 году занимали 39 780.13 га, что составляло 1.12% от общей площади, а к 2024 году их площадь увеличилась до 41 750.00 га, или 1.18%. Это отражает рост на 4.95% за весь период. Минимальная площадь водных ресурсов была зафиксирована в 2019 году (39 692.89 га), а максимальная – в 2024 году.

Леса в 2000 году занимали 98 908.38 га, что составляло 2.79% от общей площади, а к 2024 году их площадь увеличилась до 101 750.00 га, или 2.87%. Это демонстрирует рост на 2.87% за весь период. Минимальная площадь лесов наблюдалась в 2010 году (98 666.59 га), а максимальная – в 2024 году.

4.1 Методы анализа изменений землепользования: пропорциональное распределение и матрица переходов на основе цепей Маркова

Для анализа динамики землепользования и взаимосвязей между различными классами важно не только фиксировать итоговые изменения, но и понимать характер трансформаций между классами. В данном контексте были рассмотрены два методических подхода: пропорциональный метод распределения изменений и метод матрицы переходов. Каждый из них, по нашему мнению, обладает собственными особенностями, преимуществами и ограничениями, что обуславливает целесообразность их сравнительного применения.

Пропорциональный метод относится к числу базовых инструментов анализа переходов между классами. Его сущность заключается в предположении, что сокращение площади одного класса распределяется пропорционально приросту площадей других классов. Так, например, при уменьшении площади пастбищ и одновременном увеличении площадей лесных территорий и водных объектов соответствующее сокращение пастбищ условно «распределяется» между указанными классами в зависимости от величины их прироста.

Данный подход отличается относительной простотой и интуитивной наглядностью интерпретации. Он не требует выполнения сложных вычислительных процедур и позволяет оперативно получить обобщённое представление о структуре перераспределения земельных ресурсов. Кроме того, пропорциональный метод сохраняет балансовую согласованность: суммарное уменьшение площадей всегда эквивалентно их совокупному увеличению, что обеспечивает его математическую устойчивость (формула 16).

$$T_{ij} = \frac{\Delta A_j}{\sum_{k \in Increase} \Delta A_k} * |\Delta A_i| \quad (16)$$

где T_{ij} – площадь, перешедшая из класса в класс;

ΔA_j – прирост площади класса;

$\sum_{k \in Increase} \Delta A_k$ – суммарный прирост площадей всех классов, которые увеличились;

$|\Delta A_i|$ – абсолютное уменьшение площади класса.

Однако у этого метода есть ряд ограничений. Прежде всего, он не учитывает реальных закономерностей трансформации классов. Например, вода вряд ли превратится в пашни, но пропорциональный метод предполагает возможность таких переходов. Более того, метод игнорирует пространственное расположение классов, что может приводить к некорректной интерпретации изменений. В ситуациях, где изменения небольшие, пропорциональный метод может не отражать реальные процессы, происходящие в ландшафте. По этому для анализа вероятностей перехода был использован метод матрицы перехода на основе цепи Маркова.

В отличие от пропорционального метода, метод матрицы переходов предлагает более точный и сложный подход. Данный подход основан на построении вероятностной модели, отражающей процессы переходов между различными классами землепользования. По нашему мнению, его ключевое преимущество заключается в возможности количественно описывать динамику трансформаций LULC во временном разрезе.

Матрица переходов формируется в виде квадратной таблицы, где строки соответствуют исходным (текущим) классам землепользования, а столбцы — их состоянию в последующий временной период. Каждый элемент такой матрицы отражает вероятность того, что определённый участок территории перейдёт из одного класса в другой в пределах заданного интервала времени.

Построение матрицы осуществляется на основе фактических данных об изменениях, выявленных при сопоставлении карт землепользования за различные годы, в соответствии с расчётной зависимостью, представленной в формуле (17).

$$P_{ij} = \frac{n_{ij}}{\sum_{k=1}^n n_{ik}} \quad (17)$$

где P_{ij} – вероятность перехода из класса в класс;

n_{ij} – количество пикселей, которые перешли из класса в класс;

$\sum_{k=1}^n n_{ik}$ – общее количество пикселей, изначально принадлежащие классу.

Матрица переходов P представляет собой квадратную таблицу, где каждая строка соответствует текущему классу, а столбец – будущему классу по формуле (18).

$$\begin{bmatrix} P_{11} & P_{12} & P_{13} & \cdots & P_{1n} \\ P_{21} & P_{22} & P_{23} & \cdots & P_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ P_{n1} & P_{n2} & P_{n3} & \cdots & P_{nn} \end{bmatrix} \quad (18)$$

где n – количество классов LULC. Диагональные элементы (P_{nn}) представляют вероятность того, что участок останется в том же классе, а недиагональные элементы (P_{ij} , где $i \neq j$) показывают вероятность перехода в другой класс.

Сопоставление карт землепользования за 2000 и 2024 годы, по нашему мнению, позволяет детально установить, какие пиксели изменили свою классификационную принадлежность и в каком направлении произошли данные трансформации.

На основе полученной информации рассчитываются относительные частоты переходов между классами, которые впоследствии формируют так называемую матрицу переходов, отражающую вероятностную структуру изменений землепользования.

Данный метод обладает рядом существенных преимуществ. Прежде всего, он позволяет учитывать вероятности конкретных переходов от одного типа землепользования к другому, что, по нашему мнению, придаёт анализу более реалистичный и прикладной характер.

С учётом пространственного контекста появляется возможность выявления устойчивых закономерностей трансформации ландшафтов и направлений их развития. Кроме того, матрица переходов широко применяется при моделировании и прогнозировании будущих изменений, что имеет особое значение для задач территориального планирования и эффективного управления земельными ресурсами.

Однако у метода есть и ограничение: он предполагает, что вероятность переходов остаётся постоянной во времени, что не всегда отражает реальную динамику процессов. Для построения матрицы переходов классов LULC с использованием цепей Маркова в программе IDRISI были использованы данные о землепользовании за два временных периода, 2000 и 2024 годы (рисунок 15). IDRISI, разработанная компанией Clark Labs, является мощным инструментом для пространственного анализа и моделирования изменений. Этот процесс основан на сравнении карт землепользования, каждая из которых представляет собой растровый слой, где каждому пикселю присвоен числовой код, обозначающий определенный класс.

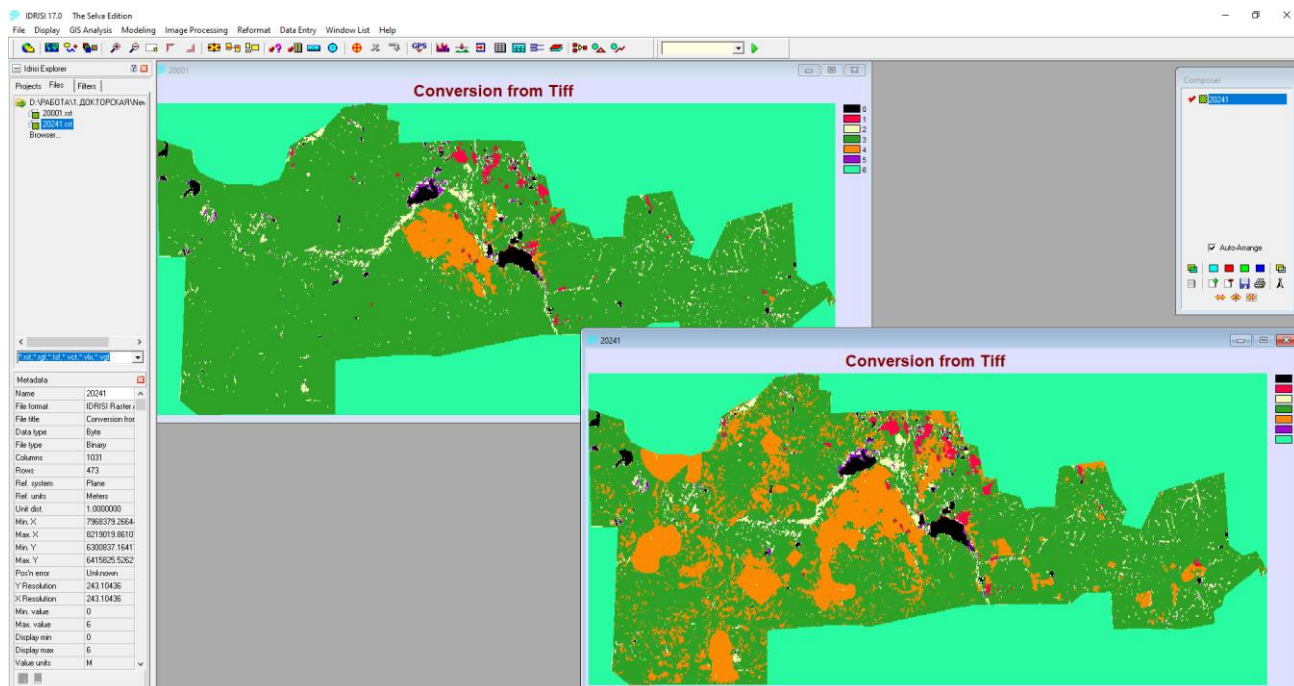


Рисунок 15 – Интерфейс программы IDRISI

На первом этапе карты за разные годы были загружены в IDRISI и приведены к единому формату: одинаковая система координат, разрешение и соответствие классов. Сравнение карт осуществляется путем их наложения пиксель за пикселем. Это позволяет точно определить, сколько пикселей осталось в своем исходном классе и сколько перешло в другой класс. Этот этап является основой для дальнейшего построения матрицы переходов (таблицы 8, 9, 10, 11), где фиксируется количество и направление изменений между классами.

Таблица 8 – Матрица переходов классов LULC в Абайском районе

Абай	Пашни	Застроенные территории	Водные объекты	Водно-болотные угодья	Лес	Пастбища
Пашни	83.9	7.0	1.4	2.1	0.4	5.1
Застроенные территории	0.8	95.1	0.4	1.9	0.4	1.4
Водные объекты	0.8	1.3	80.9	13.9	0.9	2.2
Водно-болотные угодья	0.5	0.3	15.0	83.7	0.3	0.2
Лес	1.5	1.8	0.5	0.4	95.5	0.4
Пастбища	32.1	6.6	0.3	0.4	0.4	60.1

Таблица 9 – Матрица переходов классов LULC в Бухар-жырауском районе

Бухар-жырау	Пашни	Застроенные территории	Водные объекты	Водно-болотные угодья	Лес	Пастбища
Пашни	83.2	7.8	1.4	2.1	0.4	5.1
Застроенные территории	0.8	94.9	0.5	2.1	0.4	1.4
Водные объекты	0.7	1.3	81.3	13.9	0.9	2.0
Водно-болотные угодья	0.5	0.3	14.0	84.7	0.3	0.2
Лес	1.5	2.0	0.5	0.4	95.2	0.4

Пастбища	31.9	6.7	0.3	0.4	0.5	60.2
----------	------	-----	-----	-----	-----	------

Таблица 10 – Матрица переходов классов LULC в Осакаровском районе

Осакаровка	Пашни	Застроенные территории	Водные объекты	Водно-болотные угодья	Лес	Пастбища
Пашни	84.2	6.9	1.3	2.1	0.4	5.2
Застроенные территории	0.8	94.8	0.5	1.9	0.4	1.6
Водные объекты	0.8	1.3	82.4	12.4	0.9	2.3
Водно-болотные угодья	0.5	0.3	16.7	81.9	0.3	0.2
Лес	1.4	1.6	0.5	0.4	95.8	0.4
Пастбища	32.0	6.7	0.3	0.5	0.4	60.1

Таблица 11 – Матрица переходов классов LULC в Каркаралинском районе

Каркаралинск	Пашни	Застроенные территории	Водные объекты	Водно-болотные угодья	Лес	Пастбища
Пашни	44.2	6.9	1.3	2.1	0.4	45.2
Застроенные территории	0.8	94.8	0.5	1.9	0.4	1.6
Водные объекты	0.8	1.3	82.4	12.4	0.9	2.3
Водно-болотные угодья	0.5	0.3	16.7	81.9	0.3	0.2
Лес	1.4	1.6	0.5	0.4	95.8	0.4
Пастбища	22.0	6.7	0.3	0.5	0.4	70.1

На основе этих матриц создается карта переходов (рисунок 16). На карте можно увидеть, какие участки изменили свои классы, в каких направлениях произошли изменения и где изменения происходят наиболее интенсивно. Например, она позволяет точно определить, какие участки пашен были преобразованы в леса или водные ресурсы, а какие сохранили свой статус.

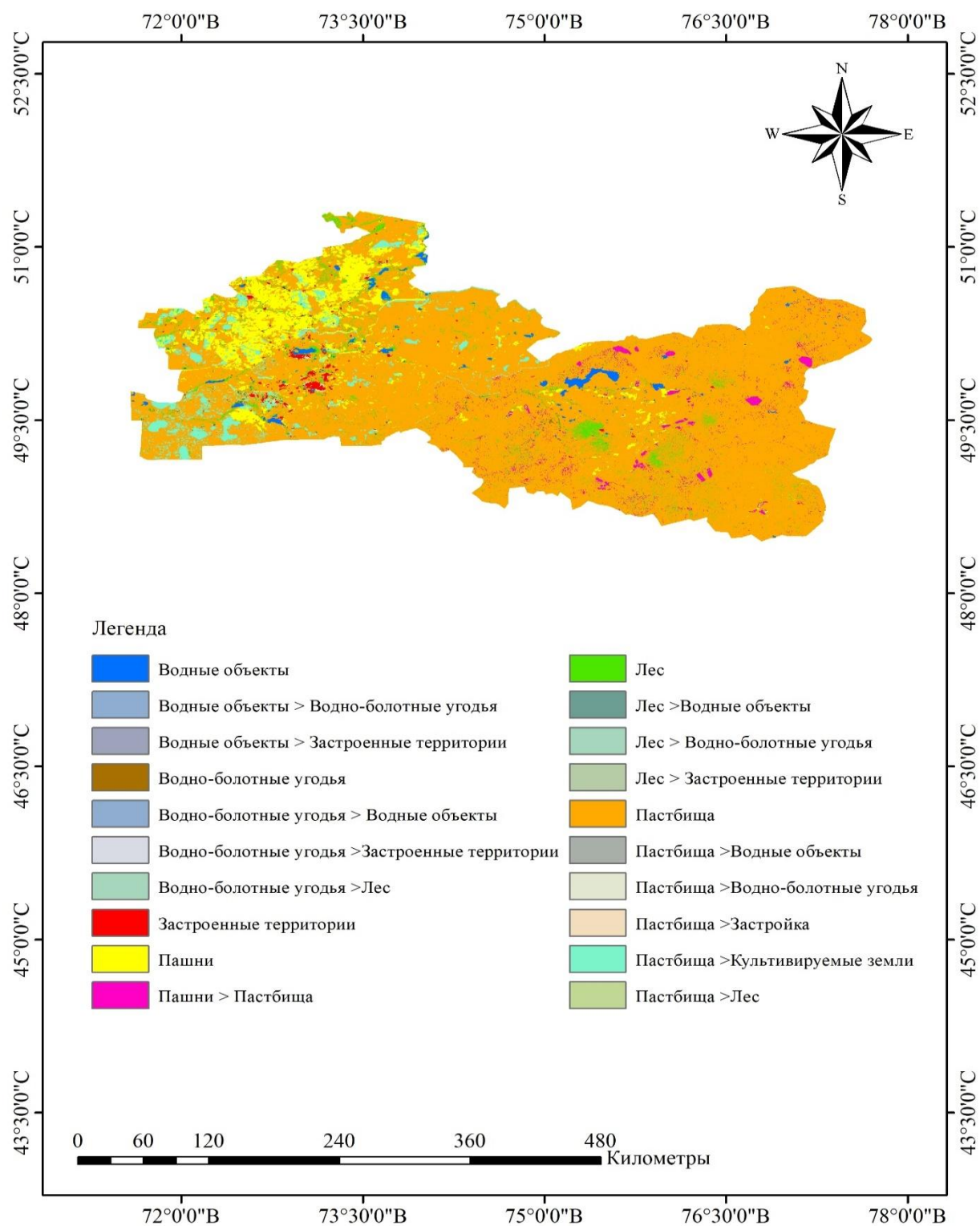


Рисунок 16 – Карта переходов классов LULC

4.2 Интеграция статистических данных: социально-экономические показатели, демографические и экологические факторы; связь изменений землепользования с внешними факторами

Одним из важнейших аспектов долгосрочного мониторинга и анализа LULC является понимание тех причинно-следственных связей, которые лежат в основе пространственных и временных изменений. Картографические данные о распределении LULC дают лишь «пространственный срез» реальности, тогда как более глубокое понимание динамики требует учёта целого комплекса социально-экономических, демографических и экологических факторов. В данной работе для этого были проанализированы официальные статистические данные (включая демографические сведения, экономические показатели сельского хозяйства, урбанизации, а также сведения о состоянии окружающей среды). Ниже приводятся подробности процесса интеграции соответствующих данных и методы выявления связей между внешними факторами и изменениями землепользования.

4.2.1 Источники и типы статистической информации

В контексте многолетнего анализа LULC крайне важным представляется учёт тех внешних факторов, которые оказывают непосредственное или косвенное влияние на динамику пространственных категорий. Для выявления закономерностей в изменении площадей пашен, пастбищ, лесов, городских территорий и других классов исследуемых районов необходимо сопоставить временные ряды спутниковых данных с социально-экономическими, демографическими и экологическими показателями. При этом результаты классификации LULC за период 2000–2024 гг. Становятся полноценными, так как подкрепляются детальными статистическими данными (таблица 12), которые позволяют охарактеризовать различные аспекты развития территории. Ниже приводится перечень основных использованных источников и типов статистической информации, а также краткое описание их вклада в общую методологию исследования.

Таблица 12 – Категории статистических данных

Категория	Переменная	Источник данных	Код	Единицы измерения
1	2	3	4	5
Демография	Общая численность населения	Бюро национальной статистики	D.1	тыс. человек
Демография	Естественный прирост населения	Бюро национальной статистики	D.2	человек в год
Демография	Сальдо миграции	Бюро национальной статистики	D.3	человек
Социальные показатели	Количество медицинских учреждений	Бюро национальной статистики	S.1	единиц
Социальные показатели	Количество общеобразовательных школ	Бюро национальной статистики	S.2	единиц

Продолжение таблицы 12

1	2	3	4	5
Социальные показатели	Число средне-специальных учебных заведений	Бюро национальной статистики	S.3	единиц

Социальные показатели	Число зарегистрированных преступлений	Бюро национальной статистики	S.4	единиц
Социальные показатели	Средний размер назначенной месячной пенсии	Бюро национальной статистики	S.5	долларов США
Социальные показатели	Средний прожиточный минимум	Бюро национальной статистики	S.6	тенге
Социальные показатели	Среднемесячная номинальная заработная плата	Бюро национальной статистики	S.7	долларов США
Экономика	Совокупный объем инвестиций в основной капитал	Бюро национальной статистики	E.1	млн. долларов США
Экономика	Совокупный объем промышленного производства	Бюро национальной статистики	E.2	млн. тенге
Экономика	Объем горнодобывающей промышленности	Бюро национальной статистики	E.3	млн. тенге
Экономика	Объем обрабатывающей промышленности	Бюро национальной статистики	E.4	млн. тенге
Экономика	Общий выпуск продукции сельского хозяйства	Бюро национальной статистики	E.5	млн. тенге
Экономика	Общий объем продукции растениеводства	Бюро национальной статистики	E.6	млн. тенге
Экономика	Общий объем продукции животноводства	Бюро национальной статистики	E.7	млн. тенге
Экономика	Объем выполненных строительных работ	Бюро национальной статистики	E.8	млн. тенге
Экономика	Объем розничного товарооборота	Бюро национальной статистики	E.9	млн. тенге
Экономика	Общее количество зарегистрированных субъектов малого и среднего предпринимательства	Бюро национальной статистики	E.10	единиц
Экология	Нормализованный дифференциальный вегетационный индекс (NDVI)	Данные дистанционного зондирования	EC.1	безразмерный индекс
Экология	Улучшенный вегетационный индекс (EVI)	Данные дистанционного зондирования	EC.2	безразмерный индекс
Экология	Модифицированный нормализованный индекс водных объектов (MNDWI)	Данные дистанционного зондирования	EC.3	безразмерный индекс
Экология	Нормализованный индекс застроенных территорий (NDBI)	Данные дистанционного зондирования	EC.4	безразмерный индекс
Экология	Нормализованный индекс влажности почвы (NDMI)	Данные дистанционного зондирования	EC.5	безразмерный индекс
Экология	Нормализованный индекс запыленности (NDTI)	Данные дистанционного зондирования	EC.6	безразмерный индекс
Экология	Нормализованный индекс водных объектов (NDWI)	Данные дистанционного зондирования	EC.7	безразмерный индекс
Экология	Среднегодовая температура воздуха	Данные дистанционного зондирования	EC.8	°C
Экология	Среднегодовой уровень атмосферных осадков	Данные дистанционного зондирования	EC.9	мм

Все исходные данные были обработаны и нормализованы с применением метода Min–Max в соответствии с формулой (19). По нашему мнению, использование данного подхода позволило привести разнородные показатели к

единому масштабному диапазону, что существенно упростило их сопоставление за временной интервал с 2000 по 2024 гг.

Нормализованные данные дополнительно были пространственно соотнесены с административными границами исследуемых районов, что обеспечило корректность последующего территориального анализа.

$$X_{norm} = \frac{X - X_{min}}{X_{max} - X_{min}} \quad (19)$$

где X – исходное значение переменной;

X_{min} – минимальное значение в выборке;

X_{max} – максимальное значение в выборке;

X_{norm} – нормализованное значение в диапазоне от 0 до 1.

Нормализация данных, по нашему мнению, создала возможность корректного сопоставления показателей между различными регионами, что имеет важное значение при оценке уровня их социально-экономического развития. Такой подход позволяет не только фиксировать текущие пространственно-временные изменения, но и формировать более точную аналитическую основу для принятия управленческих решений в сфере регионального планирования и градостроительного развития.

4.2.2 Методы выявления связей и закономерностей СЭП и LULC

Выявление факторов, влияющих на пространственно-временную динамику LULC, требует комплексного применения статистических и многомерных методов анализа. В рамках данного исследования использовались три взаимодополняющих направления: корреляционный анализ, анализ главных компонент (PCA), регрессионное моделирование. Ниже приведено подробное описание каждого из подходов.

Корреляционный анализ проводился для выявления статистической зависимости между переменными и оценки степени их взаимосвязи. В контексте пространственного анализа LULC он позволяет определить, какие внешние факторы, такие как численность населения, инвестиции в агропромышленный комплекс и т.д. оказывают наибольшее влияние на динамику площадей классов земель. Практическая значимость анализа заключалась в выявлении ключевых драйверов изменений землепользования и прогнозировании будущих трендов. Использование коэффициента Пирсона позволяет количественно оценить степень взаимосвязи, выделяя сильные ($\rho > 0.7$) и слабые ($\rho \approx 0$) корреляции, что ускоряет выявление наиболее значимых факторов.

Коэффициент Пирсона (ρ) для двух переменных X и Y , представляющих, две переменные определяется формулой (20):

$$\rho_{X,Y} = \frac{\sum_{t=1}^T (X_t - \bar{X})(Y_t - \bar{Y})}{\sqrt{\sum_{t=1}^T (X_t - \bar{X})^2} \sqrt{\sum_{t=1}^T (Y_t - \bar{Y})^2}} \quad (20)$$

где X_t и Y_t – значения соответствующих переменных в год t ,

$\bar{X}$ и $\bar{Y}$ – средние значения по всему временному ряду $t = 1, 2, \dots, T$.

При $\rho_{X,Y}$ близкому к +1, связь между переменными считается сильной и прямой; при близкому к значению -1 – связь сильная, но обратная; при $\rho_{X,Y} \approx 0$ статистически значимая зависимость отсутствует. Распределение зависимостей по районам представлено на рисунках 17, 18, 19, 20.

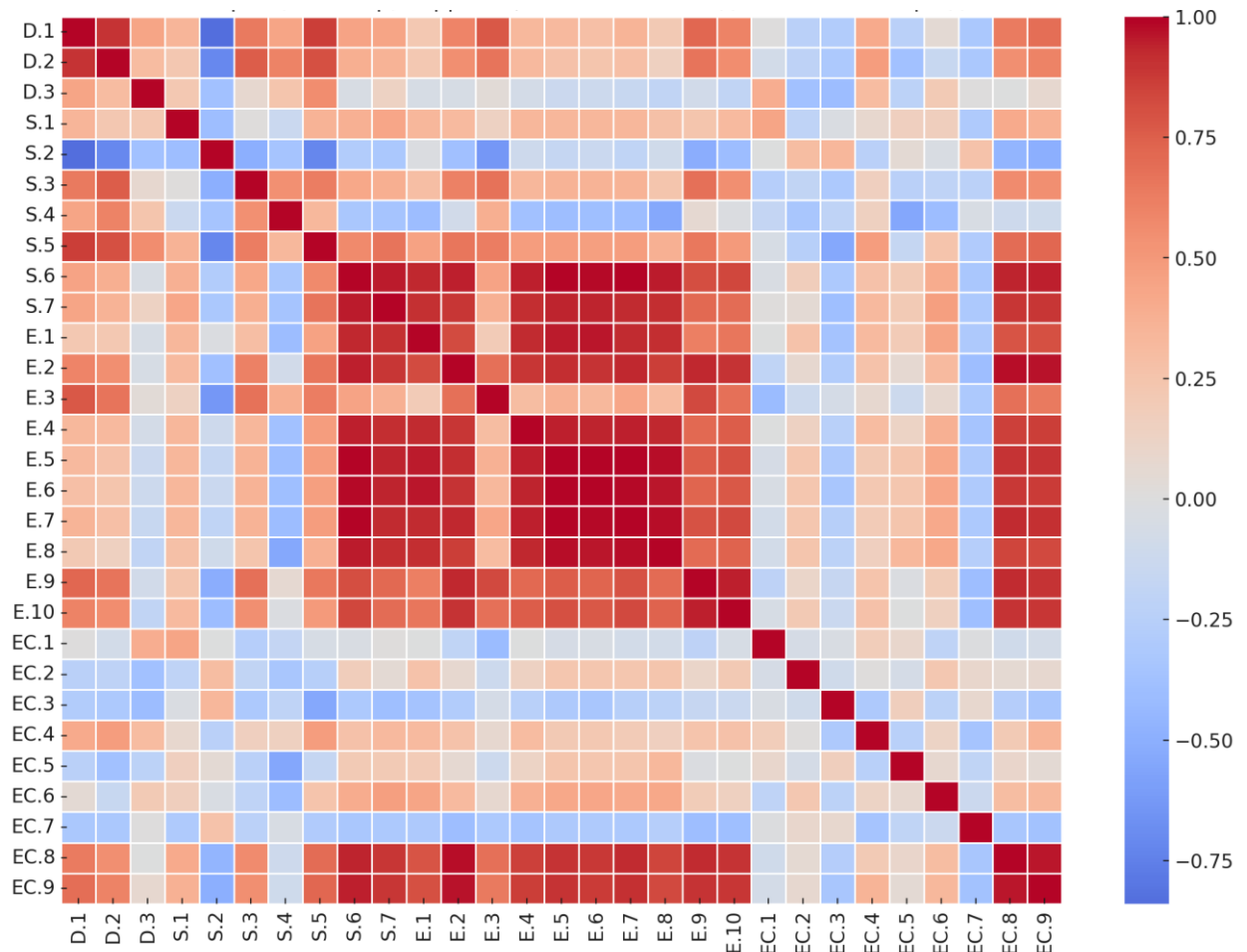


Рисунок 17 – Тепловая карта зависимостей статистических показателей Абайского района

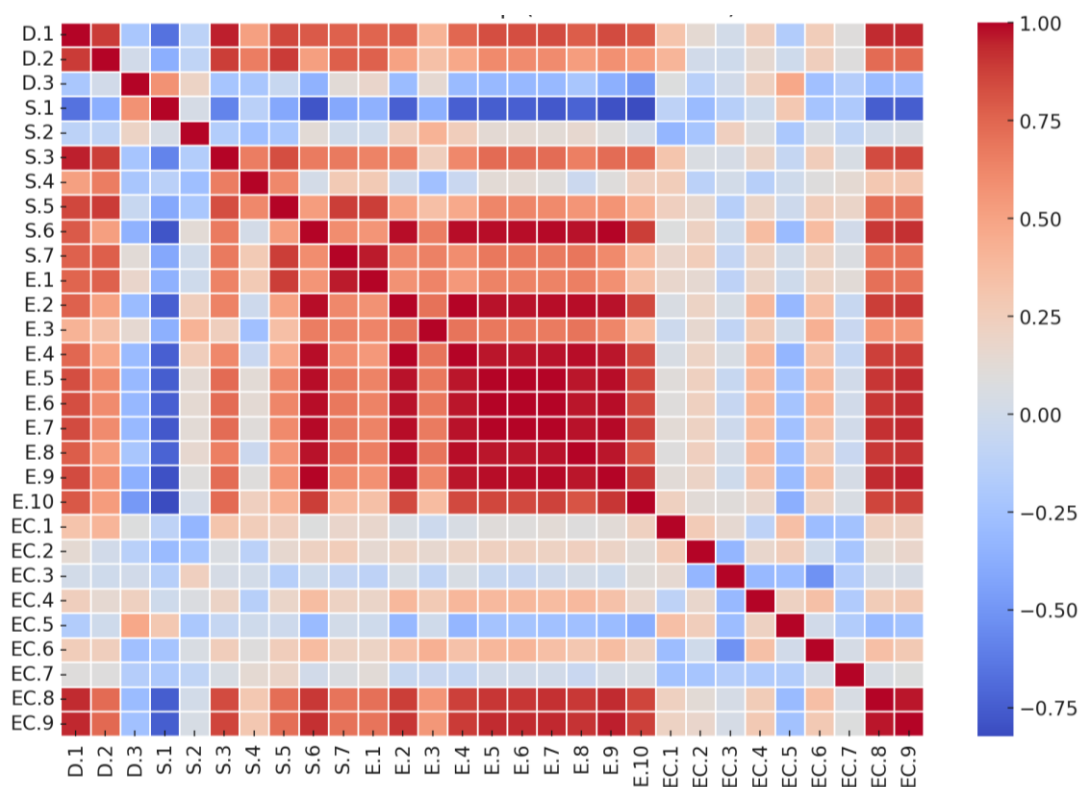


Рисунок 18 – Тепловая карта зависимостей статистических показателей Бухар-Жырауского района

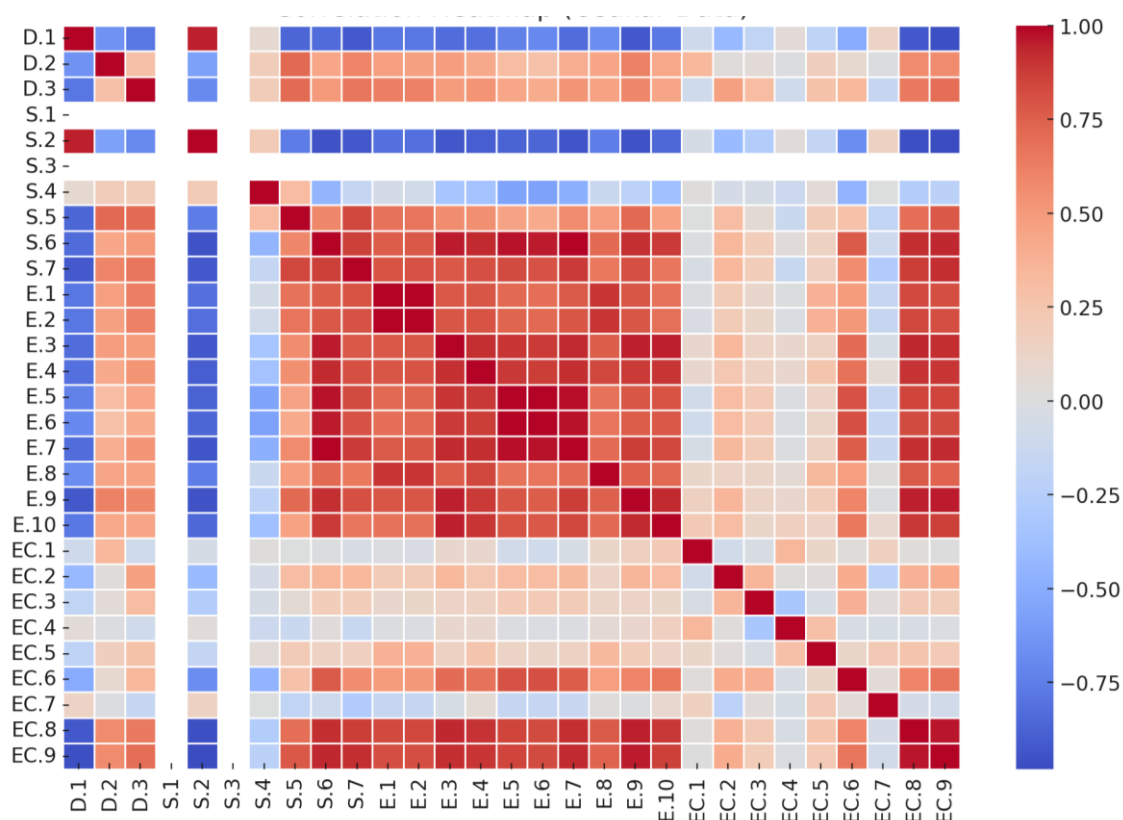


Рисунок 19 – Тепловая карта зависимостей статистических показателей Осакаровского района

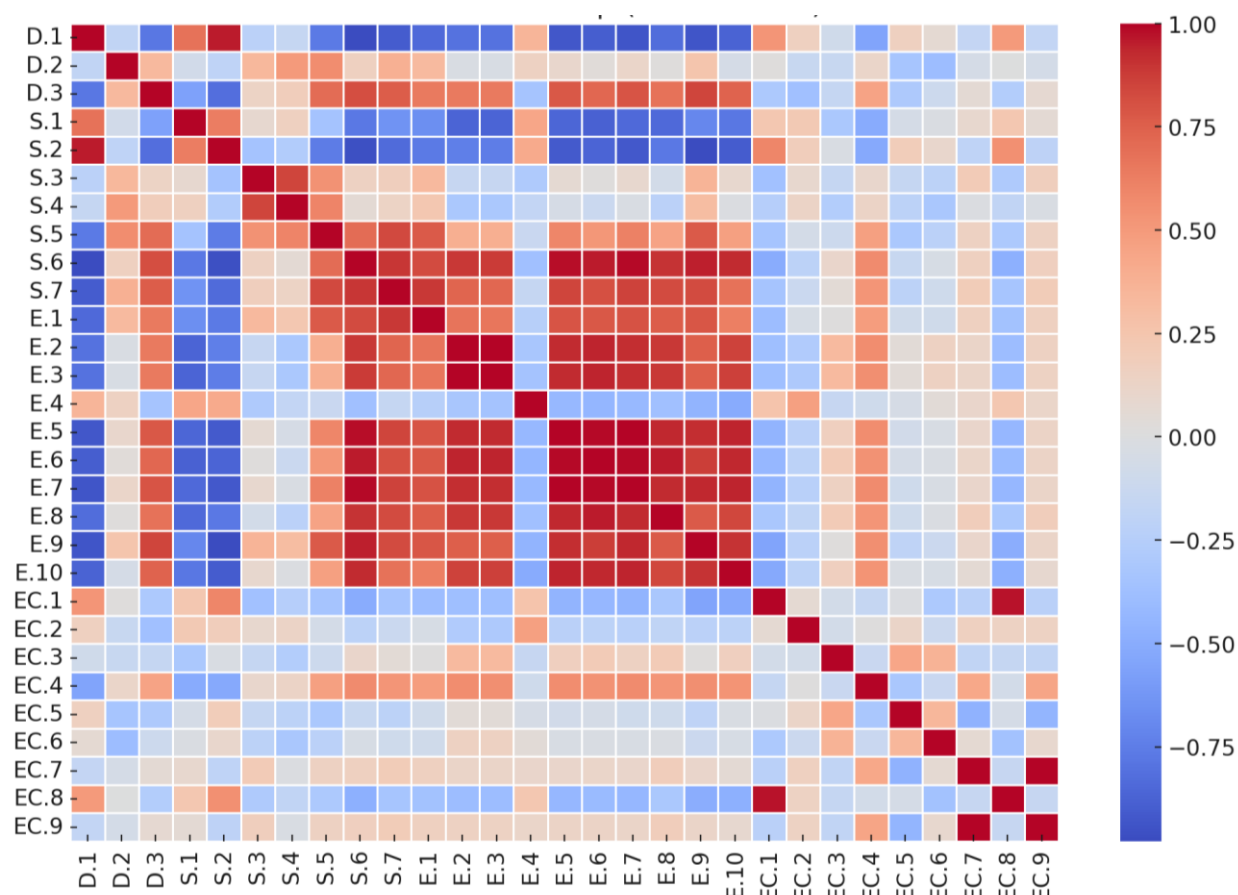


Рисунок 20 – Тепловая карта зависимостей статистических показателей Каркаралинского района

В рамках исследования был проведён анализ главных компонент (PCA) с целью выявления ключевых факторов, определяющих динамику социально-экономических, демографических и экологических параметров, а также их связь с изменением площадей классов LULC. Метод главных компонент позволил сократить размерность исходного многомерного пространства, выделив несколько латентных факторов, наиболее существенно объясняющих общую дисперсию данных. Данный подход позволил выявить скрытые закономерности, сгруппировать взаимосвязанные переменные и устранить избыточность в исходных данных, что критически важно для построения интерпретируемых регрессионных моделей прогнозирования изменений землепользования.

Изначально анализ включал широкий спектр переменных, характеризующих экономическое развитие, демографическую ситуацию и климатические условия. Среди них численность населения, уровень урбанизации, валовой региональный продукт, инвестиции в основной капитал, уровень заработной платы, а также климатические показатели, такие как осадки, температура воздуха и вегетационные индексы. Все эти параметры варьировались во времени и по территориям, формируя многомерную матрицу данных. Перед проведением PCA был выполнен процесс стандартизации, чтобы переменные с разными шкалами измерения имели одинаковый вес в итоговом анализе. Это позволило исключить ситуацию, при которой переменные с

большим разбросом значений (например, численность населения) доминировали бы над переменными с меньшим разбросом (например, индексами растительности).

Стандартный PCA требует центрирования и нормировки входных переменных, после чего ищет векторы-направления $\omega_1, \omega_2, \dots$, при проекции на которые дисперсия максимальна. Первая главная компонента PC_1 даёт наибольший вклад, вторая (PC_2) - следующий по величине вклад и т.д. Формально, для первой компоненты решается задача по формуле (21):

$$\max_{\|\omega\|=1} Var(X\omega) \quad (21)$$

где ω – искомое направление в пространстве переменных.

Коэффициенты вектора ω_i показывают, какие исходные переменные сильнее всего влияют на соответствующую компоненту PC_i . После определения PC_1, PC_2 каждая точка X_t имеет координаты по двум компонентам ($PC_1(t), PC_2(t)$). Эти новые переменные коррелируются с изменениями в площадях классов LULC.

Метод главных компонент предполагает поиск направлений в пространстве переменных, при проекции на которые дисперсия максимальна. Первая главная компонента (P1) объясняет наибольшую долю вариативности данных, вторая компонента (P2) объясняет следующую по величине часть дисперсии, и так далее. Результаты анализа PCA приведены ниже в таблицах 13, 14, 15, 16.

Таблица 13 – распределение переменных в главных компонентах по Абайскому району

Абайский район		
переменная	компонента P1	компонента P2
Совокупный объем промышленного производства	0.342	0.094
Общий объем продукции животноводства	0.356	-0.073
Общий выпуск продукции сельского хозяйства	0.355	-0.067
Общий объем продукции растениеводства	0.351	-0.065
Объем выполненных строительных работ	0.344	-0.155
Объем обрабатывающей промышленности	0.342	0.049
Среднегодовой уровень атмосферных осадков	0.337	0.157
Среднегодовая температура воздуха	0.337	0.051
Нормализованный индекс запыленности (NDTI)	0.159	-0.039
Улучшенный вегетационный индекс (EVI)	0.072	-0.020
Нормализованный индекс застроенных территорий (NDBI)	0.099	0.640
Нормализованный индекс влажности почвы (NDMI)	0.069	-0.716

Таблица 14 – распределение переменных в главных компонентах по Бухар-Жыраускому району

Бухар-Жырауский район		
переменная	компонента P1	компонента P2
Общий объем продукции животноводства	0.350	-0.021
Общий выпуск продукции сельского хозяйства	0.350	0.009
Общий объем продукции растениеводства	0.350	0.016
Совокупный объем промышленного производства	0.345	0.024
Объем выполненных строительных работ	0.344	-0.020
Объем розничного товарооборота	0.349	-0.015
Среднегодовой уровень атмосферных осадков	0.336	-0.100
Среднегодовая температура воздуха	0.331	-0.076
Нормализованный индекс застроенных территорий (NDBI)	0.146	0.314
Нормализованный дифференциальный вегетационный индекс (NDVI)	0.046	-0.691
Улучшенный вегетационный индекс (EVI)	0.085	-0.346
Нормализованный индекс запыленности (NDTI)	0.142	0.536

Таблица 15 – распределение переменных в главных компонентах по Осакаровскому району

Осакаровский район		
переменная	компонента P1	компонента P2
Объем горнодобывающей промышленности	0.337	0.096
Объем обрабатывающей промышленности	0.329	0.159
Общий объем продукции животноводства	0.343	0.039
Объем розничного товарооборота	0.324	0.123
Общий выпуск продукции сельского хозяйства	0.334	0.006
Общий объем продукции растениеводства	0.329	-0.004
Среднегодовая температура воздуха	0.333	0.085
Среднегодовой уровень атмосферных осадков	0.335	0.057
Нормализованный индекс запыленности (NDTI)	0.277	-0.252
Улучшенный вегетационный индекс (EVI)	0.147	-0.527
Нормализованный индекс влажности почвы (NDMI)	0.073	0.332
Модифицированный нормализованный индекс водных объектов (MNDWI)	0.091	-0.697

Таблица 16 – распределение переменных в главных компонентах по Каркаралинскому району

Каркаралинский район		
переменная	компонента P1	компонента P2
Общий выпуск продукции сельского хозяйства	0.377	-0.096
Общий объем продукции животноводства	0.376	-0.094
Общий объем продукции растениеводства	0.376	-0.099
Совокупный объем промышленного производства	0.369	-0.093
Объем горнодобывающей промышленности	0.369	-0.097
Объем выполненных строительных работ	0.358	-0.059
Среднегодовой уровень атмосферных осадков	0.104	0.631
Нормализованный индекс водных объектов (NDWI)	0.097	0.636
Улучшенный вегетационный индекс (EVI)	-0.096	0.265
Среднегодовая температура воздуха	-0.203	-0.062
Нормализованный индекс застроенных территорий (NDBI)	0.242	0.243
Нормализованный дифференциальный вегетационный индекс (NDVI)	-0.210	-0.125

В рамках исследования было установлено, что первая главная компонента объясняет от 55.5 до 67.7% общей изменчивости данных в зависимости от региона, тогда как вторая главная компонента описывает от 10.8 до 18.1%, суммарно обеспечивая объяснение от 73.6 до 79.1% общей дисперсии. Это, по нашему мнению, подтверждает, что построенная модель адекватно отражает внутреннюю структуру исходных данных и позволяет корректно интерпретировать влияние ключевых факторов на динамику LULC.

Результаты анализа главных компонент (PCA) показали, что первая главная компонента (P1) тесно связана с экономическими переменными, включая инвестиции в основной капитал, объём промышленного производства, уровень заработной платы и численность населения. Высокие значения загрузочных коэффициентов по данным показателям свидетельствуют о том, что P1 отражает уровень экономического развития региона и степень антропогенного воздействия на территорию.

Анализ загрузок продемонстрировал, что P1 оказывает наиболее существенное влияние на категории сельскохозяйственных земель (пашни и пастбища), а также на урбанизированные территории. В отдельных районах её воздействие на застроенные площади достигало значений $R^2 = 0.96$, а на пашни до 0.91, что указывает на высокую степень детерминированности указанных типов землепользования экономическими процессами.

Вторая главная компонента (P2), в свою очередь, демонстрирует выраженную связь с климатическими факторами, такими как среднегодовая температура, объём атмосферных осадков, а также вегетационные индексы (NDVI, EVI). Значительные загрузочные коэффициенты данных переменных позволяют интерпретировать P2 как климато-экологический фактор, определяющий изменчивость природных классов землепользования, включая лесные территории, водные объекты и болота.

Результаты регрессионного анализа подтвердили, что P2 оказывает существенное влияние на динамику лесных экосистем, особенно в пределах Каркаралинского района ($R^2 = 0.91$). Влияние на водно-болотные угодья оказалось менее выраженным: коэффициенты детерминации варьировались в диапазоне от 0.03 до 0.42, что, по нашему мнению, указывает на более сложную и многофакторную природу их пространственно-временной изменчивости.

Полученные результаты демонстрируют, что объединение социально-экономических и климатических факторов через метод главных компонент позволяет выявить глубинные взаимосвязи в структуре данных, что, в свою очередь, даёт возможность строить более точные модели прогнозирования изменений землепользования. Выделенные компоненты не только позволяют сократить размерность данных без потери значимой информации, но и формируют основу для построения интерпретируемых регрессионных уравнений, связывающих динамику LULC с ключевыми факторами, управляющими её изменениями. Это открывает возможности для дальнейшего применения метода PCA в комплексном анализе пространственной динамики, а

также в разработке стратегий управления земельными ресурсами, устойчивого планирования территорий и экологического мониторинга.

4.3 Регрессионный анализ

Была проведена регрессионная оценка влияния главных компонент, полученных с помощью анализа PCA, на динамику классов LULC в четырёх исследуемых районах. Основной целью данного этапа анализа являлось формализованное определение количественных зависимостей между социально-экономическими, демографическими и экологическими факторами и изменением площадей различных классов LULC. Применение метода множественной линейной регрессии позволило не только установить наличие статистически значимых связей, но и оценить вклад каждой главной компоненты в объяснение динамики землепользования.

Концептуальная модель основана на стандартной форме множественной линейной регрессии, в которой приращение и сокращение площади конкретного класса LULC моделируется как функция главных компонент. В данной постановке P1 и P2 выступают в роли предикторов, а регрессионные коэффициенты β отражают степень их влияния. Если коэффициент β статистически значимо отличен от нуля, это указывает на наличие сильной взаимосвязи между компонентой и соответствующим классом землепользования рассчитываемой по формуле (22).

$$\Delta LULC_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \dots + \beta_n X_n + \varepsilon \quad (22)$$

где β_n – искомые регрессионные коэффициенты,
 ε – случайная ошибка.

Так β_1 показывает, насколько единичное изменение фактора X_1 в среднем приводит к изменению $\Delta LULC_i$.

Если β_n значительно отличен от нуля при статистически значимом уровне, это говорит о существенном вкладе фактора X_n .

Результаты регрессионного моделирования показали, что наиболее детерминированными категориями землепользования являются пашни, пастбища и застроенные территории. В районах Абай, Бухар-Жырау и Осакаровка пашни демонстрируют высокую предсказуемость модели, что подтверждается коэффициентами детерминации $R^2 = 0.91$ в Абаете, $R^2 = 0.66$ в Бухар-Жырау и $R^2 = 0.82$ в Осакаровке. Вклад P1 в эти регрессионные модели значителен, что подтверждает сильное влияние экономических факторов на изменение сельскохозяйственных площадей. Например, в Абаете уравнение (23) для пашен принимает вид:

$$\text{Пашни} = 5437.95 \times PC1 + 200.79 \times PC2 + 55228.27 \quad (23)$$

что свидетельствует о том, что рост экономической активности (отражённый в P1) приводит к увеличению площади пашен. Аналогичная ситуация наблюдается для пастбищ, где значения $R^2 = 0.91$ (Абай), 0.66 (Бухар-Жырау) и 0.82 (Осакаровка) подтверждают значительное влияние социально-экономических факторов на этот класс LULC.

Категория застроенных территорий демонстрирует наиболее высокую предсказательную силу модели во всех районах. В Осакаровке регрессионное уравнение (24):

$$\text{Застроенные территории} = 12.92 \times PC1 - 1.15 \times PC2 + 5428.11 \quad (24)$$

с коэффициентом детерминации $R^2=0.96$ показывает, что изменение урбанизированных земель почти полностью объясняется первой главной компонентой, что связано с интенсивной урбанизацией и ростом инвестиций в инфраструктуру. В Бухар-Жырау аналогичное уравнение (25):

$$\text{Застроенные территории} = 82.80 \times PC1 - 21.26 \times PC2 + 30492.88 \quad (25)$$

с $R^2=0.93$ подтверждает сильную зависимость застройки от экономических факторов, таких как численность населения и уровень урбанизации.

Прогнозируемая динамика лесных массивов оказалась неоднородной в разных районах. В Каркаралинске леса показали наиболее высокую степень объяснения с $R^2=0.91$, что позволяет с высокой точностью предсказывать изменения лесных площадей на основе P1 и P2. Уравнение (26):

$$\text{Лес} = 315.29 \times PC1 - 63.20 \times PC2 + 99657.42 \quad (26)$$

свидетельствует о значительном вкладе как экономических, так и климатических факторов. В других регионах уровень объясняющей способности ниже, например, в Бухар-Жырау $R^2 = 0.54$, что может быть связано с большей зависимостью лесных экосистем от локальных климатических условий, не учтённых в текущих компонентах.

Категории водных объектов и болот продемонстрировали наименьшую предсказательную силу модели. В Осакаровском районе водные объекты показали $R^2=0.03$, а болота $R^2=0.05$, что указывает на высокую пространственную изменчивость и сложные природные детерминанты этих классов LULC. В Бухар-Жырау и Каркаралинске водные объекты предсказываются несколько лучше ($R^2=0.25-0.37$), однако показатели всё равно остаются низкими. Для болот регрессионные уравнения (27), (28) варьируются по регионам:

$$\text{Водно-болотные угодья} = -41.2 \times PC1 - 1.50 \times PC2 + 3103.73 \quad (27)$$

в Абее и Каркаралинске:

$$\text{Водно-болотные угодья} = -23.01 \times \text{PC1} - 16.46 \times \text{PC2} + 4840.14 \quad (28)$$

Это подтверждает, что для корректного прогнозирования водных и болотных экосистем требуется учёт дополнительных переменных, таких как гидрологические факторы и сезонные колебания влажности.

Проведённый регрессионный анализ, по нашему мнению, позволил не только количественно оценить степень влияния ключевых факторов на трансформацию структуры землепользования, но и выявить различия в предсказательной способности построенных моделей для отдельных классов LULC. Высокие значения коэффициента детерминации (R^2), полученные для пашен, пастбищ и урбанизированных территорий, подтверждают, что динамика данных категорий в значительной степени обусловлена экономическими процессами. В то же время лесные экосистемы, водные объекты и болотные территории демонстрируют большую чувствительность к воздействию климатических и гидрологических факторов, что указывает на иную природу их пространственно-временной изменчивости.

Выводы по разделу 4

Данная работа, основанная на интеграции спутниковых данных землепользования (LULC) с социально-экономическими, демографическими и экологическими показателями за период 2000–2024 гг., позволила не только описать пространственно-временные изменения ландшафтов, но и проанализировать их причинно-факторную обусловленность. Использование официальной статистической информации о численности населения, инвестиционной активности, климатических параметрах, уровне урбанизации и промышленного развития способствовало выявлению ключевых драйверов трансформации землепользования и количественной оценке степени их влияния. По нашему мнению, подобный интеграционный подход имеет важное значение для задач территориального планирования и принятия управленческих решений в сфере рационального природопользования.

Применение метода анализа главных компонент (РСА) позволило выделить базовые факторы, определяющие изменения структуры землепользования. Первая главная компонента (P1) объясняет от 55.5% вариации данных в Каркаралинском районе до 67.7% в Осакаровском, тогда как вклад второй компоненты (P2) варьируется от 10.8% (Осакаровский) до 18.1% (Каркаралинский). В совокупности обе компоненты охватывают от 73.6% (Каркаралинский) до 79.1% (Бухар-Жырауский) общей вариации, что, на наш взгляд, подтверждает высокую эффективность применённого метода. Использование РСА позволило существенно сократить сложность исходного массива данных и выделить наиболее значимые закономерности их изменчивости.

Полученные результаты показали, что P1 демонстрирует тесную связь с экономическими параметрами, включая инвестиции в промышленность и

сельское хозяйство, а также уровень заработной платы. Данное влияние особенно выражено для категорий пашен и пастбищ, характеризующихся высокими коэффициентами детерминации: $R^2 = 0.91$ для пашен в Абайском районе и $R^2 = 0.88$ в Бухар-Жырауском. Пастбищные территории также демонстрируют значительную зависимость от экономических факторов ($R^2 = 0.91$ в Абайском и $R^2 = 0.82$ в Осакаровском районах). В Каркаралинском районе указанная зависимость выражена слабее ($R^2 = 0.51$ и $R^2 = 0.47$ соответственно), что, по нашему мнению, может быть обусловлено более существенным влиянием природно-экологических условий.

Категория застроенных территорий демонстрирует наивысшую предсказательную точность модели во всех рассматриваемых регионах. Значения коэффициента детерминации составили $R^2 = 0.96$ в Осакаровском районе, $R^2 = 0.93$ в Бухар-Жырауском и $R^2 = 0.82$ в Каркаралинском, что подтверждает тесную взаимосвязь урбанизированных пространств с экономическими процессами. Основной вклад в формирование данной зависимости вносит первая главная компонента (P1), что, по нашему мнению, объясняется прямой связью застройки с инвестициями в инфраструктуру и ростом численности населения. Влияние второй компоненты (P2) остаётся минимальным, указывая на слабую зависимость урбанизированных территорий от климатических факторов.

Анализ лесных массивов показал, что уровень их предсказуемости существенно варьируется по регионам. Наиболее высокие значения зафиксированы в Каркаралинском районе ($R^2 = 0.91$), что свидетельствует о сильной связи с выделенными главными компонентами. В Осакаровском и Бухар-Жырауском районах показатели ниже ($R^2 = 0.67$ и $R^2 = 0.54$ соответственно). В Абайском районе коэффициент детерминации составил лишь $R^2 = 0.28$, что, как представляется, указывает на преобладание природных факторов над экономическими в формировании динамики лесных территорий.

Вклад P1 в моделирование лесных изменений варьируется от 125.92 (Абайский район) до 315.29 (Каркаралинский), тогда как влияние P2 выражено сильнее по сравнению с другими классами LULC и колеблется в диапазоне от -63.20 до -24.99, что подчёркивает значимость климато-экологических условий.

Категории водных объектов и болот продемонстрировали наименьшую предсказательную способность модели, особенно в Осакаровском районе, где значения составили $R^2 = 0.03$ (водные объекты) и $R^2 = 0.05$ (болота). В Бухар-Жырауском и Каркаралинском районах точность несколько выше ($R^2 = 0.25$ – 0.42), однако и эти показатели свидетельствуют о слабой зависимости от главных компонент.

Подобная ситуация, по нашему мнению, объясняется высокой чувствительностью водных экосистем и болотных территорий к сезонным колебаниям, изменениям режима осадков и уровню почвенной влажности. Коэффициенты влияния P1 на водные объекты варьируются от 5.83 (Осакаровский район) до 150.99 (Каркаралинский), тогда как вклад P2 находится

в диапазоне от 22.49 до 63.25, что указывает на доминирующее воздействие природно-климатических факторов.

Интеграция статистических данных значительно расширила возможности анализа и принятия решений в области планирования землепользования. Ранее карты LULC фиксировали только фактические изменения в структуре территорий, теперь же стало возможным не только объяснять причины этих изменений, но и прогнозировать их развитие. Полученные результаты могут использоваться для оценки устойчивости природопользования, выявления территорий с высоким риском потери сельскохозяйственного потенциала или экологической деградации.

ЗАКЛЮЧЕНИЕ

Диссертация представляет собой завершённую научно-квалификационную работу, в которой содержится новое решение актуальной научной и практической задачи – анализ и управление динамикой землепользования с использованием методов дистанционного зондирования Земли, геоинформационных технологий и машинного обучения. Основные научные и практические результаты диссертационной работы заключаются в следующем:

1. Проведён комплексный анализ изменений землепользования на территории Карагандинской области за период 2000–2024 гг. с использованием спутниковых данных Landsat и инструментов ГИС. Это, по нашему мнению, позволило выявить ключевые тенденции трансформации ландшафтов, включая рост урбанизированных территорий, восстановление растительного покрова и структурные изменения сельскохозяйственных угодий.

2. Разработаны и апробированы подходы к интеграции данных дистанционного зондирования, предусматривающие расчёт вегетационных индексов (NDVI, EVI), а также анализ ночной светимости. Применение данных методов существенно повысило точность классификации и уровень детализации создаваемых карт землепользования.

3. Алгоритм Random Forest был адаптирован и оптимизирован для решения задач классификации землепользования. В процессе моделирования использовалась стратифицированная обучающая выборка, а также дополнительные предикторы, включая цифровую модель рельефа (SRTM) и вегетационные индексы. В результате достигнут высокий уровень точности классификации — до 92% при коэффициенте Карра, равном 0.88, что подтверждает надёжность полученных результатов. На основе моделирования сформированы классификационные карты землепользования, отражающие пространственно-временную динамику ландшафтов, а также выявлены ключевые драйверы изменений, среди которых урбанизация, экономическая активность и климатические факторы.

4. Разработана методология оценки устойчивости территорий к антропогенному воздействию, основанная на интеграции статистических показателей социально-экономического и демографического развития с пространственными геоданными.

5. Определены основные направления трансформации землепользования в пределах исследуемых районов. Так, в Осакаровском районе зафиксировано сокращение площадей пастбищ в пользу пахотных земель, тогда как в Каркаралинском районе отмечен рост лесных массивов, обусловленный реализацией природоохранных мероприятий.

6. Сформулированы практические рекомендации, направленные на обеспечение устойчивого управления земельными ресурсами, включая оптимизацию использования сельскохозяйственных и урбанизированных территорий, восстановление деградированных земель и сохранение биологического разнообразия.

7. Внедрение результатов исследования в производственную практику продемонстрировало их эффективность для мониторинга и планирования землепользования, включая сценарное моделирование и прогнозирование изменений в природных и антропогенных ландшафтах.

8. Экономический эффект от применения предложенных методов подтверждается повышением точности и эффективности оценки землепользования, что создает основу для принятия обоснованных решений в области территориального планирования и управления природными ресурсами.

Результаты исследований реализованы в рамках территориального планирования и мониторинга динамики землепользования, а также интегрированы в образовательные и исследовательские программы.

Таким образом, выполненная работа вносит значительный вклад в развитие методов анализа землепользования с использованием современных технологий дистанционного зондирования, машинного обучения и ГИС, обеспечивая основу для устойчивого управления земельными ресурсами.

СПИСОК ИСПОЛЬЗОВАННЫХ ИСТОЧНИКОВ

- 1 Журкин И.Г., Шайтура С.В. Геоинформационные системы. – М.: КУИЦ-ПРЕСС, 2009. – 272 с.
- 2 Liping C., Yujun S., Saeed S. Monitoring and predicting land use and land cover changes using remote sensing and GIS techniques – A case study of a hilly area, Jiangle, China // PloS ONE. – 2018. – Vol. 13. – P. e0200493.
- 3 Alipbeki O. et al. A spatiotemporal assessment of land use and land cover changes in peri-urban areas: A case study of Arshaly District, Kazakhstan // Sustainability. – 2020. – Vol. 12, Issue 4. – P. 1556-1-1556-15.
- 4 Aplin P. Remote sensing: Land cover // Progress in Physical Geography. – 2003. – Vol. 27, Issue 2. – P. 201-211.
- 5 Bartholome E., Belward A.S. GLC2000: A new approach to global land cover mapping from Earth observation data // Int. J. Remote Sens. – 2005. – Vol. 26, Issue 9. – P. 1959-1977.
- 6 Gorelick N., Hancher M., Dixon M. et al. Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone // Remote Sensing of Environment. – 2017. – Vol. 202. – P. 18-27.
- 7 Halder A., Ghosh A., Ghosh S. Supervised and unsupervised landuse map generation from remotely sensed images using ant based systems // Applied Soft Computing. – 2011. – Vol. 11, Issue 8. – P. 5770-5781.
- 8 Braun A., Hochschild V. A SAR-based index for landscape changes in African savannas // Remote Sensing. – 2017. – Vol. 9, Issue 4. – P. 359-1-359-23.
- 9 DeFries R.S., Townshend J.R.G., Hansen M.C Continuous Fields of Vegetation Characteristics at the Global Scale at 1-km Resolution // Journal of Geophysical Research-Atmospheres. – 1999. – Vol. 104. – P. 16911-16923.
- 10 Singh S.K., Mustak S., Srivastava P.K. et al. Predicting spatial and decadal LULC changes through cellular automata Markov chain models using Earth observation datasets and geo-information // Environmental Processes. – 2015. – Vol. 2. – P. 61-78.
- 11 Lateef F., Ruichek Y. Survey on semantic segmentation using deep learning techniques // Neurocomputing. – 2019. – Vol. 338. – P. 321-348.
- 12 Alipbeki O., Alipbekova Ch., Mussaif G., Grossul P.P., Zhenshan D., Muzyka O., Turekeldiyeva R., Yelubayev D., Rakhimov D., Kupidura P., Aliken E. Analysis and Prediction of Land Use/Land Cover Changes in Korgalzhyn District, Kazakhstan // Agronomy. – 2024. – Vol. 14, Issue 2. – P. 268-1-268-28. – <https://doi.org/10.3390/agronomy14020268>.
- 13 Zhu Y., Geiß C., So E. et al. Multitemporal relearning with convolutional LSTM models for land use classification // J. Sel. Top. Appl. Earth Obs. Remote Sens. – 2021. – Vol. 14. – P. 3251-3265.
- 14 Wang Z., Yang P., Liang H. et al. Semantic segmentation and analysis on sensitive parameters of forest fire smoke using Smoke-Unet and Landsat-8 imagery // Remote Sens. – 2022. – Vol. 14. – P. 45-1-45-19.

- 15 Merchant C.J., Paul F., Popp T. et al. Uncertainty information in climate data records from Earth observation // *Earth Syst. Sci. Data*. – 2017. – Vol. 9. – P. 511-527.
- 16 Yuan X., Shi J., Gu L. A review of deep learning methods for semantic segmentation of remote sensing imagery // *Expert Syst. Appl.* – 2021. – Vol. 169. – P. 114417.
- 17 Wei H., Wang J., Cheng K. et al. Desertification information extraction based on feature space combinations on the Mongolian Plateau // *Remote Sens.* – 2018. – Vol. 10. – P. 1614-1-1614-17.
- 18 Kolios S., Mitrakos S., Stylios C. Detection of areas susceptible to land degradation in Cyprus using remote sensed data and environmental quality indices // *Land Degrad. Dev.* – 2018. – Vol. 29. – P. 2338-2350.
- 19 van der Meer F.D. et al. Multi- and hyperspectral geologic remote sensing: A review // *Int. J. Appl. Earth Obs. Geoinf.* – 2012. – Vol. 14. – P. 112-128.
- 20 Remer L.A., Kaufman Y.J., Tanre D. et al. The MODIS aerosol algorithm, products, and validation // *J. Atmos. Sci.* – 2005. – Vol. 62. – P. 947-973.
- 21 Ramirez-Reyes C., Brauman K. et al. Reimagining the potential of Earth observations for ecosystem service assessments // *Sci. Total Environ.* – 2019. – Vol. 665. – P. 1053-1063.
- 22 Yao X.C., Li G.Q., Xia J.S. et al. Enabling the big Earth observation data via cloud computing and DGGS: Opportunities and challenges // *Remote Sens.* – 2020. – Vol. 12. – P. 62-1-62-15.
- 23 Rivera-Marin D., Dash J., Ogutu B. The use of remote sensing for desertification studies: A review // *J. Arid. Environ.* – 2022. – Vol. 206. – P. 104829.
- 24 Alipbeki O., Rakhimov D., Kupidura P., Alipbekova Ch., Musaif G., Turekeldiyeva R., Augambaev K., Begaliyeva M., Grossul P.P. Ecosystem Health Assessment of the Zerendy District, Kazakhstan // *Sustainability*. – 2025. – Vol. 17, Issue 1. – P. 277-1-277-25. – <https://doi.org/10.3390/su17010277>.
- 25 Vasavi S., Sri Somagani H., Sai Y. Classification of buildings from VHR satellite images using ensemble of U-Net and ResNet // *Egypt. J. Remote Sens. Space Sci.* – 2023. – Vol. 26. – P. 937-953.
- 26 Cai D., Wang X., Hua T. et al. Baseline and status of desertification in Central Asia // *Land Degrad. Dev.* – 2022. – Vol. 33. – P. 771-784.
- 27 Zhang C., Sargent I., Pan X. et al. Joint deep learning for land cover and land use classification // *Remote Sensing of Environment*. – 2019. – Vol. 221. – P. 173-187.
- 28 Weng L., Pang K., Xia M. et al. Sgformer: A local and global features coupling network for semantic segmentation of land cover // *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* – 2023. – Vol. 16. – P. 6812-6824.
- 29 Segarra J., Buchaillot M.L., Araus J.L. et al. Remote sensing for precision agriculture: Sentinel-2 improved features and applications // *Agronomy*. – 2020. – Vol. 10, Issue 5. – P. 641-1-641-18.
- 30 Kadri N., Jebari S., Augusseau X. et al. Analysis of four decades of land use and land cover change in semiarid Tunisia using Google Earth Engine // *Remote Sens.* – 2023. – Vol. 15. – P. 3257-1-3257-16.

- 31 Nasiri E., Matkan A.A., Azarakhsh Z. et al. Land use/land cover mapping using Sentinel-2 data and machine learning algorithms in Google Earth Engine // J of the Indian Society of Remote Sensing. – 2022. – Vol. 50, Issue 2. – P. 369-381.
- 32 Xia J.Z. et al. Building a spatiotemporal index for Earth observation big data // Int. J. Appl. Earth Obs. Geoinf. – 2018. – Vol. 73. – P. 245-252.
- 33 SDG Indicator Metadata (2022) // <https://docs.yandex.kz/docs>. 10.10.2025.
- 34 Wang Z., Ma Y., Zhang Y. et al. Review of remote sensing applications in grassland monitoring // Remote Sens. – 2022. – Vol. 14. – P. 2903-1-2903-40.
- 35 Wu L., Zhu X., Lawes R. et al. Comparison of machine learning algorithms for classification of LiDAR points for characterization of canola canopy structure // International Journal of Remote Sensing. – 2019. – Vol. 40, Issue 16. – P. 5973-5991.
- 36 Hu Y., Yang G. Analysis of land use and land cover change in Kazakhstan from 1992 to 2015 // Remote Sensing. – 2019. – Vol. 11, Issue 22. – P. 2630.
- 37 Alipbeki O., Mussaif G., Grossul P. et al. Untangling the integral impact of land use change, economic, ecological and social factors on the development of Burabay District (Kazakhstan) during the period 1999–2021 // Sustainability. – 2023. – Vol. 15. – P. 7548-1-7548-36.
- 38 Alipbeki O., Rakhimov D., Alikin E., Augambayev K., Alipbekova Ch., Begaliyeva M., Ravshanov A., Grossul P.P. Spatiotemporal analysis and identifying the driving forces of land use change in the Abay district (Karagandy Region, Kazakhstan) // E3S Web of Conferences. – 2024. – Vol. 590. – P. 04007-1-04007-11. – <https://doi.org/10.1051/e3sconf/202459004007>.
- 39 Land Use, Land-Use Change, and Forestry: Special Report of the Intergovernmental Panel on Climate Change / IPCC. – Cambridge, 2000. – 377 p.
- 40 Rahman A., Kumar S., Fazal S. et al. Assessment of land use/land cover change in the North-West District of Delhi using remote sensing and GIS techniques // Journal of the Indian Society of Remote Sensing. – 2012. – Vol. 40, Issue 4. – P. 689-697.
- 41 Prince S.D. Challenges for remote sensing of the Sustainable Development Goal SDG 15.3.1 productivity indicator // Remote Sens. Environ. – 2019. – Vol. 234. – P. 111428.
- 42 Wang D., Yang R., Zhang Z. et al. P-Swin: Parallel Swin Transformer multi-scale semantic segmentation network for land cover classification // Comput. Geosci. – 2023. – Vol. 175. – P. 105340.
- 43 Vali A., Comai S., Matteucci M. Deep learning for land use and land cover classification based on hyperspectral and multispectral Earth observation data: A review // Remote Sens. – 2020. – Vol. 12. – P. 2495-1-2495-31.
- 44 Гук А.П., Евстратова Л.Г., Алфёрова А.С. Комплексный анализ космических снимков для мониторинга территорий // Сб. науч. тр. СГГА. – Новосибирск, 2009. – С. 181-185.
- 45 AbdelRahman M.A.E. An overview of land degradation, desertification and sustainable land management using GIS and remote sensing applications // Rend. Fis. Acc. Lincei. – 2023. – Vol. 34. – P. 767-808.

- 46 Deng X., Huang J., Rozelle S. et al. Economic growth and the expansion of urban land in China // *Urban Studies*. – 2010. – Vol. 47. – P. 813-843.
- 47 Carvalho-Santos C. et al. Assessing the effects of land cover and future climate conditions on the provision of hydrological services in a medium-sized watershed of Portugal // *Hydrological Processes*. – 2016. – Vol. 30. – P. 720-738.
- 48 Әліпбеки О.А., Алипбекова Ч.А., Гроссул П.П. Прогнозирование изменений землепользования и земного покрова на основе комбинированного подхода клеточных автоматов и искусственных нейронных сетей // *Материалы VI Международной научно-практической конференции «Геодезия, землеустройство и кадастры: проблемы и перспективы развития»*. – Омск, 2024. – С. 200-204.
- 49 Cerovski-Darriau C., Roering J.J. Influence of anthropogenic land-use change on hillslope erosion in the Waipaoa River Basin, New Zealand // *Earth Surface Processes and Landforms*. – 2016. – Vol. 41, Issue 14. – P. 2167-2176.
- 50 Dong R., Fang W., Fu H. et al. High-resolution land cover mapping through learning with noise correction // *IEEE Trans. Geosci. Remote Sens.* – 2022. – Vol. 60. – P. 4402013.
- 51 Olofsson P., Foody G.M., Herold M. et al. Good practices for estimating area and assessing accuracy of land change // *Remote Sensing of Environment*. – 2014. – Vol. 148. – P. 42-57.
- 52 Pal S., Talukdar S. Assessing the role of hydrological modifications on land use/land cover dynamics in Punarbhaba river basin of Indo-Bangladesh // *Environmental Development and Sustainability*. – 2020. – Vol. 22, Issue 1. – P. 363-382.
- 53 Алипбеки О.А., Гроссул П.П. Современные вызовы и технологии в определении динамики землепользования и земного покрова // *Материалы Международной научной конференции «Фараби әлемі»*. – Алматы, 2024.
- 54 Molina A.A., Huang Y., Jiang Y. A review of unmanned aerial vehicle applications in construction management: 2016–2021 // *Standards*. – 2023. – Vol. 3. – P. 95-109.
- 55 Barthakur M., Sarma K.K., Mastorakis N. Modified semi-supervised adversarial deep network and classifier combination for segmentation of satellite images // *IEEE Access*. – 2020. – Vol. 8. – P. 117972-117985.
- 56 Yu Y., Guo Z., Wang Y.-C. Spatial patterns and driving forces of land change in Tibetan-inhabited Three Rivers Headwaters region, China // *J. Mt. Sci.* – 2019. – Vol. 16. – P. 207-225.
- 57 Prakash M. et al. Open Earth observations for sustainable urban development // *Remote Sens.* – 2020. – Vol. 12. – P. 1646-1-1646-16.
- 58 Lamaamri M., Lghabi N., Ghazi A. et al. Evaluation of desertification in the Middle Moulouya Basin (North-East Morocco) using Sentinel-2 images and spectral index techniques // *Earth Syst. Environ.* – 2023. – Vol. 7. – P. 473-492.
- 59 Smith A.M., Wooster M.J., Drake N.A. et al. Testing the potential of multi-spectral remote sensing for retrospectively estimating fire severity in African savanna environments // *Remote Sensing of Environment*. – 2003. – Vol. 88, Issue 1-2. – P. 161-173.

- 60 Sun P., Lu Y., Zhai J. Mapping land cover using a developed U-Net model with weighted cross entropy // *Geocarto Int.* – 2022. – Vol. 37. – P. 9355-9368.
- 61 Li J., Wang H., Zhang A. et al. Semantic segmentation of hyperspectral remote sensing images based on PSE-Unet model // *Sensors.* – 2022. – Vol. 22. – P. 9678-1-9678-18.
- 62 Talukdar S., Pal S. Effects of damming on the hydrological regime of Punarbhaba river basin wetlands // *Ecological Engineering.* – 2019. – Vol. 135. – P. 61-74.
- 63 Shen X. et al. Multi-scale feature aggregation network for semantic segmentation of land cover // *Remote Sens.* – 2022. – Vol. 14. – P. 6156-1-6156-20.
- 64 Song Y., Wu P. Earth observation for sustainable infrastructure: A review // *Remote Sens.* – 2021. – Vol. 13. – P. 1528-1-1528-20.
- 65 Kazantseva V.V., Ozhigin D.S., Dolgonosov V.N., Ozhigina S.B., Grossul P.P. Assessment of the accuracy of the geometric scheme of GCPs when creating DSM using UAV // *NEWS of the National Academy of Sciences of the Republic of Kazakhstan. Series of Geology and Technical Sciences.* – 2025. – Vol. 2, No. 470. – P. 110-124. – <https://doi.org/10.32014/2025.2518-170X.494>.
- 66 Weng Q., Xu B., Hu X. et al. Use of Earth observation data for applications in public health // *Geocarto Int.* – 2013. – Vol. 29. – P. 3-16.
- 67 Zhang R., Chen J., Feng L. et al. A refined pyramid scene parsing network for polarimetric SAR image semantic segmentation in agricultural areas // *IEEE Geosci. Remote Sens. Lett.* – 2022. – Vol. 19. – P. 4014805.
- 68 Turner B.L., Skole D., Sanderson S. et al. *Land-Use and Land-Cover Change: Science/Research Plan.* – Stockholm, 1994. – 132 p.
- 69 Chen S., Wu C., Mukherjee M. et al. Ha-Mppnet: Height aware-multi path parallel network for high spatial resolution remote sensing image semantic segmentation // *ISPRS Int. J. Geoinf.* – 2021. – Vol. 10. – P. 672-1-672-18.
- 70 Liu R., Tao F., Liu X. et al. RAANet: A residual ASPP with attention framework for semantic segmentation of high-resolution remote sensing images // *Remote Sens.* – 2022. – Vol. 14. – P. 3109-1-3109-18.
- 71 Pontius R.G., Jr., Boersma W., Castella J.-C. et al. Comparing the input, output, and validation maps for several models of land change // *Annals of Regional Science.* – 2008. – Vol. 42. – P. 11-47.
- 72 Paris Agreement (2015) / UN // <https://ru.ruwiki.ru/wiki>. 10.09.2025.
- 73 Wang Q., Luo X., Feng J. et al. CCENet: Cascade class-aware enhanced network for high-resolution aerial imagery semantic segmentation // *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* – 2022. – Vol. 15. – P. 6943-6956.
- 74 Liu S., Wang T., Kang W. et al. Several challenges in monitoring and assessing desertification // *Environ. Earth Sci.* – 2015. – Vol. 73. – P. 7561-7570.
- 75 Kiani-Harchegani M., Sadeghi S.H. Practicing land degradation neutrality (LDN) approach in the Shazand Watershed, Iran // *Sci. Total Environ.* – 2020. – Vol. 698. – P. 134319.

76 Dao P.D., Liou Y.A. Object-based flood mapping and affected rice field estimation with Landsat 8 OLI and MODIS data // Remote Sensing. – 2015. – Vol. 7, Issue 5. – P. 5077-5097.

77 Zhao Y., Yang S., Liu L. et al. Variations and driving mechanisms of desertification in the southeast section of the China-Mongolia-Russia Economic Zone // Sci. Total Environ. – 2023. – Vol. 887. – P. 164004.

78 Алипбеки О.А., Гроссул П.П. Использование искусственных нейронных сетей (ANN) для автоматической классификации типов землепользования в городских районах на основе данных дистанционного зондирования (ДЗЗ) // Материалы XIX Международной научной конференции студентов, магистрантов и молодых ученых «ЛОМОНОСОВ–2024», посвященной 270-летию Московского университета. – Астана, 2024. – С. 216-218.

79 Wang Y., Shi H., Zhuang Y. et al. Bidirectional grid fusion network for accurate land cover classification of high-resolution remote sensing images // IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. – 2020. – Vol. 13. – P. 5508-5517.

80 Xu Y., Wu L., Xie Z. et al. Building extraction in very high resolution remote sensing imagery using deep learning and guided filters // Remote Sens. – 2018. – Vol. 10. – P. 144-1-144-18.

81 Weier J., Herring D. Measuring Vegetation: NDVI and EVI // <https://science.nasa.gov/earth/earth-observatory/measuring-vegetation>. 10.09.2025.

82 Кондратьев К.Я. Актинометрия. – Л.: Гидрометеиздат, 1965. – 690 с.

83 Hou H., Wang R., Murayama Y. Scenario-based modelling for urban sustainability focusing on changes in cropland under rapid urbanization: A case study of Hangzhou from 1990 to 2035 // Science of the Total Environment. – 2019. – Vol. 661. – P. 422-431.

84 Chen G., Zhang X., Wang Q. et al. Symmetrical dense-shortcut deep fully convolutional networks for semantic segmentation of very-high-resolution remote sensing images // IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. – 2018. – Vol. 11. – P. 1633-1644.

85 Li A., Jiao L., Zhu H. et al. Multitask semantic boundary awareness network for remote sensing image segmentation // IEEE Trans. Geosci. Remote Sens. – 2022. – Vol. 60. – P. 5400314.

86 Xoxo S., Mantel S., De Vos A. et al. Towards SDG 15.3: The biome context as the appropriate degradation monitoring dimension // Environ. Sci. Policy. – 2022. – Vol. 136. – P. 400-412.

87 Xie H. et al. A bibliometric analysis on land degradation: Current status, development, and future directions // Land. – 2020. – Vol. 9. – P. 28-1-28-37.

88 Hagen A. Fuzzy set approach to assessing similarity of categorical maps // International Journal of Geographical Information Science. – 2003. – Vol. 17, Issue 3. – P. 235-249.

89 Mbow C., Brandt M., Ouedraogo I. et al. What four decades of Earth observation tell us about land degradation in the Sahel? // Remote Sens. – 2015. – Vol. 7. – P. 4048-4067.

- 90 Degife A.W., Zabel F., Mauser W. Assessing land use and land cover changes and agricultural farmland expansions in Gambella Region, Ethiopia, using Landsat 5 and Sentinel 2a multispectral data // *Heliyon*. – 2018. – Vol. 4. – P. e00919.
- 91 Lambin E.F., Turner B.L., Geist H.J. et al. The causes of land-use and land-cover change: Moving beyond the myths // *Global Environmental Change*. – 2001. – Vol. 11. – P. 261-269.
- 92 Liou Y.A., Nguyen K.A., Li M.H. Assessing spatiotemporal eco-environmental vulnerability by Landsat data // *Ecological Indicators*. – 2017. – Vol. 80. – P. 52-65.
- 93 Li W., Chen K., Chen H. et al. Geographical knowledge-driven representation learning for remote sensing images // *IEEE Trans. Geosci. Remote Sens.* – 2022. – Vol. 60. – P. 5405516.
- 94 Kumar P., Srivastava P.K., Tiwari P. et al. Application of GPS and GNSS technology in geosciences // In book: *GPS and GNSS Technology in Geosciences*. – Amsterdam: Elsevier, 2021. – P. 415-427.
- 95 Дейвис Ш., Ландгребе Д.А., Филипс Т.Л. Дистанционное зондирование: количественный подход. – М.: Недра, 1989. – 415 с.
- 96 Lillesand T.M., Kiefer R.W. *Remote Sensing and Image Interpretation*. – Ed. 6-e. – NY.: Wiley, 2006. – 804 p.
- 97 Alipbeki O., Grossul P.P. Spatio-time analysis of land use of the Korgalzhyn district of the Akmola region in the framework of the concept of sustainable development // *Материалы международной научно-теоретической конференции «Сейфуллинские чтения–19: посвященной 110-летию М. А. Гендельмана»*. – Астана, 2023. – С. 13-15.
- 98 Lambin E.F., Geist H.J., Lepers E. Dynamics of land-use and land-cover change in tropical regions // *Annual Review of Environment and Resources*. – 2003. – Vol. 28. – P. 205-241.
- 99 Chen X., Wang Y., Chen Y. et al. NDVI-based assessment of land degradation trends in Balochistan, Pakistan, and analysis of the drivers // *Remote Sens.* – 2023. – Vol. 15. – P. 2388-1-2388-17.
- 100 Liu J., Zhang Z., Xu X. et al. Spatial patterns and driving forces of land use change in China during the early 21st century // *Journal of Geographical Sciences*. – 2010. – Vol. 20. – P. 483-494.
- 101 Behera M.D., Borate S.N., Panda S.N. et al. Modelling and analyzing the watershed dynamics using Cellular Automata (CA)–Markov model – A geo-information-based approach // *Journal of Earth System Science*. – 2012. – Vol. 121. – P. 1011-1024.
- 102 Petropoulos G., Srivastava P., Piles M. et al. Earth observation-based operational estimation of soil moisture and evapotranspiration for agricultural crops in support of sustainable water management // *Sustainability*. – 2018. – Vol. 10. – P. 181-1-181-20.
- 103 Sogno P., Traidl-Hoffmann C., Kuenzer C. Earth observation data supporting non-communicable disease research: A review // *Remote Sens.* – 2020. – Vol. 12. – P. 2541-1-2541-34.

- 104 Sengani D., Ramoelo A., Archer E. A review of fusion framework using optical sensors and Synthetic Aperture Radar imagery to detect and map land degradation and sustainable land management in the semi-arid regions // *Geocarto Int.* – 2023. – Vol. 38. – P. 2278325.
- 105 Sterk G., Stoorvogel J.J. Desertification – Scientific versus political realities // *Land.* – 2020. – Vol. 9. – P. 156-1-156-18.
- 106 Zhang H., Liu M., Wang Y. et al. Automated delineation of agricultural field boundaries from Sentinel-2 images using recurrent residual U-Net // *Int. J. Appl. Earth Obs. Geoinf.* – 2021. – Vol. 105. – P. 102557.
- 107 Seitkazy M., Beisekenov N., Taukebayev O. et al. Forecasting land use dynamics in Talas District, Kazakhstan, using Landsat data and the Google Earth Engine (GEE) platform // *Sustainability.* – 2024. – Vol. 16. – P. 6144-1-6144-18.
- 108 Che T., Xiao L., Liou Y.A. Changes in glaciers and glacial lakes and the identification of dangerous glacial lakes in the PumHua River Basin, Xizang (Tibet) // *Advances in Meteorology.* – 2014. – Vol. 2014. – P. 903709.
- 109 Li W., He C., Fang J. et al. Semantic segmentation-based building footprint extraction using very high-resolution satellite images and multi-source GIS data // *Remote Sens.* – 2019. – Vol. 11. – P. 403-1-403-19.
- 110 Hao S., Zhou Y., Guo Y. A brief survey on semantic segmentation with deep learning // *Neurocomputing.* – 2020. – Vol. 406. – P. 302-321.
- 111 Pontius R.G., Millones M. Death to Kappa: Birth of quantity disagreement and allocation disagreement for accuracy assessment // *International Journal of Remote Sensing.* – 2011. – Vol. 32, Issue 15. – P. 4407-4429.
- 112 Roy D.P., Wulder M.A., Loveland T.R. et al. Landsat-8: Science and product vision for terrestrial global change research // *Remote Sens. Environ.* – 2014. – Vol. 145. – P. 154-172.
- 113 Van den Broeck W.A.J., Goedemé T., Loopmans M. Multiclass land cover mapping from historical orthophotos using domain adaptation and spatio-temporal transfer learning // *Remote Sens.* – 2022. – Vol. 14. – P. 5911-1-5911-34.
- 114 Abd El-Hamid H.T., El-Alfy M.A., Elnaggar A. Prediction of future situation of land use/cover change and modeling sensitivity to pollution in Edku Lake // *GeoJournal.* – 2020. – Vol. 86. – P. 1895-1913.
- 115 Guo Y., Liu Y., Oerlemans A. et al. Deep learning for visual understanding: A review // *Neurocomputing.* – 2016. – Vol. 187. – P. 27-48.
- 116 Levy R.C., Remer L.A., Dubovik O. Global aerosol optical properties and application to Moderate Resolution Imaging Spectroradiometer aerosol retrieval over land // *J. Geophys. Res.-Atmos.* – 2007. – Vol. 112. – P. D13210.
- 117 Li T., Chen Y.-Z., Han L.-J. et al. Shortened duration and reduced area of frozen soil in the Northern Hemisphere // *Innovation.* – 2021. – Vol. 2. – P. 100146.
- 118 Hua L., Tang L., Cui S. et al. Simulating urban growth using the SLEUTH model in a coastal peri-urban district in China // *Sustainability.* – 2014. – Vol. 6. – P. 3899-3914.
- 119 Shi W., Qin W., Chen A. Towards robust semantic segmentation of land covers in foggy conditions // *Remote Sens.* – 2022. – Vol. 14. – P. 4551-1-4551-22.

- 120 Houghton R.A. The annual net flux of carbon to the atmosphere from changes in land use 1850-1990 // *Tellus B: Chemical and Physical Meteorology*. – 1999. – Vol. 51, Issue 2. – P. 298-313.
- 121 Sepuru T.K., Dube T. An appraisal on the progress of remote sensing applications in soil erosion mapping and monitoring // *Remote Sens. Appl. Soc. Environ.* – 2018. – Vol. 9. – P. 1-9.
- 122 Mayr S., Kuenzer C., Gessner U. et al. Validation of Earth observation time-series: A review for large-area and temporally dense land surface products // *Remote Sens.* – 2019. – Vol. 11. – P. 2616-1-2616-34.
- 123 Rogan J., Franklin J., Stow D. et al. Mapping land-cover modifications over large areas: A comparison of machine learning algorithms // *Remote Sensing of Environment*. – 2008. – Vol. 112, Issue 5. – P. 2272-2283.
- 124 Sirous A., Satari M., Shahraki M.M. et al. A conditional generative adversarial network for urban area classification using multi-source data // *Earth Sci. Inf.* – 2023. – Vol. 16. – P. 2529-2543.
- 125 Ustin S.L., Middleton E.M. Current and near-term advances in Earth observation for ecological applications // *Ecol. Process.* – 2021. – Vol. 10. – P. 1-57.
- 126 Wang L., Li R., Zhang C. et al. UnetFormer: A Unet-like transformer for efficient semantic segmentation of remote sensing urban scene imagery // *ISPRS J. Photogramm. Remote Sens.* – 2022. – Vol. 190. – P. 196-214.
- 127 Maxwell A.E., Warner T.A., Fang F. Implementation of machine-learning classification in remote sensing: An applied review // *International Journal of Remote Sensing*. – 2018. – Vol. 39, Issue 9. – P. 2784-2817.
- 128 Zhang Z., Huang X., Li J. Dwin-HRFormer: A high-resolution transformer model with directional windows for semantic segmentation of urban construction land // *IEEE Trans. Geosci. Remote Sens.* – 2023. – Vol. 61. – P. 5400714.
- 129 Rimal B., Zhang L., Keshtkar H. et al. Land use/land cover dynamics and modeling of urban land expansion by the integration of Cellular Automata and Markov chain // *ISPRS International Journal of Geo-Information*. – 2018. – Vol. 7. – P. 154-1-154-21.
- 130 Li Z., Zhang H., Lu F. et al. Breaking the resolution barrier: A low-to-high network for large-scale high-resolution land-cover mapping using low-resolution labels // *ISPRS J. Photogramm. Remote Sens.* – 2022. – Vol. 192. – P. 244-267.
- 131 Chen H., Pontius R.G. Diagnostic tools to evaluate a spatial land change projection along a gradient of an explanatory variable // *Landscape Ecology*. – 2010. – Vol. 25, Issue 9. – P. 1319-1331.
- 132 Justice C.O. et al. The Moderate Resolution Imaging Spectroradiometer (MODIS): Land remote sensing for global change research // *IEEE Trans. Geosci. Remote Sens.* – 1998. – Vol. 36. – P. 1228-1249.
- 133 Liu Q., Kampffmeyer M., Jenssen R. et al. Dense dilated convolutions merging network for land cover classification // *IEEE Trans. Geosci. Remote Sens.* – 2020. – Vol. 58. – P. 6309-6320.

- 134 Pandey B.K., Khare D. Analyzing and modeling of a large river basin dynamics applying integrated Cellular Automata and Markov model // *Environmental Earth Sciences*. – 2017. – Vol. 76. – P. 779.
- 135 United Nations. Political Declaration of the High-Level Political Forum on Sustainable Development Convened under the Auspices of the General Assembly: https://data.opendevlopmentmekong.net/laws_record/political. 10.09.2025.
- 136 Jensen J.R. Introductory Digital Image Processing: A Remote Sensing Perspective. – Ed. 3-e. – Upper Saddle River, NJ, 2005. – 544 p.
- 137 Thien B.B., Phuong V.T., Huong D.T.V. Forest cover depletion and land use/land cover change in Bac Kan province, Vietnam // *GeoFocus. International Review of Geographical Information Science and Technology*. – 2024. – Vol. 34. – P. 29-44.
- 138 Алипбеки О.А., Гроссул П.П. Применение перспективных методик в определении открытости и доступности пространственных данных // *Материалы Первой международной научно-практической конференции «Цифровая реальность: новые вызовы в картографии, ГИС и ДЗЗ»*. – Алматы, 2024. – С. 169-174.
- 139 Guo H.D., Liu Z., Jiang H. et al. Big Earth Data: A new challenge and opportunity for Digital Earth's development // *Int. J. Digit. Earth*. – 2017. – Vol. 10. – P. 1-12.
- 140 Shan L., Wang W. DenseNet-based land cover classification network with deep fusion // *IEEE Geosci. Remote Sens. Lett.* – 2022. – Vol. 19. – P. 6001705.
- 141 Simms D.M. Fully convolutional neural nets in-the-wild // *Remote Sens. Lett.* – 2020. – Vol. 11. – P. 1080-1089.
- 142 Ronneberger O., Fischer P., Brox T. U-Net: Convolutional networks for biomedical image segmentation // *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*: col. of scient. pap. – Munich, 2015. – P. 234-241.
- 143 Anderson K., Ryan B., Sonntag W. et al. Earth observation in service of the 2030 Agenda for Sustainable Development // *Geo-Spat. Inf. Sci.* – 2017. – Vol. 20. – P. 77-96.
- 144 Liu W., Liu J., Luo Z. et al. Weakly supervised high spatial resolution land cover mapping based on self-training with weighted pseudo-labels // *Int. J. Appl. Earth Obs. Geoinf.* – 2022. – Vol. 112. – P. 102931.
- 145 Jia W., Gu X., Mi X. et al. Multi-scale spatiotemporal pattern analysis and simulation (MSPAS) model with driving factors for land cover change and Sustainable Development Goals: A case study of Nepal // *Remote Sens.* – 2022. – Vol. 14. – P. 6295-1-6295-26.
- 146 Gardner R.H., Urban D.L. Model validation and testing: Past lessons, present concerns, future prospects // In book: *Models in Ecosystem Science*. – Princeton, 2003. – P. 184-203.

ПРИЛОЖЕНИЕ А

Акты внедрения

УТВЕРЖДАЮ

Член Правления-проректор
по научной и инновационной
деятельности
НАО «Казахский агротехнический
исследовательский университет
имени С.Сейфуллина»
Сыргалиев Е.О.

«__» _____ 2025 г.



АКТ

внедрения результатов научно-исследовательской работы (НИР) в учебный процесс

Мы нижеподписавшиеся, декан факультета Управления земельными ресурсами, архитектуры и дизайна Толеубекова Ж.З., председатель Совета по академическому качеству, факультета УЗРАиД, старший преподаватель кафедры «Кадастр» Жагипарова Т.Т., заведующий кафедрой «Геодезия и картография» Кусаинова Г.Д. составили настоящий акт о том, что результаты НИР докторанта ОП «8D07302 - Геодезия» Гроссула П.П. «Геодезические методы анализа землепользования с применением машинного и глубокого обучения» (учебное пособие) внедрены и используются в лекционных курсах и практических занятиях по следующим специальным дисциплинам:

- «Дистанционное зондирование Земли»;
- «Фотограмметрия и дистанционное зондирование»;
- «Использование БПЛА в разных отраслях экономики»;
- «Геоинформационные системы»;

а также при выполнении научно-исследовательской работы обучающихся и при подготовке выпускных квалификационных работ по образовательной программе «6B07302 - Геодезия и картография».

Декан факультета УЗРАиД

Председатель Совета
по АК факультета УЗРАиД

Заведующая кафедрой
«Геодезии и картографии»



Ж.З. Толеубекова

Т.Т. Жагипарова

Г.Д. Кусаинова

УТВЕРЖДАЮ

Член Правления-проректор
по научной и инновационной
деятельности
НАО «Казахский агротехнический
исследовательский университет
имени С.Сейфуллина»
Сыргалиев Е.О.

«__» _____



АКТ

**внедрения результатов научно-исследовательской работы (НИР)
в учебный процесс**

Мы нижеподписавшиеся, декан факультета Управления земельными ресурсами, архитектуры и дизайна Толеубекова Ж.З., председатель Совета по академическому качеству, факультета УЗРАиД, старший преподаватель кафедры «Кадастр» Жагипарова Т.Т., заведующий кафедрой «Геодезия и картография» Кусаинова Г.Д. составили настоящий акт о том, что результаты НИР докторанта ОП «8D07302 - Геодезия» Гроссула П.П. «Геодезические методы анализа землепользования с применением машинного и глубокого обучения» (учебное пособие) внедрены и используются в лекционных курсах и практических занятиях по следующим специальным дисциплинам:

- «Дистанционное зондирование Земли»;
- «Фотограмметрия и дистанционное зондирование»;
- «Использование БПЛА в разных отраслях экономики»;
- «Геоинформационные системы»;

а также при выполнении научно-исследовательской работы обучающихся и при подготовке выпускных квалификационных работ по образовательной программе «6B07302 - Геодезия и картография».

Декан факультета УЗРАиД

Председатель Совета
по АК факультета УЗРАиД

Заведующая кафедрой
«Геодезии и картографии»



Ж.З. Толеубекова

Т.Т. Жагипарова

Г.Д. Кусаинова

ПРИЛОЖЕНИЕ Б

Программный код для GEE

```
// Определяем регион интереса (ROI)
var roi = ee.table;

// Временные интервалы
var startDate = '2020-05-01';
var endDate = '2020-09-30';

// Функция маскирования облаков для Landsat 8 Collection 2 TOA
// Reflectance данных
function maskCloudsL8(image) {
  var qa = image.select('QA_PIXEL');
  // Биты для определения облаков (в документации: bit 3 и 5 для облаков
  // и облачной тени)
  var cloudMask = qa.bitwiseAnd(1 << 3).eq(0)
    .and(qa.bitwiseAnd(1 << 5).eq(0));
  return image.updateMask(cloudMask);
}

// Загрузка коллекции Landsat 8
var l8_col = ee.ImageCollection("LANDSAT/LC08/C02/T1_L2")
  .filterDate(startDate, endDate)
  .filterBounds(roi)
  .map(function(image){
    var rad = image.multiply(0.0000275).add(-0.2);
    return rad.copyProperties(image, image.propertyNames());
  })
  .map(maskCloudsL8);

// Аналогично Landsat 7
var l7_col = ee.ImageCollection("LANDSAT/LE07/C02/T1_L2")
  .filterDate(startDate, endDate)
  .filterBounds(roi)
  .map(function(image){
    var rad = image.multiply(0.0000275).add(-0.2);
    return rad.copyProperties(image, image.propertyNames());
  })
  .map(maskCloudsL8);

// Объединяем Landsat 7 и Landsat 8 коллекции
var merged_col = l8_col.merge(l7_col);

// Создаём медианный композит
var medianComposite = merged_col.median().clip(roi);

// Расчёт NDVI для медианного композита:
// Для Landsat 8: NIR = B5, RED = B4
```

```

// Для Landsat 7: NIR = B4, RED = B3
var ndvi = medianComposite.expression(
  '(NIR - RED) / (NIR + RED)', {
    'NIR': medianComposite.select('B5'), // Для Landsat 8
    'RED': medianComposite.select('B4')
  });

// Аналогично NDWI:
var ndwi = medianComposite.expression(
  '(GREEN - NIR) / (GREEN + NIR)', {
    'GREEN': medianComposite.select('B3'), // Для L8 - B3 это Green
    'NIR': medianComposite.select('B5')
  });

// EVI = 2.5 * ((NIR - RED)/(NIR + 6*RED - 7.5*BLUE + 1))
var evi = medianComposite.expression(
  '2.5 * ((NIR - RED) / (NIR + 6 * RED - 7.5 * BLUE + 1))', {
    'NIR': medianComposite.select('B5'),
    'RED': medianComposite.select('B4'),
    'BLUE': medianComposite.select('B2')
  });

// Интеграция данных SRTM:
var srtm = ee.Image("USGS/SRTMGL1_003").clip(roi);
var elevation = srtm.select('elevation');

// Добавление данных ночной светимости (VIIRS):
var viirs = ee.ImageCollection("NOAA/VIIRS/DNB/MONTHLY_V1/VCMSFG")
  .filterDate('2020-01-01', '2020-12-31')
  .filterBounds(roi)
  .median()
  .clip(roi);

// Объединение в общий стек:
var stacked = medianComposite.addBands(ndvi.rename('NDVI'))
  .addBands(ndwi.rename('NDWI'))
  .addBands(evi.rename('EVI'))
  .addBands(elevation.rename('ELEVATION'))
  .addBands(viirs.select('avg_rad').rename('VIIRS_LIGHT'));

// Теперь у нас есть многоуровневый набор данных, пригодный для
последующей классификации и анализа
Map.centerObject(roi, 7);
Map.addLayer(ndvi, {min: -1, max: 1, [81]ette:
['blue','white','green']}, 'NDVI');
Map.addLayer(stacked, {}, 'Stacked bands');

// Подготовка эталонных данных
// ee.FeatureCollection с эталонными точками `trainingPoints`
// Каждый объект имеет поле 'class' - целевую метку LULC

```

```

//                                     trainingPoints
ee.FeatureCollection("users/username/training_data")

// Извлекаем пиксельные значения в точках
var trainingData = finalComposite.sampleRegions({
  collection: trainingPoints,
  properties: ['class'],
  scale: 30,      // Масштаб Landsat
  tileScale: 2
});

// Разделение на обучающую и тестовую выборки
// Перемешиваем точки
trainingData = trainingData.randomColumn('random');

// 70% для обучения
var split = 0.7;
var trainingSet = trainingData.filter(ee.Filter.lt('random', split));
var testSet = trainingData.filter(ee.Filter.gte('random', split));

// Настройка и тренировка Random Forest
var rfClassifier = ee.Classifier.smileRandomForest({
  numberOfTrees: 300,
  variablesPerSplit: null, // Использование sqrt(число признаков)
  minLeafPopulation: 1,
  maxDepth: 20,
  seed: 42
}).train({
  features: trainingSet,
  classProperty: 'class',
  inputProperties: finalComposite.bandNames()
});

// Оценка модели на тестовой выборке
var classified = testSet.classify(rfClassifier);
var confusionMatrix = classified.errorMatrix('class',
'classification');
print('Confusion Matrix:', confusionMatrix);
print('Overall Accuracy:', confusionMatrix.accuracy());
print('Kappa:', confusionMatrix.kappa());

// Применение модели ко всему композиту
var classifiedImage = finalComposite.classify(rfClassifier);

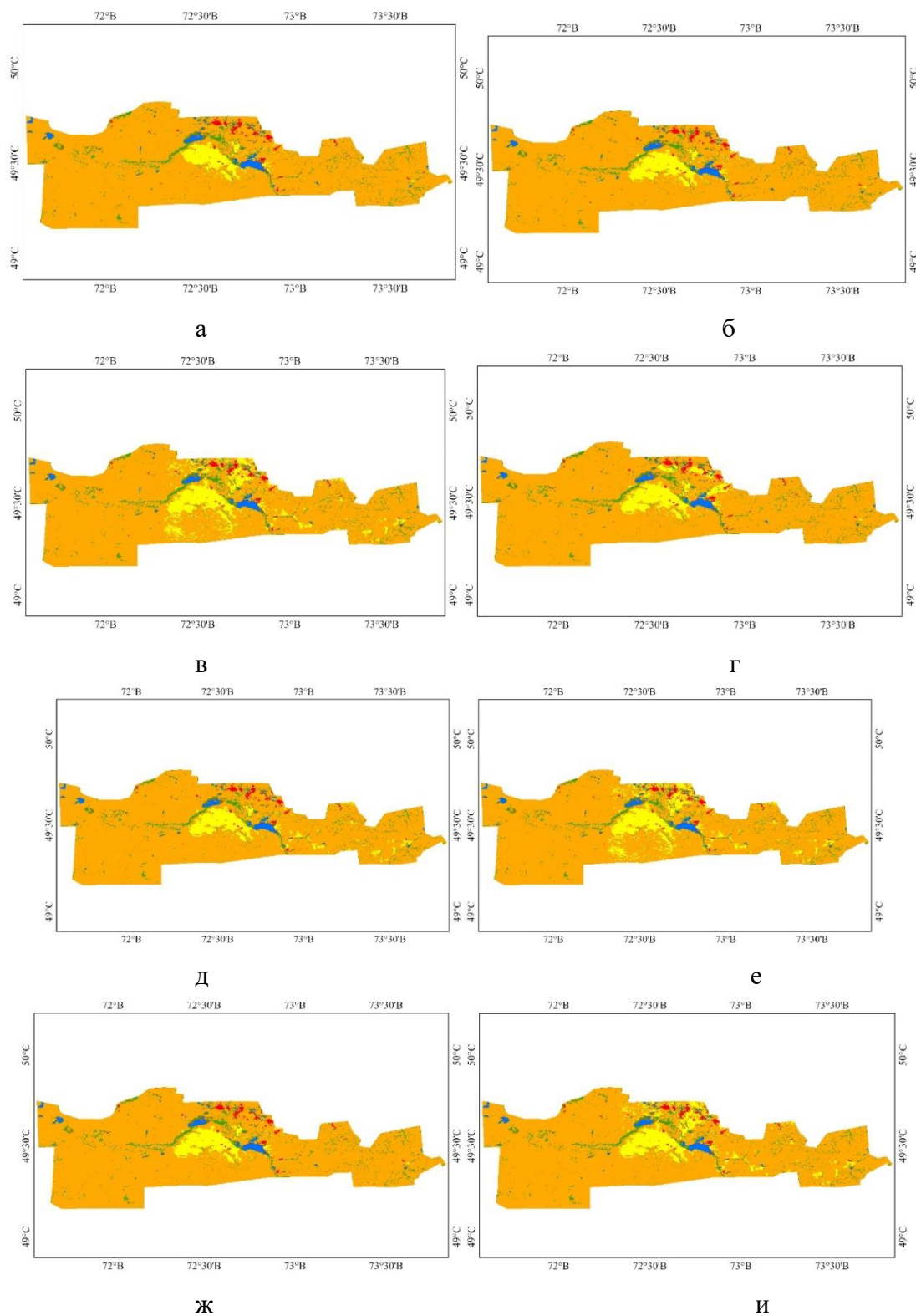
// Визуализация классифицированного изображения
Map.centerObject(roi, 8);
Map.addLayer(classifiedImage, {
  min: 0, max: 5, // Количество классов
  [81]ette: ['yellow', 'green', 'blue', 'brown', 'red', 'purple']
}, 'Classified LULC');

```

```
// Экспорт классифицированного изображения
Export.image.toDrive({
  image: classifiedImage,
  description: 'LULC_Classification',
  scale: 30,
  region: roi,
  fileFormat: 'GeoTIFF',
  maxPixels: 1e13
});
```

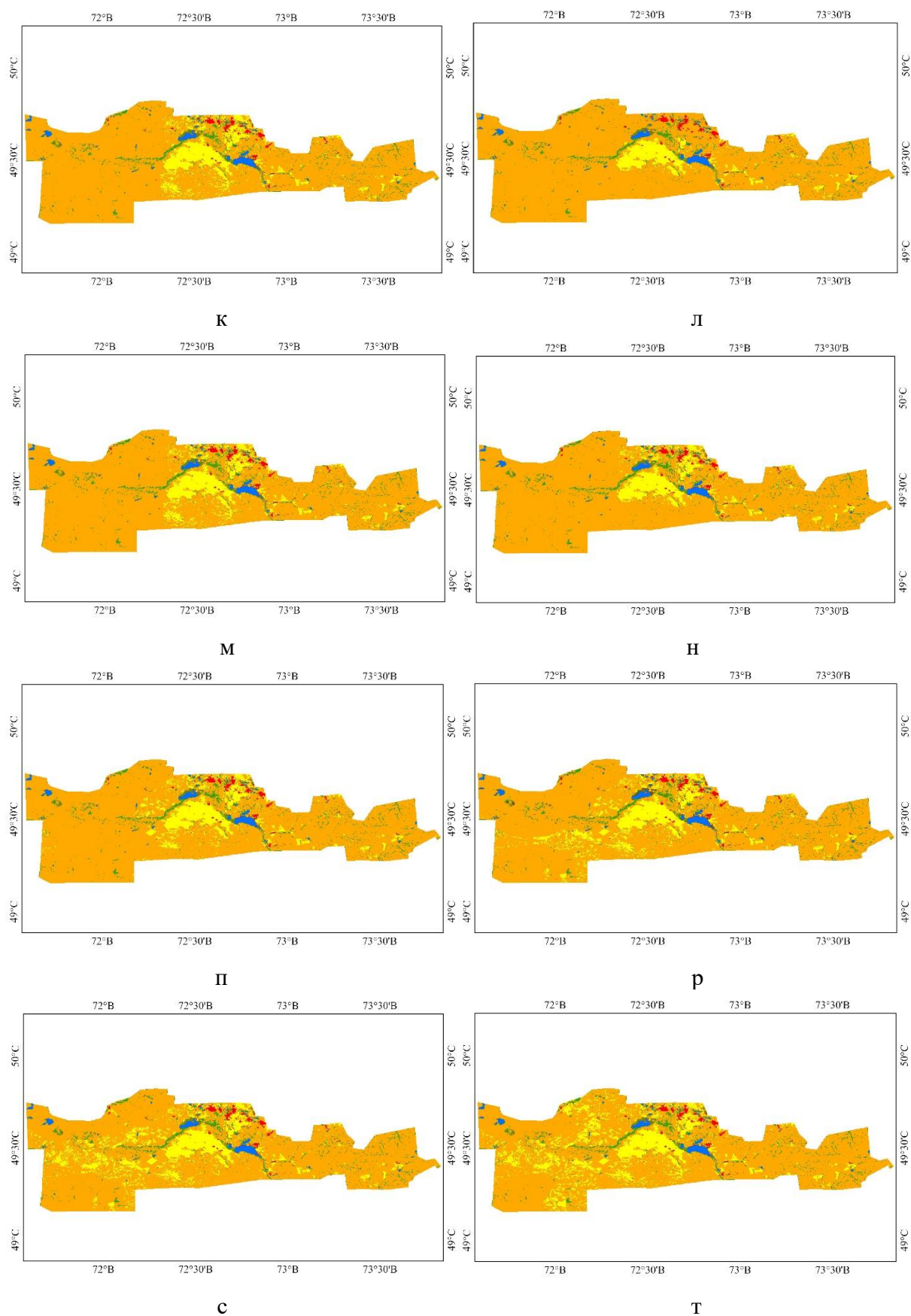
ПРИЛОЖЕНИЕ В

Карты LULC по 4 районам Карагандинской области



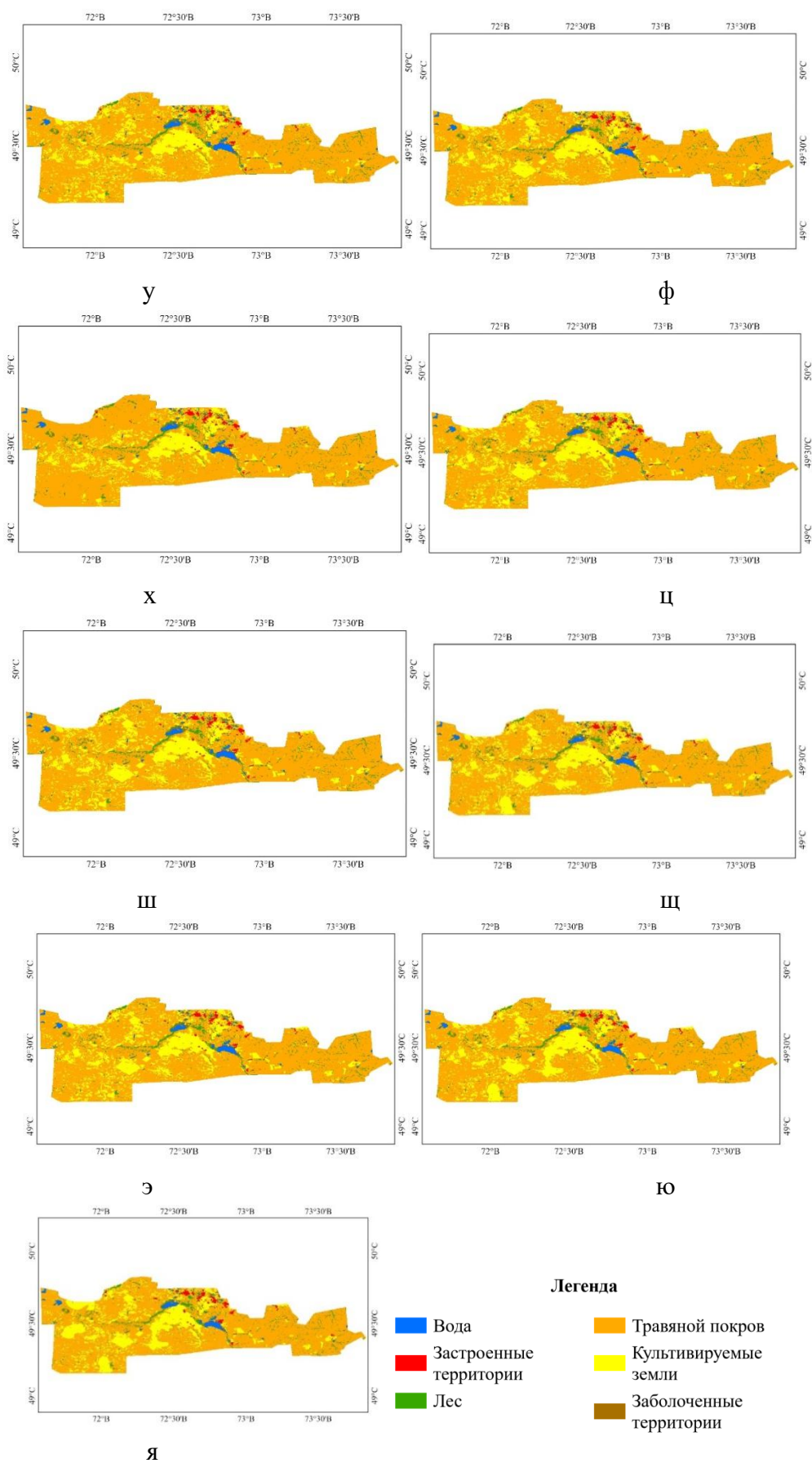
а – 2000; б – 2001; в – 2002; г – 2003; д – 2004; е – 2005; ж – 2006; и – 2007

Рисунок В.1 – Карты LULC по 4 районам Карагандинской области в период с 2000 по 2024 годы, лист 1



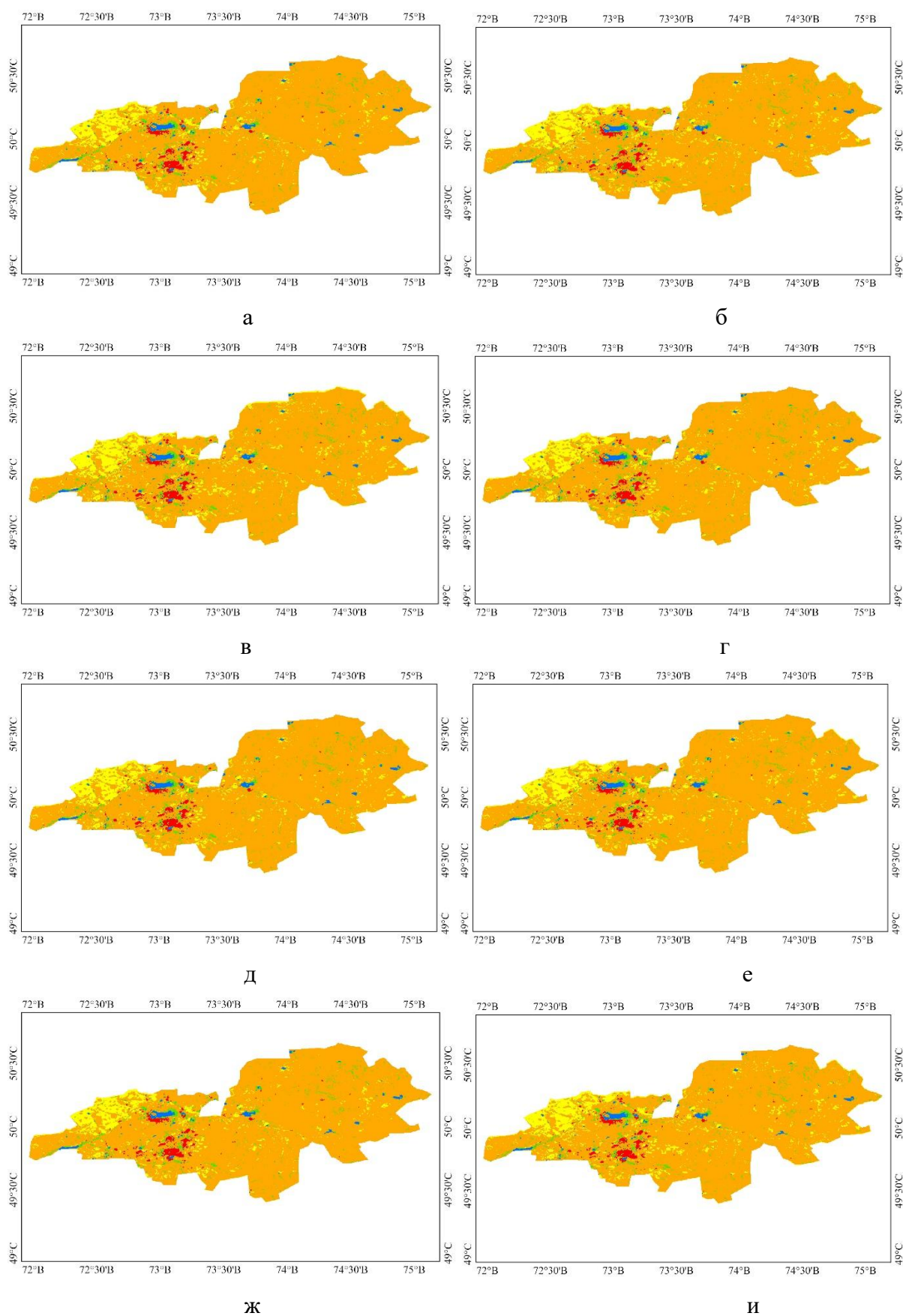
к – 2008; л – 2009; м – 2010; н – 2011; п – 2012; р – 2013; с – 2014; т – 2015

Рисунок В.1, лист 2



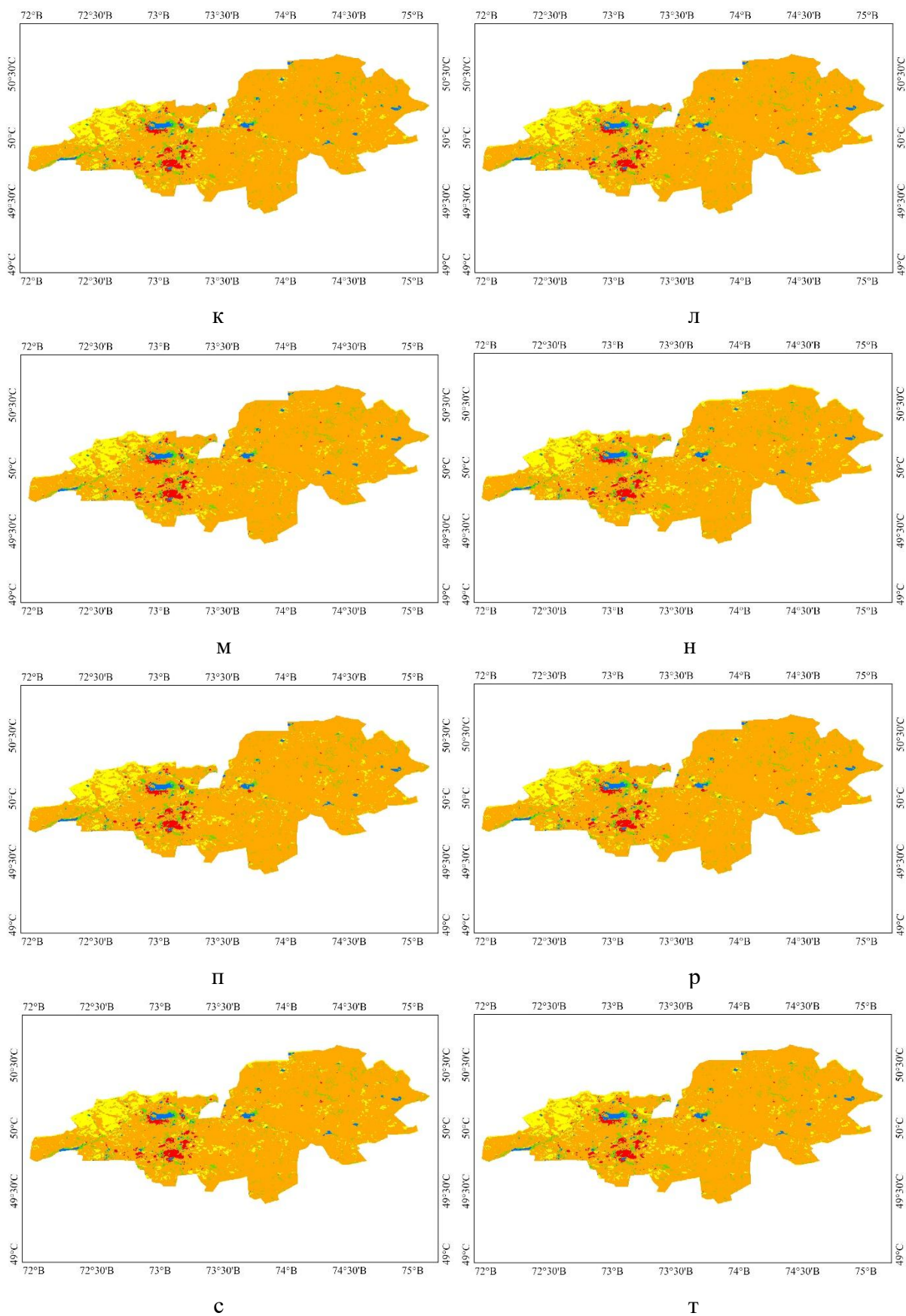
у – 2016; ф – 2017; х – 2018; ц – 2019; ш – 2020; щ – 2021; э – 2022; ю – 2023; я – 2024

Рисунок В.1, лист 3



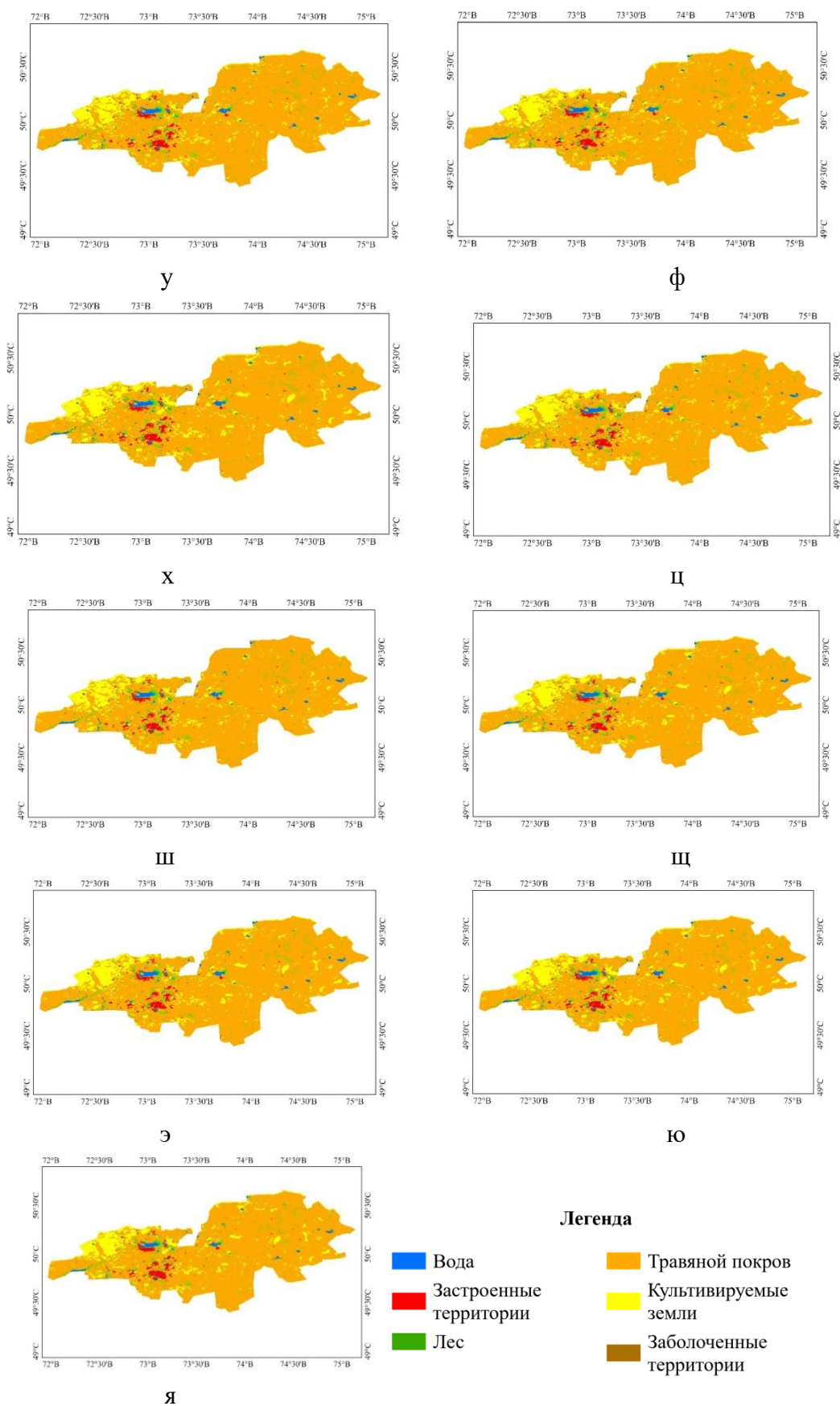
а – 2000; б – 2001; в – 2002; г – 2003; д – 2004; е – 2005; ж – 2006; и – 2007

Рисунок В.2 – Бухар-Жырауский район, лист 1



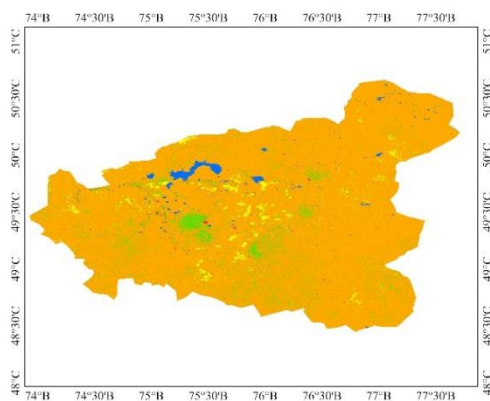
к – 2008; л – 2009; м – 2010; н – 2011; п – 2012; р – 2013; с – 2014; т – 2015

Рисунок В.2, лист 2

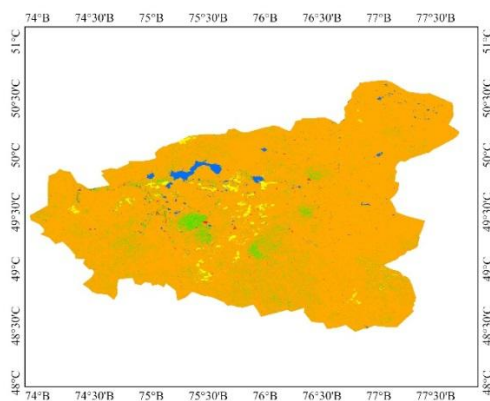


у – 2016; ф – 2017; х – 2018; ц – 2019; ш – 2020; щ – 2021; э – 2022; ю – 2023; я – 2024

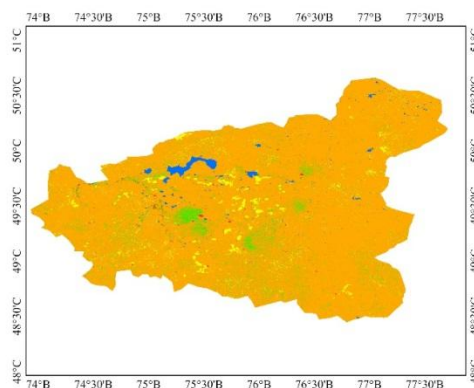
Рисунок В2, лист 3



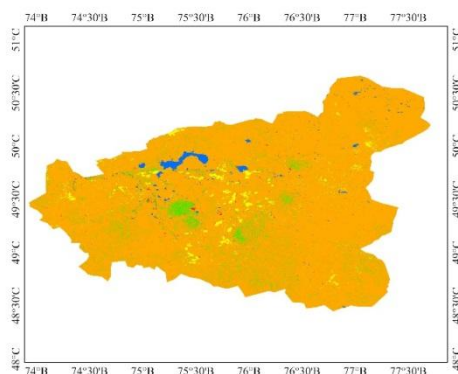
а



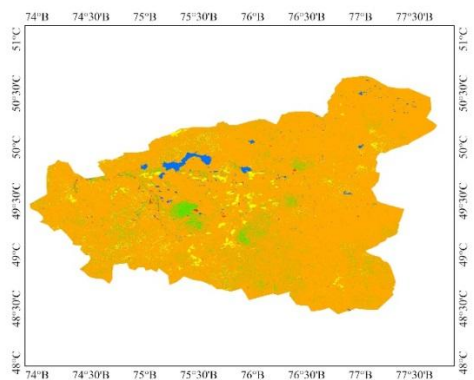
б



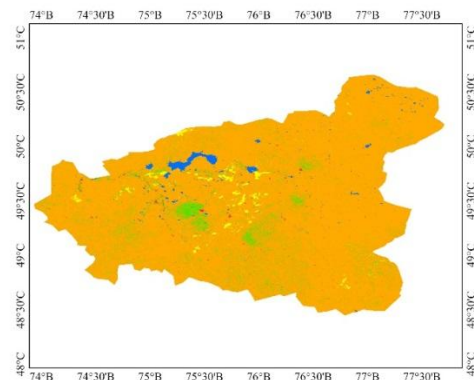
в



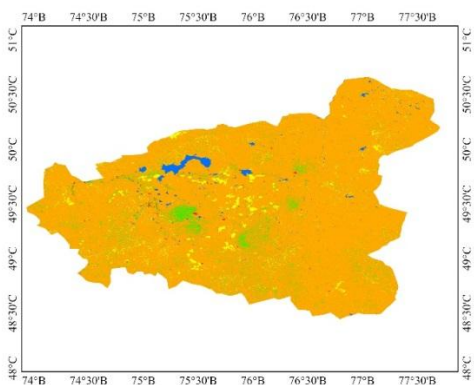
г



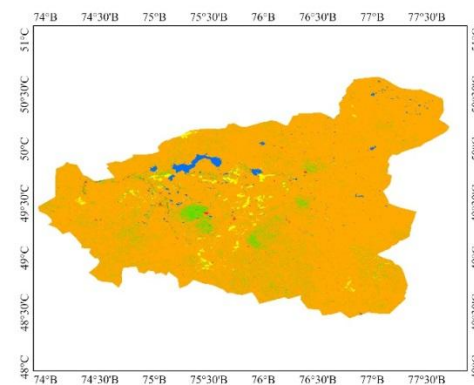
д



е



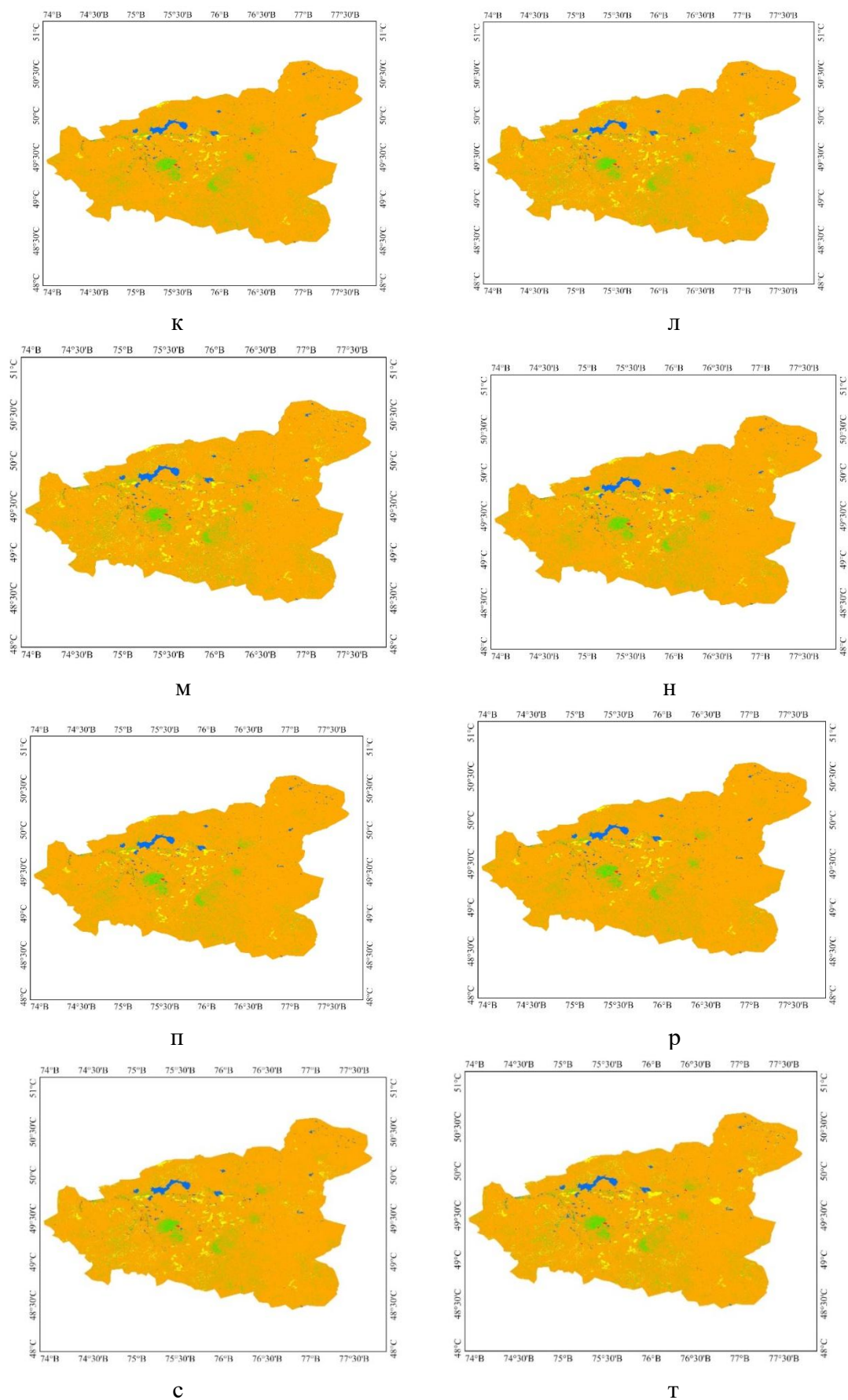
ж



и

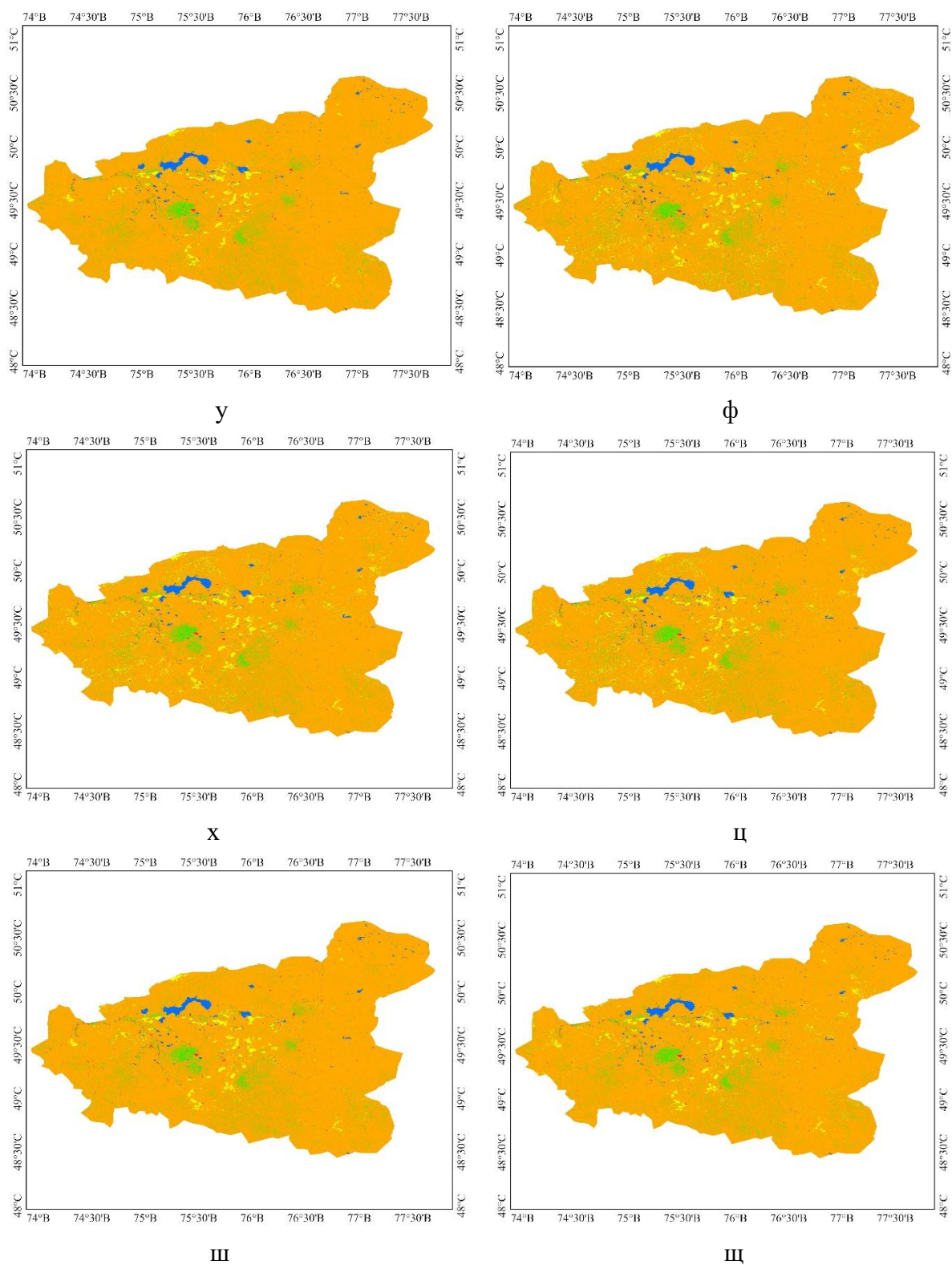
а – 2000; б – 2001; в – 2002; г – 2003; д – 2004; е – 2005; ж – 2006; и – 2007

Рисунок В.3 – Каркаралинский район, лист 1



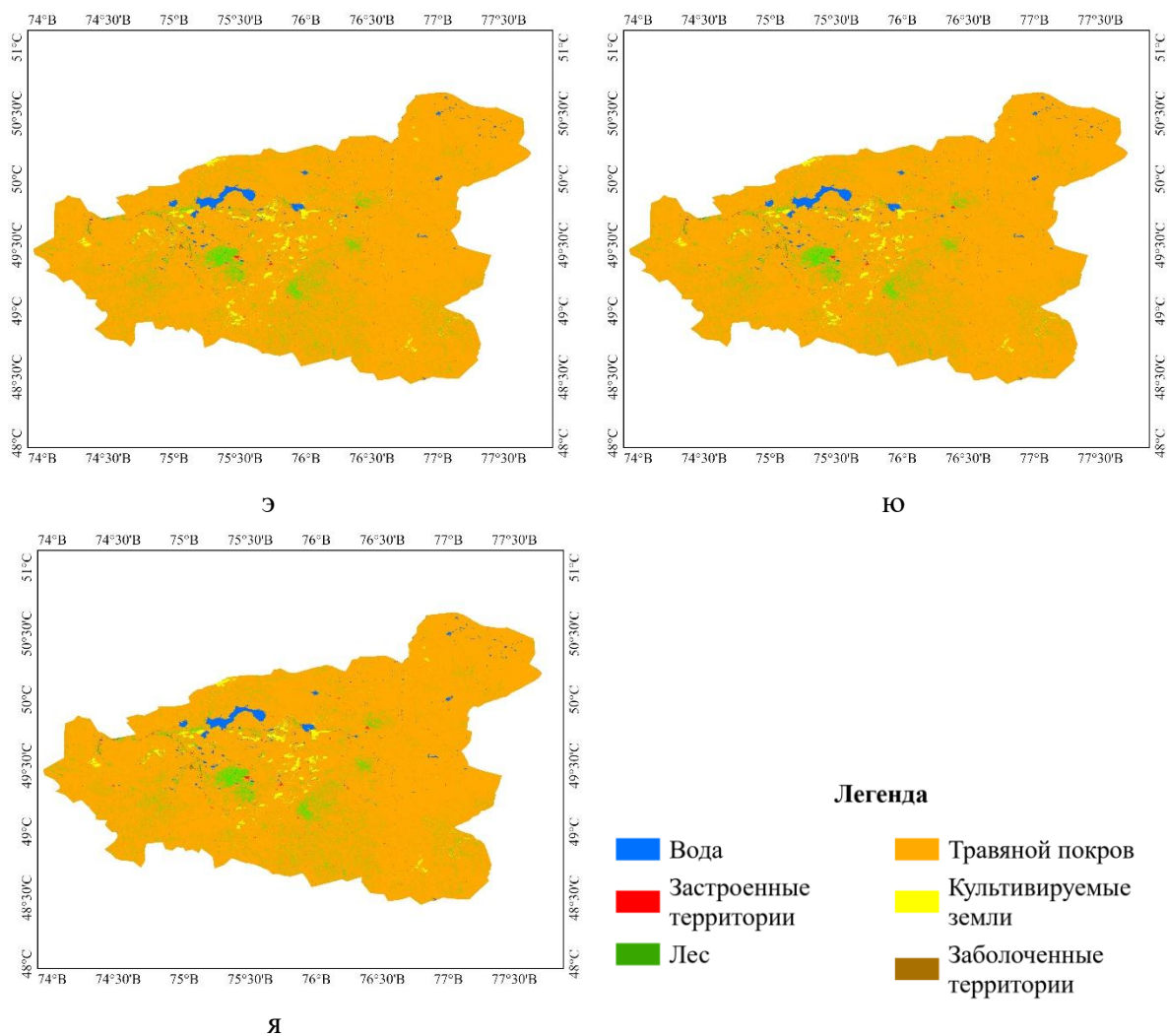
к – 2008; л – 2009; м – 2010; н – 2011; п – 2012; р – 2013; с – 2014; т – 2015

Рисунок В.3, лист 2



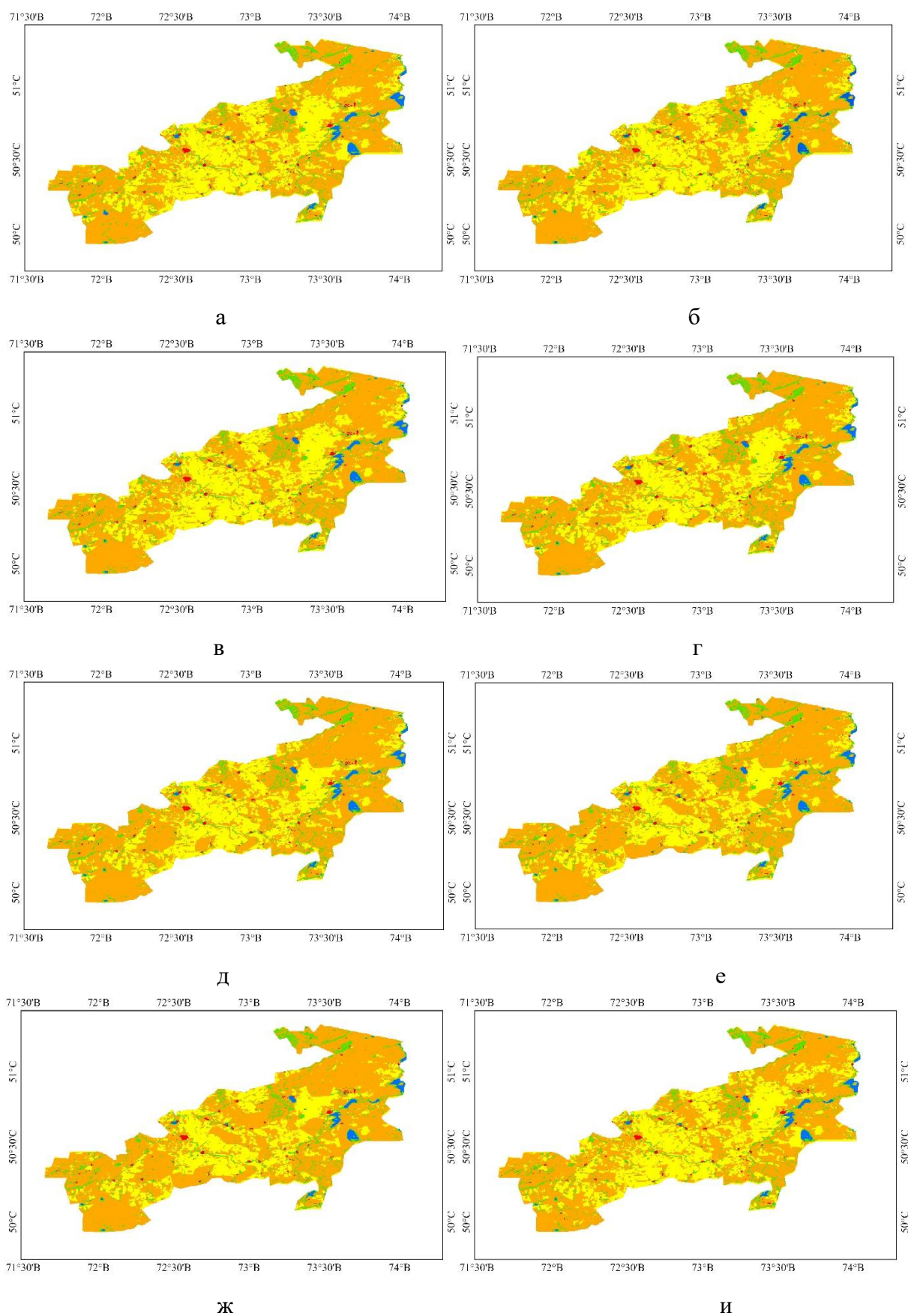
у – 2016; ф – 2017; х – 2018; ц – 2019; ш – 2020; щ – 2021

Рисунок В.3, лист 3



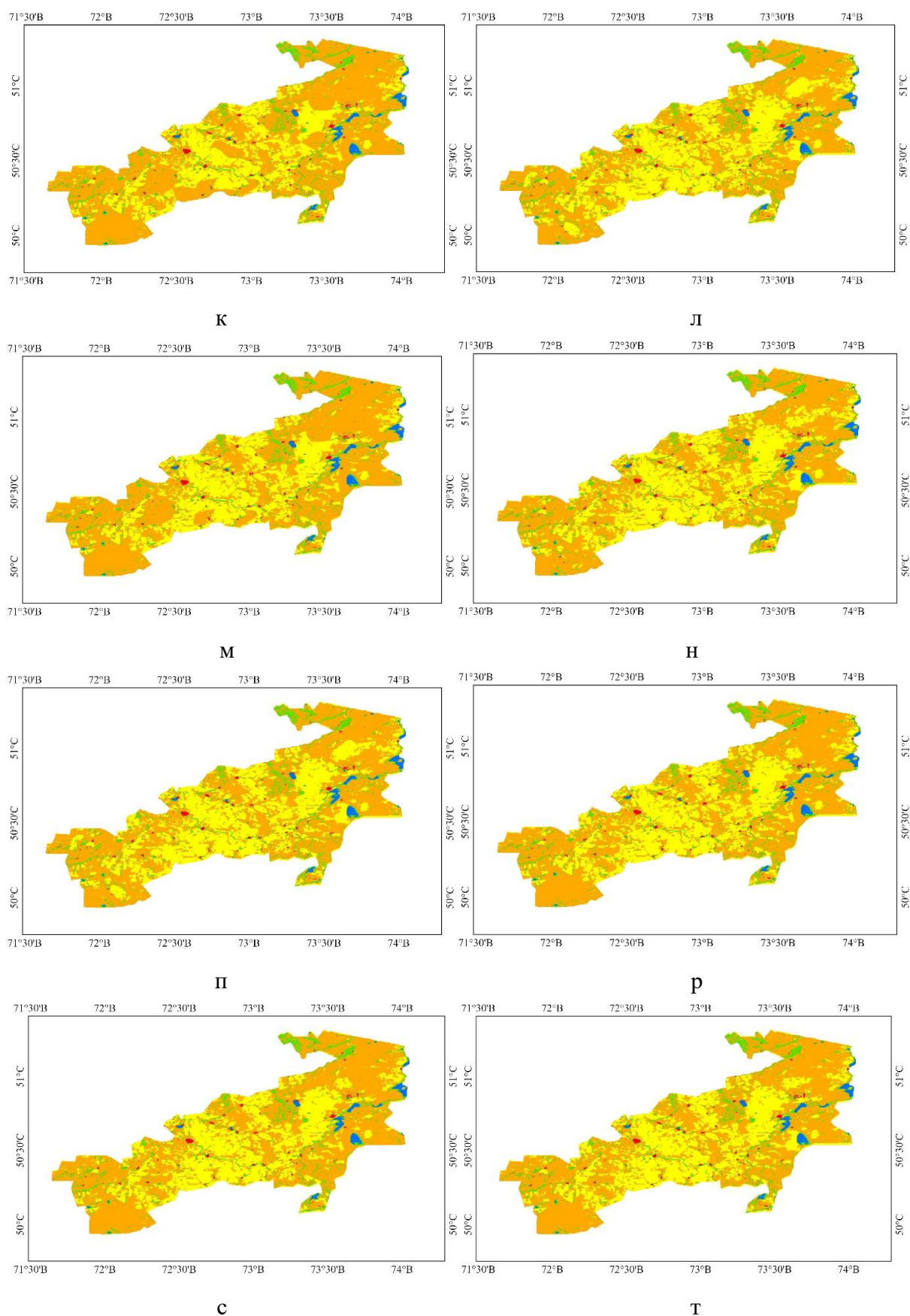
э – 2022; ю – 2023; я – 2024

Рисунок В.3, лист 4



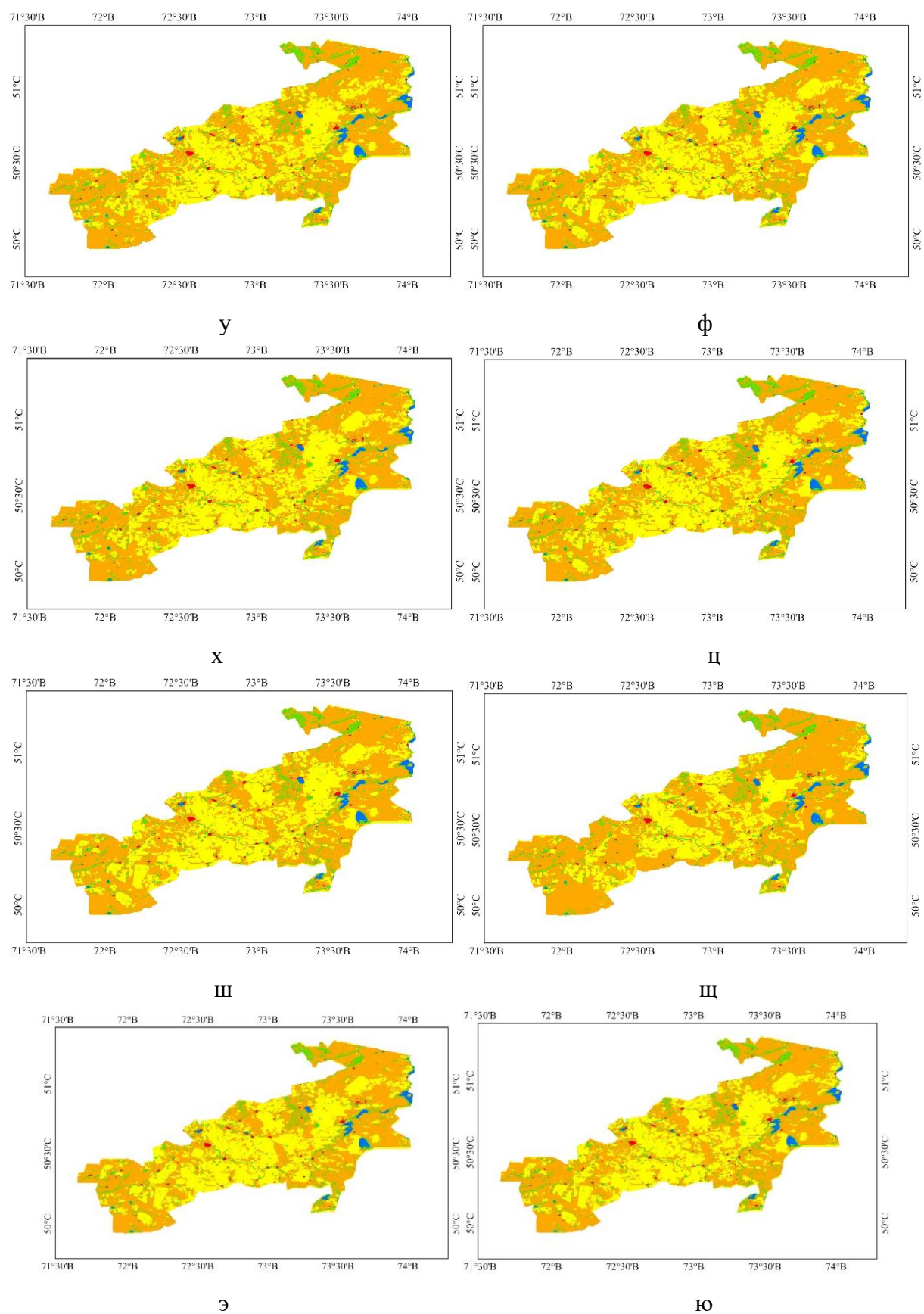
а – 2000; б – 2001; в – 2002; г – 2003; д – 2004; е – 2005; ж – 2006; и – 2007

Рисунок В.4 – Осакаровский район, лист 1



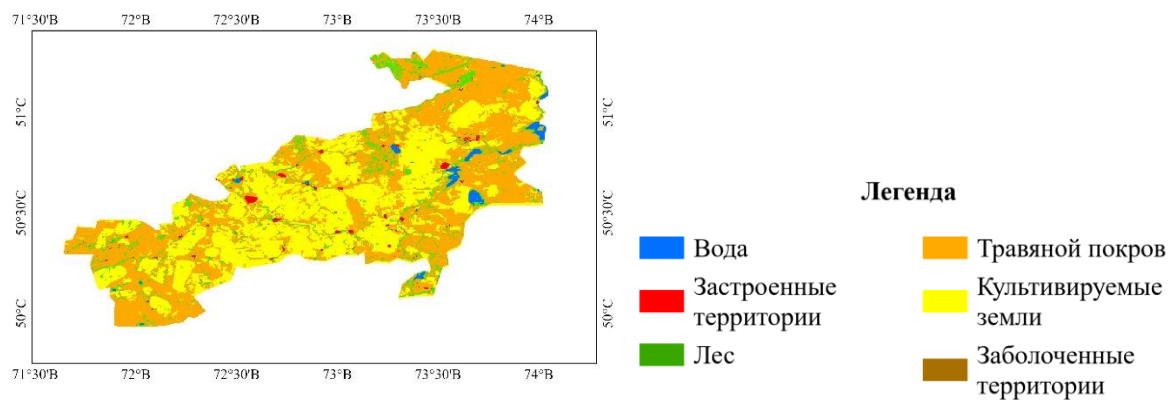
к – 2008; л – 2009; м – 2010; н – 2011; п – 2012; р – 2013; с – 2014; т – 2015

Рисунок В.4, лист 2



у – 2016; ф – 2017; х – 2018; ц – 2019; ш – 2020; щ – 2021; э – 2022; ю – 2023

Рисунок В.4, лист 3



я

я – 2024

Рисунок В.4, лист 4