



Научная статья

УДК 631.4:004.8:528.8

doi: 10.55186/25876740_2026_69_6_847

ПРИМЕНЕНИЕ МАШИННОГО ОБУЧЕНИЯ И ДИСТАНЦИОННОГО ЗОНДИРОВАНИЯ В ПОЧВОВЕДЕНИИ: СОВРЕМЕННОЕ СОСТОЯНИЕ И ПРАКТИЧЕСКИЕ ВОЗМОЖНОСТИ

С.С. Огородников^{1,2}, В.А. Блынский²

¹Московский авиационный институт, Москва, Россия

²МГУ имени М.В. Ломоносова, Москва, Россия

Аннотация. В статье рассмотрены современные подходы к оценке и картографированию свойств почв на основе интеграции данных дистанционного зондирования, наземных наблюдений, лабораторных измерений и методов машинного обучения. Цель исследования заключается в систематизации источников данных, этапов их предварительной обработки и алгоритмов анализа, применяемых для повышения точности прогноза почвенных характеристик при ограниченном объеме полевых измерений. Методологическую основу работы составили обзор и сравнительный анализ современных научных публикаций, обобщение сведений о мультиспектральных, гиперспектральных, радиолокационных и беспилотных данных, а также оценка практики применения случайного леса, метода опорных векторов, градиентного бустинга, нейронных сетей и вероятностных моделей. Показано, что наибольшая устойчивость почвенных моделей достигается при совместном использовании разнородных источников данных, корректном выделении участков открытой почвы, расчете спектральных, рельефных и климатических признаков и пространственно обоснованной проверке качества моделей. Установлено, что ансамблевые и нелинейные алгоритмы эффективны для цифрового почвенного картографирования, методы глубокого обучения перспективны при обработке пространственных и временных структур данных, а вероятностные подходы позволяют дополнительно учитывать неопределенность прогноза. Обосновано, что применение машинного обучения и дистанционного зондирования в почвоведении обеспечивает переход от точечных измерений к пространственно непрерывным оценкам свойств почв и формирует информационную основу для точного земледелия, агрохимического обследования и устойчивого управления земельными ресурсами.

Ключевые слова: почвоведение, цифровое почвенное картографирование, свойства почвы, точное земледелие

Благодарности: исследование выполнено за счет гранта Российского научного фонда № 23-74-01050.

Original article

APPLICATION OF MACHINE LEARNING AND REMOTE SENSING IN SOIL SCIENCE: CURRENT STATE AND PRACTICAL OPPORTUNITIES

S.S. Ogorodnikov^{1,2}, V.A. Blynsky²

¹Moscow Aviation Institute, Moscow, Russia

²Lomonosov Moscow State University, Moscow, Russia

Abstract. The article examines current approaches to the assessment and mapping of soil properties based on the integration of remote sensing data, ground observations, laboratory measurements, and machine learning methods. The aim of the study is to systematize data sources, preprocessing stages, and analytical algorithms used to improve the accuracy of soil property prediction under conditions of limited field measurements. The methodological basis of the study includes a review and comparative analysis of recent scientific publications, a synthesis of information on multispectral, hyperspectral, radar, and unmanned aerial vehicle data, and an assessment of the practical use of random forest, support vector machines, gradient boosting, neural networks, and probabilistic models. It is shown that the highest stability of soil models is achieved through the combined use of heterogeneous data sources, correct identification of bare soil areas, calculation of spectral, terrain, and climatic predictors, and spatially justified model validation. Ensemble and nonlinear algorithms are effective for digital soil mapping, deep learning methods are promising for processing spatial and temporal data structures, and probabilistic approaches additionally allow prediction uncertainty to be taken into account. It is substantiated that the application of machine learning and remote sensing in soil science enables a transition from point-based measurements to spatially continuous assessments of soil properties and provides an information basis for precision agriculture, agrochemical surveys, and sustainable land resource management.

Keywords: soil science, digital soil mapping, soil properties, precision agriculture

Acknowledgments: this study was supported by the Russian Science Foundation, grant No. 23-74-01050.

Введение. Почвоведение относится к числу дисциплин, в которых точечные наземные измерения остаются основой достоверной оценки свойств почвы, однако пространственная и временная полнота таких наблюдений ограничена высокой стоимостью полевых работ, неоднородностью схем отбора и недостаточной плотностью наблюдательной сети. По этой причине в последние годы одним из основных источников пространственно непрерывной

информации для почвенных исследований становятся данные дистанционного зондирования, включая мультиспектральную, гиперспектральную и радиолокационную съемку, а также материалы беспилотных и наземных сенсорных измерений. Совместное использование этих данных с результатами лабораторного анализа позволяет переходить от локальных описаний почвенных свойств к цифровому картографированию, пространственному прог-

нозированию и поддержке решений в точном земледелии.

Рост объема доступных данных привел к широкому внедрению методов машинного обучения и искусственного интеллекта в задачи оценки органического углерода почвы, влажности, гранулометрического состава, кислотности, электропроводности, плодородия, эрозионной опасности и засоления. Наиболее активно в почвенных исследованиях применяются методы

Figure 1. Generalized framework for the application of machine learning and remote sensing in soil science, developed on the basis of the literature review

Результатом проведенного литературного обзора стала обобщающая структурная схема применения машинного обучения и дистанци-

Исследование носит обзорно-аналитический характер. Методологическую основу составили анализ научной литературы, сравнительная оценка возможностей мультиспектральных, гиперспектральных, радиолокационных, беспилотных и наземных сенсоров данных, а также систематизация опубликованных результатов

Характеристика данных дистанционного зондирования в почвенных исследованиях. Особенность современного почвоведения заключается в том, что одним из основных источников пространственно непрерывных данных для исследований служат материалы дистанционного зондирования. Полевые обследования остаются необходимыми для верификации и обучения моделей, однако они требуют значительных затрат и редко формируют большие открытые базы данных с едиными схемами отбора и анализа. Поэтому в цифровом почвенном картографировании данные дистанционного зондирования используются совместно с результатами лабораторных испытаний почвенных образцов и наземных наблюдений.



В почвенных исследованиях выделяют несколько основных групп данных: спутниковые мультиспектральные изображения, гиперспектральные данные, радиолокационные изображения, материалы беспилотной съемки и результаты ближнего сенсорного обследования почвы. Эти источники различаются пространственным разрешением, временной повторяемостью, спектральной детализацией, чувствительностью к влажности и состоянию поверхности, а также пригодностью для оценки отдельных почвенных свойств.

Наиболее часто в прикладных исследованиях используются мультиспектральные данные Sentinel-2. Спутники Sentinel-2 регистрируют информацию в 13 спектральных каналах: четыре канала имеют пространственное разрешение 10 м (B2 — blue, B3 — green, B4 — red, B8 — near infrared), шесть каналов — 20 м (B5 — red edge 1, B6 — red edge 2, B7 — red edge 3, B8A — narrow near infrared, B11 — short-wave infrared 1, B12 — short-wave infrared 2), три канала — 60 м (B1 — coastal aerosol, B9 — water vapour, B10 — cirrus). Ширина полосы съемки составляет около 290 км, а период повторного покрытия в двухспутниковой конфигурации — около 5 суток. Такая конфигурация обеспечивает баланс между детализацией изображения и частотой наблюдений, поэтому Sentinel-2 широко применяют для картографирования органического углерода почвы, оценки засоления, выделения участков открытой почвы и подготовки многовременных композитов.

Данные Landsat-8 и Landsat-9 применяются преимущественно в задачах многолетнего мониторинга. Эти спутники обеспечивают съемку в видимом, ближнем инфракрасном, коротковолновом инфракрасном и тепловом диапазонах. В видимом, ближнем и коротковолновом инфракрасном диапазонах пространственное разрешение составляет 30 м, панхроматический канал имеет разрешение 15 м, а тепловые каналы — 100 м. Преимущество Landsat связано не только с характеристиками отдельных сцен, но и с длительным архивом наблюдений, который позволяет выполнять ретроспективную оценку изменений почвенного покрова, деградиационных процессов и динамики открытой поверхности почвы.

Гиперспектральные данные формируются в сотнях соседних узких спектральных каналов. Такая детализация позволяет выявлять слабо выраженные особенности спектральной отражательной способности почв, связанные с органическим веществом, карбонатами, оксидами железа, глинистыми минералами и влажностью. В современных исследованиях используются данные космических платформ EnMAP, PRISMA и EMIT, а также материалы авиационной и беспилотной гиперспектральной съемки.

Радиолокационные спутниковые данные используются в задачах, где оптическая съемка ограничена облачностью или требуется информация о влажности и состоянии поверхности почвы. Sentinel-1 работает в С-диапазоне и обеспечивает съемку независимо от облачности и освещенности. В почвенных исследованиях такие данные применяют для оценки влажности верхнего слоя почвы, шероховатости поверхности после механической обработки и состояния поверхности поля. Для регионального мониторинга влажности дополнительно используются

специализированные миссии SMAP и SMOS, однако их пространственное разрешение существенно грубее, чем у Sentinel-1.

Для полевого и внутрихозяйственного масштаба исследований применяют беспилотные летательные аппараты. Они позволяют получать изображения с пространственным разрешением от нескольких сантиметров на пиксель и использовать обычные цифровые камеры, мультиспектральные, гиперспектральные, тепловые и лазерные сенсоры. Такие данные пригодны для оценки внутриполевой неоднородности, микрорельефа, локальных зон уплотнения и подготовки решений для дифференцированного применения агротехнологий. Основное ограничение беспилотной съемки связано с небольшим охватом территории и высокой зависимостью от организации полевых работ.

Отдельную группу составляют методы ближнего сенсорного обследования почвы. Портативная рентгенофлуоресцентная съемка используется для оценки элементного состава, спектроскопия среднего инфракрасного диапазона — для оценки органического вещества, содержания глины и ряда химических свойств, а спектроскопия видимого и ближнего инфракрасного диапазонов — для оценки органического углерода, влажности и части агрохимических показателей. Эти методы связывают лабораторный анализ и дистанционную съемку, поскольку одна точка наблюдения может включать лабораторное определение свойства, спектральную характеристику образца и координаты участка.

Наземные и лабораторные данные остаются обязательной частью почвенных исследований. Спутниковая съемка обеспечивает сплошное покрытие территории, но не позволяет напрямую измерять содержание органического углерода, кислотность, емкость катионного обмена, содержание глины или электропроводность почвы. Поэтому модели обучаются на данных почвенных проб, для которых известны координаты точки отбора, глубина, дата и метод анализа. Для крупных территорий в модели дополнительно включают климатические данные, сведения о типах земного покрова, геологическом строении и показателях рельефа.

Предварительная обработка данных. Предварительная обработка данных в почвенных исследованиях включает несколько последовательных этапов. На первом этапе спутниковые изображения приводят к сопоставимому виду: выполняют атмосферную коррекцию, выравнивают яркостные значения и совмещают данные, полученные разными сенсорами. На следующем этапе выделяют участки открытой почвы и маскируют пиксели, занятые облаками, тенями, растительностью и растительными остатками. На завершающем этапе рассчитывают показатели, которые затем используются в модели: спектральные индексы, характеристики рельефа и климатические признаки. От качества этой последовательности зависит, будет ли модель обучаться по сигналу самой почвы или по смешанному сигналу почвы, растительности и поверхностной влажности.

Для оптических данных основным этапом является выделение участков открытой почвы. В пахотных ландшафтах один спутниковый снимок обычно не содержит достаточного числа таких участков, поскольку значительная часть

поверхности в течение вегетационного сезона закрыта растительностью, растительными остатками, облаками или тенями. Поэтому в последние годы для почвенных исследований чаще используют многовременные изображения открытой почвы, собранные по сериям снимков Sentinel-2 и Landsat. В исследовании Xue et al. [12] для 10-метровых данных Sentinel-2 применялась схема выделения открытой почвы на основе двух индексов. В работах по оценке органического углерода почвы и цифровому почвенному картографированию многовременные обобщенные изображения дают более устойчивый спектральный отклик, чем снимки, полученные в одну дату.

После выделения участков открытой почвы рассчитывают спектральные показатели для дальнейшего моделирования. Для этого используют значения отражения по отдельным каналам, отношения каналов и спектральные индексы. В обзорах по картографированию органического углерода почвы показано, что отдельные широкополосные индексы часто менее информативны, чем полный набор спектральных каналов, однако при малом числе почвенных проб индексы могут повышать устойчивость модели и уменьшать число признаков. В работе Xu et al. [13] новый широкий спектральный индекс для MODIS, Landsat OLI и Sentinel MSI показал корреляцию 0,72, 0,74 и 0,77 с измеренным содержанием органического углерода почвы. Поэтому индексный подход целесообразно рассматривать как базовый уровень описания данных, но не как замену многомерной модели.

Вторую обязательную группу признаков составляют характеристики рельефа. На основе цифровой модели рельефа рассчитывают абсолютную высоту, крутизну склона, экспозицию, кривизну поверхности, площадь водосбора и комплексные топографические индексы. Эти показатели используются в цифровом почвенном картографировании более 20 лет, поскольку свойства почвы изменяются по элементам рельефа вместе с условиями стока, накопления вещества и освещенности склонов. В современных исследованиях такие показатели либо включают в модель по отдельности, либо объединяют в группы, отражающие типы локального рельефа.

Для спектральных данных почвы предварительная обработка включает сглаживание шума, нормирование спектра, расчет производных и удаление фоновой составляющей. Для этого чаще всего применяют сглаживание по Савицкому-Голею [15], первую и вторую производные, стандартное нормирование и коррекцию рассеяния. Эти операции улучшают выделение спектральных признаков, связанных с органическим веществом, глиной и химическими свойствами почвы. В отдельных исследованиях такие процедуры снижали ошибку прогноза по сравнению с использованием исходных спектров без обработок.

После спектральной и пространственной обработки выполняют отбор наиболее информативных входных переменных и сокращают их число. В цифровом почвенном картографировании это необходимо из-за высокой взаимосвязи между спектральными каналами, индексами, климатическими показателями и характеристиками рельефа. При большом числе переменных используют отбор по корреляции, оценку





значимости признаков, метод главных компонент и автоматический отбор в составе модели. Цель этих процедур состоит в уменьшении риска переобучения и мультиколлинеарности при сохранении наиболее информативных признаков.

Конечный набор данных для почвенной модели обычно включает четыре группы переменных: полевые и лабораторные анализы как измеренные значения, спутниковые данные или данные ближнего сенсорного обследования как основной спектральный блок, характеристики рельефа как пространственный блок и климатические показатели как фоновые переменные. Такая схема преобладает в цифровом почвенном картографировании крупных территорий и в задачах оценки органического углерода почвы, поскольку ни один источник данных отдельно не обеспечивает одновременно детализацию порядка 10 м, информацию по глубине почвенного профиля, сезонную динамику и химическую специфичность почвы.

Методы машинного обучения и искусственного интеллекта. В обзоре Lamichhane et al. [2] по моделированию влажности почвы проанализированы 144 статьи, опубликованные в 2010-2024 годах. В 67 исследованиях наилучшие результаты чаще всего получал метод случайного леса, в 39 — метод опорных векторов для регрессии, в 27 — искусственные нейронные сети. Авторы показали, что использование данных из нескольких источников обеспечивает более высокую точность, чем применение только одного источника данных. Для оценки влажности в корнеобитаемом слое более надежными входными переменными оказались свойства почвы и влажность ее поверхностного слоя, чем спутниковые признаки.

В обзоре Padarian et al. [1], посвященном применению машинного обучения в почвоведении, показан рост числа публикаций за последние 10 лет и выделены 12 основных тематических направлений, включая дистанционное зондирование, органический углерод почвы, воду, загрязнение, спектроскопию, ансамблевые методы, нейронные сети и метод опорных векторов. Авторы пришли к выводу, что методы машинного обучения дают более точные результаты, чем простые линейные и нелинейные статистические модели, если почвенные данные имеют сложные нелинейные связи.

Традиционные методы машинного обучения. Метод случайного леса остается одним из наиболее распространенных методов цифрового картографирования почвенных свойств. Он устойчив к шуму, способен работать с большим числом входных переменных и не требует линейной связи между исходными данными и прогнозируемым свойством почвы. В сравнении 15 алгоритмов для картографирования органического углерода почвы метод случайного леса занял первое место для материковой Франции и второе место для Чехии. В этих двух случаях использовались 2514 и 400 почвенных образцов [4]. Коэффициент детерминации составил 0,411 для Франции и 0,249 для Чехии. Эти результаты показывают, что метод может быть устойчивым как при большом, так и при меньшем числе наблюдений.

В исследовании Radočaj et al. [4] метод ближайших соседей и метод градиентного бустинга дали сопоставимые результаты, но в научных

работах используются реже, чем метод случайного леса. Авторы сделали вывод о том, что эти методы недооценены и их следует чаще применять в сравнительных исследованиях. Это актуально для задач, характеризующихся небольшим количеством полевых наблюдений и значительным числом рассматриваемых переменных.

Метод опорных векторов и его регрессионный вариант часто используют в почвенной спектроскопии, особенно при небольшом числе образцов. В обзоре Shin et al. [3] по спектроскопии видимого и ближнего инфракрасного диапазонов показано, что регрессия опорных векторов и искусственные нейронные сети лучше линейных моделей описывают сложные связи между спектром и свойствами почвы. В одном из приведенных примеров для общего азота регрессия опорных векторов показала более высокую точность, чем метод частных наименьших квадратов: коэффициент детерминации составил 0,810 против 0,634, а отношение стандартного отклонения к ошибке прогноза — 2,129 против 1,838.

Метод частных наименьших квадратов сохраняет значение базовой модели в почвенной спектроскопии, поскольку устойчиво работает с большим числом спектральных переменных и дает воспроизводимые результаты на выборках среднего объема. Ограничение этого подхода связано с его линейной природой. В случаях, когда спектральный отклик одновременно зависит от влажности, гранулометрического состава и органического вещества, более сложные нелинейные методы, включая регрессию опорных векторов, искусственные нейронные сети и ансамблевые алгоритмы, часто дают более высокий коэффициент детерминации.

Глубокое обучение. Методы глубокого обучения применяются в трех основных группах почвенных задач: при анализе спектров, пространственном картографировании и обработке временных рядов. В почвенной спектроскопии сверточные нейронные сети используют для автоматического выделения информативных участков спектра, что снижает необходимость ручного отбора признаков. В обзоре Padarian et al. [1] такие модели рассматриваются как более сложные по сравнению с деревьями решений и линейными моделями. Их преимущества связаны с возможностью работать с большим числом входных переменных и анализировать изображение или спектр как связанную структуру, а не как простой набор отдельных каналов.

Сверточные нейронные сети используют в цифровом почвенном картографировании тогда, когда свойства почвы зависят не только от значения в одной точке, но и от соседних участков. В работе Padarian et al. [5] показано, что учет окружающего пространства позволяет уменьшить ошибку прогноза примерно на 30% по сравнению с моделями, которые используют только значения входных переменных в самой точке. При картографировании с разрешением 100 м наилучшие результаты были получены при анализе соседних участков в окне от 3 до 9 пикселей. Это соответствует площади примерно 300-900 м вокруг точки прогноза и показывает, что такие модели лучше учитывают пространственные связи между почвой и условиями рельефа.

Рекуррентные нейронные сети применяются для процессов, которые меняются во времени и зависят от последовательности наблюдений. В обзоре Lamichhane et al. [2] такие модели названы перспективными для повышения устойчивости моделей влажности почвы при переносе на новые условия. В глобальном исследовании AW-LSTM при прогнозе влажности почвы коэффициент корреляции составил 0,9456, а коэффициент детерминации — 0,8489. Эти результаты показывают, что модели типа LSTM хорошо подходят для сложных временных рядов, в которых на влажность почвы одновременно влияют осадки, температура, испарение и растительность.

Модели типа LSTM особенно полезны для оценки влажности почвы на нескольких глубинах и для прогноза на несколько шагов вперед. В работе Azad et al. [6] этот тип модели использовали для оценки влажности почвы по профилю, а не только в верхнем слое 0-5 см. В исследовании были задействованы данные на глубинах 5, 10, 20, 50 и 100 см, что показало пригодность таких моделей для описания вертикального распределения влаги в почве.

Вероятностные и статистические методы. Регрессия на основе гауссовских процессов используется в тех случаях, когда требуется получить не только одно прогнозируемое значение, но и оценку его неопределенности. В исследовании Wang et al. [7] по трехмерному почвенному картографированию этот метод описан как модель, включающая среднюю функцию и ядро пространственной зависимости. В отличие от многих других методов, он позволяет получать не только точечные оценки, но и распределение вероятных значений. На этой основе авторы построили трехмерные карты кислотности и электропроводности почвы для фермерского участка в Новом Южном Уэльсе и показали, что такой метод позволяет прогнозировать свойства не только в отдельных точках, но и для объемных участков, а также учитывать данные с нескольких глубин.

Для почвенных исследований этот метод дает два основных преимущества. Первое состоит в том, что он позволяет оценивать не только само прогнозируемое значение, но и степень его надежности в каждой точке или на каждом участке. Второе связано с тем, что глубину можно учитывать как непрерывно изменяющуюся величину, а не только как набор отдельных слоев. В задачах трехмерного почвенного картографирования это делает метод удобным для оценки кислотности, электропроводности и ряда гидрофизических свойств, особенно в тех случаях, когда измерения выполнены по почвенным горизонтам или на неравных глубинах.

Байесовские сети используются в почвенных исследованиях в тех случаях, когда часть информации задается измерениями, а часть — экспертными знаниями. В работе Hassall et al. [8] такая модель применена для количественной оценки качества и состояния почвы. Авторы опирались на экспертные представления о функциях почвы, измеряемые свойства и данные длительных полевых опытов продолжительностью 100 лет и более. Такой подход полезен тогда, когда требуется связать способы землепользования, функции почвы и вероятности перехода почвенной системы из одного состояния в другое, а не только получить одно численное значение прогноза.



Методы с размытыми границами между классами применяют тогда, когда границы между классами заданы нечетко и бинарное деление на пригодные и непригодные участки оказывается слишком грубым. В исследовании Venkateswarlu et al. [9] модель включала 80 правил и 250 почвенных образцов. В нее вошли кислотность, электропроводность вытяжки, органический углерод, азот, фосфор, калий, сера, железо, марганец, цинк, бор и медь. Итоговые оценки плодородия составили 41,55-52,60. Значения пригодности по урожайности составили 0,919 для проса жемчужного, 0,890 для горчицы, 0,863 для пшеницы и 0,861 для ячменя. Это показывает, что такие методы подходят для комплексной оценки плодородия почвы по совокупности показателей.

Ансамблевые методы и бустинг. После 2015 года ансамблевые методы заняли одно из ведущих мест в цифровом почвенном картографировании. В сравнении 15 алгоритмов для картографирования органического углерода почвы метод случайного леса показал наилучший результат, а метод градиентного бустинга вошел в число моделей, эффективность которых оказалась выше, чем можно было ожидать по частоте их применения в публикациях. Это показывает, что методы, основанные на объединении нескольких моделей, сегодня обеспечивают одно из наиболее удачных сочетаний точности, устойчивости результатов и приемлемой вычислительной сложности при работе с табличными экологическими данными.

Методы XGBoost, LightGBM и CatBoost чаще используют в тех случаях, когда число входных переменных исчисляется десятками, а связи между ними имеют сложный нелинейный характер. В исследовании Alonso-Sarria et al. [10] при оценке содержания органического углерода, глины, ила и песка метод XGBoost входил в набор из четырех моделей вместе с методом случайного леса, методом опорных векторов и многослойной нейронной сетью. Авторы отдельно отметили, что подбор настроек модели может повышать точность прогноза, но одновременно увеличивает риск переобучения и снижает устойчивость модели при переносе на новые данные. Для почвенных данных с высокой пространственной связностью этот риск особенно велик, если использовать обычную перекрестную проверку вместо пространственной.

Практические возможности применения результатов. В прикладном отношении рассмотренные методы позволяют переходить от разрозненных полевых, лабораторных и дистанционных данных к цифровым картам свойств почвы, схемам внутривополевого зонирования, оценке пространственного распределения плодородия и влажности, а также к обоснованию дифференцированного внесения удобрений, полива и других агротехнологических мероприятий.

Информация об авторах:

Огородников Сергей Сергеевич, кандидат биологических наук, доцент, Московский авиационный институт, sir.ogorod@yandex.ru
Блынский Виктор Анатольевич, аспирант, МГУ имени М.В. Ломоносова

Information about the authors:

Sergey S. Ogorodnikov, candidate of biological sciences, associate professor, Moscow Aviation institute, Moscow, Russia, sir.ogorod@yandex.ru
Victor A. Blynsky, postgraduate student, Lomonosov Moscow State University

Для цифрового почвенного картографирования на уровне района, области или страны в качестве основных моделей сравнения чаще всего используют метод случайного леса, XGBoost и метод опорных векторов. В обзорах 2024-2026 годов именно эти методы чаще других показывали близкую по уровню и достаточно высокую точность. Для спектроскопии видимого, ближнего и среднего инфракрасного диапазонов в базовый набор обычно включают метод частных наименьших квадратов, регрессию опорных векторов, метод случайного леса и искусственные нейронные сети. Метод частных наименьших квадратов используют как устойчивую исходную модель, а более сложные нелинейные методы позволяют лучше описывать спектральные зависимости.

Для временных рядов влажности почвы и изменений состояния поля чаще выбирают модели типа LSTM. Для задач, где важно оценить не только прогноз, но и степень его надежности, в набор сравниваемых методов обычно включают регрессию на основе гауссовских процессов и байесовские модели. В результате выбор алгоритма должен определяться не только формальной точностью на обучающей выборке, но и типом исходных данных, пространственным масштабом задачи, числом полевых наблюдений, требуемой интерпретируемостью и возможностью переноса модели на новые территории.

Выводы.

1. Современные почвенные исследования основаны на совместном использовании спутниковых, радиолокационных, беспилотных, сенсорных, наземных и лабораторных данных, поскольку ни один источник по отдельности не обеспечивает одновременно высокую пространственную детализацию, сведения о глубине почвенного профиля, сезонную динамику и химическую специфичность почвы.

2. Точность прогноза определяется не только выбором алгоритма, но и качеством предварительной обработки данных, включая выделение участков открытой почвы, расчет спектральных, рельефных и климатических показателей, а также отбор наиболее информативных входных переменных.

3. Наиболее устойчивые результаты в задачах цифрового почвенного картографирования и оценки свойств почвы дают метод случайного леса, метод опорных векторов, методы градиентного бустинга, модели глубокого обучения для пространственных и временных данных, а также вероятностные подходы, если требуется оценка надежности прогноза.

Список источников, References

1. Padarian J., Minasny B., McBratney A.B. Machine learning and soil sciences: a review aided by machine learning tools // SOIL. 2020. Vol. 6. P. 35-52. DOI: 10.5194/soil-6-35-2020.

2. Lamichhane M., Mehan S., Mankin K.R. Soil Moisture Prediction Using Remote Sensing and Machine Learning Algorithms: A Review on Progress, Challenges, and Opportunities // Remote Sensing. 2025. Vol. 17. No. 14. Article 2397. DOI: 10.3390/rs17142397.

3. Shin S.K., Lee S.J., Park J.H. Prediction of Soil Properties Using Vis-NIR Spectroscopy Combined with Machine Learning: A Review // Sensors. 2025. Vol. 25. No. 16. Article 5045. DOI: 10.3390/s25165045.

4. Radočaj D., Jug D., Jug I., Jurišić M. A Comprehensive Evaluation of Machine Learning Algorithms for Digital Soil Organic Carbon Mapping on a National Scale // Applied Sciences. 2024. Vol. 14. No. 21. Article 9990. DOI: 10.3390/app14219990.

5. Padarian J., McBratney A.B., Minasny B. Game theory interpretation of digital soil mapping convolutional neural networks // SOIL. 2018. Vol. 4. P. 389-397. DOI: 10.5194/soil-4-389-2018.

6. Azad N., Sheikhabglou A., Zvomuya F., He H., Moulin A.P., Zeleke T.B., Gameda S. Applying LSTM to Model Multi-Depth Soil Moisture Under Various Land Covers, Climates and Soils // European Journal of Soil Science. 2025. DOI: 10.1111/ejss.70142.

7. Wang J., et al. Gaussian process regression for three-dimensional soil mapping over multiple spatial supports // Geoderma. 2024. Vol. 446. Article 116899.

8. Hassall K.L., Zawadzka J., Milne A.E., Corstanje R., et al. Putting numbers to a metaphor: A Bayesian Belief Network with which to infer Soil Quality and Health // European Journal of Agronomy. 2025. Vol. 168. Article 127537. DOI: 10.1016/j.eja.2025.127537.

9. Venkateswarlu M., Rallapalli S., Singh A., Chalapati G.S.S., Kumar S., Katpatal Y.B., Sujatha G. Macro and micronutrient based soil fertility zonation using fuzzy logic and geospatial techniques // Scientific Reports. 2025. Vol. 15. Article 24689. DOI: 10.1038/s41598-025-12184-3.

10. Alonso-Sarria F., Blanco-Bernardeau A., Romero-Diaz A. Estimation of soil properties using machine learning techniques to improve hydrological modeling in a semiarid environment: Campo de Cartagena (Spain) // Earth Science Informatics. 2025. DOI: 10.1007/s12145-025-01833-w.

11. Bohn M.P., Miller B.A. Digital soil mapping via machine learning of agronomic soil survey properties from 0-200 cm in depth in Illinois // Agronomy Journal. 2025. Vol. 117. No. 5. Article e70144. DOI: 10.1002/agj2.70144.

12. Xue J., et al. A two-dimensional bare soil separation framework using multi-temporal Sentinel-2 images across China // International Journal of Applied Earth Observation and Geoinformation. 2024. Vol. 134. Article 104181.

13. Xu Z., Hou D., Lin N., et al. Satellite remote sensing enables monitoring of soil organic carbon decline in croplands of Jilin, China // Scientific Reports. 2026. Vol. 16. Article 6966. DOI: 10.1038/s41598-026-38386-x.

14. Thompson J.A., Bell J.C., Butler C.A. Digital elevation model resolution: effects on terrain attribute calculation and quantitative soil-landscape modeling // Geoderma. 2001. Vol. 100. No. 1-2. P. 67-89. DOI: 10.1016/S0016-7061(00)00081-1.

15. Savitzky A., Golay M.J.E. Smoothing and Differentiation of Data by Simplified Least Squares Procedures // Analytical Chemistry. 1964. Vol. 36. No. 8. P. 1627-1639. DOI: 10.1021/ac60214a047.

