

УДК: 528.92

DOI: 10.35595/2414-9179-2022-1-28-613-628

А.А. Ямашкин¹, С.А. Ямашкин²

АНАЛИЗ МЕТАГЕОСИСТЕМ ПОЛИГОНА «ИНЕРКА» ПОСРЕДСТВОМ АНСАМБЛЕЙ МОДЕЛЕЙ МАШИННОГО ОБУЧЕНИЯ

АННОТАЦИЯ

В статье дана характеристика геоинформационных методов и алгоритмов интерпретации данных дистанционного зондирования Земли, основанных на формировании ансамбля классификаторов на базе методологии Ensemble Learning. Предлагаемое решение может быть использовано с целью оценки устойчивости геосистем и прогнозирования стихийных процессов. Отличие создаваемого подхода определяется принципиально новой схемой организации метаклассификатора, как блока принимающего решение, а также использованием геосистемного подхода к подготовке данных для автоматизированного анализа посредством моделей глубокого обучения. В статье показано, что использование ансамблей, выстраиваемых по предложенной методике, позволяет проводить оперативный автоматизированный анализ пространственных данных для решения задачи тематического картографирования метагеосистем и природных процессов. При этом объединение моделей в ансамбль на основе предложенной архитектуры метаклассификатора позволяет повысить устойчивость анализирующей системы: точность решений, принимаемых ансамблем, стремится к точности наиболее эффективного моноклассификатора системы. Интеграция отдельных классификаторов в ансамбли позволяет подойти к решению научной проблемы поиска гиперпараметров классификаторов за счет комбинированного использования моделей одного типа с разными конфигурациями, при этом формирование метаклассификатора по предложенному алгоритму представляет собой возможность добавить элемент предсказуемости и контроля в использование нейронных сетей, традиционно представляющих собой «черный ящик». Картографирование геосистем тестового полигона «Инерка» показывает их слабую устойчивость к рекреационному освоению. Основными лимитирующими факторами являются состав четвертичных отложений, характер рельефа, механический состав и влажность почвы, генезис и состав растительности.

КЛЮЧЕВЫЕ СЛОВА: ансамбли, машинное обучение, метагеосистемы, картография, устойчивое развитие, ГИС

¹ ФГБОУ ВО «Национальный исследовательский Мордовский государственный университет им. Н. П. Огарева», ул. Большевикская, д. 68, 430005, Саранск, Россия, *e-mail*: yamashkinsa@mail.ru

² ФГБОУ ВО «Национальный исследовательский Мордовский государственный университет им. Н. П. Огарева», ул. Большевикская, д. 68, 430005, Саранск, Россия, *e-mail*: yamashkin56@mail.ru

Anatoliy A. Yamashkin¹, Stanislav A. Yamashkin²

ANALYSIS OF THE INERKA POLYGON METAGEOSYSTEMS BY MEANS OF ENSEMBLES OF MACHINE LEARNING MODELS

ABSTRACT

The article describes a geoinformation algorithm for interpreting Earth remote sensing data based on the Ensemble Learning methodology. The proposed solution can be used to assess the stability of geosystems and predict natural (including exogeodynamic) processes. The difference of the created approach is determined by a fundamentally new organization scheme of the metaclassifier as a decision-making unit, as well as the use of a geosystem approach to preparing data for automated analysis using deep neural network models. The article shows that the use of ensembles, built according to the proposed method, makes it possible to carry out an operational automated analysis of spatial data for solving the problem of thematic mapping of metageosystems and natural processes. At the same time, combining models into an ensemble based on the proposed architecture of the metaclassifier makes it possible to increase the stability of the analyzing system: the accuracy of decisions made by the ensemble tends to tend to the accuracy of the most efficient monclassifier of the system. The integration of individual classifiers into ensembles makes it possible to approach the solution of the scientific problem of finding classifier hyperparameters through the combined use of models of the same type with different configurations. The formation of a metaclassifier according to the proposed algorithm is an opportunity to add an element of predictability and control to the use of neural network models, which are traditionally a “black box”. Mapping of the geosystems of the Inerka test site shows their weak resistance to recreational development. The main limiting factors are the composition of Quaternary deposits, the nature of the relief, the mechanical composition of soils, soil moisture, the thickness of the humus horizon of the soil, the genesis and composition of vegetation.

KEYWORDS: ensembles, machine learning, metageosystems, cartography, sustainable development, GIS

ВВЕДЕНИЕ

Прогресс в области технологий цифрового картографирования и анализа геопространственных данных и материалов дистанционного зондирования Земли, а также развитие методического и алгоритмического обеспечения процесса анализа структуры земель обусловили повышение спроса на географическую информацию [Тикунов и др., 2021]. Актуальность решения научной проблемы разработки новых методов и алгоритмов интеллектуального анализа пространственных данных на основе технологий машинного обучения для поддержки процесса принятия управленческих решений в области анализа состояния и структуры систем землепользования определена необходимостью перехода к передовым цифровым технологиям для обеспечения эффективного решения стратегических задач устойчивого пространственного развития и территориального планирования [Han et al., 2021; Heaton et al., 2019].

Цель представленного в статье исследования – разработка и апробация методов и алгоритмов построения ансамблей моделей машинного обучения для решения задачи анализа структуры и состояния метагеосистем (на примере тестового полигона «Инерка»).

¹ National Research Mordovia State University, ul. Bolshevistskaya 68, 430005, Saransk, Russia, *e-mail:* yamashkinsa@mail.ru

² National Research Mordovia State University, ul. Bolshevistskaya 68, 430005, Saransk, Russia, *e-mail:* yamashkin56@mail.ru

Анализ отечественных и зарубежных научных публикаций показывает, что методы и алгоритмы машинного обучения могут быть эффективно использованы для интерпретации геопространственных данных, для которых характерны свойства пространственной зависимости, пространственной неоднородности и масштабируемости. Одновременно с этим, применение методов и алгоритмов глубокого машинного обучения к анализу геопространственных данных сталкивается с рядом открытых проблем, требующих научно обоснованного решения [LeCun et al., 2015; Lee et al., 2017]. К числу наиболее актуальных следует отнести следующие [Kim et al., 2015]: разработка системы методов и алгоритмов интеграции и предварительной обработки пространственных данных на основе новых методов машинного обучения и цифровой обработки сигналов данных [Сергеев, Юзькив, 2016]; формирование методологического, алгоритмического и программного обеспечения для построения моделей глубокого обучения, позволяющих интерпретировать многомерные массивы пространственных данных [Кузнецов, Мясников, 2021]; разработка методологического подхода к решению задачи проектирования, итерационной разработки и внедрения геопортальных систем [Gkonos et al., 2019], как точек доступа к распределенным массивам пространственной информации, и оптимизация при решении практических задач в области анализа систем землепользования.

МАТЕРИАЛЫ И МЕТОДЫ ИССЛЕДОВАНИЯ

В исследованиях геосистем выделяется спектр взаимосвязанных подходов, определяющих формирование систем электронных карт и баз данных: ландшафтный, включающий: 1) морфологическое – диагностика элементов геосистем и структурных отношений между ними; 2) парагенетическое – анализ динамически пространственно взаимосвязанных между собой по происхождению и процессами современного метаболизма вещества и энергии геосистем; 3) синтетическое – исследование взаимодействия процессов метаболизма вещества и энергии со структурой географических объектов; 4) исследование систем техногенез – отклик, в которых ключевым элементом является изменение состояния эколого-социально-экономических систем под влиянием хозяйственной деятельности человека.

Моделирование иерархической системы таксонов геосистем ориентируется на выделение разрядов, классов, групп, типов, родов и видов по В.А. Николаеву [Николаев, 1978].

1) Разряд геосистем – выделяется по особенностям макро- и мезоклимата. Среди климатических факторов в процессе формирования геосистем основную роль играют суммарная поглощенная солнечная радиация, температура воздуха, осадки, их режим и интенсивность, а также условия испарения – влажность, скорость ветра. Функционирование геосистем проявляется в сменах сезонов года

2) Класс геосистем – картографируется по орографическим признакам. Роль рельефа в процессах формирования геосистем определяется высотой, крутизной и экспозицией склонов, которые перераспределяют водные массы и изменяют величину испарения. От вершины склона к подножью за счет склонового стока происходит перераспределение поступившей атмосферной влаги, что может существенно влиять на ее впитывание в почву и инфильтрацию.

3) Группа геосистем – диагностируются по типам водного и геохимического режима. Основным объектом исследования является зона свободного водообмена, содержащая пресные и слабоминерализованные воды.

4) Тип геосистем – диагностируется по почвенно-биотическим признаками, с выделением зональных, интразональных, экстразональных и азональных геосистем. При составлении карты типов геосистем особое внимание уделялось анализу элементарных почвообразовательных, экзогеодинамических, геохимических и др. При геоэкологическом анализе почв как фактора формирования водного баланса, кроме водно-физических свойств, учитывались их устойчивость к развитию эрозионных процессов, техногенному загрязнению.

5) Род геосистем (GS_5) – в региональном масштабе исследований отражает морфоскульптурные формы рельефа и слагающие их отложения, что позволяет интегрально учитывать природное многообразие литогеогенных систем и их элементов (области питания, транзита и разгрузки подземных вод и т.п.).

Для решения задачи автоматизированной классификации метагеосистем территории актуально использование методов и алгоритмов машинного обучения. Важное направление в данной области представляет применение ансамблей классификаторов, выстраиваемых на базе методологии Ensemble Learning и объединяющих в систему различные модели, и позволяющие достичь повышения точности и устойчивости моделей машинного обучения.

Для апробации нового алгоритма интерпретации данных космической съемки с целью анализа состояния и структуры метагеосистем сформирован тестовый полигон «Инерка», расположенный на юго-востоке Русской равнины.

В схеме физико-географического районирования России это провинция лесостепи Приволжской возвышенности. Исследованный регион отличается древностью хозяйственного освоения, что обусловило значительную трансформацию геосистем, проявление широкого спектра деструктивных геоэкологических процессов, в первую очередь плоскостной и линейной эрозии, усиление поверхностного стока, уменьшение биологического разнообразия.

Разряд геосистем. Радиационный баланс составляет за год 1 638,29 Мдж/м², в декабре он равен - 16,76 Мдж/м², в июне 339,39 Мдж/м². Годовая сумма прямой солнечной радиации, поступающей на поверхность, перпендикулярную солнечным лучам, равна 3 536,36 Мдж/м², в июне она равна 536,32 Мдж/м², в декабре 8,38 Мдж/м². На условия, созданные притоком солнечной радиации, оказывает воздействие циркуляция атмосферы, определяющаяся взаимодействием воздушных масс Атлантики, Арктического бассейна и Евразийского материка. Климат умеренно континентальный, с хорошо выраженной сменой сезонов года.

Классы геосистем. Особенности тектоники определяют врезание русла реки в толщу горных пород и меандрирование на юго-восток. Долина Суры имеет правостороннюю асимметрию бортов. Правый борт представлен уступом высотой до 20–40 м и крутизной 10–20 град., и местами обрывистый, расчлененный глубокими долинами рек, оврагами, балками и лощинами, в которых вскрываются палеогеновые и верхнемеловые горные породы. Кремнисто-карбонатные породы сызранская свита палеоцена – опоки, мергели, с прослоями песчаников и песков, мощностью до 40 м, перекрывают кампан-маастрихтские – белый пясчий мел, белые и светло-серые мергели с прослоями кварц-глауконитового песка; сантонские – мергели с прослоями кварц-глауконитового песка; туронские – светло-серые известковистые глины, мергели. В нижней части долины Суры вскрываются породы альбского яруса – глины с прослоями опоквидных песчаников и песков.

Надпойменные террасы значительно преобразованы эоловыми процессами и характеризуются распространением песчаных дюнообразных возвышений протяженностью 500–700 м, высотой до 15 м, разделяемых заболоченными понижениями и дефляционными котловинами. По левобережью древнеаллювиальная равнина пересекается левыми притоками Суры – Синяш, Семилей, Сярдолатка, Урлейпотмо. Вблизи оз. Инерка, являющегося озером старичного типа р. Суры, расположены озера Татарка, Ишак, Журка. Абсолютные отметки изменяются от 105 м в русле р. Сура до 128,8 м на поверхности надпойменной террасы. Абсолютная отметка уреза воды оз. Инерка составляет 108,7 м, оз. Татарка – 109,4 м.

Группы геосистем. Зона свободного водообмена, определяющая структуру и функционирование групп геосистем, включает водоносные горизонты в четвертичных отложениях и коренных породах. Опоки и опоквидные песчаники сызранской свиты, залегающие в верхних водораздельных частях рельефа, сильно трещиноваты и не имеют нижнего водоупора. В связи с этим воды дренируются в нижележащие кампан-маастрих-

тские отложения. Кампан-маастрихтский водоносный горизонт имеет свободное зеркало воды на глубинах от первых метров (в местах эрозионного среза данных отложений) до 20–40 м (на участках полных разрезов кампан-маастрихтских образований и перекрытия их палеогеновыми опоками). Вмещающие породы – трещиноватый мел и мергели. Мощность горизонта от первых метров до 20,0 м. Дебит горизонта колеблется от 0,2 до 0,76 л/сек. Минерализация изменяется от 0,07 до 0,3 г/дм³. Воды гидрокарбонатно-кальциево-магниевого. Туронский водоносный горизонт развит более широко, часто питая многочисленные разнопорядковые водотоки. Вмещающими породами являются трещиноватые мергели. Подстилающим водоупором – альбская глинистая толща. Мощность горизонта от первых метров (в местах эрозионного среза) до 40 м (на участках полного разреза верхнемеловых и палеогеновых отложений). Дебит горизонта аналогичен кампан-маастрихтскому. Воды имеют минерализацию в пределах 0,4–0,9 г/дм³ и по составу относятся к гидрокарбонатно-кальциево-магниевым.

Верхняя часть альбских отложений, которые вскрываются под современными аллювиальными отложениями р. Суры, представлены темно-серыми тонкими глинами, являющимися водоупором для туронского водоносного горизонта. Водоносные горизонты I и II надпойменной террасы, как правило, плохо разделимы и имеют мощность 11,5–12,9 м. Горизонт не напорный. В связи с преобладанием в разрезе мелкозернистых песков и уменьшения роли галечников и песчано-гравийных отложений, водообильность его значительно меньше, чем в современном аллювиальном водоносном горизонте (0,008–0,01 л/сек). Минерализация воды – 0,28–0,76 г/дм³. Воды по химическому составу гидрокарбонатно-кальциево-магниевого.

Начало наибольшего подъема уровня приходится на апрель и первую половину мая, наиболее низкие – перед началом весеннего снеготаяния, а также в августе–сентябре. Минерализация изменялась от 0,3 до 0,7 г/дм³. Воды преимущественно гидрокарбонатные кальциево-магниевого и кальциево-натриевого. Горизонт имеет тесную гидравлическую связь с поверхностными водами реки Сура и озера Инерка. Так, химический состав воды в оз. Инерка весьма близок к химическому составу подземных вод современного аллювиального водоносного горизонта. Минерализация составляет 0,43 г/дм³, жесткость 5,0–5,1 мг-экв. По составу воды гидрокарбонатные кальциево-магниевого. Река Сура является областью разгрузки водоносного горизонта. Химический анализ воды показал сходимость ее с подземными водами современного аллювиального водоносного горизонта и водами озера Инерка. Минерализация воды из реки Сура составляет 0,67 г/дм³, жесткость – 8,0 мг-экв. По составу воды гидрокарбонатные кальциево-магниевого.

Наиболее крупными озерами являются Инерка, Татарка, которое серповидно вытянуто вдоль Инерки (юго-западнее). В 500 м к юго-западу от озера Инерка расположено торфяное болото, которое представляет собой бессточную впадину у основания песчаной гряды. Сюда впадают реки Синяш, Саксаурка, Семилей. Рельеф поймы изрезан многочисленными старицами, озерами, протоками, прирусловыми валами. В настоящее время озеро Инерка и озеро Татарка интенсивно заиливаются и зарастают.

Ледовые явления начинаются обычно с образованием сала и заберегов, и поступают в среднем 6 ноября. Средняя дата начала ледостава 22 ноября, продолжительность ледостава в среднем 138 суток. Толщина льда в марте достигает 60–80 см. В конце марта с появлением воды на льду закраин начинается разрушение ледового образования. Средняя дата начала весеннего ледохода 8 апреля, средняя продолжительность весеннего ледохода – 4 суток. По опросу местных жителей 10 лет на рассматриваемом участке реки Сура весеннего ледохода не наблюдалось – лед таял на месте. Поступающая вода частично испаряется летом, определенное ее количество в летне-зимние периоды года просачивается в дно и берега и стекает в Суру. Явных признаков значительного участия подземных вод в процессе водного питания озера нет. Крупные родники не обнаружены. Подпитка подземными водами осуществляется равномерно через толщу песчано-илистых отложений берегов и дна озера. Подземные воды слабоминерализованные (200–300 мг/дм³), гидрокарбонатно-кальциево-магниевого состава, нейтральные и слабоще-

лочные. Воды, стекающие с заболоченных участков, содержат повышенное количество органических веществ.

Типы, роды и виды геосистем. Особенности присурского ландшафта, в котором расположен памятник природы «Озеро Инерка», определяются его положением в лесостепи Приволжской возвышенности. Характерной чертой природы этого района Мордовии является хорошо выраженная склоновая смена типов геосистем. На возвышенных останцово-водораздельных массивах распространены лесные типы ландшафтов, сменяющиеся на нижних придолинных участках склонов лугово-степными геоконструкциями. На тестовом полигоне на древнеаллювиальных песках, сильно переработанных эоловыми процессами, распространены интразональные геосистемы хвойных и смешанных лесов.

Современные аллювиальные отложения, слагающие пойму р. Суры, представлены песками мелкозернистыми и среднезернистыми с значительным содержанием гравия и гальки опоки и опокovidного песчаника. Аллювий малых рек представлен мелкозернистыми и тонкозернистыми песками, с прослоями илистых глин. Мощность отложений в пойме Суры до 10–15 м. Древнеаллювиальные отложения надпойменных террас р. Суры представлены песками мелкозернистыми и среднезернистыми, внизу с гравием и уплотненной галькой опок и опокovidного песчаника. Мощность достигает 15–19 м.

Нижние части бортов долины и надпойменные террасы плащеобразно перекрываются элювиально-делювиальными и покровными образованиями проблематичного генезиса – лессовидные суглинки и глины, с прослоями и линзами мелкозернистого песка. Мощность отложений колеблется от первых метров до 12–13 м.

Для решения задачи синтеза цифровой карты метагеосистем тестового полигона «Инерка» разработан алгоритм формирования ансамблей классификатора на базе концепции Ensemble Learning [Boucher et al., 2009; Dong et al., 2020], основанный на применении геосистемного подхода для решения задачи классификации ландшафтов посредством технологий глубокого обучения и разработке моделей глубокого обучения, способной эффективно эти данные интерпретировать. Под «интерпретацией» будем понимать операцию, позволяющую отнести пространственный объект к определенному классу.

Алгоритм проектирования моноклассификаторов ансамбля. Модели машинного обучения, используемые для решения задачи классификации метагеосистем, могут иметь различную архитектуру, гиперпараметры, а также обучаться на разных наборах данных, однако в общем случае, при вовлечении геосистемного подхода, задача классификации территории X_j сводится к поиску функциональной зависимости вида [Yamashkin et al., 2020]:

$$y = f(X_j) = f(\langle \mathbf{x}_{land}, \mathbf{x}_{Geosystem} \rangle) \quad (1)$$

Данные о территории \mathbf{x}_{land} могут быть многомерными, включая информацию о спектральных свойствах территории, особенностях ее пространственной организации, иную информацию атрибутивного, пространственного и временного характера. Набор признаков локального территориального объекта (в том числе пространственных, спектральных и радиометрических характеристик) может быть сформирован на основе материалов космической съемки или иных источников, при этом пространственную область можно классифицировать на основе попиксельного анализа, либо путем выделения признаков из фрагментов области разного размера. Совокупность признаков локального объекта, которые сами по себе могут быть представлены в форме тензоров, определяет базовый уровень создаваемой геопространственной модели местности.

Под $\mathbf{x}_{Geosystem}$ понимаются данные, характеризующие вмещающие иерархические геосистемы. В соответствии с геосистемным подходом, вмещающая геосистема оказывает существенное влияние на свойства иерархически подчиненных образований. Полезным источником информации о ней являются данные дистанционного зондирования Земли (ДЗЗ).

Процесс интеграции расширенных данных о геосистемах при определенных допущениях подлежит потенциально полной автоматизации: имея данные о географических координатах (широте и долготе) классифицируемой области, можно запросить API поставщика геопространственных данных о получении фрагмента спутникового снимка местности, включающего эти координаты и характеризующегося необходимым масштабом отображения снимка и разрешением. Таким образом, мы получаем возможность алгоритмического расширения обучающей выборки за счет импорта фрагментов данных ДЗЗ, описывающих геосистемы более высокого иерархического уровня и вмещающих классифицируемую территорию.

Следует отметить, что не только данные дистанционного зондирования определенного масштаба могут характеризовать геосистемы разного иерархического уровня, эта функция вполне может быть возложена и на синтетические цифровые карты, а также цифровые модели рельефа. Карты систем землепользования, электронные ландшафтные и другие тематические карты, традиционно являющиеся конечным продуктом интерпретации данных ДЗЗ, содержат значительный объем информации о свойствах территорий, входящих в его состав, и, как следствие, их можно использовать для создания входных тензоров дополнительной информации в наборе. Несмотря на то, что эти карты зачастую не характеризуют территорию в крупном масштабе, их высокий уровень абстракции позволяет говорить о высоком потенциале для обогащения информации по классификации небольшого по размерам региона. Синтетические цифровые карты формируют еще один уровень геопространственной модели местности и становятся важным источником семантического расширения набора вспомогательных данных. Множество $X_{Geosystem}$ может быть пустым, в этом случае можно говорить о классической классификации без учета данных о вмещающих геосистемах.

Оптимизационный алгоритм построения модели классификации метагеосистем и поиска ее гиперпараметров, сформирован и апробирован в ходе выполнения проектных работ и включает следующую последовательность этапов: 1) формирование системы требований к модели: определение входов и выходов, показателей производительности и точности; 2) определение базовой архитектуры модели на основе модульного подхода, описывающего общую организацию классификатора, декомпозиция единицы верхнего уровня на линейные или ветвящиеся структуры; 3) решение проблемы снижения точности классификации и переобучения за счет эвристической конфигурации гиперпараметров глубокой модели и добавления слоев нормализации и регуляризации; 4) оптимизация модели по принципу «меньшее лучше большего», согласно которому процесс обучения больших моделей требует больших вычислительных ресурсов и, что еще важнее, такие глубокие нейронные сети склонны к переобучению; 5) обучение модели с тестированием различных мер точности, алгоритмов оптимизации, функций потерь и количества эпох обучения; анализ процесса обучения модели путем расчета зависимости математического ожидания и стандартного отклонения точности классификации от эпохи обучения на основе серии экспериментов; 6) оценка качества полученного результата путем построения матриц ошибок и расчета метрик точности и ошибочности по результатам функционирования модели; вывод о соответствии полученной модели объективным и субъективным требованиям к ней.

Предлагаемая цепочка действий приводит к получению экземпляра модели с определенными свойствами, при соответствии которых проектным требованиям поиск может быть завершен. Процесс поиска эффективной модели классификации может быть формализован в виде дерева, корневой узел которого предшествует первому этапу алгоритма поиска и соответствует решению задачи постановки задачи исследования.

На каждом этапе на основании эвристик, паттернов, а также автоматизированных алгоритмов формируется пространство гипотез, в рамках которого осуществляется поиск решения по выбору направления детализации и оптимизации модели. Принятие решения определяет выбор реализации одной из гипотез, при котором осуществляется переход к узлу дерева, определяющий вариант состояния модели на определенном эта-

пе алгоритма поиска эффективной модели. Если свойства спроектированной модели на определенном этапе не соответствуют предъявляемым к ней требованиям, необходимо принять решение о возвращении на несколько этапов назад в рамках алгоритма построения глубокой модели (вплоть до первого этапа, если сформулированные требования оказались недостижимыми) и осуществить новый поиск в эвристически скорректированном направлении.

Конечные узлы (листья) дерева соответствуют частному решению задачи поиска оптимальной модели, готовой к использованию. Полученные частные решения могут быть сопоставлены с помощью сплит-тестирования, основанного на сопоставлении показателей эффективности модели с субъективной экспертной оценкой качества классификации.

Ансамбль моделей машинного обучения E формируется на основе моноклассификаторов M_i (отдельных обученных моделей) и метаклассификатора MC , осуществляющего принятие результирующего решения при решении задачи классификации территориального объекта X_j для определения его принадлежности к классу метагеосистем.

Набор моноклассификаторов ансамбля BC состоит из обученных моделей M_i , осуществляющих функцию определения принадлежности территориального объекта X_j классу метагеосистем C_k . Отличие моноклассификаторов ансамбля может заключаться как в их архитектурной и топологической организации, так и в данных, использованных для обучения модели.

$$BC = \{M_i: i \in \mathbb{N} \wedge i \leq N\} \quad (2)$$

Алгоритм проектирования метаклассификатора ансамбля. При классификации территориального объекта X_j моноклассификаторы ансамбля формируют вектор гипотез P_j относительно принадлежности данного объекта определенному классу метагеосистем C_k из номенклатуры классов C мощностью K :

$$\begin{cases} P_j = \{f(M_i, X_j): i, j \in \mathbb{N} \wedge i \leq N\} \\ P_j \in C = \{c_k: k \in \mathbb{N} \wedge k \leq K\} \end{cases} \quad (3)$$

При этом, функция $f_i(X_j) = f(M_i, X_j)$ порождается в процессе обучения моноклассификатора M_i при накоплении опыта решения задачи классификации метагеосистем на основе меры качества P и возвращает набор байесовских вероятностей, определяющих степень уверенности моноклассификатора M_i в истинности факта принадлежности территории X_j классу метагеосистем C_k .

$$f_i(X_j) = \{p_{jik}: k \in \mathbb{N} \wedge k \leq K \wedge 0 \leq p_k \leq 1\} \quad (4)$$

Решение Y_{ij} моноклассификатора M_i о принадлежности определенного территориального объекта X_j конкретному классу метагеосистемы может быть вынесено путем выбора класса c_i , для которого рассчитанная байесовская вероятность максимальна.

$$Y_{ij} = \operatorname{argmax}_k p_{ik \in f_i(X_j)} p_{jik} \quad (5)$$

Результирующая гипотеза Y_E об отнесении территории X_j конкретному классу метагеосистем c_i выносится метаклассификатором MC ансамбля E . В этом случае, принятие результирующего решения на основе выходных данных моделей глубокого обучения целесообразно принять на основе взвешенного голосования, обобщенное представление которого имеет следующий вид:

$$Y_{Ej} = \operatorname{argmax}_k \varphi(M_i, k) \cdot \psi(p_{jik}) \quad (6)$$

Расчет мер эффективности моноклассификаторов. В данной формуле параметр ω_{ik} представляет собой весовой коэффициент, представляющий собой меру эффективности моноклассификатора M_i в детекции метагеосистем класса k . Функция определяет преобразование вида $\varphi: M_i \rightarrow \Lambda_{ik}$, при котором весовой коэффициент и мера эффективности Λ_{ik} определяется путем математических преобразований экспериментально полученных данных матрицы ошибок M_i моноклассификатора M_i .

Предлагается следующий алгоритм расчета меры эффективности Λ_{ik} :

1) Построение матрицы ошибок M_i для каждого моноклассификатора M_i системы.
 2) Вычисление абсолютных метрик точности для классификатора M_i при определении метагеосистем класса k : попаданий (TP_{ik}), верных отклонений (TN_{ik}), ошибок I (FP_{ik}) и II (FN_{ik}) типов.

3) Расчет относительной метрики \mathcal{R}_{ik} , определяющей точности классификации метагеосистем класса k , позволяющей осуществить интегральную оценку полученной матрицы ошибок M_i числом в интервале $[0; 1]$. Так, оценка F_β , комплексно учитывающие показатели точности (*precision*) и полноты (*recall*), а, следовательно, ошибки I и II типов, а также число верных попаданий моноклассификатора. Кроме этого, метрика настраивается посредством конфигурирования параметра β , позволяющего расставить акцент влияния точности и полноты на результат.

$$\mathcal{R}_{ik} = F_{i\beta k} = (1 + \beta^2) \cdot \frac{precision_{ik} \cdot recall_{ik}}{(\beta^2 \cdot precision_{ik}) + recall_{ik}} = \frac{(1 + \beta^2) \cdot TP_{ik}}{(1 + \beta^2) \cdot TP_{ik} + \beta^2 \cdot FN_{ik} + FP_{ik}} \quad (7)$$

При решении задачи классификации метагеосистем важное значение может быть уделено максимизации метрики верных отклонений. Данное обстоятельство учитывается посредством расчета корреляции Мэтьюса:

$$\mathcal{R}_{ik} = |\Phi_{ik}| = \left| \frac{TP_{ik} \cdot TN_{ik} - FP_{ik} \cdot FN_{ik}}{\sqrt{(TP_{ik} + FP_{ik})(TP_{ik} + FN_{ik})(TN_{ik} + FP_{ik})(TN_{ik} + FN_{ik})}} \right| \quad (8)$$

При необходимости может быть спроектирована какая-либо иная метрика, удовлетворяющая требованиям построения ансамбля.

4) Деактивация неэффективных классификаторов относительно порогового значения может быть осуществлена по следующему принципу:

$$\tilde{\mathcal{R}}_{ik} = (\mathcal{R}_{ik} > \epsilon) ? \frac{\mathcal{R}_{ik} - \epsilon}{1 - \epsilon} : 0 \quad (9)$$

В результате приведенной условной тернарной операции решаются две задачи: во-первых, значения мер эффективности меньше порога ϵ обнуляются, убирая неэффективные моноклассификаторы из системы принятия решений; во-вторых, результирующее значение вновь нормируется в интервале $[0; 1]$. При $\epsilon = 0$ реализуется возможность отказа от использовании деактивации.

5) Активация метрики, посредством осуществления дополнительного нелинейного нормализованного монотонного преобразования θ метрики $\tilde{\mathcal{R}}_{ik}$

$$\Lambda_{ik} = \theta(\tilde{\mathcal{R}}_{ik}) \quad (10)$$

Операция активации метрики обеспечивает минимизацию или ускорение роста величины метрики у ее граничных значений. В качестве активационной функции θ может быть взята логистическая кривая. При тождественном отображении $id_{\tilde{\mathcal{R}}_{ik}}$ реализуется возможность отказа от использовании активации.

Результирующее значение метрики Λ_{ik} может быть использовано для определения меры эффективности моноклассификатора M_i для осуществления детекции метагеосистем класса k .

Расчет меры голоса моноклассификатора. Функция ψ представляет собой преобразование вида $\psi: p_{jik} \rightarrow \mathbb{Q}_{jik}$, при котором байесовская вероятность p_{jik} , определяю-

шая степень уверенности моноклассификатора M_i в истинности факта принадлежности территории X_j классу метагеосистем C_k трансформируется в меру голоса Q_{jik} .

Мера голоса может быть определена по принципу «победитель забирает все», при котором моноклассификатор M_i выставляет 1 наиболее вероятному решению и 0 всем остальным:

$$Q_{jik} = (Y_{Ej} = k)? 1: 0 \quad (11)$$

При тождественном отображении id_ψ при голосовании будет учитываться байесовские вероятности принадлежности территории X_j классу метагеосистем k . Наконец, может быть осуществлена активация метрики Q_{jik} , посредством осуществления дополнительного нелинейного нормализованного монотонного преобразования $\theta(Q_{jik})$, обеспечивающего изменение меры голоса у граничных значений.

РЕЗУЛЬТАТЫ ИССЛЕДОВАНИЯ И ИХ ОБСУЖДЕНИЕ.

Для решения задачи классификации геосистем тестового полигона «Инерка» использованы пространственные данные трех различных источников, предоставляющие информацию о динамических и инвариантных свойствах исследуемой территории. Источником об изменяющихся во времени спектральных характеристиках территории были выбраны данные дистанционного космического мониторинга, получаемые со спутника Sentinel-2. Анализируемое изображение было получено 30 августа 2021 года (по системе MGRS, рассчитываемой на основе универсальной поперечной проекции Меркатора квадрат съемки определяется положением 38UNF), пространственное разрешение съемки в видимых зонах спектра и ближнем инфракрасном диапазоне составляет 10 метров на 1 атомарный участок. Данные о морфометрических свойствах территории, характеризующие ее инвариантное (медленное, но безвозвратно изменяемое) состояние, получены на основе материалов SRTM, скорректированных и распространяемых компанией Mapzen на основе облачных прикладных программных интерфейсов Amazon. Средняя квадратическая ошибка относительно фиксируемой высоты для исследуемого тестового полигона измеряется в пределах 1,4 метров, а пространственное разрешение составляет 30 метров. Третьим источником информации об анализируемой территории выступили данные региональной ГИС «Мордовия», используемые для построения цифровой ландшафтной карты и содержащие информацию о региональных системах землепользования. Не характеризуясь высоким пространственным разрешением, они обладают значительной информационной емкостью, храня синтетические показатели анализируемой территории: классы геосистем и типы систем землепользования.

Предварительная подготовка изображения Sentinel-2 уровня 2A выполнена на основе процессора Sen2Cor и включила в себя нормализацию значений спектральной яркости, коррекцию данных на основе атмосферных параметров с поправкой на отражательную способность рельефа и перистых облаков. Для повышения точности анализа был рассчитан набор территориальных дескрипторов, характеризующих свойства вмещающей для атомарного участка окрестности. В число синтезированных признаков вошли локальная энтропия вмещающей геосистемы (окрестности), разброс спектральных яркостей относительно среднего значения, гистограммы оттенков, статистические данные карт высот, экспозиции и крутизны склонов. Расчет территориальных дескрипторов позволяет достичь эффекта снижения размерности анализируемых данных (в сравнении с анализом фрагмента космической съемки) при неизбежной утрате определенного количества информации об анализируемой территории. При этом был достигнут баланс между максимально возможным облегчением допустимой емкости модели машинного обучения, повышением ее устойчивости к переобучению и недопущением значительного снижения точности классификации в рамках решаемой задачи классификации геосистем тестового полигона. Рассчитанные дескрипторы обладают свойством наглядности: картограммы, построенные на их основе информативны для специалистов в области

анализа данных и наук о Земле. Перед машинным анализом метрики территории были нормализованы путем масштабирования вектора данных по значению его стандартного отклонения. Для каждого территориального класса подготовлено около сотни размеченных выборок.

На основе спектральных и морфометрических свойств территории, а также синтетических дескрипторов вмещающих геосистем сформирован вектор параметров, пригодный для обучения искусственными нейронными сетями невысокой глубины, основанными на использовании плотно связанных слоев. Такие модели, в отличие от широко используемых сверточных сетей, могут обучаться на более низком количестве размеченных данных, характеризуются устойчивостью к проблеме переобучения, менее требовательны к аппаратному обеспечению. В качестве функции активации слоев нейронной сети использована операция линейной ректификации (ReLU). Для повышения устойчивости модели к переобучению введен блок субдискретизации на основе взятия максимальной величины. Для решения проблемы снижения точности классификации и переобучения в структуру нейронной сети введены слои нормализации. Принятие решения о принадлежности территории определенному классу осуществляется на основе принципа «победитель забирает все», путем отбора гипотезы, для которой оценочная вероятность максимальна. Общее число полносвязных слоев ограничено двумя.

В ходе эксперимента спроектировано три искусственные нейронные сети (ИНС), которые также были объединены в ансамбль согласно описанной методике. Первая модель основана на одном плотно связанном слое в 10 нейронов, вторая и третья – на двух, мощностью 10-10, и 10-20 нейронов соответственно. Результаты расчета F_1 метрики на основе матрицы ошибок представлены на рисунке 1.

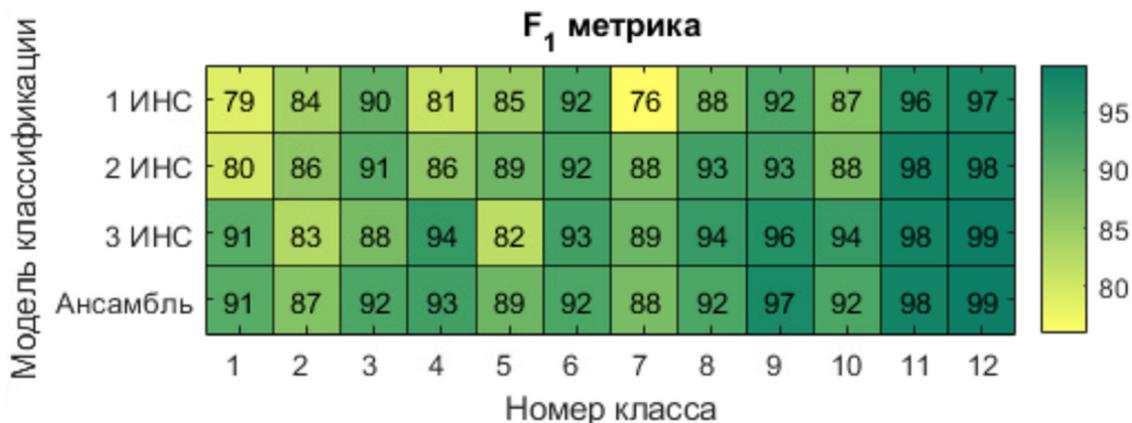


Рис. 1. Значение F_1 метрики для классификаторов при классификации геосистем
 Fig. 1. The F_1 score for classifiers in the classification of geosystems

Видно, что увеличение емкости моделей не приводит к однозначному улучшению результата, так как более мощные модели могут быть более неустойчивы к переобучению, а также требуют большего объема размеченных данных для обучения. При объединении в ансамбль результирующая гипотеза стала применяться на основе взвешенного голосования на основе меры эффективности, что позволило избежать грубых ошибок в классификации, свойственных для каждого классификатора в отдельности. При этом ансамбль лишь незначительно проигрывает в точности отдельным классификаторам системы, сохраняя общую устойчивость к ошибкам первого и второго рода, свойственным для отдельного классификатора при определении объектов конкретного класса территории.

Перейдем к обсуждению результатов анализа состояния и классификации геосистем тестового полигона «Инерка». Особенности структуры геосистем исследуемой территории определяются правосторонней асимметрией долины р. Суры. На рисунке 2 приведены соответствующие спектральные профили, отражающие изменение спектральной яркости геосистем долины р. Сура, а также приведен соответствующий профиль рельефа.

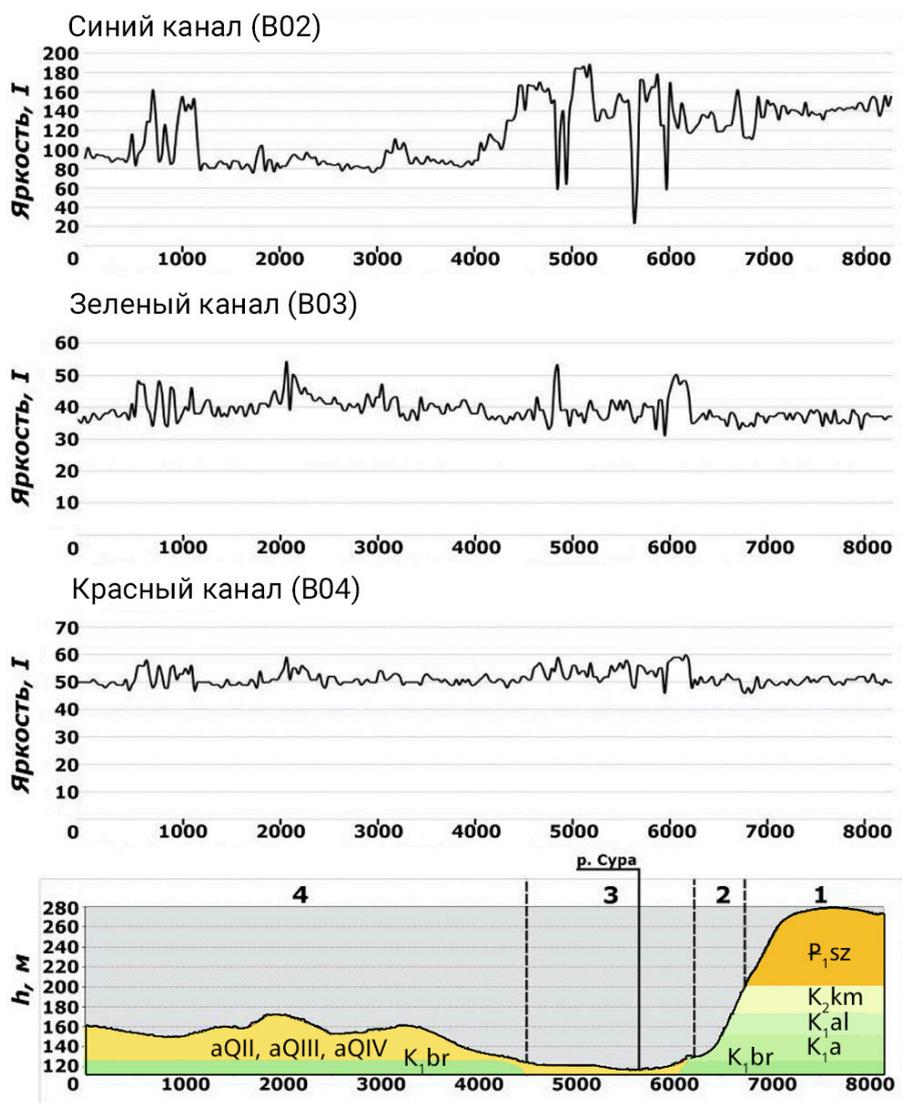


Рис. 2. Изменение спектральной яркости ландшафтов долины р. Сура
 Fig. 2. Changes in the spectral radiance of the landscapes of the valley of the Sura river

Цифрами на рисунке обозначены территориальные комплексы: 1 – останцово-водораздельный массив; 2 – склон; 3 – пойма; 4 – надпойменные террасы. Территория представлена следующими геологическими отложениями: P_1sz – опоки, пески, песчаники, алевриты (зеландский ярус); K_2km – пески, глины, мергели, опоки, алевриты (кампанский ярус); K_1al – глины, алевриты, пески, песчаники, в основании с галькой фосфоритов (альбский ярус); K_1a – пески, песчаники, алевриты, глины (аптский ярус); K_1br – пески, песчаники, алевриты, глины (барремский ярус); $aQII, aQIII, aQIV$ – аллювиальные отложения.

Значения яркостей пикселей весьма неравномерно колеблются, в особенности в левобережье Суры, и, всплески значений яркости соответствуют неравномерно распределенным обнажениям песков, появившимся в результате вырубок сосновых лесов. Общее и наиболее заметное повышение значений яркости спектральной кривой в ближнем инфракрасном диапазоне свойственно при переходе к пойменным ландшафтам Суры, занятым луговыми и болотными растительными сообществами. Спектральная кривая «выравнивается» при переходе к лесным геосистемам правобережья Суры. В целом, этот участок Сурского долинного ландшафта имеет достаточно контрастный характер, четко прослеживается склоновая смена ПТК и особенности хозяйственного освоения. Морфологическая структура геосистем территории полигона «Озеро Инерка» определяется правосторонней асимметрией долины р. Суры, значительной трансформацией надпой-

менных террас эоловыми процессами. Наиболее ярко смена геосистем выражена в ближнем инфракрасном канале, в котором также ярко проявляются и последствия хозяйственной деятельности, выражающиеся в появлениях открытых участков почв (песчаных пустошей) на месте сосновых лесов надпойменных террас долины р. Сура. Характерен «всплеск» значений яркости пикселей и при переходе в луговые, заболоченные пойменные геосистемы. При приближении к правому берегу (останцово-водораздельный массив с крутыми склонами) значения яркости пикселей падают и геосистема приобретает более однородный характер. На участке профиля антропогенные объекты в спектральной составляющей занимают около 30 % и, как было сказано выше, связаны с вырубкой леса. Хозяйственное освоение этих геосистем носило мелкоочаговый характер, что прослеживается и в спектральных кривых снимка.

Результаты классификации пространственных данных позволили выделить на территории полигона «Инерка» следующие роды и виды геосистем (рис. 3).

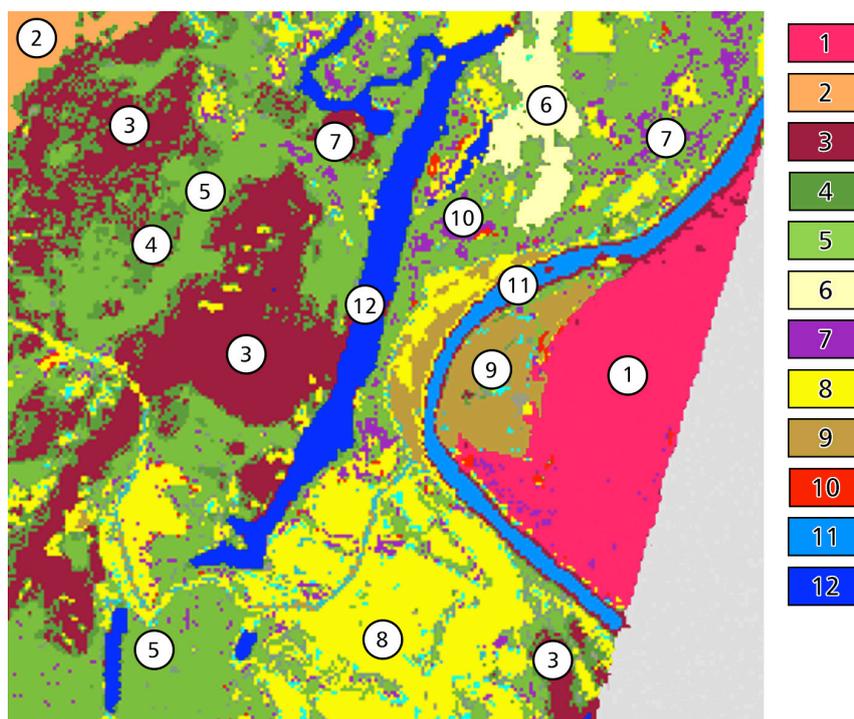


Рис. 3. Карта геосистем полигона «Инерка»
Fig. 3. Map of geosystems of the test polygon "Inerka"

1. Коренной борт долины Суры с абсолютными отметками до 265 м, сложенный мергелями, опоками, псичим мелом с выходами коренных горных пород на крутых склонах на дневную поверхность. Активные склоновые процессы обуславливают доминирование в структуре почвенного покрова серых лесных щебнистых почв под дубравами волосисто-осоковыми, снытево-осоковыми, снытевыми, снытево-пролесниковыми, злаково-разнотравными. Растительные сообщества формируют *Quercus robur*, *Acer platanoides*, *Tilia cordata*, *Ulmus glabra*, *Corylus avellana*, *Lonicera axylosteum* и др.

Надпойменные террасы.

2. Выровненные поверхности, сложенные песками и супесями, с дерново-слабоподзолистыми почвами под сосняками дубово-снытево-осоковыми и сосняками липово-широкотравными: *Corylus avellana*, *Euonymus verrucosa*, *Sorbus aucuparia*, *Rubus idaeus*, *Festuca gigantea*, *Poa nemoralis*, *Convallaria majalis*, *Asarum europaeum* и др.

3. Дюнообразные останцы внутрипойменных террасовых комплексов с котловинами выдувания со слаборазвитыми подзолистыми и дерново-слабоподзолистыми почвами под сосняками лишайниковыми, зеленомошными, сухотравно-злаковыми: *Pleurozium schreberi*, *Dicranum undulatum*, *Sorbus aucuparia*, *Chamaecytisus ruthenicus*.

4. Нижние части надпойменных террас с невысокими песчаными гривами осложненными неглубокими ложбинами, западинами и котловинами в структуре почвенного покрова характерна мозаика дерново-подзолистых, дерново-подзолистых глееватых и дерново-подзолистых глеевых почв под сосняками дубово-снытево-осоковыми, сосняками липово-широкотравными, сырыми березняками и осинниками: *Betula pubescens*, *Polytrichum commune*, *Carex*, *Eriophorum vaginatum*, *Ledum palustre*, *Andromeda polifolia*.

Пойма.

5. Слабоволнистые поверхности, сложенные суглинками с прослоями песков, осложненные староречьями и руслами небольших водотоков с серыми лесными оподзоленными, пойменными дерновыми зернистыми и иловато-перегнойными с избыточным увлажнением; в естественной растительности преобладают дубняки пойменные, ольшаники; встречаются небольшие по площади злаково-разнотравно-лисохвостовые луга; *Alnus glutinosa*, *Frangula alnus*, *Salix cinerea*, *Padus avium*, *Rubus idaeus*, *Urtica dioica*.

6. Комплексы крупных и мелких грив с дерновыми слабообразованными песчаными и дерново-луговыми маломощными легкосуглинистыми почвами, с мятликово-лугово-овсянниковыми лугами, иногда дубравами: *Dianthus superbus*, *Ranunculus polyphyllus*, *Orchis militaris*, *Gratiola officinalis*, *Pedicularis sceptrum-carolinum*.

7. Межгривные понижения и староречья с лугово-болотными тяжелосуглинистыми почвами под осоково-манниково-канареечниковыми лугами и ивняковыми зарослями; дерново-глеевыми и перегнойно-глеевыми почвами под осоковой, влажнотравно-крупнозлаковой растительностью, зарослями ивы и ольхи: *Coeloglossum sumviride*, *Botrychium virginianum*, *Eriophorum latifolium*, *Carex hartmanii*, *Malaxis monophyllos*.

8. Прирусловые песчаные сырые поймы со злаково-крупнотравными лугами и ивняками;

9. Прирусловые песчаные отмели (пляжи).

10. Топи ольховые с торфянисто-глеевыми и торфяные иловато-суглинистыми почвами. *Salix cinerea*, *Calamagrostis canescens*, *Carex lasiocarpa*, *Comarum palustre*, *Scutellaria riagalericulata*, *Filipendula ulmaria*, *Cicuta virosa*, *Phragmites australis*.

Природные аквальные комплексы.

11. Природно-аквальные комплексы Суры;

12. Природно-аквальные комплексы старичных озер: *Salvinia natans*, *Potamogeton praelongus*, *Potamogeton obtusifolius*, *Najas major*, *Caulinia minor*, *Cossonet Germ*, *Trapa natans*.

В настоящее время процесс трансформации русла Суры находится в стадии развития, и в связи с этим актуальна организация мониторинга экосистем в районе Инерки. Исследования показывают, что питание озера происходит за счет талых, дождевых и грунтовых вод. Высокие половодья р. Суры, которые могли бы пополнять и чистить озеро, из-за перехвата талых вод многочисленными водохранилищами, становятся очень редкими.

ВЫВОДЫ

Анализ эффективности методики построения ансамблей классификаторов при решении задачи исследования структуры метагеосистем тестового полигона «Инерка» позволяет сделать следующие выводы:

– использование ансамблей, выстраиваемых по предложенной методике, позволяет проводить оперативный автоматизированный анализ пространственных данных для решения задачи тематического картографирования метагеосистем и природных процессов;

– объединение моделей в ансамбль на основе предложенной архитектуры мета-классификатора позволяет повысить устойчивость анализирующей системы: точность решений, принимаемых ансамблем, имеет тенденцию стремиться к точности наиболее эффективного моноклассификатора системы;

– ошибочность системы в большинстве случаев не превышает ошибочность наиболее эффективного классификатора, избегая при этом грубых систематических ошибок, допускаемых отдельными моноклассификаторами;

– использование ансамблей позволяет подойти к решению проблемы подготовки данных для обучения моделей за счет интеграции в единую систему моделей, обученных на различных комбинациях обучающих и валидационных выборок для снижения влияния ошибок, возникающих при формировании наборов данных;

– интеграция отдельных классификаторов в ансамбли позволяет подойти к решению научной проблемы поиска гиперпараметров классификаторов за счет комбинированного использования моделей одного типа с разными конфигурациями;

– формирование метаклассификатора по предложенному алгоритму представляет собой возможность добавить элемент предсказуемости и контроля в использование нейронных сетей, традиционно представляющих собой «черный ящик»;

– построение эффективных ансамблей может быть основано на моделях относительно небольшой ширины и глубины, что позволяет проектировать высокоточные классификаторы, обучение которых менее требовательно к вычислительным мощностям в сравнении с классическими глубокими моделями.

Картографирование геосистем тестового полигона «Инерка» показывает их слабую устойчивость к рекреационному освоению. Основными лимитирующими факторами являются состав четвертичных отложений, характер рельефа, механический состав и влажность почв, мощность гумусового горизонта, генезис и состав растительности.

БЛАГОДАРНОСТИ

Исследование выполнено за счет гранта Российского научного фонда № 22-27-00651, <https://rscf.ru/project/22-27-00651/>.

ACKNOWLEDGEMENTS

The study was supported by the Russian Science Foundation grant No. 22-27-00651, <https://rscf.ru/en/project/22-27-00651/>.

СПИСОК ЛИТЕРАТУРЫ

1. Кузнецов А.В., Мясников В.В. Сравнение алгоритмов управляемой поэлементной классификации гиперспектральных изображений. Компьютерная оптика, 2014, Т. 38, ¹ 3. С. 494–502.
2. Николаев В.А. Классификация и мелкомасштабное картографирование ландшафтов. М.: Изд-во Моск. ун-та, 1978. 62 с.
3. Сергеев В.В., Юзькив Р.Р. Параметрическая модель автокорреляционной функции космических гиперспектральных изображений. Компьютерная оптика, 2016, Т. 40, ¹ 3, С. 416–421.
4. Тикунов В.С., Котова Т.В., Белоусов С.К. Экологическое состояние: определение, показатели, картографирование. ИнтерКарто. ИнтерГИС. 2021, Т. 27. Ч. 1. С. 165–194. DOI: 10.35595/2414-9179-2021-1-27-165-194.
5. Boucher M., Perreault L., Ancil F. Tools for the assessment of hydrological ensemble forecasts obtained by neural networks. Journal of Hydroinformatics, 2009, V. 11, No. 3–4. P. 297–307. DOI: 10.2166/hydro.2009.037.
6. Dong X., Yu Z., Cao W., Shi Y., Ma Q. A survey on ensemble learning. Frontiers of Computer Science, 2020, V. 14, No. 2. P. 241–258. DOI: 10.1007/s11704-019-8208-z.
7. Gkonos C., Iosifescu Enescu I., Hurni L. Spinning the wheel of design: evaluating geoportals Graphical User Interface adaptations in terms of human-centred design. International Journal of Cartography, 2019, V. 5, No. 1, P. 23–43. DOI: 10.1080/23729333.2018.1468726.
8. Han Q., Zhao N., Xu J. Recognition and location of steel structure surface corrosion based on unmanned aerial vehicle images. Journal of Civil Structural Health Monitoring, 2021, V. 11, No. 5, P. 1375–1392. DOI: 10.1007/s13349-021-00515-7.
9. Heaton J., Datta A., Finley A.O. A case study competition among methods for analyzing large spatial data. Journal of Agricultural, Biological and Environmental Statistics, 2019, V. 24, No. 3, P. 398–425. DOI: 10.1007/s13253-018-00348-w.

10. *Kim S.E., Seo I.W.* Artificial neural network ensemble modeling with exploratory factor analysis for streamflow forecasting. *Journal of Hydroinformatics*, 2015, V. 17, No. 4. P. 614–639. DOI: 10.2166/hydro.2015.033.
11. *LeCun Y., Bengio Y., Hinton G.* Deep learning, *Nature*, 2015, V. 521, No. 7553, P. 436–444. DOI: doi.org/10.1038/nature14539.
12. *Lee J., Kang M.* Geospatial Big Data: Challenges and Opportunities, *Big Data Research*, 2017, V. 2, No. 2, P. 74–81. DOI: 10.1016/j.bdr.2015.01.003.
13. *Yamashkin S. A., Yamashkin A. A., Zanozin V. V., Radovanovic M. M., Barmin A. N.* Improving the efficiency of deep learning methods in remote sensing data analysis: geosystem approach. *IEEE Access*, 2020, V. 8. P. 179516–179529. DOI: 10.1109/ACCESS.2020.3028030.

REFERENCES

1. *Boucher M., Perreault L., Anctil F.* Tools for the assessment of hydrological ensemble forecasts obtained by neural networks. *Journal of Hydroinformatics*, 2009, V. 11, No. 3–4. P. 297–307. DOI: 10.2166/hydro.2009.037.
2. *Dong X., Yu Z., Cao W., Shi Y., Ma Q.* A survey on ensemble learning. *Frontiers of Computer Science*, 2020, V. 14, No. 2. P. 241–258. DOI: 10.1007/s11704-019-8208-z.
3. *Gkonos C., Iosifescu Enescu I., Hurni L.* Spinning the wheel of design: evaluating geoportal Graphical User Interface adaptations in terms of human-centred design. *International Journal of Cartography*, 2019, V. 5, No. 1, P. 23–43. DOI: 10.1080/23729333.2018.1468726.
4. *Han Q., Zhao N., Xu J.* Recognition and location of steel structure surface corrosion based on unmanned aerial vehicle images. *Journal of Civil Structural Health Monitoring*, 2021, V. 11, No. 5, P. 1375–1392. DOI: 10.1007/s13349-021-00515-7.
5. *Heaton J., Datta A., Finley A.O.* A case study competition among methods for analyzing large spatial data. *Journal of Agricultural, Biological and Environmental Statistics*, 2019, V. 24, No. 3, P. 398–425. DOI: 10.1007/s13253-018-00348-w.
6. *Kim S.E., Seo I.W.* Artificial neural network ensemble modeling with exploratory factor analysis for streamflow forecasting. *Journal of Hydroinformatics*, 2015, V. 17, No. 4. P. 614–639. DOI: 10.2166/hydro.2015.033.
7. *Kuznetsov A.V., Myasnikov V.V.* Comparison of algorithms for controlled element-by-element classification of hyperspectral images. *Computer Optics*, 2014, V. 38, No. 3. P. 494–502. (in Russian).
8. *LeCun Y., Bengio Y., Hinton G.* Deep learning, *Nature*, 2015, V. 521, No. 7553, P. 436–444. DOI: doi.org/10.1038/nature14539.
9. *Lee J., Kang M.* Geospatial Big Data: Challenges and Opportunities, *Big Data Research*, 2017, V. 2, No. 2, P. 74–81. DOI: 10.1016/j.bdr.2015.01.003.
10. *Nikolaev V.A.* Classification and small-scale landscape mapping. Moscow: Moscow State University Press, 1978. 62 p. (in Russian).
11. *Sergeev V.V., Yuzkiv R.R.* Parametric model of the autocorrelation function of space hyperspectral images. *Computer Optics*, 2016, V. 40, No. 3, P. 416–421. (in Russian).
12. *Tikunov V.S., Kotova T.V., Belousov S.K.* Environmental conditions: definition, indicators, mapping. *InterCarto. InterGIS*, 2021, V. 27. Part 1. P. 165–194 (in Russian). DOI: 10.35595/2414-9179-2021-1-27-165-194.
13. *Yamashkin S.A., Yamashkin A.A., Zanozin V.V., Radovanovic M.M., Barmin A.N.* Improving the efficiency of deep learning methods in remote sensing data analysis: geosystem approach. *IEEE Access*, 2020, V. 8. P. 179516–179529. DOI: 10.1109/ACCESS.2020.3028030.