

Моделирование зон затопления на основе прогнозирования временных рядов и ГИС-технологий

Е.В. Пальчевский¹, В.В. Антонов², Л.Е. Родионова², Л.А. Кромина², А.Р. Фахруллина²

¹ Финансовый университет при Правительстве Российской Федерации,
109456, Россия, г. Москва, ул. 4-й Вешняковский проезд, д. 4;

² Уфимский университет науки и технологий,
450008, Россия, г. Уфа, ул. Карла Маркса, д. 12

Аннотация

Предлагается специализированная веб-ГИС, реализованная за счет интеграции искусственной нейронной сети и геотехнологий, которая обеспечивает раннее прогнозирование и моделирование зон затопления на 5 суток вперед.

Реализованные в рамках данной веб-ГИС методы и алгоритмы позволяют ежедневно прогнозировать временные ряды на основе ретроспективных данных об уровнях воды и общего водного притока, температуре воздуха и воды, толщине снежного покрова и количестве осадков, скорости ветра и атмосферного давления. При этом только на основе полученных прогнозных значений об уровнях воды реализована возможность раннего моделирования и визуализации разлива рек. Это предоставит возможность специализированным организациям и службам, а также управляющим органам заблаговременно и в кратчайшие сроки принимать решения, связанные с противопаводковыми мероприятиями.

Ключевые слова: геоинформационная система, моделирование зон затопления, прогнозирование временных рядов, искусственные нейронные сети.

Цитирование: Пальчевский, Е.В. Моделирование зон затопления на основе прогнозирования временных рядов и ГИС-технологий / Е.В. Пальчевский, В.В. Антонов, Л.Е. Родионова, Л.А. Кромина, А.Р. Фахруллина // Компьютерная оптика. – 2024. – Т. 48, № 6. – С. 913-923. – DOI: 10.18287/2412-6179-CO-1418.

Citation: Palchevsky EV, Antonov VV, Rodionova LE, Kromina LA, Fakhrullina AR. Modeling flood zones on the basis of time series forecasting and GIS-technologies. Computer Optics 2024; 48(6): 913-923. DOI: 10.18287/2412-6179-CO-1418.

Введение

Паводковые ситуации являются неизбежным природным явлением, которое приводит не только к подтоплению прибрежных территорий, но и к человеческим жертвам, а также финансовому ущербу населенному пункту, региону или стране: например, в Российской Федерации среднегодовой ущерб за последние 20 лет достигает 10 миллиардов рублей [1], а общемировые ежегодные потери составляют более 300 миллиардов долларов США [2]. Более того, с учетом прогнозов изменения климата ожидается, что паводки станут более частыми и масштабными [3].

Рамочная конвенция ООН об изменении климата делает упор на глобальную политику в отношении убытков и ущерба [4]. Парижское соглашение подтвердило, что Варшавский международный механизм по потерям и ущербу является основным механизмом в рамках процесса РКИК ООН для предотвращения, минимизации и устранения потерь и ущерба, связанных с последствиями изменения климата, включая экстремальные погодные явления и явления с медленным течением, в том числе и паводками. При этом глобальная политика в отношении человеческих потерь и экономического ущерба важна для реагирования на климатические бедствия и природные явления,

к которым и относится паводок. Таким образом, раннее прогнозирование паводковых ситуаций с последующим моделированием и визуализацией зон затопления требует обширных исследований с целью минимизации рисков человеческих жертв и финансово-го ущерба.

Существует несколько разновидностей моделей прогнозирования уровней воды (УВ) и визуализации зон затопления (ЗЗ), основанных на использовании временных рядов [5–9]: гидрологические [10], статистические [11] и гибридные [12] модели, а также модели машинного обучения, включая глубокое обучение с подкреплением [13].

Гидрологические модели прогнозирования УВ и расчета ЗЗ. Ядром данной модели прогнозирования является гидрологическая (HL) модель, представляющая собой пространственно неоднородные свойства реальной гидрологической системы [14]. Так, например, в [15] использовали модель WRF-Hydro (Weather Research and Forecasting-Hydro) для моделирования ЗЗ энергосистем. В [16] оценивались 4 набора (ERA5, ERA-Interim, CFSR и JRA55) данных повторного анализа для использования в гидрологической модели прогнозирования УВ GR4J с целью расчета ЗЗ. Классическая математическая модель прогнозирования УВ и визуализации ЗЗ представлена в [15].

Основным ограничением данных гидрологических моделей прогнозирования является существенная погрешность расчета как УВ на водных объектах, так и ЗЗ. Для повышения прогнозной точности возникает необходимость в большом количестве данных, получение которых требует серьезных финансовых затрат.

Статистические модели прогнозирования УВ.

Среди данных моделей широко распространены две модели: ARIMA (AutoRegressive Integrated Moving Average) [17], SVR (Support Vector Regression) [18] и модели экспоненциального сглаживания [19]: Хольта, Брауна, простые модели, демпфированный тренд, мультипликативная и аддитивная модели Уинтерса, а также простая сезонная.

Модель ARIMA. Стандартный математический аппарат данной модели может быть представлен формулой (1):

$$\Delta^d WL_t = \alpha + \beta_1 \Delta^d WL_{t-1} + \beta_2 \Delta^d WL_{t-2} + \dots + \beta_p \Delta^d WL_{t-p} + \varepsilon_t + \phi_1 \varepsilon_{t-1} + \dots + \phi_q \varepsilon_{t-q}, \tag{1}$$

в которой h_t – прогнозируемое значение УВ; t – коэффициент задержки временного ряда; $h_{t-1}, h_{t-2}, \dots, h_{t-p}$, – запаздывание каждого временного ряда, подаваемого на вход для прогнозирования; ε_t – коэффициент ошибки прогнозирования; $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$ – коэффициент ошибки прогнозирования с учетом задержки временного ряда; $\beta_1, \beta_2, \dots, \beta_p$ – лаговый оператор сдвига, позволяющий получить значения временного ряда на основании предыдущих элементов; ϕ_1, \dots, ϕ_q – авторегрессионный многочлен; p – порядок авторегрессии, позволяющий добавить предыдущие значения временного ряда, d – порядок интегрирования – разностей исходного временного ряда; q – порядок скользящего среднего, позволяющий установить погрешность модели как линейную комбинацию наблюдавшихся ранее значений ошибок $\varepsilon_{t-1}, \varepsilon_{t-2}, \dots, \varepsilon_{t-q}$.

Классическая модель SVR при прогнозировании УВ hl_t представлена в выражении

$$hl_t = \sum_h w_h (hl_{ht} - \overline{hl_{ht}}),$$

где hl_{ht} – значение измеренного УВ h в определенный период t ; $\overline{hl_{ht}}$ – средневзвешенное значение УВ h по всей выборке в определенный период t ; w – относительный вес, придаваемый hl_{ht} в hl_t .

Модели экспоненциального сглаживания зачастую применяются при прогнозировании временных рядов. Например, простая модель экспоненциального сглаживания схожа с моделью ARIMA с нулевым порядком авторегрессии. При этом одним из ключевых факторов успешного прогнозирования данной моделью является отсутствие тренда и сезонности. Более того, из-за данной особенности для прогнозирования УВ наиболее точно подходит модель демпфированного тренда (2), являющаяся одной из самых популярных

моделей экспоненциального сглаживания в рамках прогнозирования УВ.

$$\begin{cases} h_0 = \left(\frac{1}{\phi}\right) \cdot \frac{(h_k - h_1)}{(k-1) \cdot p}, \\ hp_0 = \frac{(h_1 - p) \cdot h_0}{2}, \end{cases} \tag{2}$$

где ϕ – параметр сглаживания временного ряда УВ; k – число полных сезонных циклов; h_0 – начальный тренд временного ряда УВ; h_1 и h_k – средние значения УВ на первом и последнем сезонных циклах; p – длина сезонного цикла, hp_0 – прогнозируемое значение УВ.

Основными преимуществами статистических моделей являются простота реализации, возможность обработки больших структурированных данных и доступность среди готовых решений в виде специализированных библиотек, написанных на различных языках программирования. В качестве основного ограничения необходимо выделить невозможность корректного расчета прогнозных значений временных рядов при высоком значении горизонта прогнозирования, т.е. при долгосрочном прогнозировании, как правило, более 4 суток.

Гибридные модели прогнозирования УВ и визуализации ЗЗ. Данные модели внедрены во многих мировых компаниях, в том числе и в Российской Федерации, и сочетают в себе линейные и нелинейные модели прогнозирования и обработки больших данных. Например, в [20] интегрировали LDBNN, CEEMDAN и VMD в единую модель, позволяющую моделировать ЗЗ. В [21] разработан новый гибридный подход, основанный на машине экстремального обучения для прогнозирования УВ и моделирования ЗЗ.

Стандартный алгоритм, представляемый в виде интеграции модели машинного и глубокого обучения, основанной на топологической структуре многослойного перцептрона, имеет следующий вид: $\hat{Y} = W_2 \sigma(W_1 x)$, где W_1 – матрица весов между входом и скрытым слоем, σ – функция активации, W_2 – матрица весовых коэффициентов между скрытым слоем и выходом, x – вектор входного сигнала (входные данные, $x = [x_1, x_2, \dots, x_n]$). Суть такой интеграции заключается в повышении скорости подбора весовых коэффициентов синаптических связей за счет одного скрытого слоя нейронов, в котором веса между нейронами распределяются случайным образом и никогда не обновляются.

Однако большинство существующих гибридных моделей прогнозирования временных рядов не позволяют моделировать и визуализировать ЗЗ. При этом существенным их недостатком отмечается отсутствие достаточной универсальности и скорости прогнозирования.

Модели машинного и глубокого обучения для прогнозирования УВ и визуализации ЗЗ. Важно отметить, что по сравнению со статистическими и

гибридными моделями прогнозирования УВ и визуализации 3З модели глубокого обучения очень популярны, поскольку они хорошо справляются со сложностью и неоднородностью исходных данных, имеют широкий спектр возможностей расчета различных параметров, необходимых для моделирования 3З.

Например, в [22] разработана простая модель прогнозирования УВ FDPRE (flood prediction) с последующим моделированием 3З. В [23] сделан упор на устранение неопределенностей в геопространственных подходах на основе машинного обучения для прогнозирования паводковых ситуаций.

Тем не менее, к недостаткам указанных классических и современных моделей следует отнести отсутствие фиксации характеристик, например, асимметричных циклов и случайных выбросов (аномалий) наблюдений, существующих в реальных временных рядах, а также невозможность визуализации 3З. Классическая математическая модель рекуррентной нейронной сети для прогнозирования временных рядов представлена в [24].

При этом у всех данных методов прогнозирования временных рядов УВ и моделирования 3З имеется серьезное общее ограничение: отсутствует универсальность относительно прогноза и его визуализации. Т.е. визуализация 3З выполняется не только на основе полученных статистических данных, но и с использованием цифровых моделей рельефа (ЦМР), что усложняет применение этих методов на практике из-за частой недоступности качественной ЦМР для той или иной местности.

Таким образом, основная цель данного исследования состоит в том, чтобы с учетом вышеописанных недостатков разработать универсальное научно-техническое специализированное решение в виде веб-ГИС «Паводок» [25, 26] для раннего прогнозирования УВ и моделирования 3З с использованием минимального количества данных. Это позволит в кратчайшие сроки получать необходимые паводковые сведения. Как следствие, получение этих сведений предоставит возможность специализированным организациям и службам, а также управляющим органам в кратчайшие сроки принимать решения, связанные с противопаводковыми мероприятиями.

Выбор технологии реализации «клиент-сервер» в рамках веб-ГИС «Паводок» связан не только с современным трендом перехода на облачные вычисления, но и с техническими преимуществами: не требуется установка на персональный компьютер и смартфон; для прогнозирования и моделирования используются заранее зарезервированные вычислительные мощности, что исключает необходимость дополнительной покупки оборудования; адаптивный интерфейс под все устройства в зависимости от разрешения и диагонали экрана.

1. Математическое моделирование зон затопления на основе прогнозирования временных рядов

Для достижения поставленной цели в виде моделирования 3З на основе прогнозирования временных рядов необходимо решить две задачи: (1) прогнозирование УВ на стационарных гидрологических постах и (2) моделирование 3З на основе полученных прогнозных значений временных рядов.

Первая задача решена в работе [27], где на основе искусственной нейронной сети выполняется ранний прогноз (в сантиметрах) УВ на 5 суток вперед. Для прогнозирования УВ использовались ретроспективные данные за последние 20 лет, предоставленные ФГУП «Центр регистра и кадастра»: временные ряды УВ; температуры и прогноза температуры воздуха и воды; толщина снежного покрова и количества осадков; скорость ветра и атмосферное давление; общий приток воды. Используемые метрики ошибок в модели искусственной нейронной сети: MAE (Mean Absolute Error) и RMSE (Root Mean Square Error).

Вторую задачу предлагается решать с использованием прогнозных значений УВ (h), полученных в результате решения **первой задачи**, а также численных методов и геоинформационных технологий моделирования.

Опираясь на задействованный в [27] метод прогноза весеннего половодья в виде нейросетевой модели неравномерного распределения водоудерживающей емкости, а также основываясь на представлении в виде функции зависимости слоя воды выше глубины емкостей, образуемых рельефом, от данных многолетних наблюдений с гидрологических постов, можем вычислить динамику изменения распределения действующей площади с изменением слоя поступившей воды. Следовательно, приходим к возможности построения топологической модели (графа) рассматриваемого информационного пространства [28] в виде множества вершин (узлов) и множества соединяющих их ребер. При этом в качестве вершин (узлов) используем множество данных гидрологических постов и точек наблюдения, а вершины будут соединены между собой отрезками, длины которых равны длинам соответствующих участков русла.

Применяя алгоритм Дейкстры, при условии заблаговременного вычисления пропускной способности дуг, можем определить максимальный поток, т.е. максимальный объем, который может протекать через дугу на графе из вершины (узла) в требуемую точку (геоточку) – виртуальную вершину за счет наличия ретроспективных данных [29, 30]. Таким образом, можем дополнить датасет данными из новых виртуальных точек и, вместо вычисления расчетной формулы для конкретного бассейна, будем использовать возможности ИНС.

Каждой области, образованной ребрами и вершинами нашей модели, поставим в соответствие высо-

ту – N-координату. Группу объектов, образующих в пространстве компактную область, расстояние которых от узла не превышает заданную величину, и N-координату, также отличающуюся от узла не более чем на заданную величину, будем называть кластером. При неполной или нечеткой информации об объекте кластеры могут иметь размытые границы и перекрытия друг с другом.

В данной области каждый объект содержит ссылку – отношение с центром кластера (узлом) и с объектами смежных кластеров.

В качестве исходной информации для расчета прогнозных значений УВ и визуализации ЗЗ используется массив темпоральных данных, предоставленных ФГУП «Центр регистра и кадастра»: временные ряды УВ, общего притока воды, температуры и прогноза температуры воздуха и воды, толщина снежного покрова и количества осадков. Наблюдение за уровнем разлива ведется на основе данных стационарных гидрологических постов и автоматических станций Республики Башкортостан (рис. 1).

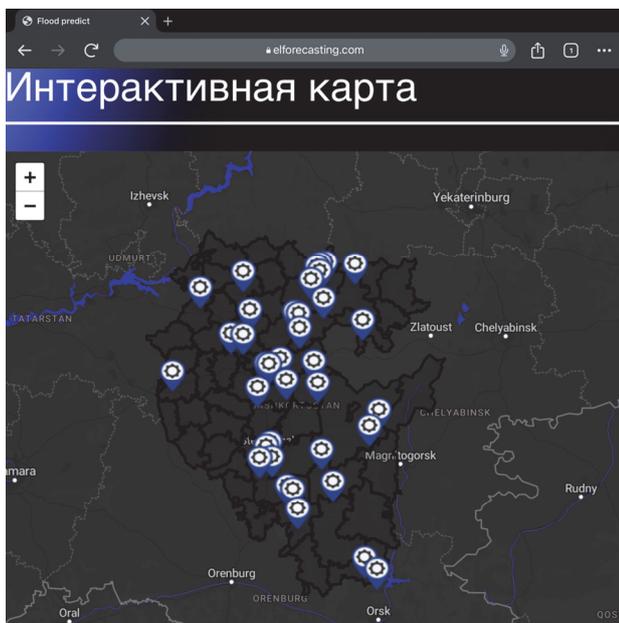


Рис. 1. Карта гидрологических постов и автоматических станций из веб-ГИС «Паводок»

Данные о гидрологических постах и автоматических станциях (координатные и атрибутивные данные), а также все компоненты их пространственных данных представлены в виде логически единой БД с последующим установлением связи между ними – геокодированием.

Таким образом, имеем множество (обозначим Z) гидрологических постов и автоматических станций, являющихся основным источником формирования информационного пространства. Любая точка на геокарте обладает рядом свойств (обозначим y^i , где i – номер объекта), которые могут быть использованы для прогнозирования требуемых контролируемых показателей. В частности, к таким показателям относятся предпола-

гаемые ЗЗ, а также не связанные с руслом рек направления движения масс воды и время прихода паводка в контролируемые точки. В результате ко множеству основных источников можно добавить новое множество объектов – временные ряды, формирующиеся на выходе искусственной нейронной сети (ИНС).

Соответственно, представим исследуемую предметную область множеством взаимодействующих объектов и множеством отношений этих объектов. Получаем возможность визуализации множества таких объектов в виде геометрического покрытия всей предметной области некоторыми областями – кластерами с учетом гидрологических характеристик: УВ, переноса речных наносов и эрозии речных русел, площадь бассейна реки и ее уклон.

Следовательно, каждый основной объект – гидрологический пост вместе с некоторой территорией, представленной точками на геокарте. Каждая такая территория является ограниченной областью. Множество таких областей образуют пространство предметной области. Эти области динамически меняются во времени, т.е. при изменении некоторых атрибутов, происходит изменение границ гидропоста. При попадании двух объектов в одну область происходит их взаимодействие через отношения с последующим изменением определенных параметров. Таким образом, возникает новая информационная единица, однозначно соответствующая процессу взаимодействия, в которой присутствуют новые характеристики этого образовавшегося объекта – виртуального процесса, например, перемещение паводковых вод. Каждый объект нашего информационного пространства характеризуется атрибутивными данными с географически организованной информацией. В качестве атомарной модели используется двумерный элемент – пиксель (ячейка). Упорядоченная совокупность атомарных моделей образует растр, являющийся моделью карты или геообъекта. Т.е. данные представляются в виде: X, Y – пространственные координаты, применяемые при моделировании зон затопления, Q – зависящая от них переменная. В общем случае каждую точку сетки с координатами (x, y) характеризует некоторый вектор состояний (q_1, \dots, q_p) , а пространственное положение может быть выражено через относительные единицы, например, как обратно пропорциональное квадрату расстояний между объектами. Показатели состояния формируются из входных переменных $C_i (i=1, \dots, l)$, полученных тем или иным способом, и выходных $D_j (j=1, \dots, mj)$, выражающихся через входные, т.е. $D_j = F(C_1, \dots, C_l), j=1, \dots, mj, l+mj=p$.

Рассматривая данный процесс на разных уровнях, начиная от физического взаимодействия и заканчивая абстрактными составляющими, получаем схему взаимодействия объектов – гидрологических постов с новыми точками, т.е. виртуальными объектами, а также схему взаимодействия этих виртуальных объектов с новыми виртуальными объектами. Такая схе-

ма представляет собой многомерные поликубы. Таким образом, производя классификацию гидрологических постов и автоматических станций по некоторым значимым для вычисления уровней паводка показателям, получаем возможность обеспечить прогноз необходимых сведений даже по отсутствующим с некоторых постов /станций своевременных данных.

Выделяя объекты одного класса, можем определить взаимное соответствие некоторого множества свойств, например:

- нахождение на одной и той же высоте над уровнем моря;
- принадлежность к руслу одной реки;
- возможности вычисления расстояний между точками одного кластера, но не более размера кластера, и т.д.

Все это позволяет прогнозировать некоторые временные параметры прихода паводковых вод, их уровней и динамики.

Под процедурой формирования классов будем понимать упорядочивание объектов по их схожести методом кластеризации.

Таким образом, вектор состояния объекта является базой разбиения объектов на кластеры, и множество объектов предметной области будем использовать как множество объектов кластеризации.

Это множество определяется множеством атрибутов, представляющих собой характеристику объектов анализируемой предметной области. За счет использования нечеткой логики утверждаем, что для каждого объекта в определенной количественной шкале измерены все значения атрибутов, то есть соотнесен вектор с координатами из значений указанной шкалы. Исходя из этого введем обозначения: $Z = \{z_1, \dots, z_n\}$ – множество объектов предметной области (гидрологические посты и автоматические станции), где n – количество объектов; $q^i = \{q^i_1, \dots, q^i_{p_i}\}, i = 1, \dots, n$ – множество свойств (в виде вектора параметров i -го объекта), характеризующих объекты кластеризации, т.е. являющихся базой для объединения объектов в кластеры; p_i – количество атрибутов, характеризующих i -й объект. В настоящее время одним из самых распространенных алгоритмов осуществления нечеткой кластеризации считается алгоритм нечетких s -средних. При этом основным отличием методов нечеткой кластеризации от четкой является введение для каждого объекта функций принадлежности к различным кластерам.

Введем обозначения: F – множество нечетких кластеров рассматриваемой предметной области. Каждый кластер будет включать один или более объектов предметной области, тогда $F = \{F_1, \dots, F_k\}$ – множество нечетких кластеров предметной области, где k – их количество. Введем обозначение: $\mu^F = \{(\mu^1_1, \dots, \mu^1_k), \dots, (\mu^n_1, \dots, \mu^n_k)\}$ – множество функций принадлежности объектов нечетким кластерам. Также введем ограничение

$$\sum_{j=1}^k \mu_j^i = 1,$$

означающее, что сумма значений функций принадлежности ко всем кластерам для каждого объекта i равна 1. Так как нас интересует покрытие кластерами всего множества объектов, можем ввести дополнительное ограничение

$$\sum_{i=1}^n \mu_j^i \geq 1,$$

т.е. выполнение данного условия для каждого кластера $j = 1, \dots, k$. Тогда оценка i -го объекта по атрибуту $y_j^i, j = 1, \dots, k$, может быть представлена функцией принадлежности $\mu_j^i(y^i)$.

При использовании метода кластеризации s -средних (FCM) центр кластера будет определяться по значениям всех пространственных атрибутов объектов, взвешенных по степени их принадлежности к кластеру, и может быть вычислен по формуле (3):

$$v^j = \frac{\sum_{i=1}^n (\mu_j^i(y^i))^m \times \sum_{l=1}^{p_i} y_l^i}{\sum_{i=1}^n \mu_j^i(y^i)^m}, \tag{3}$$

где m – гиперпараметр, определяющий нечеткость кластера: чем больше значение m , тем более нечетким будет кластер, а

$$\sum_{l=1}^{p_i} y_l^i (l = 1, \dots, p_i)$$

для каждого объекта i вычисляет сумму свойств – атрибутов вектора параметров. Таким образом, в результате предложенной кластеризации для любого объекта системы, представленного в виде вершины (узла) – реального гидрологического поста или геоточки, описанной выше, определяется степень принадлежности объекта к кластеру максимальным расстоянием от центра кластера. При отсутствии информации на таком объекте (например, пропуск показателя УВ) он может быть сгенерирован для формирования датасета.

Для системы визуализации ЗЗ может быть сформировано некоторое количество показателей, в том числе и объектов, которые выходят за рамки представленной системы. Тогда каждый такой объект системы, представленный в виде вершины (узла) – реального гидрологического поста или геоточки, описанной выше, может быть дополнен множеством возможных управляющих воздействий, определяемых кортежем $\{Sv, Sp, Spv, Sq\}$:

1. Множество возможностей – маршрутов $Sv = Fv \{Fv^1, \dots, Fv^k\}$.
2. Множество ограничений по параметрам основных объектов и параметров геоточек, например,

нет условий для перемещения, встречный напор воды и т.д.: $Sp = Fo\{Fo^1, \dots, Fo^m\}$.

3. Система управления инцидентами (появлением нештатных ситуаций и т.д.) в виде множества $Spv = Sv \circ Sp$ и $Spv^j = Sv^j \circ Sp^j$ ($j = 1, \dots, l$), где j – номер инцидента, а l – количество инцидентов.

4. Система управления прецедентами (рис. 2) – сбор информации о том, что система должна делать, последовательность действий, приводящих к значимому результату и т.д.

На первом этапе в системе управления прецедентами выявляется проблема/инцидент, далее она сравнивается с базой прецедентов, определяя наиболее подходящий прецедент. Затем происходит адаптация старого прецедента к новому и предлагается решение. После этого оценивается и утверждается предложенное решение с последующим сохранением прецедента в базу прецедентов.

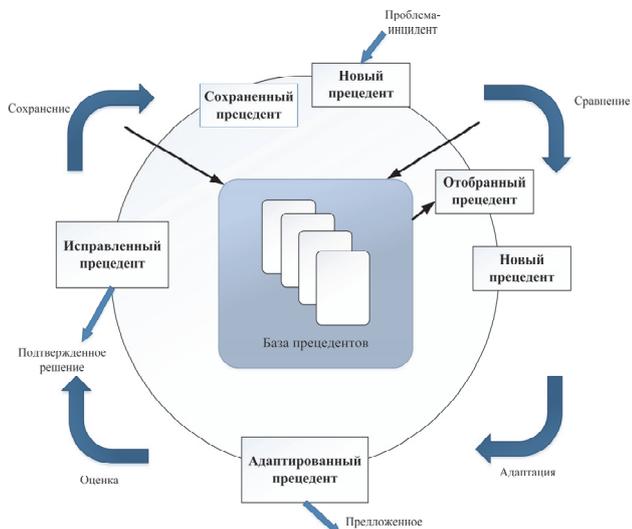


Рис. 2. Система управления прецедентами в виде цикла PDCA

Учитывая, что прецедент – описание и решение проблемы с дальнейшим применением полученных результатов, система управления прецедентами представима в виде множества $Sq = Sv \circ Spv$ (множество исключительных ситуаций и т.д.) и $Sp^j = Sv^j \circ Spv^j$ ($j = 1, \dots, r$), где j – номер прецедента, а r – количество прецедентов.

Очевидно, что $F = Fo \cup Fv$. Таким образом, приходим к многоуровневой системе веб-ГИС «Паводок» с возможностью прогнозирования инцидентов на основе базы правил (знаний), перевода их в прецеденты, выстраивания цепочек возможных состояний системы (рис. 3), основанных на функции принадлежности. Конкретный вид функций принадлежности определяется только с учетом специфики имеющейся неопределенности, реальной ситуации на объекте и числа степеней свободы в функциональной зависимости, что определяет вычисление самой функции при ИНС на основе обработки темпорального набора данных.

Из многомерной схемы (рис. 3) получаем характеристику прецедента 1 (Spv^1), которая определяется

уравнением (4) для инцидента Spv^j (в нашем случае – паводковой ситуации):

$$Spv^j = (A_s \cdot \tan \beta) \cdot h, \tag{4}$$

где для каждого кластера, к которому принадлежит объект соответствующего инцидента, введены обозначения: A_s – площадь бассейна реки, β – уклон реки (в градусах), h – прогнозный уровень воды, вычисленный с помощью ИНС [25]. Исходя из этого получаем прецедент Sq^1 паводковой ситуации, позволяющий смоделировать зону затопления на основе построения мультиполигона (кластера) на основе библиотеки Leaflet (параметры Lat и Lng) за счет применения выпрямленной линейной функции активации (RELU), зачастую используемой в Deep Learning (5):

$$Sq^1 = RELU \cdot ((h - h_{min}) \cdot (Sq^1 - h_{max})), \tag{5}$$

где h_{min} и h_{max} – минимальный и максимальный прогнозные уровни воды, вычисленные с помощью системы «Паводок» [25]. Таким образом, может быть определено все множество прецедентов системы управления прецедентами.

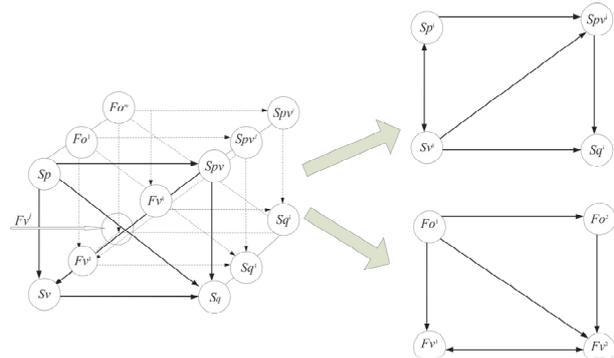


Рис. 3. Многомерная схема взаимодействия параметров геоточек (в соответствии с правилом квадрата Декарта)

2. Визуализация зон затопления

Для визуализации зон затопления необходимо реализовать разработанную в параграфе 1 математическую модель: на рис. 4 представлен адаптированный для веб-ГИС «Паводок» алгоритм, обеспечивающий обработку, визуализацию и интерпретацию темпоральных данных в виде зон затопления на примере Республики Башкортостан, а описание технологий реализации ГИС показано в табл. 1.

Моделировать 33 предлагается в три этапа (рис. 4).

На первом этапе производится авторизация в веб-ГИС «Паводок», выбор гидрологического поста/автоматической станции (рис. 1) и горизонта прогнозирования посредством веб-интерфейса с целью нейросетевого (на основе ИНС) расчета прогнозных значений УВ. На втором этапе посредством Web API через веб-интерфейс происходит обращение к заранее обученной модели (в формате .h5)

ИНС для выполнения нейросетевого прогноза временного ряда. Результаты прогноза УВ выводятся в веб-интерфейс. На третьем этапе на карте гидрологических постов и автоматических станций инициализируем модуль моделирования ЗЗ (рис. 6) с их последующей визуализацией (рис 5, 7). Значения ЗЗ рассчитываются на основе математической модели, представленной в параграфе 1.

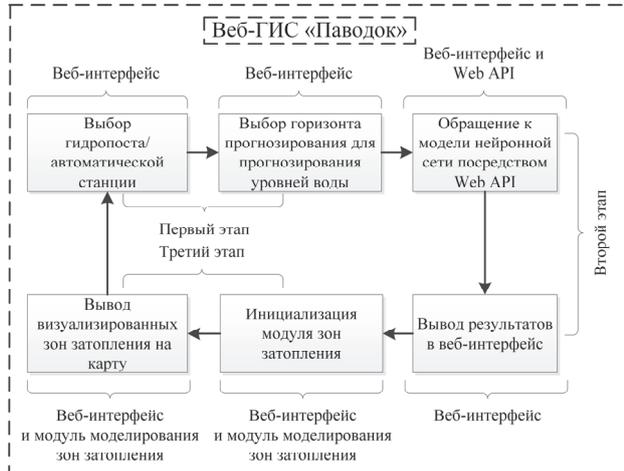


Рис. 4. Алгоритм визуализации зон затопления посредством веб-ГИС «Паводок»

Табл. 1. Описание используемых для реализации веб-ГИС технологий

#	Стек технологий	Описание и язык программирования	Платформа
1	Web API	Цель: взаимодействие пользователя посредством backend с frontend. Языки программирования и разметки: PHP.	Ubuntu 22.04, веб-сервер Nginx
2	ИНС	Цель: прогноз уровней воды. Языки программирования: Java и Python	Ubuntu 22.04
3	Веб-интерфейс	Цель: взаимодействие frontend с backend посредством архитектуры MVC и Web API. Языки программирования/разметки: HTML, CSS, Bootstrap, OpenStreetMap, JavaScript.	Linux, Windows, MacOS, IOS, iPadOS, Android
4	Модуль моделирования зон затопления	Цель: визуализация зон затопления для эвакуации людей. Языки программирования: Java, JavaScript (библиотека Leaflet).	Linux, Windows, MacOS, IOS, iPadOS, Android

Изначально (рис. 5) после выбора параметров, необходимых для визуализации зон затопления, данные о прогнозных уровнях воды подаются в виде объектов на карту OpenStreetMap с целью их масштабирования. Далее выполняется этап кластеризации, необходимый для преобразования объектов в формат GeoJSON (JSON-файл) согласно математической модели, представленной в параграфе 1. В итоге получаем новый слой, накладываемый на OpenStreetMap посредством его внедрения в HTML-файл с применением JavaScript.

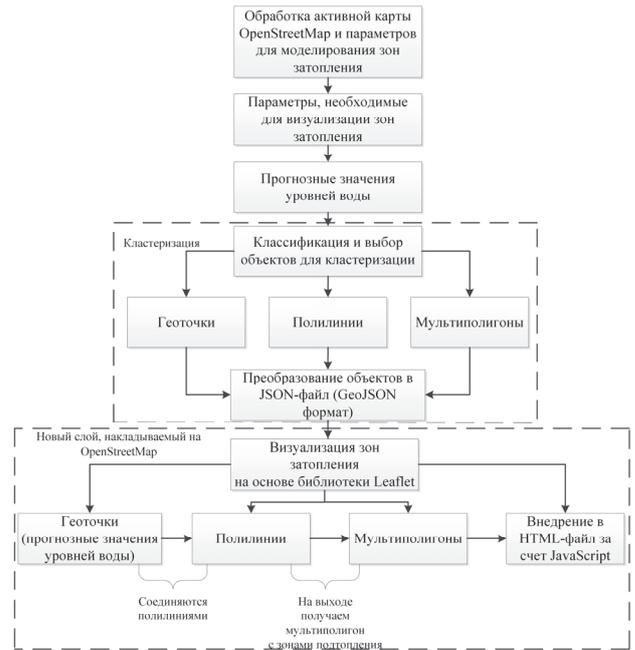


Рис. 5. Алгоритм построения зон затопления посредством веб-ГИС «Паводок»

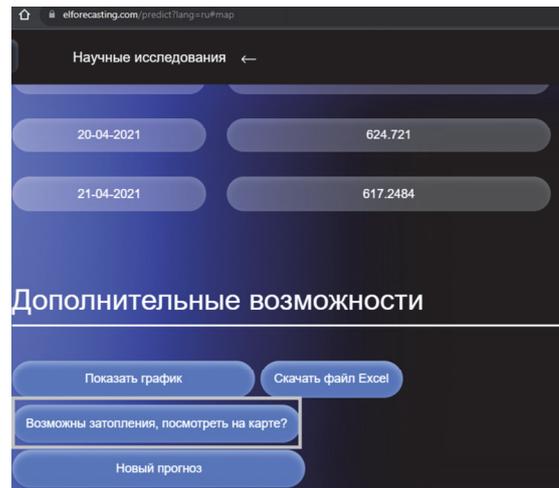


Рис. 6. Инициализация модуля моделирования зон затопления

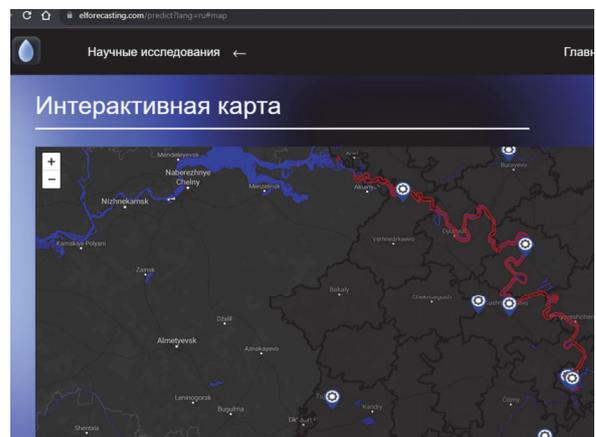


Рис. 7. Визуализированные зоны затопления на реке «Белая» во время паводка 2021 года в Республике Башкортостан

3. Сравнительный анализ результатов прогнозирования УВ и моделирования ЗЗ

В этом параграфе сравниваются модели прогнозирования УВ, реализованные в системе «Паводок», со статистическими моделями прогнозирования. Эксперимент, результаты которого представлены в табл. 2, проводился с прогнозом на 5 дней и по архитектуре моделей, а также их настройкам полностью идентичен эксперименту, представленному в первой части исследования [27]. Для удобства результаты представлены в виде средних ошибок.

Табл. 2. Сравнение моделей прогнозирования УВ системы «Паводок» со статистическими моделями прогнозирования УВ

Общая средняя ошибка при прогнозе УВ на 5 дней, см /%	
RNN (система «Паводок»)	50,96/11,55
LSTM (система «Паводок»)	175,92/39,86
PNN (система «Паводок»)	242,46/54,93
Полиномиальная регрессия (система «Паводок»)	109,28/24,78
Градиентный бустинг (библиотека XGBoost)	75,18/17,03
Простое экспоненциальное сглаживание (система «Паводок»)	156,32/35,42
ARIMA (библиотека Pmdarima)	86,64/19,63
SARIMA (библиотека Pmdarima)	72,19/16,36

Таким образом, на основании экспериментальных результатов, представленных в табл. 2, можно отметить, что модель рекуррентной нейронной сети (RNN) из системы «Паводок» является наиболее точной в прогнозировании УВ. Например, она точнее модели SARIMA на 4,81 %, в то время как разница в точности с другими моделями из табл. 2 составляет от 8,08 % до 43,38 %. Следует также отметить, что другие нейросетевые модели (табл. 2) показали значительно худшие результаты, чем статистические модели, что связано с особенностями их математического аппарата, который также будет доработан.

Далее сравниваются результаты (табл. 2) моделей прогнозирования при моделировании зон затопления. По методике, описанной в параграфе 1, были взяты прогнозные значения каждой модели и подставлены в формулы (4–5). Эксперимент проводился аналогично эксперименту по прогнозированию уровней воды. Результаты представлены в табл. 3.

Исходя из результатов, представленных в табл. 3, можно сделать вывод, что точность прогноза уровня воды напрямую влияет на прогнозную модель зон затопления. Средняя точность прогноза варьируется от 5,41 % до 27,82 %. Причем рекуррентная нейронная сеть, разработанная авторами статьи, показала наиболее точный результат при моделировании зон затопления (средняя ошибка составила 5,41 %). Градиентный бустинг, SARIMA, ARIMA, а также полиномиальная регрессия показали хорошие прогностические результаты. Высокая точность прогнозирования подтверждает эффективность разработанной системы «Паводок».

Табл. 3. Сравнение моделей прогнозирования системы «Паводок» со статистическими моделями при моделировании ЗЗ (площади затопления)

Общая средняя ошибка при прогнозе площади затопления, км ² /%	
RNN (система «Паводок»)	0,000536/5,41
LSTM (система «Паводок»)	0,000143/21,10
PNN (система «Паводок»)	0,00019/27,82
Полиномиальная регрессия (система «Паводок»)	0,000079/11,49
Градиентный бустинг (библиотека XGBoost)	0,000060/8,76
Простое экспоненциальное сглаживание (система «Паводок»)	0,000121/17,82
ARIMA (библиотека Pmdarima)	0,000065/9,56
SARIMA (библиотека Pmdarima)	0,000058/8,54

4. Обсуждение результатов

Предложенная веб-ГИС «Паводок» [25, 26] находится в постоянном развитии бэкенда (совершенствование математического аппарата искусственных нейронных сетей) и фронтенда (алгоритмы и математический аппарат моделирования зон затопления). В текущей вариации системы доступна следующая основная функциональность: визуализация стационарных гидрологических постов и автоматических станций (рис. 1), динамическое прогнозирование (рис. 8) по каждому гидрологическому посту и автоматической станции, а также моделирование зон затопления за счет прогнозирования значений уровней воды (рис. 7).

В рамках неоднократных экспериментов с использованием системы «Паводок» по ежедневному расчету прогнозных значений УВ и ЗЗ подтверждается эффективность предлагаемой системы.

К ограничениям данной системы следует отнести:

1. В текущей вариации, несмотря на эффективность веб-ГИС, все-таки невозможно качественно и точно прогнозировать и моделировать ЗЗ на 6 и более суток при сильных скачках УВ.
2. Специфичность модели: нейронная сеть, разработанная в [27], может работать только с ограниченным количеством параметров. Это связано с особенностями ее математического аппарата. В настоящий момент ведутся исследования, направленные на устранение данного недостатка. По результатам исследований модель станет более универсальной, что позволит добавлять параметры и данные напрямую из веб-интерфейса системы. Это предоставит возможность применять разработанную модель ИНС и в других направлениях, а не только в паводковых ситуациях. Частично данное решение применялось в электроэнергетике [24].
3. В рамках разработанной веб-ГИС реализована библиотека машинного и глубокого обучения, одним из методов прогнозирования в которой является реализованная в [27] ИНС. При этом присутствует невозможность сравнения точности прогнозных значений УВ нейросетевых и статистиче-

ских методов, как следствие, отсутствует необходимая функциональность по сравнению с 33 при прогнозировании различными моделями. Это связано с тем, что большинство методов прогнозирования не реализованы в рамках нашей библиотеки машинного и глубокого обучения. Устранение данного ограничения позволит использовать реализованную систему не только в рамках паводковых ситуаций, но и в других фундаментальных, а также прикладных исследованиях.

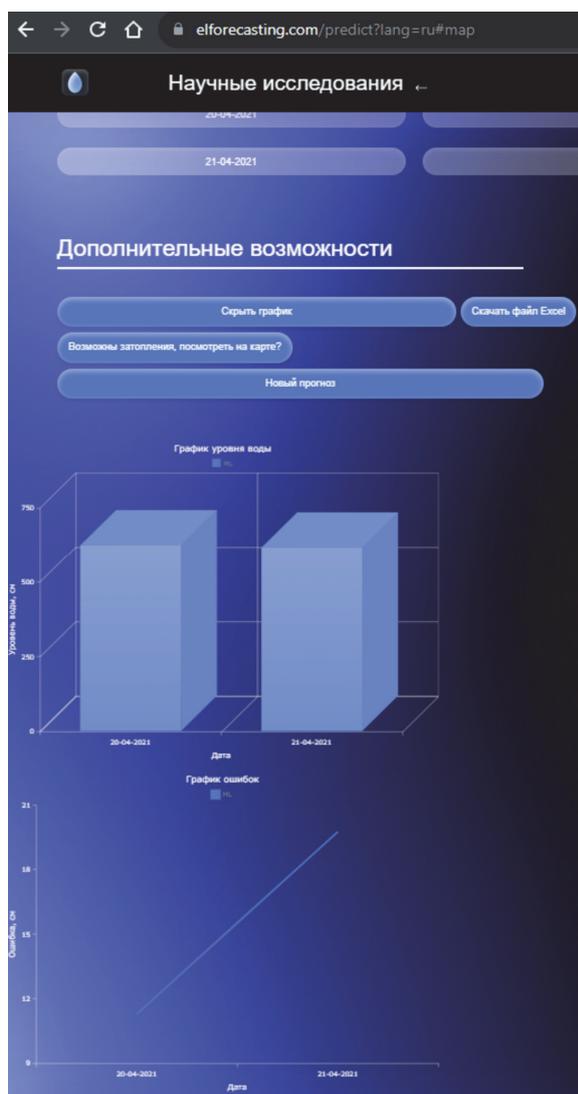


Рис. 8. Визуализация, прогнозирование и сравнение прогнозных /фактических значений уровней воды по гидрологическому посту 76289 г. Уфа

Заключение

Задача прогнозирования уровней воды и моделирования зон затопления с применением геоинформационных технологий является актуальной и необходимой, поскольку материальный ущерб от затоплений растет с каждым годом. В результате проведенных исследований была реализована веб-ГИС, позволяющая прогнозировать уровни воды по стационар-

ным гидрологическим постам и автоматическим станциям, а также моделировать зоны затопления с их последующей визуализацией.

Основной особенностью данной системы является то, что при наличии только статистических (числовых) данных и отсутствии цифровых моделей рельефа рассчитываются достаточно точные значения временных рядов, на основе которых моделируются зоны затопления.

Одним из перспективных направлений развития данной веб-ГИС является возможность прямого получения данных с автоматических станций мониторинга уровней воды каждые 15 минут, что увеличит количество прогнозов (до 5 в сутки) и их точность, а также позволит повысить значение горизонта прогнозирования до 7 дней (сейчас максимум – 5 суток). Это предоставит возможность специализированным службам более гибко реагировать на паводковые ситуации, как следствие, эффективнее проводить противопаводковые мероприятия.

Благодарности

Работа поддержана Министерством науки и высшего образования Российской Федерации в рамках базовой части государственного задания для высших учебных заведений # FEUE 2023-0007.

Коллектив авторов выражает благодарность ФГУП «Центр регистра и кадастра» за предоставленный массив архивных темпоральных данных.

References

- [1] Ministry of Emergency Situations of Russia – Floods in the territory of the Russian Federation are recognized as emergency situations of a federal nature [In Russian]. 2021. Source: <<https://mchs.gov.ru/deyatelnost/press-centr/novosti/4536592>>.
- [2] Shreevastav BB, Tivari KR, Mandal RA, Bikram BS. Flood risk modeling in southern Bagmati corridor, Nepal (a study from Sarlahi and Rautahat, Nepal). Prog Disaster Sci 2022; 16: 100260.
- [3] Ekmekcioğlu O, Koc K, Özger M, Işık Z. Exploring the additional value of class imbalance distributions on interpretable flash flood susceptibility prediction in the Black Warrior River basin, Alabama, United States. J Hydrol 2022; 610: 127877.
- [4] Flood disaster hazards; Causes, impacts and management: A state-of-the-art review. Source: <<https://www.intechopen.com/chapters/74444>>.
- [5] Buckman S, Alarcon MA, Maigret J. Tracing shoreline flooding: Using visualization approaches to inform resilience planning for small Great Lakes communities. Appl Geogr 2019; 113: 102097.
- [6] Costache R, Pham QB, Avand M, Linh NT, Vojtek M, Vojteková J, Lee S, Khoi DN, Nhi PT, Dung TD. Novel hybrid models between bivariate statistics, artificial neural networks and boosting algorithms for flood susceptibility assessment. J Environ Manage 2020; 265: 110485.
- [7] Dahri N, Yousfi R, Bouamrane A, Abida H, Pham QB, Derdous O. Comparison of analytic network process and artificial neural network models for flash flood

- susceptibility assessment. *J Afr Earth Sci* 2022; 193: 104576.
- [8] Efremova OA, Kunakov YuN, Pavlov SV, Sultanov AKh. An algorithm for mapping flooded areas through analysis of satellite imagery and terrestrial relief features. *Computer Optics* 2018; 42(4): 695-703. DOI: 10.18287/2412-6179-2018-42-4-695-703.
- [9] Pavlov SV, Sokolova AV, Christodulo OI. Integration of geographic information technologies and digital image processing for responding to emergency situations at oil pipeline transportation facilities. *Computer Optics* 2022; 46(3): 483-491. DOI: 10.18287/2412-6179-CO-925.
- [10] Zhang X, Bao W, Sun Y. Enhancing the hydrologic system differential response method for flood forecasting correction. *J Hydrol* 2021; 592: 125793.
- [11] Mourato S, Fernandez P, Marques F, Rocha A, Pereira L. An interactive Web-GIS fluvial flood forecast and alert system in operation in Portugal. *Int J Disaster Risk Sci* 2021; 58: 102201.
- [12] Cai B, Yu Y. Flood forecasting in urban reservoir using hybrid recurrent neural network. *Urban Climate* 2022; 42: 101086.
- [13] Ebtehaj I, Bonakdari H. A reliable hybrid outlier robust non-tuned rapid machine learning model for multi-step ahead flood forecasting in Quebec, Canada. *J Hydrol* 2022; 614(B): 128592.
- [14] Abbaszadeh P, Muñoz DF, Moftakhari H, Jafarzadegan K, Moradkhani H. Perspective on uncertainty quantification and reduction in compound flood modeling and forecasting. *iScience* 2022; 25(10): 105201.
- [15] Souto L, Yip J, Wu WY, Austgen B, Kutanoglu E, Hasenbein J, Yang ZL, King CW, Santoso S. Power system resilience to floods: Modeling, impact assessment, and mid-term mitigation strategies. *Int J Electr Power Energy Syst* 2022; 135: 107545.
- [16] Wanzala MA, Ficchi A, Cloke HL, Stephens EM, Badjana HM, Lavers DA. Assessment of global reanalysis precipitation for hydrological modelling in data-scarce regions: A case study of Kenya. *J Hydrol: Reg Studies* 2022; 41: 101105.
- [17] Phan EE, Nguyen XH. Combining statistical machine learning models with ARIMA for water level forecasting: The case of the Red river. *Adv Water Resour* 2020; 142: 103656.
- [18] Adaryani FR, Mousavi SJ, Jafari F. Short-term rainfall forecasting using machine learning-based approaches of PSO-SVR, LSTM and CNN. *J Hydrol* 2022; 614(A): 128463.
- [19] Wang G, Ren H-L, Liu J, Long X. Seasonal predictions of sea surface height in BCC-CSM1.1m and their modulation by tropical climate dominant modes. *Atmos Res* 2023; 281: 106466.
- [20] Xu H, Song S, Guo T, Wang H. Two-stage hybrid model for hydrological series prediction based on a new method of partitioning datasets. *J Hydrol* 2022, 612(A): 128122.
- [21] Bui DT, Ngo PT, Pham TD, Jaafari A, Minh NQ, Hoa VH, Samui P. A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping. *Catena* 2019; 179: 184-196.
- [22] Jiang Z, Yang S, Liu Z, Xy Y, Xiong Y, Qi S, Pang Q, Xu J, Liu F, Xu T. Coupling machine learning and weather forecast to predict farmland flood disaster: A case study in Yangtze River basin. *Environ Modell Softw* 2022; 155: 105436.
- [23] Adnan MS, Siam ZS, Kabir I, Kabir Z, Ahmed MR, Hassan QK, Rahman RM, Dewan A. A novel framework for addressing uncertainties in machine learning-based geospatial approaches for flood prediction. *J Environ Manage* 2023; 326(B): 116813.
- [24] Palchevsky EV, Antonov VV, Kromina LE, Rodionova LE, Fakhruullina AR. Intelligent forecasting of electricity consumption in managing energy enterprises in order to carry out energy-saving measures. *Mekhatronika, Avtomatizatsiya, Upravlenie* 2023; 24(6): 307-316.
- [25] «Flood» system. [In Russian]. 2023. Source: <<https://elforecasting.com/>>.
- [26] Palchevsky EV, Antonov VV. Module for storage, processing and preparation of spatial information for flood forecasting [In Russian]. Certificate of state registration of the computer program No 2023662065 of May 11, 2023.
- [27] Palchevsky E, Antonov V, Enikeev R, Breikin T. A system based on an artificial neural network of the second generation for decision support in especially significant situations. *J Hydrol* 2023; 616: 128844.
- [28] Pontryagin LS. Fundamentals of combinatorial topology. [In Russian]. Moscow: "Nauka" Publisher; 1986.
- [29] Goldberg AV, Tarjan RE. Efficient maximum flow algorithms. *Commun ACM* 2014; 57(8): 82-89.
- [30] Kelner A, Lee YT, Orecchia L, Sidford A. An almost-linear-time algorithm for approximate max flow in undirected graphs, and its multicommodity generalizations. *Proc ACM SIAM Symposium on Discrete Algorithms* 2014: 217-226.

Сведения об авторах

Пальчевский Евгений Владимирович, 1994 года рождения, в 2019 году окончил магистратуру Уфимского государственного авиационного технического университета по специальности 090402 «Информационные системы и технологии», работает преподавателем в департаменте анализа данных и машинного обучения Финансового университета при Правительстве Российской Федерации. Область научных интересов: интеллектуальные вычисления, методы и технологии искусственного интеллекта, геоинформационные системы. E-mail: teelxp@inbox.ru

Антонов Вячеслав Викторович, 1956 года рождения, в 1979 году окончил Башкирский государственный университет по специальности «Математик, вычислительная математика», доктор технических наук, профессор и заведующий кафедрой автоматизированных систем управления в Уфимском университете науки и технологий. Область научных интересов: искусственный интеллект и принятие решений, интеллектуальный анализ данных, математическое моделирование технических, социальных и экономических процессов, программные модели и системы. E-mail: antonov.v@bashkortostan.ru

Родионова Людмила Евгеньевна, 1984 года рождения, в 2006 окончила Уфимский государственный авиационный технический университет по специальности «Прикладная информатика в экономике», кандидат технических наук, доцент кафедры автоматизированных систем управления Уфимского университета науки и технологий. Область научных интересов: разработка математического и программного обеспечения в области интеллектуальных информационных систем.

E-mail: RodionovaKF@yandex.ru

Кромина Людмила Александровна, 1983 г. рождения, в 2005 окончила Уфимский государственный авиационный технический университет по специальности «Прикладная информатика в экономике», кандидат технических наук, доцент кафедры автоматизированных систем управления Уфимского университета науки и технологий. Область научных интересов: управление в социально-экономических системах. E-mail: luydakra@yandex.ru

Фахруллина Альмира Раисовна, 1982 года рождения, в 2004 окончила Уфимский государственный авиационный технический университет по специальности «Информационные системы в экономике», кандидат технических наук, доцент кафедры автоматизированных систем управления. Директор филиала Уфимского университета науки и технологий в г. Кумертау. Область научных интересов: Разработка математического и программного обеспечения в области интеллектуальных информационных систем. E-mail: almirafax@mail.ru

ГРНТИ: 28.23.29

Поступила в редакцию 29 августа 2023 г. Окончательный вариант – 26 февраля 2024 г.

Modeling flood zones on the basis of time series forecasting and GIS-technologies by the example of the Republic of Bashkortostan

E.V. Palchevsky¹, V.V. Antonov², L.E. Rodionova², L.A. Kromina², A.R. Fakhrullina²

*¹ Financial University under the Government of the Russian Federation,
109456, Russia, Moscow, st. 4th Veshnyakovsky passage 4,*

² Ufa University of Science and Technology, 450008, Russia, Ufa, st. Karl Marx 12

Abstract

A specialized GIS Web is proposed, implemented by integrating an artificial neural network and geotechnologies and providing early forecasting and modeling of flood zones up to five days in advance.

The methods and algorithms implemented within this GIS Web allow daily forecasting of time series based on retrospective data on water levels and total water inflow, air and water temperature, snow cover thickness and precipitation, wind speed and atmospheric pressure. At the same time, the possibility of early modeling and visualization of river floods is realized only on the basis of the obtained predictive values of the water level. This will enable specialized organizations and services, as well as management bodies to make decisions related to flood control measures in advance and as soon as possible.

Keywords: geographic information system, flood zone modelling, time series forecasting, artificial neural networks.

Citation: Palchevsky EV, Antonov VV, Rodionova LE, Kromina LA, Fakhrullina AR. Modeling flood zones on the basis of time series forecasting and GIS-technologies. *Computer Optics* 2024; 48(6): 913-923. DOI: 10.18287/2412-6179-CO-1418.

Acknowledgements: The work was financially supported by the Ministry of Science and Higher Education of the Russian Federation within the main part of the state project for Higher Education institutions, # FEUE 2023-0007.

The authors wish to thank the Federal State Unitary Enterprise "Center for Register and Cadastre" for providing archival temporal data.

Authors' information

Evgeniy Vladimirovich Palchevsky, (b. 1994) in 2019 he graduated from the Ufa State Aviation Technical University with a master's degree in 090402 "Information Systems and Technologies", works as a teacher of Data Analysis and Machine Learning department of the Financial University under the Government of the Russian Federation. Research interests: intelligent computing, methods and technologies of artificial intelligence, geographic information systems. E-mail: teelp@inbox.ru

Vyacheslav Viktorovich Antonov, (b. 1956) graduated from Mathematics faculty of the Bashkir State University in 1979, Doctor of Technical Sciences, works as a Professor and Head of Automated Control Systems department at the Ufa University of Science and Technology. Research interests: artificial intelligence and decision making, data mining, mathematical modeling of technical, social and economic processes, software models and systems. E-mail: antonov.v@bashkortostan.ru

Lyudmila Evgenievna Rodionova, (b. 1984) in 2006 graduated from the USATU with a degree in Applied Informatics in Economics, Candidate of Technical Sciences, Associate Professor of Automated Control Systems department. Research interests: development of mathematical and software in the field of intelligent information systems. E-mail: rodionovaKF@yandex.ru

Lyudmila Alexandrovna Kromina, (b. 1983) in 2005 she graduated from the USATU with a degree in Applied Informatics in Economics, Candidate of Technical Sciences, Associate Professor of Automated Control Systems department. Research interests: management in socio-economic systems. E-mail: luyda-kr@yandex.ru

Almira Raisovna Fakhrullina, (b. 1982) in 2004 graduated from the USATU with a degree in Information Systems in Economics, Candidate of Technical Sciences, Associate Professor of Automated Control Systems department. Director of the branch of the Ufa University of Science and Technology in Kumertau. Research interests: development of mathematical and software in the field of intelligent information systems. E-mail: almirafax@mail.ru

Received August 29, 2023. The final version – February 26, 2024.
