УДК 519.68 Дата подачи статьи: 18.02.22, после доработки: 11.05.22
DOI: 10.15827/0236-235X.139.384-395 2022. Т. 35. № 3. С. 384–395

Прогнозирование на основе искусственной нейронной сети второго поколения для поддержки принятия решений в особо значимых ситуациях

Е.В. Пальчевский [1], *преподаватель, teelxp@inbox.ru*
В.В. Антонов [2], *д.т.н., профессор, antonov.v@bashkortostan.ru*
Р.Р. Еникеев [2], *к.т.н., доцент, enikeevRR@mail.ru*

[1] *Финансовый университет при Правительстве Российской Федерации, г. Москва, 109456, Россия*
[2] *Уфимский государственный авиационный технический университет, факультет ИРТ, г. Уфа, 450008, Россия*

Специализированные модели систем, реализованные на основе поддержки принятия решений с помощью искусственного интеллекта для воспроизведения, прогнозирования и предотвращения последствий исключительных ситуаций, становятся все более популярными. Под такие ситуации и состояния подпадают и паводки. Исходя из этого появляется задача прогнозирования исключительной ситуации на примере подъема уровней воды на стационарных гидрологических постах.

Для решения данной проблемы предложена система поддержки принятия решений для прогнозирования уровней подъема воды, основанная на нейросетевом анализе ретроспективных данных с целью расчета значений уровней воды на пять дней вперед. За основу самой искусственной нейронной сети берется свободно распространяемая библиотека программ машинного обучения TensorFlow. В качестве обучения применен измененный метод обратного распространения ошибки. Основным отличием метода является добавление коэффициента повышения скорости обучения искусственной нейронной сети.

Был проведен анализ эффективности предлагаемого решения в рамках прогнозирования паводковой ситуации. Он показал высокую точность при расчетах прогнозных значений уровней воды: разница между реальными и спрогнозированными значениями составила 2,10 %. Это позволит специализированным службам заблаговременно осуществлять противопаводковые мероприятия.

Таким образом, информационное сопровождение во время особых ситуаций – это абсолютный показатель качества данных, позволяющий вырабатывать и принимать решения в рамках прогнозирования возможных критических ситуаций и недопущения перевода состояния системы управления территорией в критическое.

Ключевые слова: *прогнозирование, интеллектуальный анализ данных, нейросетевой анализ, ретроспективные данные, прогнозирование уровней воды.*

Появление новых источников и технологий получения и обработки данных об объектах и системах, в том числе их местоположении, обусловило формирование принципиально новых условий для более качественного управления различными процессами и явлениями. При этом в рамках управления ими возникает некая двойственность ситуации: с одной стороны, существующие модели и методы получения и обработки данных используют достаточно большие объемы разнородной информации, на обработку которой требуется много времени, с другой, существуют ограничивающие факторы, например время, для принятия каких-либо управленческих решений на основе имеющихся данных. Соответственно, данные особенности в качестве внешнего фактора способны напрямую повлиять на принятие решений в рамках различных процессов и явлений, так как способствуют угрозам различным целостным системам, например городским поселениям, и подсистемам. Таким образом, для принятия верного управленческого решения при подобных угрозах необходимо заблаговременно получать данные о происходящих процессах и явлениях. Сложность этой проблемы определяется достаточно большим количеством управляющих параметров, нечеткостью требований и изменчивостью целей управления со стороны различных пользователей систем и участников процесса управления, необходимостью минимизировать уязвимость этих систем при наличии большого количества угроз различной природы и сложности. Для ре-

шения данной проблемы предлагается получать информацию о возможном будущем явления или процесса (на примере прогнозирования уровней воды водных ресурсов) с помощью системы, реализованной на основе искусственной нейронной сети второго поколения.

Большое распространение в прогнозировании уровней воды получили различные методы, подразделяющиеся, как правило, на концептуальные модели, физически обоснованные модели и модели черного ящика. Для первых двух категорий необходимо знать физический процесс уровня подъема воды на водных ресурсах, лежащий в фундаменте самой проблемы и описываемый с помощью каких-либо упрощенных соотношений (например, различные статистические методы расчетов [1–3]) или уравнений в частных производных (численные методы [4, 5]). При этом важно отметить, что первые два метода направлены на краткосрочное прогнозирование (от 1 до 2 суток). Что касается модели черного ящика, сюда относят машинное обучение, в том числе и с использованием искусственного интеллекта, нейронных сетей. Методы машинного обучения могут применяться и в рамках различных гибридных подходов, например, при объединении разных статистических моделей машинного обучения и методологий анализа временных рядов [6, 7]. Гибридные подходы могут использовать преимущества каждого из них по отдельности, в том числе и компонентные модели данных подходов, для достижения повышенной точности и гибкости прогнозирования уровней воды и зон затопления. Как следствие, такие модели показывают достаточно точные результаты в краткосрочном прогнозировании, но являются ресурсозатратными по времени. В рамках использования методов искусственного интеллекта, как правило, преобладают простые модели, например MLP [8–10], и сложные структурные нейросетевые решения – LSTM [11–13], CNN [14–17], WNN [18, 19]. Кроме того, применяются модели с использованием большого количества статистических методов – ARMA, ARIMA [20–22], ANFIS [23], NARX [24–26] и другие. Данные методы и модели позволяют не только достаточно точно рассчитывать прогнозные значения уровней воды при создании и обработке исключительных ситуаций, но и модифицировать их под различные задачи.

Исследование, проведенное авторами данной статьи, посвящено разработке метода анализа и направлено на прогноз состояний системы на основе показателей уровня паводковых вод.

Рассмотрим *информационную систему* (ИС) управления территорией (рис. 1), которая должна обладать рядом общесистемных свойств, в том числе устойчивостью, адаптируемостью и эмерджентностью, и обозначим ее *A*. Здесь под устойчивостью ИС понимается некая способность к возвращению ее в равновесное состояние после воздействия какого-либо внешнего возмущения, например, природного явления, нарушившего данное равно-

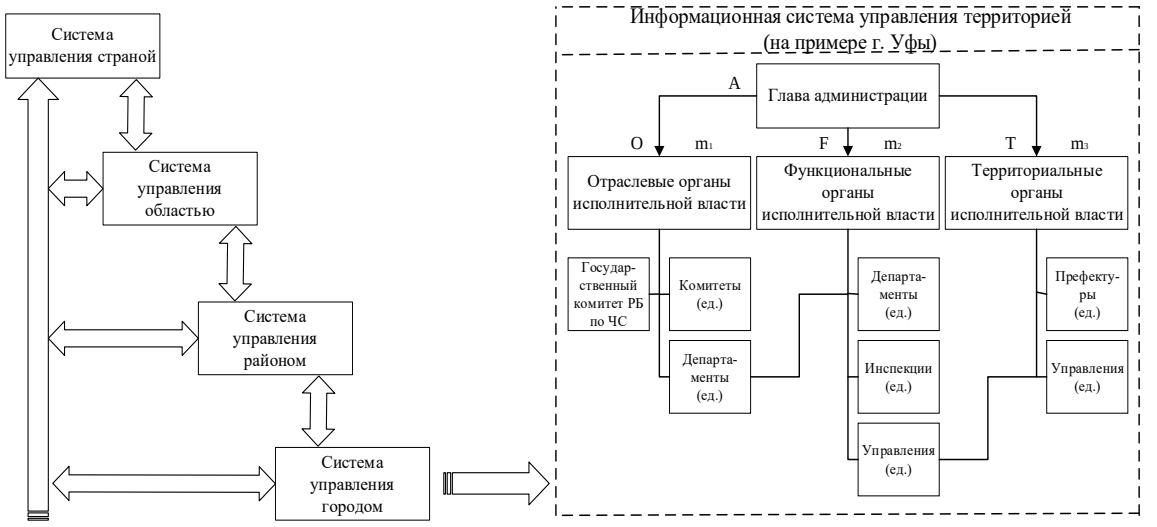

Рис. 1. Информационная система управления территорией

Fig. 1. A territory management information system

весие, под адаптируемостью – способность изменять структуру с целью сохранения показателей состояния в условиях изменения внешней среды. При этом у данной системы есть подсистемы (на рисунке 1 – *O*, *F* и *T*), изучив свойства которых, можно определить свойства самой ИС – эмерджентность. Более того, ИС и все ее подсистемы, несмотря на изменения взаимодействующей с ней внешней среды, должны сохранять свойства целостности.

Рассматриваемая ИС является сложным техническим объектом, может быть представлена в виде множества взаимодействующих подсистем *O*, *F*, *T*, и ее можно охарактеризовать множеством состояний:

$$A = \langle O, F, T \rangle. \quad (1)$$

Процесс эксплуатации такой ИС будет сопровождаться сменой состояний во времени в результате реализации отношений между ее подсистемами и внешней средой. При этом у ИС могут появляться свойства, которыми не обладают ее подсистемы. Для дальнейшего анализа следует отождествить приведенные системы с множествами состояний систем. Учитывая, что таких состояний достаточно много (в зависимости от мощности m_1, m_2, m_3 приведенных подмножеств *O*, *F*, *T*), может быть применена нечеткая кластеризация этих состояний по параметрам воздействия на предметную область. Очевидно, что не все состояния (или соответствующие кластеры) входят в подмножество разрешенных состояний, для которых возможно путем цепочки управляющих воздействий восстановить устойчивое состояние системы. Причем определение состояний, предшествующих попаданию ИС в одно из критических состояний, возможно за счет превентивной аналитики, прогнозирования или их моделирования. Речь идет об исследовании цепочек разрешенных состояний системы, приводящих к критическим состояниям, вызывающим нарушение устойчивости ИС, в том числе к состояниям, которые могут характеризоваться как новые функции и свойства ИС. Можно говорить о поведении ИС. Таким образом, множество *A* будет состоять из подмножеств *B* – разрешенных состояний и *C* – критичных состояний.

Задача перевода части состояний множества *C* в множество *B* весьма актуальна. Это множество можно классифицировать на подмножество простых отображений (алгоритмизируемых, например, методами линейного программирования или системами конверторов) и сложных, не поддающихся алгоритмизации. Большой интерес вызывают состояния системы (множество *C*), отображения которых в *B* можно реализовать экспертными системами, в том числе с использованием систем искусственного интеллекта, машинного обучения и нейронных сетей. При этом могут быть выстроены цепочки отображений, приводящие к разрешенному состоянию, включенному в *B*, и выявлены потенциально возможные ситуации или цепочки состояний, приводящие к исключительным состояниям, не имеющим решения в *B*. Следовательно, каждая из приведенных на рисунке 1 подсистем характеризуется множеством состояний, которые могут быть взяты за измерения – независимые векторы. Дальнейшая декомпозиция таких подсистем приводит к некоторому многомерному пространству, а независимые векторы могут быть взяты в качестве базиса рассматриваемого многомерного пространства. Таким образом, общая математическая модель ИС (рис. 1) представима множеством пространств – *I*, с введенными в них измерениями согласно нормативной базе каждой подсистемы. Итак, любая точка $(d_1, \ldots, d_n)$, $d_n \in D_i$, $i \in I$, вместе с мерой $(m_1, \ldots, m_k)$, $m_k \in M_i$, $i \in I$, определяют схему каждого многомерного пространства $R_i = \langle D_i, M_i \rangle$, $i \in I$. Также очевидно, что для любой точки $A \in D_i$ существует отношение (проекция) $r[A]$ и любая точка этой проекции будет одновременно значением измерения $\forall\ a \in r[A]$: $a \in ALL$. Тогда формула $Space(r)_i = \{\otimes_{A \in Di}(r[A] \cup ALL)\}$, $i \in I$, может быть взята за математическую модель многомерного пространства.

На основе приведенной совокупности измерений напрашивается вывод о возможности рассмотрения результатов управления по многомерному пространству через последовательный анализ по каждому подпространству. Полученные независимые результаты могут быть агрегированы по любым измерениям на основе ранее указанных. Далее открывается возможность установления новых отношений между уже имеющимися данными и агрегированными, что будет образовывать цепочки состояний ИС и в итоге приведет к многомерной модели, которая учитывает добавленные измерения. Приходим к многомерным информационным объектам – геоинформационным системам.

Рассмотрим применение искусственной нейронной сети второго поколения, основанной на MLP, для заблаговременного прогнозирования на примере подъема уровней воды на стационарных гидрологических постах с це-

лью недопущения перехода системы A в исключительное состояние.

Модели и методы

Одним из основных факторов, влияющих на состояние ИС управления территориями, является паводок, выражающийся в подъеме значений уровней воды (h) на водных объектах. В качестве примера можно взять стационарные гидрологические посты Республики Башкортостан. Введем обозначения: h_{ji}^{k} – значение уровня воды, измеренное на k-м гидропосту на i-ю дату j-го года. Здесь $k = 1, \ldots, n$, где n – количество гидропостов, участвующих в расчетах; j – номер года; i – конкретная дата измерения. Отметим, что задача краткосрочного прогнозирования на 1 сутки была решена в [27], и на основе предложенной методики выполнено прогнозирование на 2–5 суток. Задача заблаговременного прогнозирования заключается в том, чтобы в конкретный текущий i-й день измерения вычислить значение уровня воды на $i + 2$ день, то есть h_{ji+2}^{k}, а также на $i + 3$ (h_{ji+3}^{k}), $i + 4$ (h_{ji+4}^{k}) и $i + 5$ (h_{ji+5}^{k}) $\forall\ k = 1, \ldots, n$.

Для решения поставленной задачи предлагается использовать результаты предыдущих измерений уровней воды (массив больших ретроспективных данных) на всех стационарных гидрологических постах, находящихся на рассматриваемой территории, за все предыдущие годы. Предлагаемый метод заблаговременного прогнозирования основан на применении искусственной нейронной сети (рис. 2), реализованной с использованием свободно распространяемой библиотеки программ машинного обучения TensorFlow, настройки которой представлены далее. В данную нейронную сеть посредством API интегрирован измененный авторами статьи [28] метод обучения ИНС backpropagation (обратного распространения ошибки). Как показали результаты эксперимента, точность и время расчета прогноза доказывают целесообразность применения данного метода в рамках заблаговременного прогнозирования значений уровней воды на стационарных гидрологических постах. Таким образом, процесс прогнозирования предлагается проводить в три этапа (рис. 3): предобработка данных, формирование прогнозной модели и постобработка данных. Частично данный метод применялся в рамках обнаружения и нейтрализации киберугроз [29].

На первом этапе, выполняющемся, как правило, до развития паводковой ситуации, осуществляется выбор параметров ИНС: общее количество нейронов – 107; количество нейронов во входном ($\overrightarrow{h_{jiK}^{k}}$) – 6 и выходном ($\overrightarrow{h_{jiL}^{k}}$) – 5 слоях, количество промежуточных слоев – 3 и нейронов в них (H_{jk}^{k}) – 96. Функция актива-

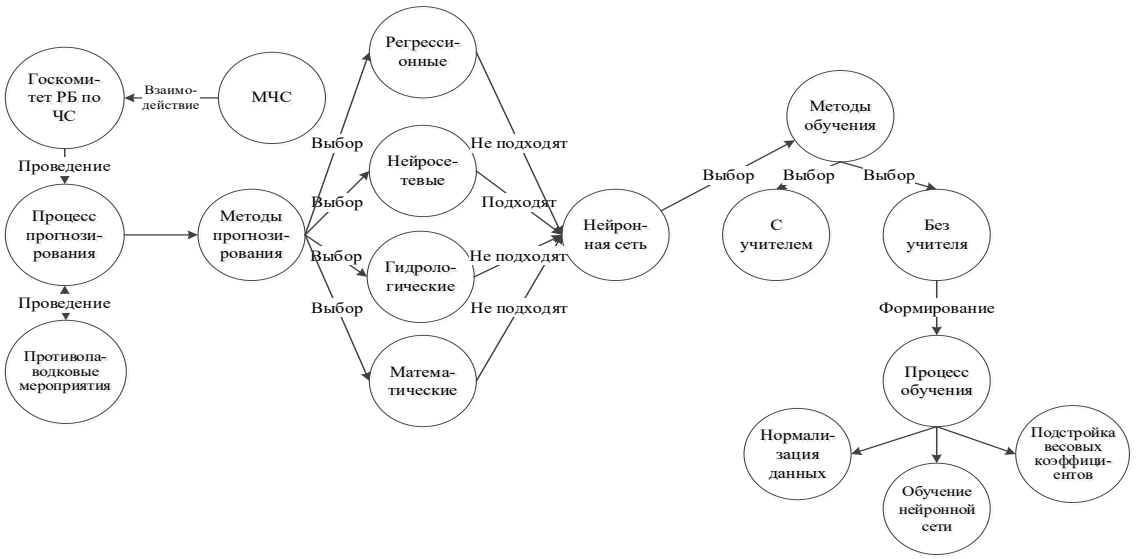

Рис. 2. Функциональные характеристики простейшего потока событий при прогнозировании паводковой ситуации

Fig. 2. Functional characteristics of the simplest flow of events when forecasting a flood situation

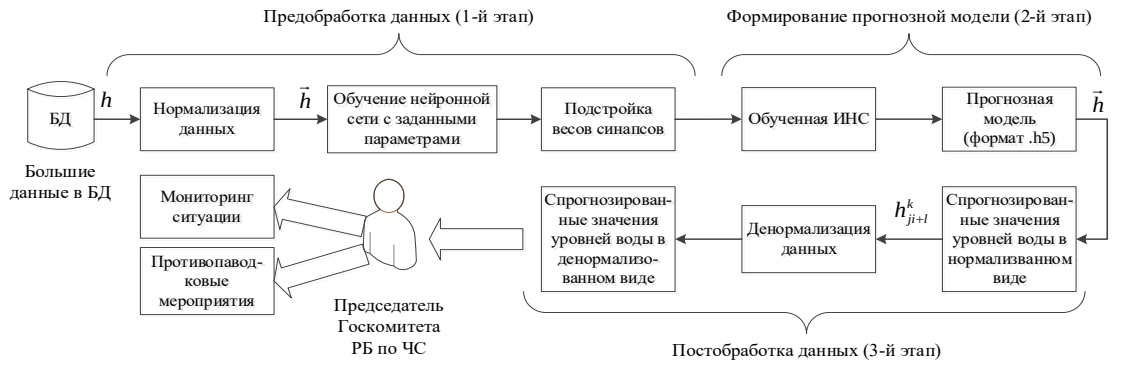

Рис. 3. Схема применения нейронной сети для прогнозирования значений уровней воды

Fig. 3. A scheme of using a neural network to predict water level values

ции на промежуточных слоях является сигмоидальной, при этом на выходном слое используется линейная функция активации по причине ее взвешенной суммы на каждом нейроне выходного слоя [29]. Количество данных по выборкам: обучающая – 60 %, валидационная – 20 % и проверочная – 20 %, количество эпох – 200; скорость обучения α – 0.90, подобранная экспериментальным путем с дальнейшим обучением искусственной нейронной сети. Стоит отметить, что при применении метода обратного распространения ошибки необходимо нормализовать входные значения уровней воды H для предотвращения исчезающих и взрывающихся градиентов:

$$\overrightarrow{h_{ji}^{k}} = \frac{\overrightarrow{h_{ji}^{k}} - \overrightarrow{h_{\min}}}{\overrightarrow{h_{\max}} - \overrightarrow{h_{\min}}}, \quad (2)$$

где $\overrightarrow{h_{\max}}$ и $\overrightarrow{h_{\min}}$ – минимальное и максимальное значения уровней воды по всем данным, подаваемым на входной слой нейронной сети для всех $\overrightarrow{h_{ji}^{k}}$, где $k = 1, \ldots, n$.

Для обучения нейронной сети (рис. 4) используется измененный авторами статьи метод обратного распространения ошибки. Изменение заключается в применении задаваемого экспериментальным путем параметра α – скорости обучения нейронной сети.

Вычисление квадратичной ошибки. Как правило, входы сети рассматриваются как входной вектор $\overrightarrow{h_{ji}^{k}}$, где $\overrightarrow{h_{ji}^{k}} = [\overrightarrow{h_{ji1}^{k}}, \overrightarrow{h_{ji2}^{k}}, \ldots, \overrightarrow{h_{jiK}^{k}}]$, а выходы сети можно представить в качестве вектора выхода $\overrightarrow{h_{ji}^{k}}$, где $\overrightarrow{h_{ji}^{k}} = [\overrightarrow{h_{ji1}^{k}}, \overrightarrow{h_{ji2}^{k}}, \ldots, \overrightarrow{h_{jiL}^{k}}]$, при этом $K = 1, \ldots, M$, а $L = 1, \ldots, N$, в которых M и N – количество значений входного и выходного векторов. Соответственно, обучающая выборка представляет собой множество пар R входных векторов $\overrightarrow{h_{jiK}^{k}}$ и желаемых (эталонных) выходных векторов $\overrightarrow{h_{jiL}^{k}}$:

$$R = \{(\overrightarrow{h_{ji1}^{k}}, \overrightarrow{h_{ji1}^{k}}), (\overrightarrow{h_{ji2}^{k}}, \overrightarrow{h_{ji2}^{k}}), \ldots, (\overrightarrow{h_{jiK}^{k}}, \overrightarrow{h_{jiL}^{k}})\}. \quad (3)$$

При каждом применении $\overrightarrow{h_{jiK}^{k}}$ от R к нейронной сети будет произведен расчет фактического выхода $\overrightarrow{h_{jiL}^{k}}$ выходного слоя:

$$\overrightarrow{h_{jiL}^{k}} = f(\sigma_L), \quad (4)$$

где σ – функция активации, определяемая по соотношению

$$\sigma(\overrightarrow{h_{jiK}^{k}}) = \frac{1}{1 + e^{-\overrightarrow{h_{jiK}^{k}}}} \quad (5)$$

и являющаяся взвешенной суммой выходов нейрона j в каждом промежуточном слое H_j:

$$\sigma_L = \sum_{j'=1}^{j} w_{j'k}^{j} H_{j'k}^{j}, \quad (6)$$

где w – вес синапса; k – k-й нейрон в промежуточном слое.

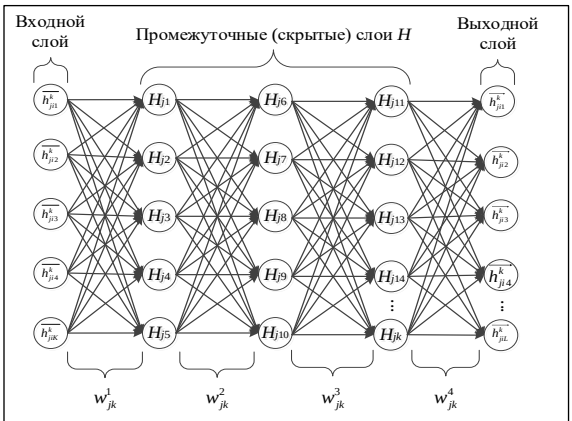

Рис. 4. Структура нейронной сети

Fig. 4. The neural network structure

Исходя из этого, можно определить квадратичную ошибку для каждой пары векторов множества R путем суммирования квадратичных ошибок в каждом выходном нейроне:

$$E_l = \frac{1}{2}\sum_{k'=1}^{L}(\overrightarrow{h^k_{jiK_{k'}}} - \overrightarrow{h^k_{jiL_{k'}}})^2. \quad (7)$$

Как следствие, вычислим полную квадратичную ошибку E путем суммирования всех пар входов и выходов в обучающей выборке:

$$E = \frac{1}{2}\sum_{l'=1}^{K}\sum_{k'=1}^{L}(\overrightarrow{h^k_{jiK_{k'l'}}} - \overrightarrow{h^k_{jiL_{k'l'}}})^2 . \quad (8)$$

Таким образом, цель обучения состоит в минимизации E за счет нахождения соответствующего набора весов w^1_{xk} и w^l_{jk}, где l изменяется от 1 до 4.

Определение частных производных по весам синапсов. Исходя из соотношения (7), можно отметить, что

$$\frac{dE}{dw^l_{jk}} = \overrightarrow{h^k_{jiK}} - \overrightarrow{h^k_{jiL}}, \quad (9)$$

при этом, опираясь на (6), получаем

$$\frac{d\sigma_k}{dw^l_{jk}} = H_j . \quad (10)$$

Исходя из (9), определяем частную производную E по весу w_{jk} для выполнения градиентного спуска для весов синапсов между промежуточным и выходным слоями:

$$\frac{dE}{dw^l_{jk}} = ((\overrightarrow{h^k_{jiK}} - \overrightarrow{h^k_{jiL}})\overrightarrow{h^k_{jiL}})\times \times((\sigma(\overrightarrow{h^k_{jiK}}) - \overrightarrow{h^k_{jiL}})H_{jk}). \quad (11)$$

При этом, если рассматривать частный случай производной E между входным и промежуточными слоями, то отдельно стоит отметить, что в этом случае все выходы зависят от w_{xk}. Таким образом, частную производную будем находить следующим образом:

$$\frac{dE}{dw^1_{xk}} = \sum_{k'=1}^{L}((\overrightarrow{h^k_{jiK_{k'}}} - \overrightarrow{h^k_{jiLk'}})\overrightarrow{h^k_{jiLk'}})\times \times(((\sigma(\overrightarrow{h^k_{jiK}}) - \overrightarrow{h^k_{jiLk'}})(w^l_{jk'}H_{jk'}))((1-H_{jk'})\overrightarrow{h^k_{jiK}}). \quad (12)$$

Следовательно, уравнения (9) и (10) дают все необходимые величины $\frac{dE}{dw}$ для применения (11) и (12) в рамках градиентного спуска для всех весов нейронной сети.

Изменение весов синапсов. Каждый вес будет изменен на dw для уменьшения E:

$$w(t+1) = w(t) + \Delta w(t), \quad (13)$$

где $\Delta w(t) = -\left.\frac{dE}{dw}\right|_t$, $w(t)$ – вес синапса во время t, $w(t + 1)$ – измененный (обновленный) вес синапса. Для увеличения скорости обучения нейронной сети добавляется ранее обозначенный параметр α:

$$\Delta w(t) = -\left.\frac{dE}{dw}\right|_t + \alpha\Delta w(t-1). \quad (14)$$

На втором этапе осуществляется непосредственно формирование прогнозной модели, а на третьем – прогноз значений уровней воды. Так как ранее все значения были нормализованы ($\overrightarrow{h^k_{ji}} \in [0;1]$), то для их дальнейшего использования в рамках противопаводковых мероприятий необходимо произвести денормализацию $\overrightarrow{h^k_{ji+l}}$ по соотношению, обратному (2):

$$\overrightarrow{h^k_{ji}} = \overrightarrow{h^k_{ji}}\left(\overrightarrow{h_{\max}} - \overrightarrow{h_{\min}}\right) + \overrightarrow{h_{\min}} . \quad (15)$$

Анализ эффективности метода в рамках прогнозирования уровней подъема воды

Для проведения анализа использовались многолетние данные измерений уровней воды на гидрологических постах, предоставленные Башкирским управлением по гидрометеорологии и мониторингу окружающей среды с 01.01.2000 г. по 02.05.2021 г. в следующем виде (БД, см. таблицу): код стационарного гидрологического поста/автоматической станции, дата, уровень воды на стационарном гидрологическом посту/автоматической станции, атмосферное давление, скорость ветра, толщина снежного покрова, количество осадков, время и температура воздуха.

В ходе эксперимента проведено множество итераций за весь период многолетних наблюдений (с 01.01.2000 по 06.05.2021) по расчету прогнозируемых значений уровней воды. Общий массив используемых в эксперименте данных составляет 1 025 815 значений, из которых 60 % (данные многолетних наблюдений в период с 01.01.2000 по 31.12.2013) подаются на обучающую выборку нейронной сети, 20 % (с 01.01.2014 по 31.12.2017) – на тестовую выборку и 20 % (с 01.01.2018 по 06.05.2021) – на проверочную, при этом глубина прогнозирования, на основе которой рассчитывается прогнозируемое значение, задается экспериментально в зависимости от стационарного гидрологического поста.

В качестве примера приведем результаты прогнозирования уровней воды, полученные на стационарном гидрологическом посту

Фрагмент исторических данных для прогнозирования

A historical data fragment for forecasting

ID	Object ID	Hydro-logical post code	Para-meter code	Date Time	Water level	Descrip-tion	Atmos-phere pressure	Wind speed	Snow thick-ness	Precipi-tation	Time	Air tempe-rature
0	11232	19123	1	2001-10-10 00:00:00	159	No anomalies	763	3	100	10	8:00	6
1	11233	19123	1	2001-10-11 00:00:00	201		762	5	110	25	8:00	3
2	11234	19123	1	2001-10-12 00:00:00	165		764	5	120	9	8:00	7
3	11235	19123	1	2001-10-20 00:00:00	141		754	4	115	12	8:00	4
4	11236	19123	1	2001-10-21 00:00:00	141		765	3	115	10	8:00	6
5	11237	19123	1	2001-10-22 00:00:00	143		763	4	125	5	8:00	3
6	11238	19123	1	2001-10-23 00:00:00	144		760	5	130	5	8:00	2

в г. Уфе (76289, река Белая) во время весеннего паводка 2021 г. в Республике Башкортостан. Первый этап заключался в сравнении реальных и прогнозных значений уровней воды на гидрологических постах в разные периоды прохождения паводка. Результаты приведены в таблице (см. http://www.swsys.ru/uploaded/image/2022-3/2022-3-dop/14.jpg), где реальный уровень воды – это фактически измеренное значение на стационарном гидрологическом посту и автоматической станции, а прогнозируемый – значение уровня воды, полученное с использованием нейронной сети. Важно отметить, что по сравнению с другими методами (рис. 2) нейросетевое прогнозирование показало более качественный результат (выходные значения оказались точнее, нежели при расчетах прогнозных значений уровней воды другими методами) [28].

Ежедневно с момента начала наблюдения за развитием паводковой ситуации (в 2021 г. – это 07 апреля, прогноз производился на 1–5 дней, соответственно, датой начала уровня подъема воды во время весеннего паводка можно считать 12.04.2021, то есть i = 1) с помощью нейронной сети осуществлялся прогноз на следующие m дней (например 1–5) по одному гидропосту (то есть n = 1; такое значение n взято для сокращения времени эксперимента). На следующий день измерялось фактическое значение уровней воды $h^k_{j,i+m}$ на этих же постах. В процессе ежедневного прогнозирования до 01 мая (соответственно введем дополнительное обозначение l – количество дней паводка, то есть значение $l_{k,i+m}$ = 25) по каждому посту получено 25 прогнозов и фактических значений уровней воды по каждому из n постов, что дает возможность определить среднюю точность прогноза по каждому k-му посту:

$$E^k_{\text{ср}} = \frac{1}{l_{k,i+m}} \sum_{i=1}^{l_{k,i+m}} E \,|\, h^k_{ji} - h^k_{j,i+m} \,|\,. \qquad (16)$$

При этом ошибка E (в сантиметрах) на i-й день j-го года каждого k-го гидропоста рассчитывается по соотношению абсолютной разности:

$$E^k_{ji} = |\, h^k_{ji} - h^k_{j.i+m} \,|\,. \qquad (17)$$

Важное отличие предлагаемого метода прогнозирования уровней воды с использованием искусственной нейронной сети по сравнению с другими известными методами заключается в быстроте получения прогноза и его корректности (более точен) при заблаговременном прогнозировании (в данном случае – на 5 дней), результаты приведены на рисунке 5.

По данным таблицы (см. http://www.swsys.ru/uploaded/image/2022-3/2022-3-dop/14.jpg) и рисунка 5 видно, что искусственная нейронная сеть дает достаточно точные результаты: средняя ошибка не превышает 55,52 см за весь паводковый период. Таким образом, реализованная искусственная нейронная сеть показала более точный и стабильный результат, что доказывает целесообразность применения предлагаемого решения в рамках прогнозирования и дальнейшего предотвращения особых/исключительных ситуаций на примере заблаговременного прогнозирования уровней воды в период паводковой ситуации.

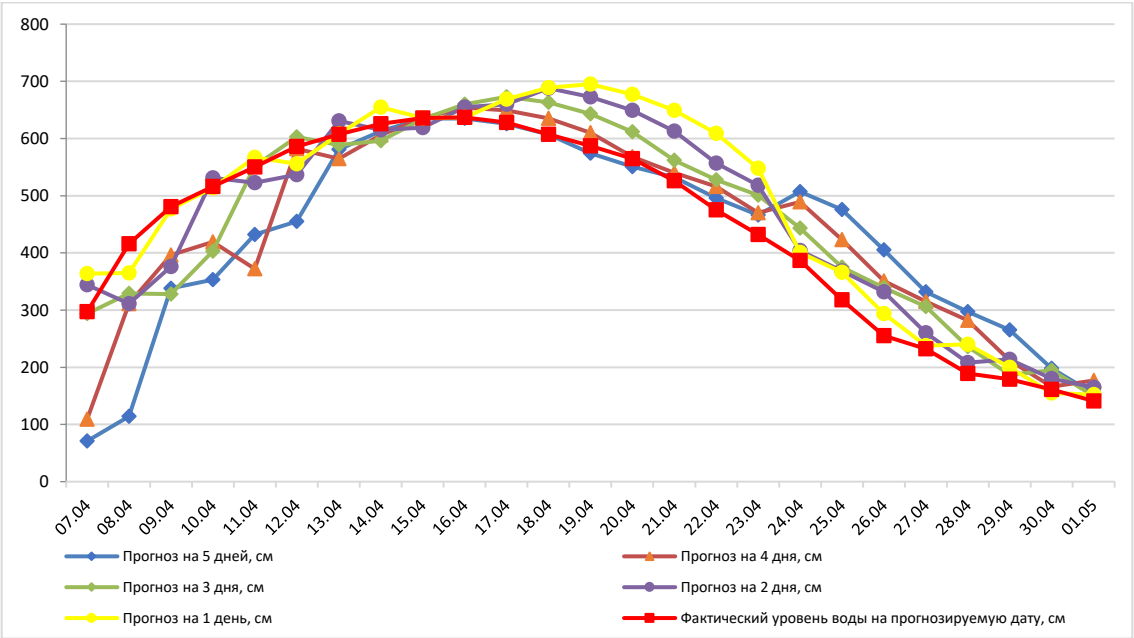

Рис. 5. Результат прогнозирования уровней воды на 5 суток на стационарном гидрологическом посту 76289 г. Уфы

Fig. 5. The result of forecasting water levels for 5 days at a stationary hydrological post 76289 in Ufa

Заключение

Предложенный анализ данных проведен на высоком уровне абстракции, полученные результаты апробированы на реальных физических объектах, что открывает возможность применения метода в широком спектре областей и задач, в том числе и при прогнозировании особых (включая критические) ситуаций с дальнейшим недопущением или минимизацией риска перевода системы в исключительное состояние.

Проведенный анализ эффективности прогнозирования на базе предлагаемой системы показал, что использование искусственной нейронной сети дает более стабильные результаты: средняя ошибка не превышает 55,52 см за весь паводковый период, средняя погрешность между реальными и рассчитанными значениями не более 2,10 %, в то время как при использовании существующих распространенных методов при прогнозировании значений уровней воды средняя ошибка между реальными и рассчитанными значениями достигает 8,36 %. Это дает специальным службам необходимое время (до 5 суток включительно) на осуществление противопаводковых мероприятий по подготовке к защите технических объектов.

Таким образом, рассчитанные заранее (за несколько дней) прогнозы уровней воды могут быть особенно полезны для предупреждения о паводковых ситуациях и принятия своевременных мер.

Исследование проводится при финансовой поддержке Минобрнауки Российской Федерации в рамках основной части государственного задания вузам № FEUE-2020-0007.

Литература

1. Supriya P., Krishnaveni M., Subbulaksmi M. Regression analysis of annual maximum daily rainfall and stream flow for flood forecasting in Vellar river basin. Aquatic Procedia, 2015, vol. 4, pp. 957–963. DOI: 10.1016/J.AQPRO.2015.02.120.
2. He Y., Yan Y., Wang X., Wang C. Uncertainty forecasting for streamflow based on support vector regression method with fuzzy information granulation. Energy Procedia, 2019, vol. 158, pp. 6189–6194. DOI: 10.1016/j.egypro.2019.01.489.
3. Luo X., Yuan X., Zhu S., Xu Z., Meng L., Peng J. A hybrid support vector regression framework for streamflow forecast. J. of Hydrology, 2019, vol. 568, pp. 184–193. DOI: 10.1016/J.JHYDROL.2018.10.064.

4. Silvestro F., Rossi L., Campo L., Parodi A., Fiori E., Rudari R., Ferraris L. Impact-based flash-flood forecasting system: Sensitivity to high resolution numerical weather prediction systems and soil moisture. J. of Hydrology, 2019, vol. 572, pp. 388–402. DOI: 10.1016/J.JHYDROL.2019.02.055.

5. Hu R., Fang F., Salinas P., Pain C.C., Domingo N.D., Mark O. Numerical simulation of floods from multiple sources using an adaptive anisotropic unstructured mesh method. Advances in Water Resources, 2019, vol. 123, pp. 173–188. DOI: 10.1016/J.ADVWATRES.2018.11.011.

6. Phan T.T., Nguyen X.H. Combining statistical machine learning models with ARIMA for water level forecasting: The case of the Red river. Advanced in Water Resources, 2020, vol. 142, art. 103656. DOI: 10.1016/j.advwatres.2020.103656.

7. Liu T., Ramirez-Marquez J.E., Jagupilla S.C., Prigiobbe V. Combining a statistical model with machine learning to predict groundwater flooding (or infiltration) into sewer networks. J. of Hydrology, 2021, vol. 603, art. 126716. DOI: 10.1016/j.jhydrol.2021.126916.

8. Zhou T., Jiang Z., Liu X., Tan K. Research on the long-term and short-term forecasts of navigable river's water-level fluctuation based on the adaptive multilayer perceptron. J. of Hydrology, 2020, vol. 591, art. 125285.

9. Zakaria M.N., Malek M.A., Zolkepli M., Ahmed A.N. Application of artificial intelligence algorithms for hourly river level forecast: A case study of Muda River, Malaysia. Alexandria Engineering J., 2021, vol. 60, no. 4, pp. 4015–4028. DOI: 10.1016/J.AEJ.2021.02.046.

10. Ahmed A.N., Lam T.V., Hung N.D., Thieu N.V., Kisi O., Ahmed A.E. A comprehensive comparison of recent developed meta-heuristic algorithms for streamflow time series forecasting problem. Applied Soft Computing, 2021, vol. 105, art. 107282. DOI: 10.1016/J.ASOC.2021.107282.

11. Nguyen D.H., Bae D.H. Correcting mean areal precipitation forecasts to improve urban flooding predictions by using long short-term memory network. J. of Hydrology, 2020, vol. 584, art. 1247100. DOI: 10.1016/j.jhydrol.2020.124710.

12. Ding Y., Zhu Y., Feng J., Zhang P., Cheng Z. Interpretable spatio-temporal attention LSTM model for flood forecasting. Neurocomputing, 2020, vol. 403, pp. 348–359. DOI: 10.1016/j.neucom.2020.04.110.

13. Kao I.F., Zhou Y., Chang L.C., Chang F.J. Exploring a long short-term memory based encoder-decoder framework for multi-step-ahead flood forecasting. J. of Hydrology, 2020, vol. 583, art. 124631. DOI: 10.1016/j.jhydrol.2020.124631.

14. Barzegar R., Aalami M.T., Adamowski J. Coupling a hybrid CNN-LSTM deep learning model with a boundary corrected maximal overlap discrete wavelet transform for multiscale lake water level forecasting. J. of Hydrology, 2021, vol. 598, art. 126196. DOI: 10.1016/J.JHYDROL.2021.126196.

15. Kabir S., Patidar S., Xia X., Liang Q., Neal J., Pender G. A deep convolutional neural network model for rapid prediction of fluvial flood inundation. J. of Hydrology, 2020, vol. 590, art. 125481. DOI: 10.1016/j.jhydrol.2020.125481.

16. Khan M.M., Muhammad N.S., El-Shafie A. Wavelet based hybrid ANN-ARIMA models for meteorological drought forecasting. J. of Hydrology, 2020, vol. 590, art. 125380. DOI: 10.1016/j.jhydrol.2020.125380.

17. Lin A., Wu H., Liang G., Cardenas-Tristan A., Wu X., Zhao C., Li D. A big data-driven dynamic estimation model of relief supplies demand in urban flood disaster. Int. J. of Disaster Risk Reduction, 2020, vol. 49, art. 101682. DOI: 10.1016/j.ijdrr.2020.101682.

18. Wagena M., Goering D., Collick A.S., Bock E., Fuka D.R., Buda A., Easton Z.M. Comparison of short-term streamflow forecasting using stochastic time series, neural networks, process-based, and Bayesian models. Environmental Modelling and Software, 2020, vol. 126, pp. 104669. DOI: 10.1016/j.envsoft.2020.104669.

19. Adikari K.E., Shrestha S., Ratnayake D.T., Budhathoki A., Mohanasundaram S., Dailey M.N. Evaluation of artificial intelligence models for flood and drought forecasting in arid and tropical regions. Environmental Modelling and Software, 2021, vol. 144, art. 105136. DOI: 10.1016/J.ENVSOFT.2021.105136.

20. Zhou Y., Guo S., Chang F. Explore an evolutionary recurrent ANFIS for modelling multi-step-ahead flood forecasts. J. of Hydrology, 2019, vol. 570, pp. 343–355. DOI: 10.1016/J.JHYDROL.2018.12.040.

21. Arora A., Arabameri A., Pandey M., Siddiqui M.A., Shukla U.K. et al. Optimization of state-of-the-art fuzzy-metaheuristic ANFIS-based machine learning models for flood susceptibility prediction mapping in the Middle Ganga Plain, India. Science of The Total Environment, 2021, vol. 750, art. 141565. DOI: 10.1016/j.scitotenv.2020.141565.

22. Lee Y.H., Kim H.I., Han K.Y., Wong W.H. Flood evacuation routes based on spatiotemporal inundation risk assessment. Water, 2020, vol. 12, no. 8, art. 2271. DOI: 10.3390/w12082271.

23. Wunsch A., Liesch T., Broda S. Forecasting groundwater levels using nonlinear autoregressive networks with exogenous input (NARX). J. of Hydrology, 2018, vol. 567, pp. 743–758. DOI: 10.1016/J.JHYDROL.2018.01.045.

24. Banihabib M.E., Ahmadian A., Jamali F.S. Hybrid DARIMA-NARX model for forecasting long-term daily inflow to Dez reservoir using the North Atlantic Oscillation (NAO) and rainfall data. GeoResJ, 2017, vol. 13, pp. 9–16. DOI: 10.1016/J.GRJ.2016.12.002.

25. Hadid B., Duviella E., Lecoeuche S. Data-driven modeling for river flood forecasting based on a piecewise linear ARX system identification. J. of Process Control, 2020, vol. 86, pp. 44–56. DOI: 10.1016/j.jprocont.2019.12.007.

26. Bukhari A.H., Sulaiman M., Islam S., Shoaib M., Kumam P., Raja M.A. Neuro-fuzzy modeling and prediction of summer precipitation with application to different meteorological stations. Alexandria Engineering J., 2020, vol. 59, no. 1, pp. 101–116. DOI: 10.1016/j.aej.2019.12.011.

27. Palchevsky E.V., Khristodulo O.I., Pavlov S.V., Kalimgulov A. Intelligent data analysis for forecasting threats in complex distributed systems. CEUR Workshop Proceedings, 2020, vol. 2744, pp. 285–296. DOI: 10.51130/graphicon-2020-2-3-79.

28. Palchevsky E.V., Khristodulo O.I., Pavlov S.V. Threat prediction in complex distributed systems using artificial neural network technology. Proc. Int. Conf. GraphiCon, 2020, p. 2, vol. 2763, art. 79.

29. Пальчевский Е.В., Христодуло О.И. Разработка импульсной нейронной сети с возможностью скоростного обучения для нейтрализации DDoS-атак // Программные продукты и системы. 2019. Т. 32. № 4. С. 613–627. DOI: 10.15827/0236-235X.128.613-627.

Software & Systems
DOI: 10.15827/0236-235X.139.384-395

Received 18.02.22, Revised 11.05.22
2022, vol. 35, no. 3, pp. 384–395

Forecasting based on the second generation artificial neural network for decision support in especially significant situations

E.V. Palchevsky [1], *Lecturer, teelxp@inbox.ru*
V.V. Antonov [2], *Dr.Sc. (Engineering), Professor, antonov.v@bashkortostan.ru*
R.R. Enikeev [2], *Ph.D. (Engineering), Associate Professor, enikeevRR@mail.ru*

[1] *Financial University under the Government of the Russian Federation, Faculty of IT and ABD, Moscow, 109456, Russian Federation*
[2] *Ufa State Aviation Technical University, Faculty of IRT, Ufa, 450008, Russian Federation*

Abstract. Nowadays, specialized system models implemented on the basis of decision support in exceptional (emergency) situations (states) using machine learning, artificial intelligence (including using neural networks) to reproduce, predict and prevent (or minimize risk) consequences) in exceptional situations are useful and becoming increasingly popular. Floods also fall under such exceptional situations and states. Therefore, there arises the problem of early forecasting of an exceptional situation using the example of rising water levels at stationary hydrological posts in order to prevent (or minimize the risk) the transition of the system under consideration into an exceptional state (emergency situation).

To solve this problem, the authors propose a decision support system for early forecasting water rise levels. It is based on a neural network (intelligent) analysis of retrospective data (code of a stationary hydrological station / automatic station, date, water level at a stationary hydrological station / automatic station, atmospheric pressure, wind speed, snow cover thickness, amount of precipitation, time and air temperature) in order to calculate the values of water levels for 5 days in advance. The artificial neural network itself is based on the freely distributed TensorFlow machine learning software library; a modified backpropagation method is used as training. Its main difference is the addition of an artificial neural network (ANN) learning rate increase factor.

An analysis of the effectiveness of the proposed solution in the framework of forecasting the flood situation has shown high accuracy in calculating the forecast values of water levels: the difference between the real and predicted values is 2.10 %. This will allow specialized services to carry out specialized anti-flood measures in advance (5 days in advance).

Thus, information support during special situations is an absolute (not relative) indicator of data quality that allows developing and making decisions in the framework of predicting possible critical situations and preventing the transfer of the state of the territory management system to critical situations.

Keywords: forecasting; intellectual analysis; neural network analysis; retrospective data, data mining; forecasting water levels.

Acknowledgements. *The study has been financially supported by the Ministry of Education and Science of the Russian Federation within the main part of the state assignment to universities no. FEUE-2020-0007.*

References

1. Supriya P., Krishnaveni M., Subbulaksmi M. Regression analysis of annual maximum daily rainfall and stream flow for flood forecasting in Vellar river basin. *Aquatic Procedia*, 2015, vol. 4, pp. 957–963. DOI: 10.1016/J.AQPRO.2015.02.120.
2. He Y., Yan Y., Wang X., Wang C. Uncertainty forecasting for streamflow based on support vector regression method with fuzzy information granulation. *Energy Procedia*, 2019, vol. 158, pp. 6189–6194. DOI: 10.1016/j.egypro.2019.01.489.
3. Luo X., Yuan X., Zhu S., Xu Z., Meng L., Peng J. A hybrid support vector regression framework for streamflow forecast. *J. of Hydrology*, 2019, vol. 568, pp. 184–193. DOI: 10.1016/J.JHYDROL.2018.10.064.
4. Silvestro F., Rossi L., Campo L., Parodi A., Fiori E., Rudari R., Ferraris L. Impact-based flash-flood forecasting system: Sensitivity to high resolution numerical weather prediction systems and soil moisture. *J. of Hydrology*, 2019, vol. 572, pp. 388–402. DOI: 10.1016/J.JHYDROL.2019.02.055.
5. Hu R., Fang F., Salinas P., Pain C.C., Domingo N.D., Mark O. Numerical simulation of floods from multiple sources using an adaptive anisotropic unstructured mesh method. *Advances in Water Resources*, 2019, vol. 123, pp. 173–188. DOI: 10.1016/J.ADVWATRES.2018.11.011.
6. Phan T.T., Nguyen X.H. Combining statistical machine learning models with ARIMA for water level forecasting: The case of the Red river. *Advanced in Water Resources*, 2020, vol. 142, art. 103656. DOI: 10.1016/j.advwatres.2020.103656.
7. Liu T., Ramirez-Marquez J.E., Jagupilla S.C., Prigiobbe V. Combining a statistical model with machine learning to predict groundwater flooding (or infiltration) into sewer networks. *J. of Hydrology*, 2021, vol. 603, art. 126716. DOI: 10.1016/j.jhydrol.2021.126916.
8. Zhou T., Jiang Z., Liu X., Tan K. Research on the long-term and short-term forecasts of navigable river's water-level fluctuation based on the adaptive multilayer perceptron. *J. of Hydrology*, 2020, vol. 591, art. 125285.
9. Zakaria M.N., Malek M.A., Zolkepli M., Ahmed A.N. Application of artificial intelligence algorithms for hourly river level forecast: A case study of Muda River, Malaysia. *Alexandria Engineering J.*, 2021, vol. 60, no. 4, pp. 4015–4028. DOI: 10.1016/J.AEJ.2021.02.046.
10. Ahmed A.N., Lam T.V., Hung N.D., Thieu N.V., Kisi O., Ahmed A.E. A comprehensive comparison of recent developed meta-heuristic algorithms for streamflow time series forecasting problem. *Applied Soft Computing*, 2021, vol. 105, art. 107282. DOI: 10.1016/J.ASOC.2021.107282.
11. Nguyen D.H., Bae D.H. Correcting mean areal precipitation forecasts to improve urban flooding predictions by using long short-term memory network. *J. of Hydrology*, 2020, vol. 584, art. 1247100. DOI: 10.1016/j.jhydrol.2020.124710.
12. Ding Y., Zhu Y., Feng J., Zhang P., Cheng Z. Interpretable spatio-temporal attention LSTM model for flood forecasting. *Neurocomputing*, 2020, vol. 403, pp. 348–359. DOI: 10.1016/j.neucom.2020.04.110.
13. Kao I.F., Zhou Y., Chang L.C., Chang F.J. Exploring a long short-term memory based encoder-decoder framework for multi-step-ahead flood forecasting. *J. of Hydrology*, 2020, vol. 583, art. 124631. DOI: 10.1016/j.jhydrol.2020.124631.
14. Barzegar R., Aalami M.T., Adamowski J. Coupling a hybrid CNN-LSTM deep learning model with a boundary corrected maximal overlap discrete wavelet transform for multiscale lake water level forecasting. *J. of Hydrology*, 2021, vol. 598, art. 126196. DOI: 10.1016/J.JHYDROL.2021.126196.
15. Kabir S., Patidar S., Xia X., Liang Q., Neal J., Pender G. A deep convolutional neural network model for rapid prediction of fluvial flood inundation. *J. of Hydrology*, 2020, vol. 590, art. 125481. DOI: 10.1016/j.jhydrol.2020.125481.
16. Khan M.M., Muhammad N.S., El-Shafie A. Wavelet based hybrid ANN-ARIMA models for meteorological drought forecasting. *J. of Hydrology*, 2020, vol. 590, art. 125380. DOI: 10.1016/j.jhydrol.2020.125380.
17. Lin A., Wu H., Liang G., Cardenas-Tristan A., Wu X., Zhao C., Li D. A big data-driven dynamic estimation model of relief supplies demand in urban flood disaster. *Int. J. of Disaster Risk Reduction*, 2020, vol. 49, art. 101682. DOI: 10.1016/j.ijdrr.2020.101682.
18. Wagena M., Goering D., Collick A.S., Bock E., Fuka D.R., Buda A., Easton Z.M. Comparison of short-term streamflow forecasting using stochastic time series, neural networks, process-based, and Bayesian models. *Environmental Modelling and Software*, 2020, vol. 126, pp. 104669. DOI: 10.1016/j.envsoft.2020.104669.

19. Adikari K.E., Shrestha S., Ratnayake D.T., Budhathoki A., Mohanasundaram S., Dailey M.N. Evaluation of artificial intelligence models for flood and drought forecasting in arid and tropical regions. *Environmental Modelling and Software*, 2021, vol. 144, art. 105136. DOI: 10.1016/J.ENVSOFT.2021.105136.

20. Zhou Y., Guo S., Chang F. Explore an evolutionary recurrent ANFIS for modelling multi-step-ahead flood forecasts. *J. of Hydrology*, 2019, vol. 570, pp. 343–355. DOI: 10.1016/J.JHYDROL.2018.12.040.

21. Arora A., Arabameri A., Pandey M., Siddiqui M.A., Shukla U.K. et al. Optimization of state-of-the-art fuzzy-metaheuristic ANFIS-based machine learning models for flood susceptibility prediction mapping in the Middle Ganga Plain, India. *Science of the Total Environment*, 2021, vol. 750, art. 141565. DOI: 10.1016/j.scitotenv.2020.141565.

22. Lee Y.H., Kim H.I., Han K.Y., Wong W.H. Flood evacuation routes based on spatiotemporal inundation risk assessment. *Water*, 2020, vol. 12, no. 8, art. 2271. DOI: 10.3390/w12082271.

23. Wunsch A., Liesch T., Broda S. Forecasting groundwater levels using nonlinear autoregressive networks with exogenous input (NARX). *J. of Hydrology*, 2018, vol. 567, pp. 743–758. DOI: 10.1016/J.JHYDROL.2018.01.045.

24. Banihabib M.E., Ahmadian A., Jamali F.S. Hybrid DARIMA-NARX model for forecasting long-term daily inflow to Dez reservoir using the North Atlantic Oscillation (NAO) and rainfall data. *GeoResJ*, 2017, vol. 13, pp. 9–16. DOI: 10.1016/J.GRJ.2016.12.002.

25. Hadid B., Duviella E., Lecoeuche S. Data-driven modeling for river flood forecasting based on a piecewise linear ARX system identification. *J. of Process Control*, 2020, vol. 86, pp. 44–56. DOI: 10.1016/j.jprocont.2019.12.007.

26. Bukhari A.H., Sulaiman M., Islam S., Shoaib M., Kumam P., Raja M.A. Neuro-fuzzy modeling and prediction of summer precipitation with application to different meteorological stations. *Alexandria Engineering J.*, 2020, vol. 59, no. 1, pp. 101–116. DOI: 10.1016/j.aej.2019.12.011.

27. Palchevsky E.V., Khristodulo O.I., Pavlov S.V., Kalimgulov A. Intelligent data analysis for forecasting threats in complex distributed systems. *CEUR Workshop Proceedings*, 2020, vol. 2744, pp. 285–296. DOI: 10.51130/graphicon-2020-2-3-79.

28. Palchevsky E.V., Khristodulo O.I., Pavlov S.V. Threat prediction in complex distributed systems using artificial neural network technology. *Proc. Int. Conf. GraphiCon*, 2020, p. 2, vol. 2763, art. 79.

29. Palchevsky E.V., Khristodulo O.I. Development of a spiking neural network with the possibility of high-speed training to neutralize DDoS attacks. *Software & Systems*, 2019, vol. 32, no. 4, pp. 613–627. DOI: 10.15827/0236-235X.128.613-627 (in Russ.).

Для цитирования

Пальчевский Е.В., Антонов В.В., Еникеев Р.Р. Прогнозирование на основе искусственной нейронной сети второго поколения для поддержки принятия решений в особо значимых ситуациях // Программные продукты и системы. 2022. Т. 35. № 3. С. 384–395. DOI: 10.15827/0236-235X.139.384-395.

For citation

Palchevsky E.V., Antonov V.V., Enikeev R.R. Forecasting based on the second generation artificial neural network for decision support in especially significant situations. *Software & Systems*, 2022, vol. 35, no. 3, pp. 384–395 (in Russ.). DOI: 10.15827/0236-235X.139.384-395.