

УДК 656.6:626.1:004.8

ПРЕДИКТИВНОЕ МОДЕЛИРОВАНИЕ УРОВНЯ ВОДЫ В РАВНИННЫХ РЕКАХ ДЛЯ ПОВЫШЕНИЯ БЕЗОПАСНОСТИ ОБЪЕКТОВ ТРАНСПОРТНОЙ ИНФРАСТРУКТУРЫ

Муконина Мария Ильинчна, старший преподаватель; e-mail: mukonina_m@bk.ru

Кафедра «Вычислительная техника и автоматизированные системы управления», Ростовский государственный университет путей сообщения, Ростов-на-Дону

В статье рассматривается задача повышения безопасности транспортной инфраструктуры посредством разработки интеллектуальной имитационной модели. Основной фокус сделан на прогнозировании уровня воды в равнинных реках для предупреждения внештатных ситуаций: подтоплений, обрушений мостов и других чрезвычайных происшествий, способных парализовать транспортное сообщение. Актуальность исследования определяется совокупностью факторов: нарастающими климатическими рисками (паводки, ливни, наводнения), техногенными угрозами, а также ограниченной эффективностью традиционных подходов к мониторингу и прогнозированию гидрологической обстановки. В ходе исследования проведен анализ реальных чрезвычайных ситуаций, наглядно демонстрирующих острую потребность в оперативном прогнозировании уровня воды. Методология разработки системы поддержки принятия решений базируется на технологиях машинного обучения. Экспериментальная часть исследования выполнена на основе данных по р. Темерник. Было протестировано 12 моделей машинного обучения, проведена оценка качества прогнозирования на основе статистических метрик, и выбраны наилучшие модели для дальнейшего использования.

Ключевые слова: транспортная инфраструктура; чрезвычайные ситуации; мониторинг уровня воды; интеллектуальное имитационное моделирование; система поддержки принятия решений; машинное обучение; прогнозирование уровня воды; AutoGluon; NeuralNetFastAI; WeightedEnsemble_L2; равнинные реки; управление рисками.

DOI: 10.20295/2412-9186-2025-11-04-342-354

▼ Введение

Надежность работы транспортной отрасли в значительной степени зависит от правильности и оперативности принятых управленческих решений. Особенно ярко это проявляется на фоне быстро изменяющегося климата, переменчивой политической обстановки, а также увеличения количества происшествий техногенного характера.

Ответом на вышеперечисленные вызовы может стать усовершенствование процесса принятия решений на основе интеллектуальной обработки данных. Предиктивное моделирование — это метод, позволяющий анализировать и прогнозировать возможные сценарии развития событий, что критически важно для обеспечения безопасности и эффективности работы транспортной инфраструктуры. Оно включает в себя разработку компьютерных моделей, способных воспроизводить поведение сложных систем в условиях внештатных

событий. Эти модели учитывают множество факторов: технические характеристики объектов, погодные условия, действия персонала и другие переменные.

В отечественной науке вопросу разработки имитационных моделей, описывающих поведение различных производственных и бизнес-процессов, посвящено большое количество научных работ. В статье [1] Е. В. Арефьева и А. В. Рыбаков предлагают инновационный подход к прогнозированию опасностей, основанный на использовании постоянно действующей ситуационно-оптимизационной модели. В отличие от разовых статических моделей, их система работает в режиме реального времени, непрерывно актуализируя данные и адаптируясь к изменениям. Ключевая идея метода заключается в том, что модель не только анализирует текущее состояние объекта, но и активно ищет оптимальные решения для предотвращения угроз. Это позволяет перейти

от простого реагирования к заблаговременному предупреждению рисков в таких областях, как промышленная и экологическая безопасность, энергетика и транспорт.

Предметом исследования статьи [2] является комплексная проблема прогнозирования и моделирования зон затопления на территории Северного Казахстана. Основное внимание уделяется оценке рисков наводнений с применением современных цифровых моделей рельефа, что позволяет получить детальное представление о потенциальных зонах подтопления. Методологическая база исследования построена на применении современных научных методов: гидрологическом цифровом прогнозном моделировании, детальных цифровых моделях рельефа, анализе космических снимков и батиметрической съемке. Ограничения исследования связаны с необходимостью регулярного обновления данных, зависимостью от качества исходных материалов и потребностью в дальнейших исследованиях для уточнения моделей.

Статья [3] посвящена актуальной проблеме обнаружения знаний в данных, где основное внимание уделяется интеграции методов data mining и машинного обучения. Авторы фокусируются на разработке комплексных подходов к извлечению полезной информации из больших массивов данных. Исследователи ставят перед собой задачу систематизировать и развить существующие методы обнаружения знаний в данных. Особое внимание уделяется созданию эффективных алгоритмов, которые можно успешно применять в различных практических областях.

Работа [4] посвящена разработке метода краткосрочного прогнозирования водного режима рек и водохранилищ Донбасса с использованием модели Брауна. Актуальность работы обусловлена тем, что традиционные трендовые регрессионные модели демонстрируют высокую погрешность при прогнозировании нестационарных и персистентных временных рядов, характерных для гидрологических процессов. В качестве решения авторы предлагают адаптивную модель Брауна — разновидность модели скользящего среднего, которая требует малого количества наблюдений, способна

адаптироваться к изменениям исходных данных и учитывает взвешенность предыдущих наблюдений. Проведенное исследование показало, что наименьшая погрешность прогнозирования (7–19 %) достигается при значениях коэффициента сглаживания $\alpha = 0$ и $\alpha = 0,25$, тогда как при $\alpha = 1,0$ погрешность возрастает до 63 %. Практическая значимость метода заключается в возможности оперативного прогнозирования изменений водного режима, что особенно важно для своевременного принятия мер по защите от паводков и регулированию водных ресурсов при условии правильного подбора коэффициента сглаживания.

На основе проведенного анализа можно сделать вывод, что выбранная тематика исследования является достаточно актуальной, так как позволяет произвести оценку реального состояния объекта без затрат на практические эксперименты. Необходимо отметить, что все исследуемые научные работы содержат либо теоретический характер, описывая процесс моделирования с точки зрения методологии, либо содержат разработку моделей для узконаправленной отрасли.

Целью исследования является разработка предикативной модели поведения равнинных рек для поддержки принятия управленческих решений в сфере обеспечения безопасности транспортной инфраструктуры с целью предотвращения внештатных ситуаций.

1. Актуальность задачи мониторинга уровня воды

Наиболее уязвимыми местами, происшествие на которых может повлечь за собой остановку движения транспорта на длительный срок [5], а также большие финансовые затраты на его возобновление — это мосты через реки. Так, 2 июня 2020 г. река Кола вышла из берегов [6] и подтопила опоры железнодорожного моста, что привело к его обрушению (рис. 1).

В Воронежской области во время паводка 2018 г. затопило 22 моста, что привело к серии обрушений мостов. Мост через р. Савала (рис. 2) упал во время проезда грузовика [7], что является следствием сочетания перегруза и ослабления опор из-за размыва.



Рис. 1. Мост через реку Кола, июнь 2020 г.
(фото: Мурманский вестник / Лев Федосеев)



Рис. 2. Мост через реку Савала во время паводка, 2018 г.
(фото: «РИА Воронеж» / Жанна Асанова)

Однако мосты — не единственные объекты транспортной инфраструктуры, подверженные негативному влиянию природных факторов. За последние 10 лет в России произошло большое количество случаев затопления железнодорожных вокзалов и станций из-за паводков, ливней и наводнений. Затопление пути и вокзальных помещений привело к остановке движения поездов на несколько часов, а общий ущерб составил 760 млн рублей.

Так, аномальные дожди в Иркутской области вызвали разливы рек Ангара и Иркут, а также затопление станций в Тулуне

и Нижнеудинске. В результате было прервано движение по Транссибирской магистрали.

В июне 2020 г. ливни в городе Ростове-на-Дону привели к подтоплению железнодорожного вокзала Ростов-Главный [8] (рис. 3).

Все вышеперечисленные примеры свидетельствуют об актуальности задачи мониторинга уровня воды в равнинных реках вблизи критически важных объектов транспортной инфраструктуры для возможности принятия предупредительных мер, направленных на снижение материального ущерба и исключение человеческих жертв.



Рис. 3. Перрон вокзала Ростов-Главный, июнь 2020 г.
(скриншот из видео: Mail.ru / Сергей Заблоцкий)

2. Методология исследования

Машинное обучение играет важнейшую роль в разработке методик прогнозирования уровня воды в реках. Используя различные алгоритмы и методы обработки данных, модели машинного обучения могут анализировать исторические данные об уровне воды в реке и предсказывать будущие уровни с высокой точностью и достоверностью.

Эти модели обладают способностью распознавать закономерности в больших массивах данных, что делает их подходящими для динамического характера гидрологических данных [9]. Для выбора подходящей модели машинного обучения был использован фреймворк AutoGluon.

AutoGluon может автоматически выполнять предварительную обработку данных, включая очистку данных, выбор признаков и операции по проектированию признаков (рис. 4). С помощью класса TabularPredictor фреймворк способен производить автоматическое обучение модели машинного обучения.

В данном исследовании AutoGluon-Tabular предлагается использовать для построения 12 моделей прогнозирования уровня воды в реке (LightGBMX, LightGBM, LightGBM-Large, RandomForestMSE, ExtraTreesMSE, CatBoost, NeuralNetTorch, XGBoost, NeuralNet-FastAI, WeightedEnsemble_L2, KNeighborsUnif, KNeighborsDist).

Для оценки эффективности предложенного подхода используются показатели стандартных метрик оценки качества предиктивной модели. Эти показатели включают среднеквадратичную ошибку (MSE), среднеквадратичную ошибку (RMSE), среднюю абсолютную ошибку (MAE), средние абсолютные процентные ошибки (MAPE). В этом исследовании была выбрана система с самыми низкими показателями RMSE и MAE [10].

3. Подготовка данных для обучения интеллектуальной модели прогнозирования

Источником экспериментальных данных будут являться метеорологические сводки и данные с датчиков измерения уровня воды.

В рамках данного исследования реализован подход, ориентированный не на прогнозирование поведения общего речного стока, а на изучение локального изменения уровня воды в рамках выбранной мостовой конструкции равнинной реки с целью применения полученных данных расчета нагрузок на опоры, позволяющего обеспечить безопасную эксплуатацию сооружения. Выбранная для исследования точка также располагается вблизи железнодорожного вокзала г. Ростова-на-Дону, что позволит не только исследовать влияние уровня воды на конструкцию мостового сооружения, но и производить

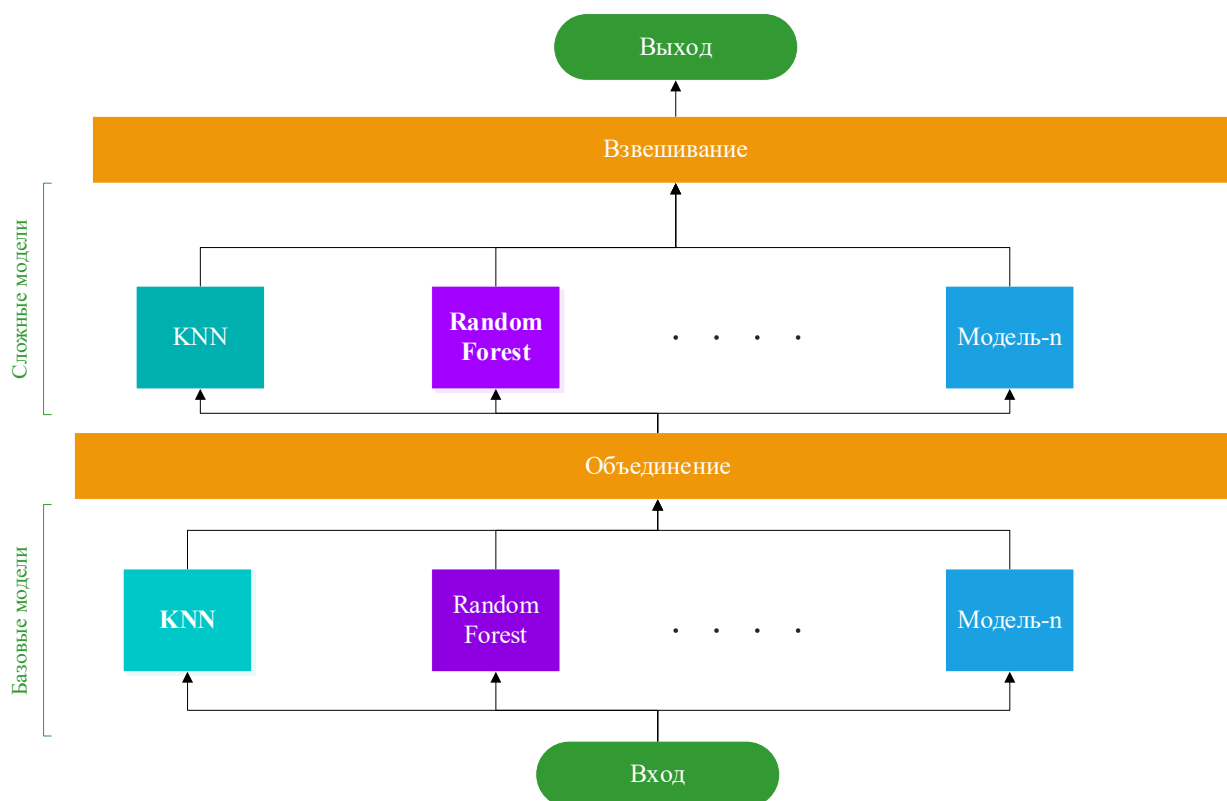


Рис. 4. Принцип работы фреймворка AutoGluon

своевременный мониторинг возникновения чрезвычайных ситуаций, связанных с затоплением железнодорожных путей.

Для проведения анализа возможностей цифровых сервисов, предоставляющих сведения о погодных условиях, в качестве источника данных для обучения модели был выбран веб-ресурс gr5.ru и выгружен архив, содержащий в себе данные о погодных условиях с 2021 по 2024 г. в Ростове-на-Дону.

Для мониторинга изменения уровня воды в реке была использована система мониторинга параметров окружающей среды «Эмерсит», в которой содержится информация со специализированных датчиков измерения воды в различных реках. Для проведения эксперимента была выбрана р. Темерник — равнинная река в Ростовской области, правый приток Дона. Ее длина составляет 33 км, из которых 18 км проходят по территории Ростова-на-Дону. Площадь водосборного бассейна — 293 км². Средний уклон реки — 2,3 %. Ширина русла в среднем до 10 м, глубина — 0,3–0,8 м [11].

Данная река неоднократно становилась причиной чрезвычайных ситуаций на объектах

транспортной инфраструктуры в г. Ростове-на-Дону. Для мониторинга уровня воды в р. Темерник была выполнена выгрузка данных за трехлетний период с датчика «ЭМЕРСИТ-0239», расположенного по координатам: 47.214991° с. ш., 39.698011° в. д. Датчик имеет следующие контрольные уровни:

- уровень датчика — 20,650 м;
- неблагоприятное явление (Н/Я) — 3,400 м;
- опасное явление (О/Я) — 3,900 м.

После выделения ключевых параметров был проведен их детальный анализ с целью выявления аномалий и определения необходимости дополнительной обработки данных (рис. 5–10).

После проведения анализа графиков распределения величин можно сделать следующие выводы:

1. Распределение уровня воды (*BSV*) имеет асимметричный вид с пиком в области низких значений уровня воды, в свою очередь, весь диапазон значений включает в себя величины от 0,3 до 20 ед. в БСВ. При этом выбросы в диапазоне до 7,5 можно считать естественными, вызванными обильным уровнем осадков или ветровыми явлениями. Выбросы, которые

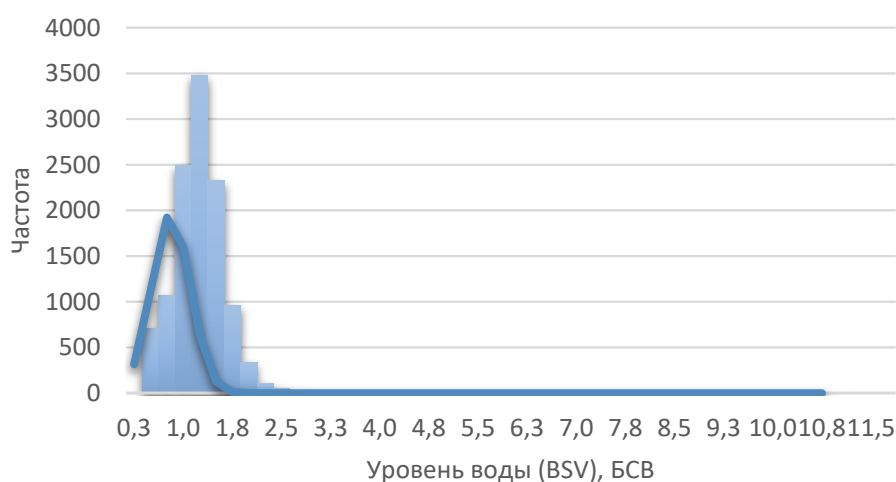


Рис. 5. Распределение уровня воды в балтийской системе высот

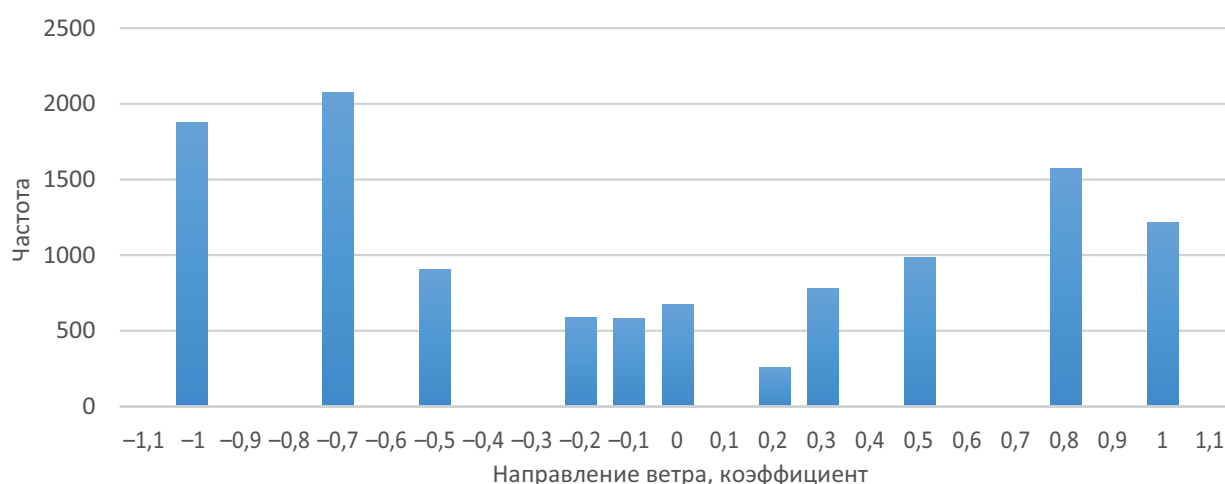


Рис. 6. Распределение величины коэффициента, указывающего направление ветра

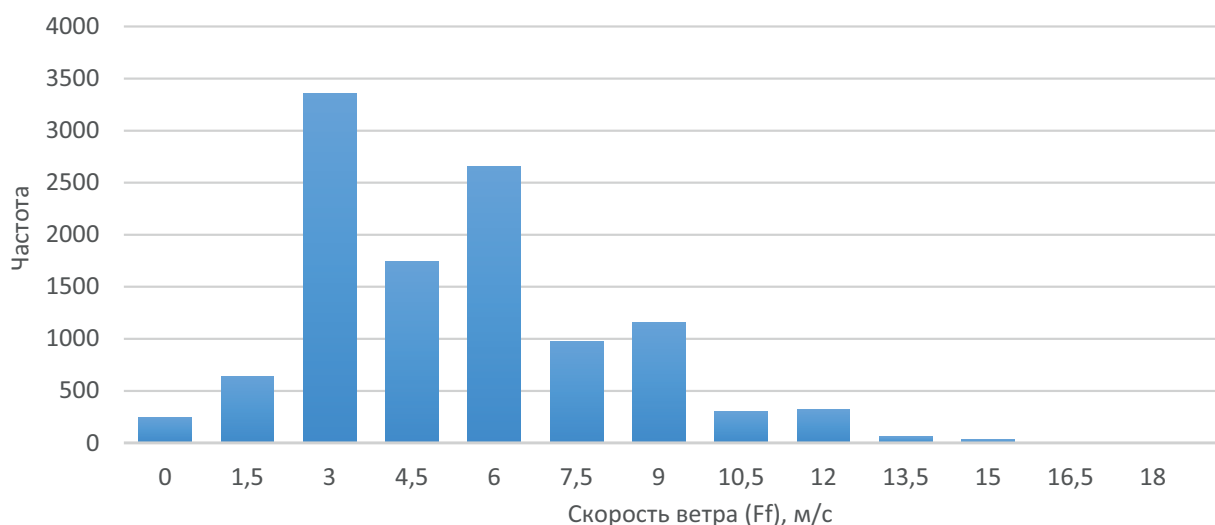


Рис. 7. Распределение величины скорости ветра

имеют значительно больший диапазон, необходимо будет удалить из выборки, так как они были получены вследствие некорректной работы измеряющего датчика.

2. При изучении гистограммы, отражающей распределение переменной DD (направление ветра), можно отметить, что форма распределения носит многопиковый (модальный)

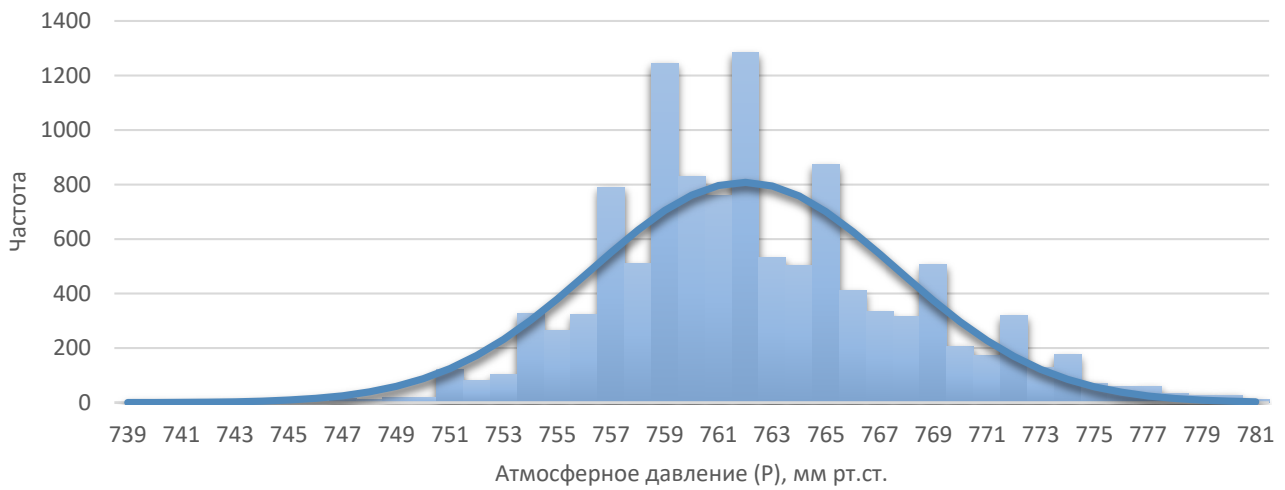


Рис. 8. Распределение величины атмосферного давления

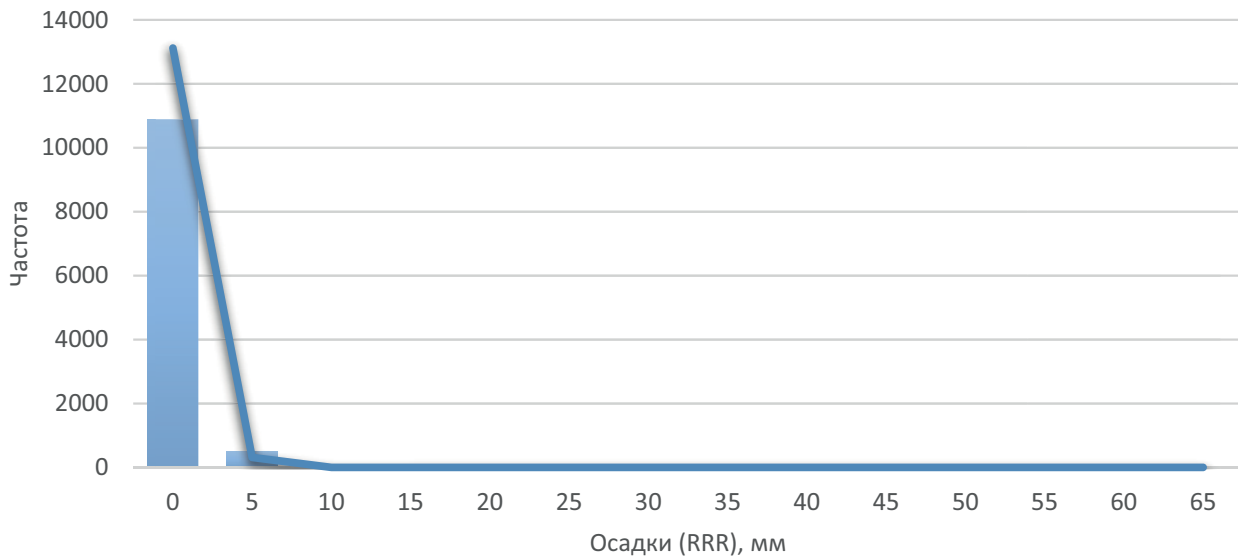


Рис. 9. Распределение уровня осадков

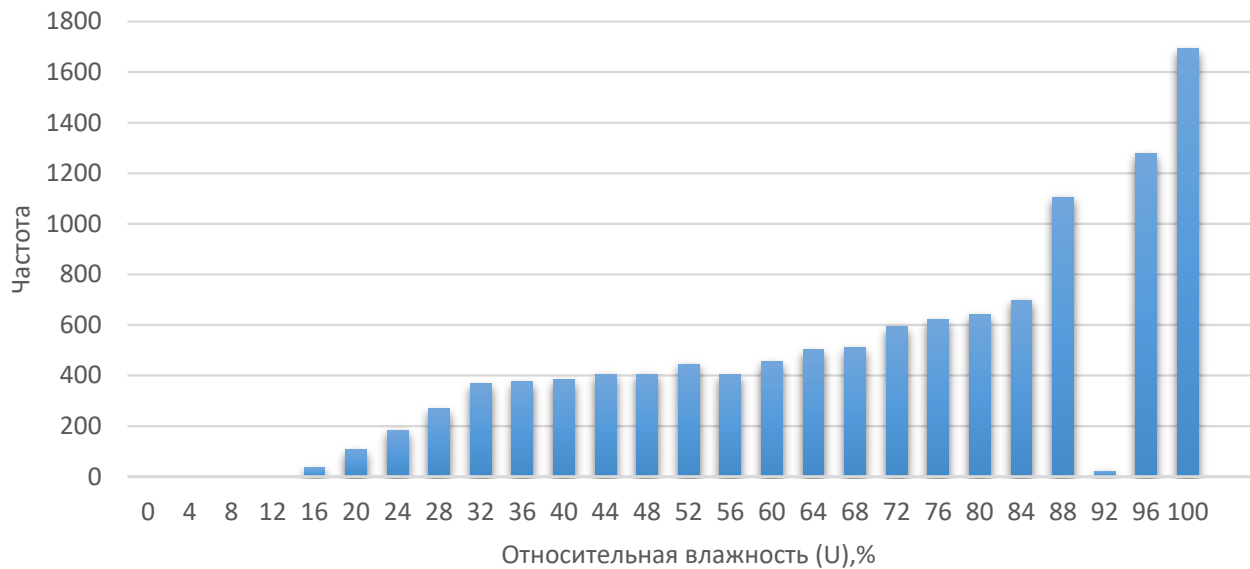


Рис. 10. Распределение величины относительной влажности

характер. Это свидетельствует о том, что в исследуемой области преобладают несколько ключевых направлений ветра — ветер чаще всего дует с определенных румбов. Диапазон значений переменной DD охватывает интервал от $-1,00$ до $1,00$, так как категориальные данные были преобразованы в числовые значения на основе анализа физических аспектов влияния направления ветра на формирование уровня воды на выбранном участке реки, а также расчета статистических характеристик для удобства анализа. Распределение значений скорости ветра имеет явный перекося вправо, что достаточно типично для данной величины.

3. Распределение скорости ветра (Ff) носит модальный (многопиковый) характер, что говорит о неравномерности ветровой активности: определенные скорости ветра заметно преобладают над другими. Это связано с локальными климатическими особенностями. Самый значительный пик наблюдается около значения $2,5$ м/с — такая скорость ветра встречается наиболее часто (частота ≈ 1750 наблюдений). Это указывает на преобладающий «фоновый» ветер в исследуемой локации. Большинство наблюдений сосредоточено на низких и умеренных скоростях ветра, а сильные порывы являются исключением.

4. Анализ гистограммы атмосферного давления (P) показал, что в исследуемой выборке наблюдается достаточно стабильная картина: центральное значение давления сосредоточено около отметки 760 мм рт. ст., что соответствует уровню нормального атмосферного давления. При этом экстремальные значения фиксируются достаточно редко, что свидетельствует о незначительной вариативности показателей в выборке. Структура распределения давления близка к нормальному распределению (симметричная «колоколообразная» форма, значения постепенно убывают от центра к «хвостам» графика). Такая закономерность типична для метеорологических наблюдений в умеренных широтах, где атмосферные условия в целом отличаются относительной стабильностью.

5. Основная масса наблюдений в гистограмме уровня осадков (RRR) сконцентрирована в нижней части диапазона — вблизи нулевой отметки. Это свидетельствует о том,

что преобладающее число измерений фиксирует либо полное отсутствие осадков, либо их крайне незначительный объем. Этот факт наглядно демонстрирует доминирование «сухих» периодов в исследуемой выборке. Еще одной важной характеристикой распределения является его асимметричность. Распределение явно скошено влево. Кроме того, распределение отличается отсутствием равномерности. На гистограмме не прослеживается наличие промежуточных «пиков» — после начального всплеска частоты на минимальных значениях RRR наблюдается практически мгновенный обрыв. Это подчеркивает крайнюю неравномерность распределения: доминирует узкий диапазон низких значений, а переходы к более высоким уровням осадков происходят крайне редко.

6. Распределение величины относительной влажности (U) явно скошено вправо в сторону высоких значений. Медиана наблюдений составляет примерно 80% и располагается ближе к верхнему квартилю. В целом распределение носит явно асимметричный характер, при этом выбросы практически отсутствуют.

Помимо анализа графика распределений был проведен статистический анализ входных данных, результаты которого представлены в табл. 1.

Необходимо отметить, что в современных условиях уровень воды в речных системах зачастую определяется не столько локальными погодными факторами, сколько антропогенным регулированием водного стока посредством гидротехнических сооружений. В связи с этим помимо данных о погоде в систему был добавлен динамический поправочный коэффициент, учитывающий водосброс на Цимлянской ГЭС (табл. 2).

4. Обучение и тестирование моделей на основе машинного обучения

В качестве средства разработки интеллектуальных моделей был выбран язык Python [12]. Выбор данного языка обусловлен наличием в его структуре большого количества библиотек и фреймворков, которые позволяют разрабатывать высокопроизводительные интеллектуальные модели на основе ML.

Таблица 1. Статистический анализ данных для ML-модели

| | date | P, мм рт. ст. | U, % | DD, коэф. | Ff, м/с | RRR, мм | BSV, БСВ |
|-------|---------------------|---------------|--------|-----------|---------|----------|----------|
| count | 11503 | 11503 | 11503 | 11503 | 11503 | 11503 | 11503 |
| mean | 2023-01-31 18:14:55 | 761,96 | 70,24 | -0,086 | 4,81 | 0,233737 | 1,123 |
| min | 2021-02-01 00:00:00 | 739,60 | 12,00 | -1,00 | 0,00 | 0,00 | 0,336 |
| 25 % | 2022-01-31 16:30:00 | 758,20 | 51,00 | -0,75 | 3,00 | 0,00 | 0,882 |
| 50 % | 2023-01-28 03:00:00 | 761,20 | 75,00 | -0,10 | 4,00 | 0,00 | 1,105 |
| 75 % | 2024-02-08 16:30:00 | 765,00 | 93,00 | 0,50 | 6,00 | 0,00 | 1,335 |
| max | 2025-02-09 15:00:00 | 781,60 | 100,00 | 1,00 | 17,00 | 60,00 | 19,650 |
| std | NaN | 5,67 | 23,47 | 0,69 | 2,73 | 1,74 | 0,586 |

Таблица 2. Зависимость корректирующего коэффициента от величины водосброса

| Водосброс, м ³ /с | 80 | 180 | 280 | 380 | 480 | 580 | 680 | 780 | 880 | 980 | 1080 | 1180 |
|------------------------------|-------|-------|-----|------|------|------|------|-----|------|------|------|------|
| Коэффициент | -0,12 | -0,06 | 0 | 0,06 | 0,12 | 0,18 | 0,24 | 0,3 | 0,36 | 0,42 | 0,48 | 0,54 |

Для построения моделей машинного обучения на основе фреймворка AutoGluon в проект был импортирован модуль *autogluon.tabular* [13].

Следующим этапом разработки является подготовка данных. На основе проведенного выше анализа были выбраны алгоритмы, позволяющие нормализовать исходный массив данных с целью обеспечения корректной работы будущей модели (корректировка выбросов, заполнение пропусков, типизация). Для учета сезонных зависимостей было произведено преобразование абсолютных временных меток в набор дискретных сезонных признаков, позволяющих алгоритмам машинного обучения распознавать и учитывать циклические закономерности в данных.

Необходимо учитывать временные задержки между выпадением осадков и их воздействием на целевую переменную, поскольку возникает эффект временного запаздывания. Это особенно критично для построения адекватных прогнозных моделей в гидрологии. В связи с этим в предобработку данных была включена операция создания лагов посредством операции сдвига временного ряда с использованием метода *shift(n)*. При 3-часовом интервале наблюдений были выбраны коэффициенты лагирования, равные 8 и 16, что соответствуют 24-часовой и 48-часовой задержкам соответственно.

В связи с тем, что величина осадков в большинстве случаев принимает нулевое значение, для улучшения интерпретируемости признака

был добавлен бинарный индикатор, фиксирующий сам факт наличия или отсутствия осадков в рассматриваемый период.

Для учета накопленного воздействия осадков на целевую переменную был применен метод скользящих статистик, позволяющий анализировать кумулятивное влияние. Техническая реализация предполагает вычисление сумм атмосферных осадков за различные временные периоды с использованием метода *rolling().sum()*, где параметр *window* определяет глубину анализируемого временного горизонта.

На практике использовались два основных временных масштаба анализа. Скользящая сумма за 24 часа (*window=8*), отражающая краткосрочное воздействие осадков и являющаяся критически важным параметром для оценки суточного водного баланса и оперативного прогнозирования гидрологической обстановки. Семидневный горизонт (*window=56*), предоставляющий информацию о среднесрочных тенденциях увлажнения, позволяет оценить степень насыщения водосборных бассейнов и формирование устойчивых гидрологических трендов.

Важным шагом в процессе обучения интеллектуальной модели является деление исходного массива данных на обучающую и тестовую выборки. Он необходим для обеспечения объективной оценки работы модели. В нашем исследовании исходная выборка данных была разделена в процентном соотношении 80 на 20.

Таблица 3. Расчет метрик качества обученных ML-моделей

| Model | MSE | RMSE | MAE | R^2 | Correlation |
|---------------------|-------|-------|-------|--------|-------------|
| NeuralNetFastAI | 0,028 | 0,166 | 0,130 | 0,692 | 0,881 |
| WeightedEnsemble_L2 | 0,028 | 0,167 | 0,131 | 0,687 | 0,863 |
| LightGBMLarge | 0,029 | 0,171 | 0,145 | 0,674 | 0,916 |
| XGBoost | 0,030 | 0,172 | 0,141 | 0,668 | 0,889 |
| ExtraTreesMSE | 0,034 | 0,184 | 0,140 | 0,623 | 0,904 |
| CatBoost | 0,034 | 0,185 | 0,147 | 0,618 | 0,880 |
| LightGBMXT | 0,045 | 0,213 | 0,149 | 0,493 | 0,874 |
| NeuralNetTorch | 0,051 | 0,227 | 0,195 | 0,426 | 0,895 |
| RandomForestMSE | 0,058 | 0,241 | 0,175 | 0,350 | 0,873 |
| LightGBM | 0,122 | 0,350 | 0,233 | -0,366 | 0,799 |
| KNeighborsUnif | 0,141 | 0,376 | 0,313 | -0,578 | 0,472 |
| KNeighborsDist | 0,142 | 0,377 | 0,314 | -0,584 | 0,473 |

Примечание: R^2 — коэффициент детерминации, Correlation — коэффициент корреляции между предсказаниями и истинными значениями.

В качестве функции, измеряющей ошибку обучения моделей, был выбран корень из среднеквадратической ошибки (RMSE). Процесс обучения модели направлен на уменьшение RMSE [14].

После задания параметров обучения (начальной скорости обучения, числа эпох, коэффициента регуляризации, штрафа регуляризации и др.) запускается процесс построения моделей с помощью метода fit(). В ходе прогноза и проверки работоспособности модели был использован метод predict() для получения предсказанных значений на основе новых данных.

В рамках проведенного исследования был выполнен сравнительный анализ предсказательной способности ряда алгоритмов машинного обучения. Оценка осуществлялась на основе общепринятых метрик качества, позволяющих всесторонне оценить точность и надежность моделей (табл. 3).

Основываясь на результатах статистических данных, можно выделить модели NeuralNetFastAI и WeightedEnsemble_L2. Модели показали лучшее качество на тестовых данных. Именно они будут использоваться в качестве моделей по умолчанию.

Наилучшие результаты продемонстрировали две модели — NeuralNetFastAI и WeightedEnsemble_L2. Их показатели практически идентичны: MSE составляет около 0,028, RMSE находится в диапазоне 0,166–0,167,

MAE — 0,130–0,131, R^2 варьируется от 0,687 до 0,692, а коэффициент корреляции — от 0,863 до 0,881. Сочетание минимальных значений ошибок (MSE, RMSE, MAE) и максимальных значений коэффициента детерминации (R^2) однозначно выделяет эти модели как наиболее эффективные в рассматриваемом наборе алгоритмов.

Для визуализации работы моделей на тестовых данных была использована графическая библиотека языка Python — *matplotlib*. На рис. 11 представлены прогнозы на основе 12 моделей. Также приведен дополнительный график, демонстрирующий работу двух лучших моделей (рис. 12).

В ходе экспериментальной проверки работоспособности модели было выявлено, что атмосферное давление (P) не вносит значительного вклада в работу моделей и является одним из факторов, влияющих на переобучение моделей. В связи с этим данный параметр был исключен из результирующего списка входных характеристик.

В процессе исследования была подтверждена эффективность применения предикативных моделей для прогнозирования локальной гидродинамической обстановки у опоры моста равнинной реки. Полученные данные свидетельствуют о том, что такие модели могут быть успешно использованы для решения задач прогнозирования в условиях неопределенности и изменчивости природных процессов.

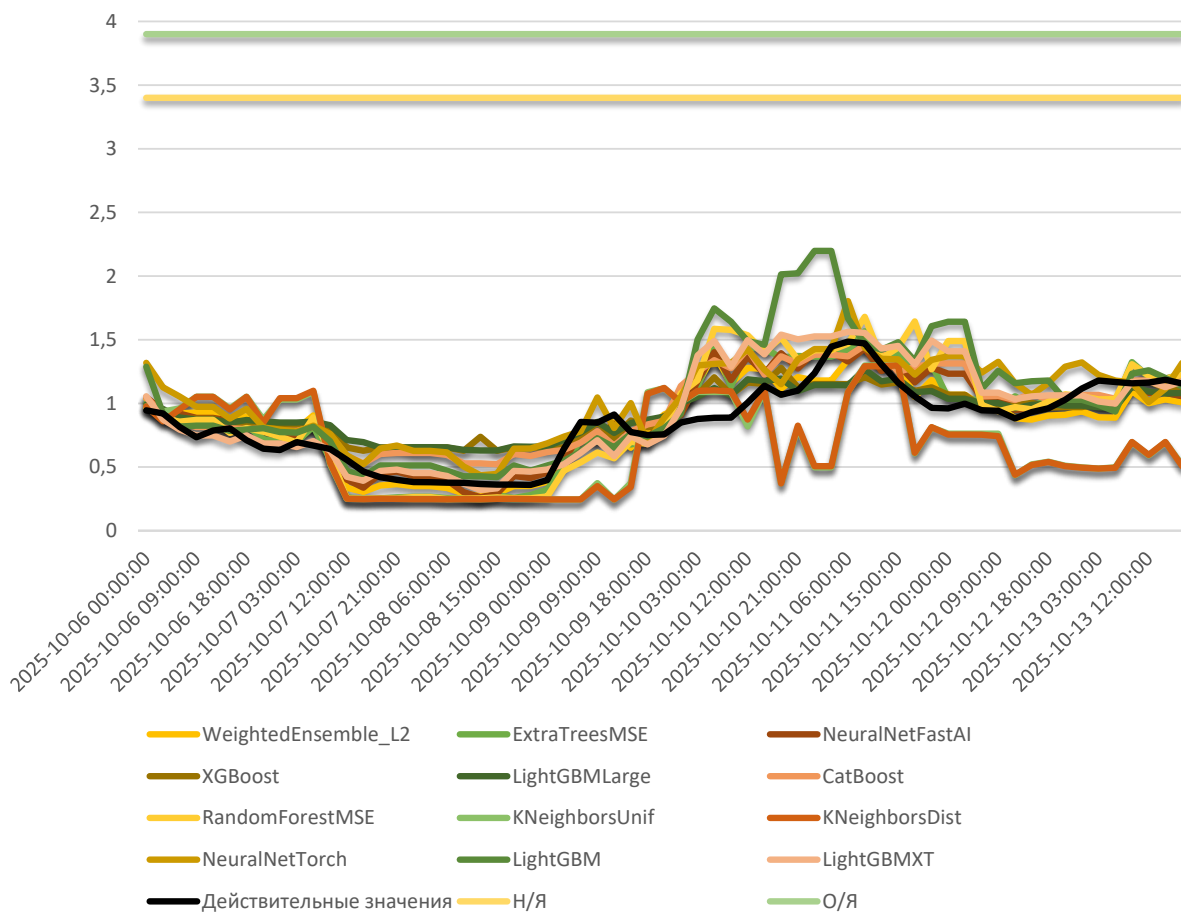


Рис. 11. Результат работы полученных ML-моделей на тестовых данных

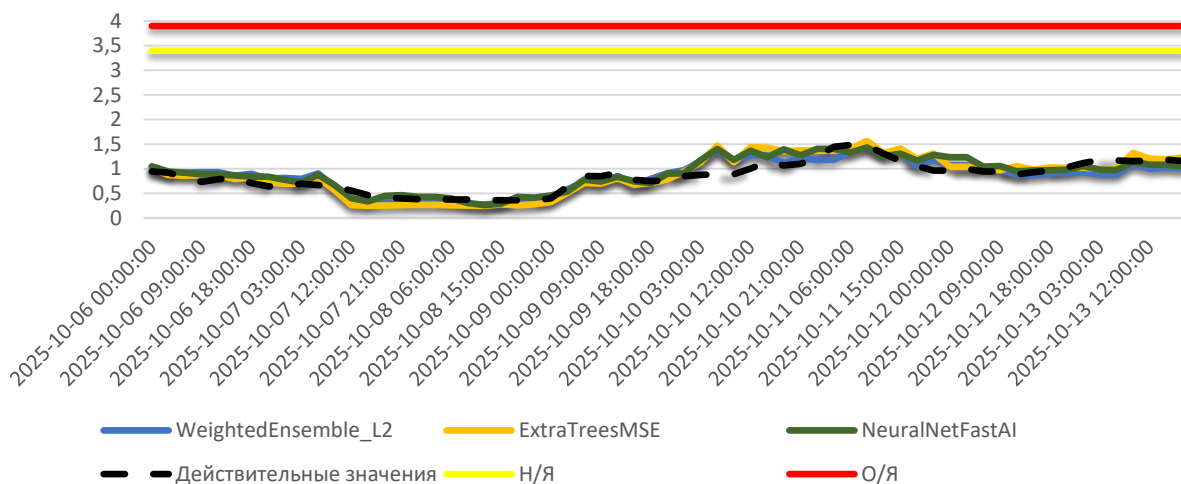


Рис. 12. Результат работы 2 лучших ML-моделей

Заключение

Проведенное исследование подтверждает реализуемость разработки предиктивных моделей для повышения надежности и безопасности транспортной инфраструктуры.

В условиях возрастающих вызовов, связанных с климатическими изменениями и техногенными рисками, традиционные методы принятия управленческих решений демонстрируют свою недостаточную эффективность.

В качестве решения данной проблемы в работе предложен подход, основанный на создании системы поддержки принятия решений, ядром которой является интеллектуальная имитационная модель для прогнозирования уровня воды в равнинных реках. Такой подход позволяет перейти от реактивного к упреждающему управлению, минимизируя материальные потери и риски для человеческих жизней.

Экспериментальная часть работы, выполненная с использованием фреймворка AutoGluon и данных по реке Темерник, продемонстрировала высокую эффективность методов машинного обучения для решения поставленной задачи. Было протестировано 12 моделей, среди которых наилучшие результаты по точности прогноза (наименьшие значения MSE, RMSE, MAE) показали модели NeuralNetFastAI и WeightedEnsemble_L2. Это доказывает, что объединение нескольких алгоритмов (ансамбли) и использование нейронных сетей позволяют достичь высокой достоверности прогнозов.

Таким образом, результаты исследования свидетельствуют о том, что разработанная методика и выбранные модели машинного обучения являются перспективным инструментом для интеграции в системы управления транспортной отраслью. Их внедрение позволит оптимизировать эксплуатацию инфраструктуры, разрабатывать эффективные планы действий в чрезвычайных ситуациях и в конечном итоге повысить устойчивость транспортной системы к негативным воздействиям природного характера. ▲

Список источников

1. Арефьева Е. В. Прогноз опасностей на основе постоянно действующей ситуационно-оптимизационной модели / Е. В. Арефьева, А. В. Рыбаков // Вестник КРСУ. — 2012. — Т. 12. — № 7. — С. 17–21.
2. Алагуджаева М. А. Гидрологическое цифровое прогнозное моделирование зон затопления территории Северного Казахстана на основе высокодетаальной цифровой модели рельефа / М. А. Алагуджаева, С. Р. Садвакасова, Б. Б. Елбасиева, А. А. Аимбетов // Гидрометеорология и экология. — 2023. — № 4. — С. 74–84. — DOI: 10.54668/2789-6323-2023-111-4-74-84.
3. Shu X. Knowledge Discovery: Methods from Data Mining and Machine Learning / X. Shu, Y. Ye // Social Science Research. — 2023. — Vol. 110. — P. 102817. — DOI: 10.1016/j.ssresearch.2022.102817.
4. Долгих В. П. Краткосрочное прогнозирование водного режима рек и водохранилищ методом Брауна / В. П. Долгих, С. В. Рыбалка, И. С. Боблева // Экологический вестник Донбасса. — 2024. — № 2(12). — С. 10–15.
5. Цыганов В. В. Климатические риски железнодорожной инфраструктуры / В. В. Цыганов, В. А. Бородин, С. А. Савушкин, А. В. Лемешкова // ИТНОУ: информационные технологии в науке, образовании и управлении. — 2024. — № 1(22)–2(23). — С. 51–57. — DOI: 10.47501/ITNOU.2024.1.51-57.
6. Восстановить разорванную нить: восстановление движения поездов до Мурманска // Мурманский вестник. — URL: <https://www.mvestnik.ru/our-home/vosstanovity-razorvannuyu-nity/> (дата обращения: 13.11.2025).
7. Обрушение моста // РИА «Воронеж». — URL: <https://riavr.ru/theme/obrushenie-mosta/> (дата обращения: 13.11.2025).
8. В Ростове-на-Дону на вокзале подтопило железнодорожные пути // ТАСС. — URL: <https://tass.ru/proisshestviya/21089611> (дата обращения: 13.11.2025).
9. Ермуханов Д. С. Машинное обучение и математика / Д. С. Ермуханов // Тенденции развития науки и образования. — 2024. — № 110-18. — С. 102–105. — DOI: 10.18411/trnio-06-2024-992.
10. Chicco D. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation / D. Chicco, M. J. Warrens, G. Jurman // PeerJ Computer Science. — 2021. — Vol. 7. — Pp. 1–24. — DOI: 10.7717/PEERJ-CS.623.
11. Пономаренко Т. С. Изучение гидрологических характеристик реки Темерник по результатам работы системы мониторинга «ЭМЕРСИТ» / Т. С. Пономаренко, А. В. Бреева // Экология и водное хозяйство. — 2021. — № 2. — URL: <https://cyberleninka.ru/article/n/izuchenie-gidrologicheskikh-harakteristik-reki-temernik-po-rezultatam-raboty-sistemy-monitoringa-emersit> (дата обращения: 30.10.2025).
12. Рашка С. Python и машинное обучение. Машинное и глубокое обучение с использованием Python, scikit-learn и TensorFlow-2 / С. Рашка, В. Мирджалили. — М.: Вильямс, 2020. — 848 с.
13. Мосин К. К. Использование методов автоматизированного машинного обучения для классификации дорожно-транспортных происшествий / К. К. Мосин, В. Э. Ковалевский, Н. А. Жукова // Интеллектуальные технологии на транспорте. — 2023. — № 2(34). — URL: <https://cyberleninka.ru/article/n/ispolzovanie-metodov-avtomatizirovannogo-mashinnogo-obucheniya-dlya-klassifikatsii-dorozhno-transportnyh-proisshestviy> (дата обращения: 30.10.2025).
14. Feng K. RMSE-minimizing confidence intervals for the binomial parameter / K. Feng, L. M. Leemis, H. Sasinowska // Computational Statistics. — 2022. — Vol. 37. — Iss. 4. — Pp. 1855–1885. — DOI: 10.1007/s00180-021-01183-3.

TRANSPORT AUTOMATION RESEARCH, 2025, Vol. 11, No. 4, pp. 342–354
DOI: 10.20295/2412-9186-2025-11-04-342-354

Predictive Modelling of Water Levels in Lowland Rivers to Enhance the Safety of Transport Infrastructure

Information about author

Mukonina M. I., Senior Lecturer. E-mail: mukonina_m@bk.ru

“Computer Technology and Automated Systems Control” Department, Rostov State Transport University, Rostov-on-Don

Abstract: The paper considers the task of improving the transport infrastructure safety through the development of an intelligent simulation model. The focus is on predicting the water levels in lowland rivers to prevent emergencies such as flooding, bridge collapses and other crises that could disrupt transportation networks. The relevance of this research is determined by a variety of factors, including the rise in climatic hazards such as floods and heavy rainfall, industrial threats, and the limited efficacy of traditional approaches to monitoring and forecasting hydrological conditions. The study involves an analysis of actual emergency incidents, which starkly illustrates the urgent need for timely water level predictions. The methodology for developing a decision support system is based on machine learning technologies. The experimental component of the research utilizes data from the Temernik river. A total of 12 machine learning models were evaluated with the objective of selecting the most accurate and efficient models for future forecasting applications. The models were subjected to rigorous statistical assessment using a variety of performance metrics. Following the analysis, the most effective models have been identified and recommended for subsequent deployment and use.

Keywords: transport infrastructure; emergencies; water level monitoring; intelligent simulation; decision support system; machine learning; water level predicting; AutoGluon; NeuralNetFastAI; WeightedEnsemble_L2; lowland rivers; risk management.

References

- Aref'eva E. V., Rybakov A. V. Prognoz opasnostey na osnove postoyanno deystvuyushchey situatsionno-optimizatsionnoy modeli [Forecast of hazards based on a continuously operating situational-optimization model]. *Vestnik KRSU* [Bulletin of KRSU]. 2012, Iss. 12(7), pp. 17–21. (In Russian)
- Alagudzhaeva M. A., Sadvakasova S. R., Elbasieva B. B., Aimbetov A. A. Gidrologicheskoe tsifrovoe prognoznnoe modelirovanie zon zatopeniya territorii Severnogo Kazakhstana na osnove vysokodetal'noy tsifrovoy modeli rel'efa [Hydrological digital predictive modeling of flood zones in Northern Kazakhstan based on a high-resolution digital elevation model]. *Gidrometeorologiya i ekologiya* [Hydrometeorology and Ecology]. 2023, Iss. 4, pp. 74–84. DOI: 10.54668/2789-6323-2023-111-4-74-84. (In Russian)
- Shu X., Ye Y. Knowledge Discovery: Methods from Data Mining and Machine Learning. *Social Science Research*. 2023, vol. 110, p. 102817. DOI: 10.1016/j.ssrsearch.2022.102817.
- Dolgikh V. P., Rybalka S. V., Bobleva I. S. Kratkosrochnoe prognozirovaniye vodnogo rezhima rek i vodokhranilishch metodom Brauna [Short-term forecasting of river and reservoir water regimes using the Brown method]. *Ekologicheskiy vestnik Donbassa* [Ecological Bulletin of Donbass]. 2024, Iss. 2(12), pp. 10–15. (In Russian)
- Tsyganov V. V., Borodin V. A., Savushkin S. A., Lemeshkova A. V. Klimaticheskie riski zheleznodorozhnoy infrastruktury [Climatic risks to railway infrastructure]. *ITNOU: informatsionnye tekhnologii v nauke, obrazovanii i upravlenii* [ITNOU: Information Technologies in Science, Education and Management]. 2024, Iss. 1(22)–2(23), pp. 51–57. DOI: 10.47501/ITNOU.2024.1.51-57. (In Russian)
- Vosstanovit' razorvannuyu nit': vosstanovlenie dvizheniya poezdov do Murmanska [Restoring the broken thread: resuming train services to Murmansk]. *Murmanskiy vestnik* [Murmansk Bulletin]. Available at: <https://www.mvestnik.ru/our-home/vosstanovit'-razorvannuyu-nit'/> (accessed: November 13, 2025). (In Russian)
- Obrushenie mosta [Bridge collapse]. RIA "Voronezh". Available at: <https://riavrn.ru/theme/obrushenie-mosta/> (accessed: November 13, 2025). (In Russian)
- V Rostove-na-Donu na vokzale podtopilo zheleznodorozhnyye puti [Railway tracks flooded at the station in Rostov-on-Don]. TASS. Available at: <https://tass.ru/proisshestiya/21089611> (accessed: November 13, 2025). (In Russian)
- Ermukhanov D. S. Mashinnoe obucheniye i matematika [Machine learning and mathematics]. *Tendentsii razvitiya nauki i obrazovaniya* [Trends in the Development of Science and Education]. 2024, Iss. 110-18, pp. 102–105. DOI: 10.18411/trnio-06-2024-992. (In Russian)
- Chicco D., Warrens M. J., Jurman G. The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*. 2021, vol. 7, pp. 1–24. DOI: 10.7717/PEERJ-CS.623.
- Ponomarenko T. S., Breeva A. V. Izucheniye gidrologicheskikh kharakteristik reki Temernik po rezul'tatam raboty sistemy monitoringa "EMERSIT" [Studying the hydrological characteristics of the Temernik River based on the results of the "EMERSIT" monitoring system]. *Ekologiya i vodnoe khozyaystvo* [Ecology and Water Management]. 2021, Iss. 2. Available at: <https://cyberleninka.ru/article/n/izucheniye-gidrologicheskikh-kharakteristik-reki-temernik-po-rezultatam-raboty-sistemy-monitoringa-emersit> (accessed: October 30, 2025). (In Russian)
- Rashka S., Mirdzhalili V. *Python i mashinnoe obucheniye. Mashinnoe i glubokoe obucheniye s ispol'zovaniem Python, scikit-learn i TensorFlow-2* [Python and machine learning. Machine and deep learning using Python, scikit-learn and TensorFlow-2]. Moscow: Vil'yams Publ., 2020, 848 p. (In Russian)
- Mosin K. K., Kovalevskiy V. E., Zhukova N. A. Ispol'zovanie metodov avtomatizirovannogo mashinnogo obucheniya dlya klassifikatsii dorozhno-transportnykh proisshestiyy [Using automated machine learning methods for classifying road traffic accidents]. *Intellektual'nye tekhnologii na transporte* [Intelligent Technologies in Transport]. 2023, Iss. 2(34). Available at: <https://cyberleninka.ru/article/n/ispolzovanie-metodov-avtomatizirovannogo-mashinnogo-obucheniya-dlya-klassifikatsii-dorozhno-transportnykh-proisshestiyy> (accessed: October 30, 2025). (In Russian)
- Feng K., Leemis L. M., Sasinowska H. RMSE-minimizing confidence intervals for the binomial parameter. *Computational Statistics*. 2022, vol. 37, Iss. 4, pp. 1855–1885. DOI: 10.1007/s00180-021-01183-3.