

Decision Support System based on Application of the Second Generation Neural Network¹

E. V. Palchevsky, Lecturer, teelpx@inbox.ru, Financial University under the Government of the Russian Federation, Moscow, 125167, Russian Federation,

V. V. Antonov, Professor, antonov.v@bashkortostan.ru, Ufa State Aviation Technical University, Ufa, 450008, Russian Federation

Corresponding author:

Palchevsky Evgeniy V., Lecturer, Financial University under the Government of the Russian Federation, Moscow, 125167, Russian Federation

E-mail: teelpx@inbox.ru

Received on October 28, 2021

Accepted on April 25, 2022

The issue of the feasibility of using existing statistical and hydrological methods for short-term and early forecasting in the framework of forecasting the levels of water rise in water bodies is considered: a comparative review is given, which describes their advantages and disadvantages. In the course of analyzing the shortcomings of these methods, the problem of operational and early (advance) forecasting of water rise levels was identified. To solve this problem, a decision support system is proposed for predicting the water rise levels in advance, based on a neural network (intelligent) analysis of retrospective data (date, water level, air temperature, atmospheric pressure and wind speed) to calculate the water level values for 5 days in advance. The artificial neural network itself is based on the freely distributed library of machine learning programs "TensorFlow", and a modified backpropagation method is used as training, the main difference of which is an increase in the learning rate of an artificial neural network. The results of the analysis of the effectiveness showed that the proposed decision support system is more accurate (the error between the real and calculated values does not exceed 2.10 %), compared to existing common methods/systems (8.36 %). This will allow to give the necessary time to special services for the implementation of flood control measures to prepare for the protection of technical facilities of enterprises.

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

Acknowledgements: *The study is conducted with financial support from the Ministry of Education and Science of the Russian Federation as part of the basic part of the state assignment to higher education educational institutions # FEUE-2020-0007.*

For citation:

Palchevsky E. V., Antonov V. V. Decision Support System based on Application of the Second Generation Neural Network, *Programmnaya Ingeneria*, 2022, vol. 13, no. 6, pp. 301—308.

УДК 004.032.26

Е. В. Пальчевский, препод., teelpx@inbox.ru, Финансовый университет при Правительстве Российской Федерации, Москва,

В. В. Антонов, д-р техн. наук, проф., antonov.v@bashkortostan.ru, Уфимский государственный авиационный технический университет

Система поддержки принятия решений на основе применения нейронной сети второго поколения

Рассмотрен вопрос о целесообразности использования существующих статистических и гидрологических методов краткосрочного и раннего прогнозирования в рамках прогнозирования уровней подъема воды в водных объектах — дан сравнительный обзор, в котором описаны

¹ The article was based on the materials of the report at the IX All-Russian Scientific Conference "Information Technologies for Intelligent Decision Support" (ITIDS'2021).

их преимущества и недостатки. В ходе анализа недостатков этих методов была выявлена проблема оперативного и раннего (заблаговременного) прогнозирования уровней подъема воды. Для решения этой задачи предлагается система поддержки принятия решений для заблаговременного прогнозирования уровней подъема воды, основанная на нейросетевом (интеллектуальном) анализе ретроспективных данных. Приведены результаты анализа эффективности предлагаемой системы поддержки принятия решений.

Ключевые слова: прогнозирование, интеллектуальный анализ, нейросетевой анализ, ретроспективные данные, интеллектуальный анализ данных, прогнозирование уровня воды

Исследования выполнены при поддержке Министерства науки и высшего образования РФ в рамках выполнения Государственного задания № FEUE-2020-0007.

Introduction

At present, various natural phenomena (in this article, we will take the level of water rise in water bodies as an example) are one of the most frequent and dangerous causes of external impact on technical objects of complex distributed systems (CDS). In this article, complex distributed systems mean an interconnected set of objects of varying complexity and physical nature (for example, water bodies of a region), and technical objects of complex distributed systems mean various potentially hazardous objects, oil and gas pipelines, while, according to [1], technical objects belong to the class of unique complex systems. For these technical objects, as a rule, it is necessary to carry out special (in our case, flood control) measures to minimize social and financial risks: for example, it is necessary to predict in advance the water levels at stationary hydrological and automatic measurement and observation posts.

Thus, based on these problems, there is a need for a sufficiently fast and early forecasting of water level values based on a decision support system based on neural network data analysis, which will allow special services to give the necessary time to implement flood control measures to prepare for the protection of technical facilities of enterprises.

1. Related Works

Statistical methods, for example, generalized regression models [2–4], the method of least squares [5], and numerical methods [6, 7], are widely used in predicting water level values based on retrospective data. But at the same time, the problem of forecasting in advance arises, since the main purpose of these methods is short-term forecasting of water level values, which is often insufficient for early implementation of flood control measures. If we consider less common methods (hydrological [8–10]) of short-term and early forecasting of water levels, then it is necessary

to note the problem of resource intensity: the duration of calculations of water level values is not permissible, especially in critical (peak) moments of flood situations. Many scientists (both domestic and foreign) are engaged in these problems, but due to the insufficient number of works describing methods for early (early) forecasting of water levels based on intelligent (neural network) analysis of retrospective data using an artificial neural network (ANN) for Better forecasting of water levels in the shortest possible time and using outdated architectures of artificial neural networks (for example, multilayer perceptron [11]), the implementation of the proposed method within the framework of early warning of a flood situation becomes relevant.

2. Formalized statement of the problem on the example of the levels of water rise in water bodies

One of the main parameters of the impact on various technical objects is the flood, which is expressed in the rise in the values of h — water levels on water bodies. As an example, we can take the stationary hydrological posts of the Republic of Bashkortostan. In this regard, we introduce the following notation: h_{ji}^k — is the value of the water level measured at the k -th gauging station on the i -th date of the j -th year. Here $k = 1, \dots, n$, where n is the number of gauging stations involved in the calculations, j is the number of the year, i is the specific date of measurement. The task of early forecasting is (it should be noted that the task of short-term forecasting for 1 day was solved in [15], and on the basis of the methodology proposed in [15], forecasting for 2–5 days is performed) in the fact that in a specific current i -th day of measurement, calculate the value of the water level for $i + 2$ days, that is h_{ji+2}^k , as well as for $i + 3$ (h_{ji+3}^k), $i + 4$ (h_{ji+4}^k) and $i + 5$ (h_{ji+5}^k) for any $k = 1, \dots, n$.

3. Application of an artificial neural network for early prediction of water levels at stationary hydropost

To solve this problem, it is proposed to use the results of previous measurements of water levels (an array of large retrospective data) at all stationary hydrological stations located on the territory under consideration for all previous years. The proposed method of early forecasting is based on the use of an artificial neural network (fig. 1), implemented using the free library of machine learning programs "TensorFlow" [12]. The method of training ANN backpropagation (backpropagation of the error), changed by the authors of the article, is integrated into this neural network via the API. As shown by the results of the experiment presented in Section 4, the accuracy and time of calculating the forecast prove the feasibility of using this method in the framework of the advance forecasting of water level values at stationary hydrological posts. Thus, the forecasting process is proposed to be carried out in three stages (fig. 2): data preprocessing, forecasting model formation and data post-processing.

At the first stage, which is performed, as a rule, before the development of a flood situation, the ANN parameters are selected: the total number of neurons (107); the number of neurons in the input (h_{jiK}^k (6)) and output (h_{jiL}^k (5)) layers; the number of intermediate layers (3) and neurons in them (H_{jk}^k (96)); the activation

function on hidden layers is sigmoidal, while a linear activation function is used on the output layer due to its weighted sum on each neuron of the output layer [13–15]; the amount of data for the samples: training (train, 60 %), validation (validation, 20 %) and checking (test, 20 %); number of epochs (200); learning rate α (0.90), selected experimentally with further training of the artificial neural network. But it is worth noting that when applying the backpropagation method, it is necessary to normalize the input values of the water levels H to prevent vanishing and exploding gradients:

$$\bar{h}_{ji}^k = \frac{\bar{h}_{ji}^k - \bar{h}_{\min}}{\bar{h}_{\max} - \bar{h}_{\min}} \quad (1)$$

where \bar{h}_{\max} and \bar{h}_{\min} are the minimum and maximum values of water levels for all data supplied to the input layer of the neural network for all \bar{h}_{ji}^k , where $k = 1, \dots, n$.

To train the neural network (fig. 3), the backpropagation method, modified by the authors of the article, is used. The essence of the change lies in the application of an experimentally set parameter (coefficient) of the neural network learning rate.

A. Calculation of the squared error

As a rule, the inputs of the network are considered as an input vector (\bar{h}_{ji}^k), where $\bar{h}_{ji}^k = [\bar{h}_{ji1}^k, \bar{h}_{ji2}^k, \dots, \bar{h}_{jiK}^k]$, and the outputs of the network can be represented as

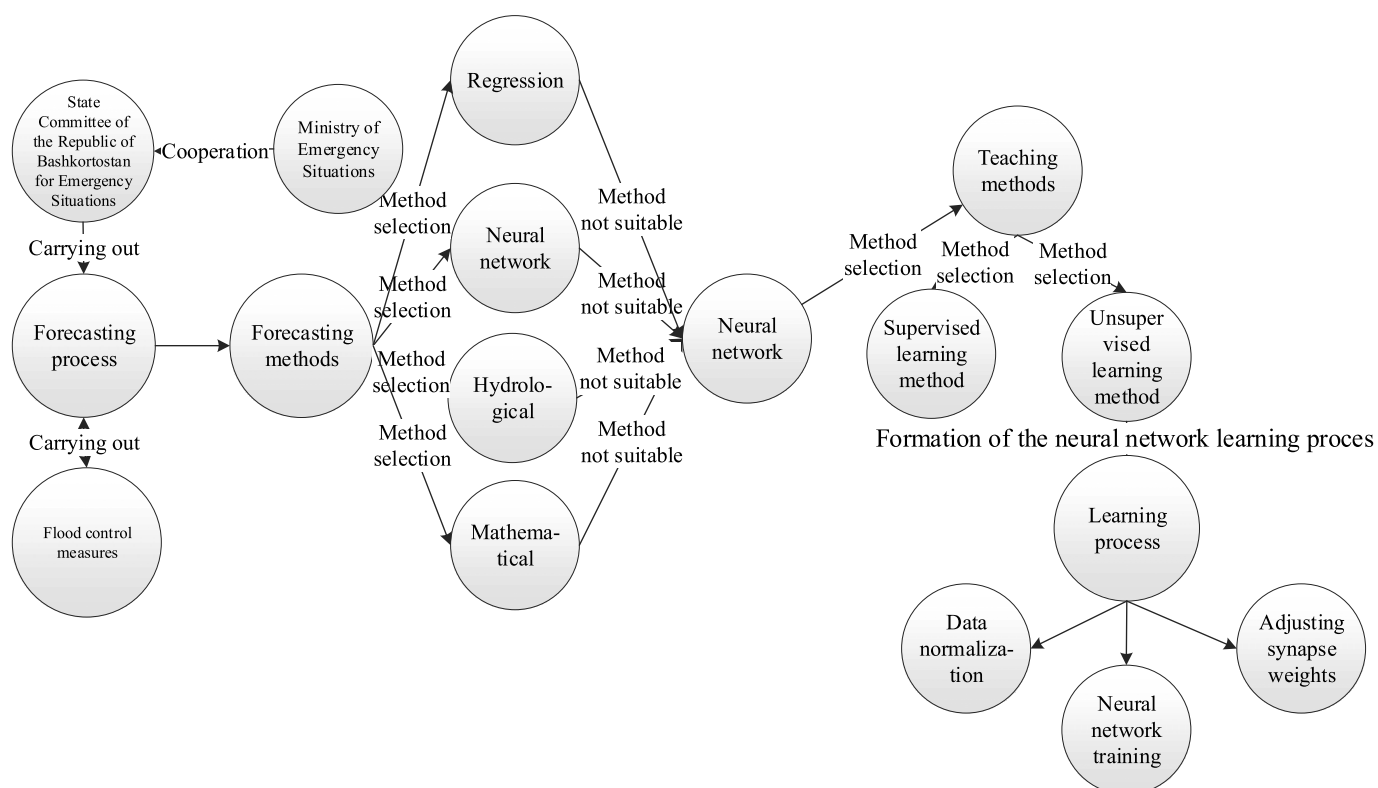


Fig. 1. Ontological model of the water level forecasting process

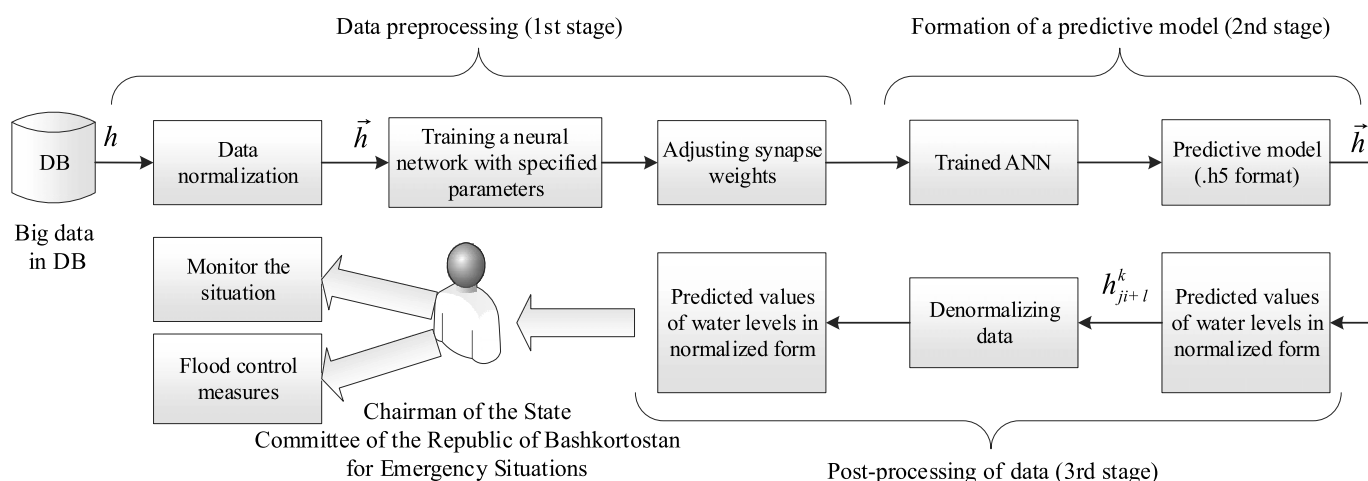


Fig. 2. Scheme of using a neural network to predict the values of water levels

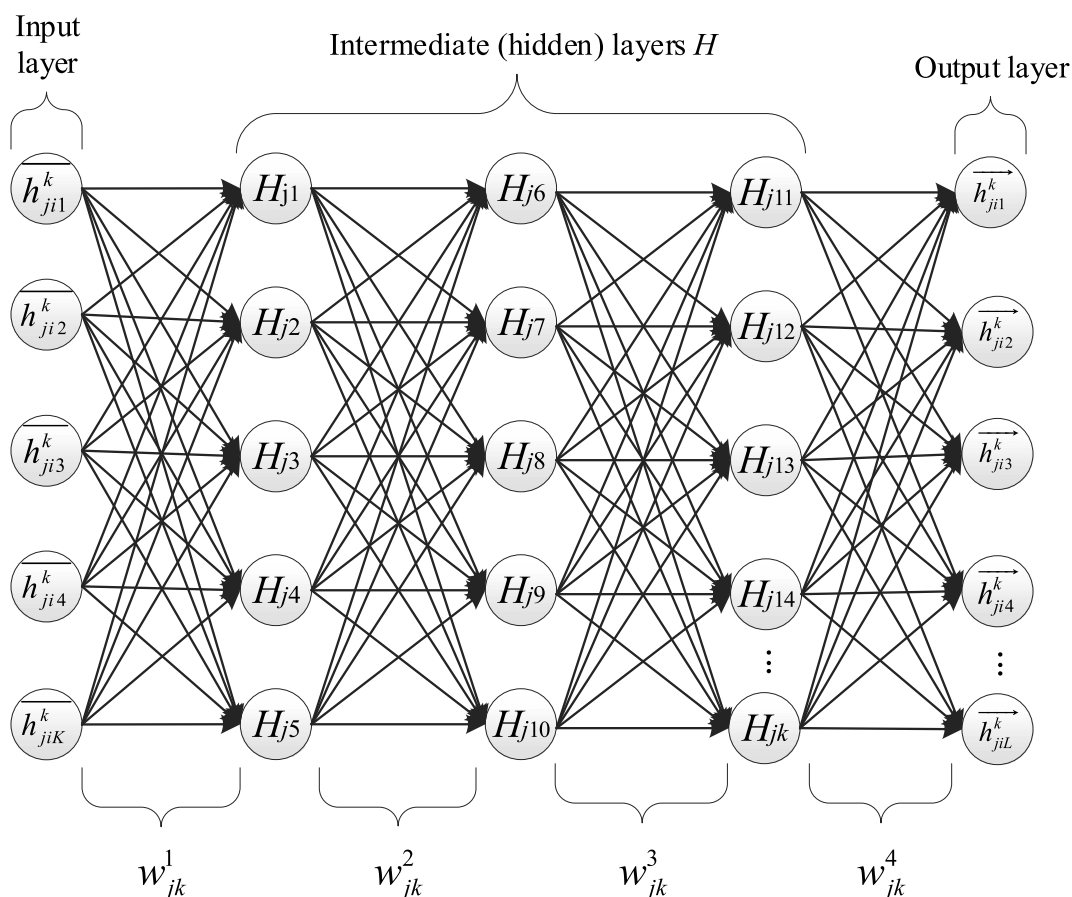


Fig. 3. Neural network structure

an output vector $\overline{(h_{ji}^k)}$, where $\overline{h_{ji}^k} = [\overline{h_{ji1}^k}, \overline{h_{ji2}^k}, \dots, \overline{h_{jiL}^k}]$, where, in this case $K = 1, \dots, M$ and $L = 1, \dots, N$, in which M and N represent the number of values input and output vectors. Accordingly, the training sample

is a set of pairs R of input vectors $\overline{h_{jiK}^k}$ and the desired output vectors: $\overline{h_{jiL}^k}$:

$$R = \{(\overline{h_{ji1}^k}, \overline{h_{ji1}^k}), (\overline{h_{ji2}^k}, \overline{h_{ji2}^k}), \dots, (\overline{h_{jiK}^k}, \overline{h_{jiL}^k})\}. \quad (2)$$

Each time $\overline{h_{jiK}^k}$ is applied from R to the neural network to the neural network, the actual output $\overline{h_{jiL}^k}$ of the output layer will be calculated:

$$\overline{h_{jiL}^k} = f(\sigma_L), \quad (3)$$

where σ is the activation function, is determined by the relation

$$\sigma(\overline{h_{jiK}^k}) = \frac{1}{1 + e^{-\overline{h_{jiK}^k}}}, \quad (4)$$

and is the weighted sum of the outputs of neuron j in each hidden layer H_j :

$$\sigma_L = \sum_{j'=1}^j w_{jk}^j H_{jk}^j, \quad (5)$$

where w is the weight of the synapse, k is the k -th neuron in the intermediate layer. Based on this, we can determine the squared error for each pair of vectors of the set R by summing the squared errors in each output neuron:

$$E_l = \frac{1}{2} \sum_{k'=1}^L (\overline{h_{jiK_{k'}}^k} - \overline{h_{jiL_{k'}}^k})^2; \quad (6)$$

and, as a consequence, the total squared error E by summing all pairs of inputs and outputs in the training set:

$$E = \frac{1}{2} \sum_{l'=1}^K \sum_{k'=1}^L (\overline{h_{jiK_{k'}}^k} - \overline{h_{jiL_{k'}}^k})^2. \quad (7)$$

Thus, the goal of training is to minimize E by finding an appropriate set of weights w_{xk}^l and w_{jk}^l , where l varies from 1 to 4.

B. Determination of partial derivatives by synapse weights

Based on relation (6), we can note that

$$\frac{dE}{dw_{jk}^l} = \overline{h_{jiK}^k} - \overline{h_{jiL}^k}, \quad (8)$$

in this case, relying on (5), we obtain

$$\frac{d\sigma_k}{dw_{jk}^l} = H_j. \quad (9)$$

Based on (9), we determine the partial derivative E with respect to the weight w_{jk} to perform gradient descent in (10) for the weights of the synapses between the intermediate and output layers:

$$\frac{dE}{dw_{jk}^l} = ((\overline{h_{jiK}^k} - \overline{h_{jiL}^k}) * \overline{h_{jiL}^k}) \cdot ((\sigma(\overline{h_{jiK}^k}) - \overline{h_{jiL}^k}) \cdot H_{jk}). \quad (10)$$

Moreover, if we consider the case between the input and hidden (intermediate) layers, then it is worth highlighting that in this case all outputs depend on w_{xk} . Thus, the partial derivative in this case will be found as follows:

$$\begin{aligned} \frac{dE}{dw_{xk}^l} = & \sum_{k'=1}^L ((\overline{h_{jiK_{k'}}^k} - \overline{h_{jiL_{k'}}^k}) * \overline{h_{jiL_{k'}}^k}) \times \\ & \times ((\sigma(\overline{h_{jiK}^k}) - \overline{h_{jiL_{k'}}^k}) \cdot (w_{jk}^l \cdot H_{jk}^l)) \cdot ((1 - H_{jk}^l) \cdot \overline{h_{jiK}^k}). \end{aligned} \quad (11)$$

Thus, equations (9) and (10) give all the necessary values of $\frac{dE}{dw}$ for applying (10) and (11) in the framework of gradient descent for all weights of the neural network.

C. Change in synapse weights

Each weight will be changed to dw to decrease E :

$$w(t+1) = w(t) + \Delta w(t), \quad (12)$$

where $\Delta w(t) = -\frac{dE}{dw}|_t$, $w(t)$ — is the weight of the synapse at time t , $w(t+1)$ is the modified (updated) weight of the synapse. And to increase the learning rate of the neural network, the previously indicated parameter α is added:

$$\Delta w(t) = -\frac{dE}{dw}|_t + \alpha \cdot \Delta w(t-1). \quad (13)$$

At the second stage, the forecast model is formed directly, and at the third stage, the water level values are forecast. But since all the values were previously normalized ($\overline{h_{ji}^k} \in [0; 1]$), then for their further use within the framework of flood control measures, it is necessary to denormalize $\overline{h_{ji+1}^k}$ (14) according to the inverse relation (1):

$$\overline{h_{ji}^k} = \overline{h_{ji}^k} \cdot (\overline{h_{\max}} - \overline{h_{\min}}) + \overline{h_{\min}}. \quad (14)$$

4. Analysis of the efficiency of the proposed method in the framework of forecasting the levels of water rise

To analyze the effectiveness of the proposed method for predicting water levels, we used long-term data on water level measurements at hydrological posts provided by the Bashkir Department for Hydrometeorology and Environmental Monitoring (Bashhydromet) from 01.01.2000 to 02.05.2021 in the form: stationary hydrological code station/automatic station, date, water level at the stationary hydrological station/automatic station.

During the experiment, many iterations were carried out for the entire period of long-term observations (from 01.01.2000 to 06.05.2021) to calculate the predicted values of water levels. The total array of data used in the experiment is 1,025,815 values, of which 60 % are data of long-term observations in the period from 01.01.2000 to 31.12.2013, fed to the training sample of the neural network, to the validation sample — 20 % (01.01.2014 — 31.12.2017) and for the verification — 20 % (01.01.2018 — 06.05.2021), while the forecasting depth, on the basis of which the predicted value is calculated, is set experimentally depending on the stationary hydrological post.

The article, as an example, presents the results of an experiment, predicting water levels in the stationary

hydrological city of Ufa (76289, Belaya River) during the spring flood of 2021 in the Republic of Bashkortostan. The first stage consisted in comparing the real and predicted values of water levels at hydrological stations in different periods of the passage of the flood. The results are shown in Table, where the real water level is the actual measured value at the stationary hydrological station and automatic station, and the predicted water level is the water level value obtained using the neural network. It is important to note that comparison with other forecasting methods (see fig. 1) was made in [14], according to the results of which neural network forecasting showed a better result (the output values turned out to be more accurate than when calculating the predicted values of water levels by other methods).

Forecast values of water levels at stationary hydrological station (76289) Ufa obtained based on an artificial neural network

Forecast date	Forecast for 5 days, cm	Forecast for 4 days, cm	Forecast for 3 days, cm	Forecast for 2 days, cm	Forecast for 1 day, cm	Actual water level for the forecast date, cm
07.04.2021	71	109	294	344	364	Forecast for 12.04.21: 297
08.04.2021	114	311	329	311	365	Forecast for 13.04.21: 416
09.04.2021	338	396	328	376	477	Forecast for 14.04.21: 481
10.04.2021	353	419	403	531	514	Forecast for 15.04.21: 516
11.04.2021	432	372	552	523	567	Forecast for 16.04.21: 550
12.04.2021	455	582	603	537	556	Forecast for 17.04.21: 586
13.04.2021	581	565	589	631	608	Forecast for 18.04.21: 607
14.04.2021	613	606	596	615	655	Forecast for 19.04.21: 626
15.04.2021	635	636	634	619	636	Forecast for 20.04.21: 636
16.04.2021	635	654	660	655	637	Forecast for 21.04.21: 637
17.04.2021	626	649	673	660	669	Forecast for 22.04.21: 628
18.04.2021	607	635	663	688	689	Forecast for 23.04.21: 607

19.04.2021	574	610	643	673	695	Forecast for 24.04.21: 587
20.04.2021	551	568	612	649	677	Forecast for 25.04.21: 565
21.04.2021	532	540	562	613	649	Forecast for 26.04.21: 526
22.04.2021	496	516	528	557	609	Forecast for 27.04.21: 475
23.04.2021	466	470	501	518	548	Forecast for 28.04.21: 432
24.04.2021	507	489	443	404	401	Forecast for 29.04.21: 387
25.04.2021	476	423	375	368	366	Forecast for 30.04.21: 318
26.04.2021	405	351	339	332	294	Forecast for 01.05.21: 255
27.04.2021	332	315	306	260	238	Forecast for 02.05.21: 232
28.04.2021	297	282	236	208	240	Forecast for 03.05.21: 189
29.04.2021	265	212	187	214	200	Forecast for 04.05.21: 179
30.04.2021	198	166	195	180	155	Forecast for 05.05.21: 161
01.05.2021	154	177	150	165	152	Forecast for 06.05.21: 141
Average error, cm	12,84	0,76	14,68	23,88	37,08	—

Every day, since the start of monitoring the development of the flood situation (in 2021 it is April 7, but the forecast was made for 1–5 days, respectively, the date of the beginning of the water rise level during the spring flood can be considered 12.04.2021, that is, $i = 1$), using a neural network, a forecast was made for the next m (for example, 1–5) days at 1 gauging station (that is, $n = 1$; this value of n was taken to reduce the experiment time). The next day, the actual value of water levels was measured at the same posts $h_{j,i+m}^k$. Forecasting daily until May 01 (respectively, we introduce the additional designation l — the number of flood days, that is, the value of $l_{k,i+m} = 25$) for each post, we obtain 25 forecasts and actual values of water

levels for each of the n posts, which makes it possible to determine the average forecast accuracy for each k -th post:

$$E_{\text{cp}}^k = \frac{1}{l_{k,i+m}} \sum_{i=1}^{l_{k,i+m}} E |h_{ji}^k - h_{j,i+m}^k|.$$

In this case, the error E (in centimeters) on the i -th day of the j -th year of each k -th gauging station is calculated by the ratio of the absolute difference:

$$E_{ji}^k = |h_{ji}^k - h_{j,i+m}^k|.$$

An important difference between the proposed method for predicting water levels using an artificial

neural network in comparison with other known methods is the speed of obtaining a forecast and its correctness (more accurate) with advance forecasting (in this case, by 5 days), the results are shown in fig. 4, see the 4th side of cover. According to Table and fig. 4, it can be seen that the artificial neural network gives fairly accurate results: the average error does not exceed 37.08 cm for the entire flood period. Thus, the implemented artificial neural network showed a more accurate and stable result, which proves the feasibility of using the proposed method in the framework of early forecasting of a flood situation.

Conclusion

This method was carried out at a high level of abstraction without reference to physical objects (this allows us to reduce the construction of a decision support system to a formal algorithm), which opens up the possibility of using it in a wide range of areas and tasks, including forecasting various natural case — water levels on water bodies) phenomena, man-made disasters, economic phenomena, etc. The analysis of efficiency based on the proposed method showed that the use of an artificial neural network showed more stable results when predicting water level values using the example of a spring flood in the Republic of Bashkortostan, which will give the necessary time for special services to implement flood control measures to prepare for the protection of technical facilities.

Thus, performing a sequence of data normalization, training a neural network with specified parameters, and adjusting the synapse weights, we obtain a method for constructing a decision support system for a wide range of tasks.

References

1. **Reimers N. F.** Nature management, Dictionary-reference book, 1990, 637 p.
2. **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.

3. **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.

4. **Luo X., Yuan X., Zhu S., Xu Z., Meng L., Pengm J.** A hybrid support vector regression framework for streamflow forecast, *Journal of Hydrology*, 2019, vol. 568, pp. 184–193.

5. **Strupczewski W. G., Kaczmarek Z.** Non-stationary approach to at-site flood frequency modelling II Weighted least squares estimation, *Journal of Hydrology*, 2001, vol. 248, no. 1, pp. 143–151.

6. **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, *Journal of Hydrology*, 2019, vol. 572, pp. 388–402.

7. **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.

8. **Tian J., Liu J., Yan D., Ding L., Li C.** Ensemble flood forecasting based on a coupled atmospheric-hydrological modeling system with data assimilation, *Atmospheric Research*, 2019, vol. 224, pp. 127–137.

9. **Gomez M., Sharma S., Reed S., Mejia A.** Skill of ensemble flood inundation forecasts at short- to medium-range timescales, *Journal of Hydrology*, 2019, vol. 568, pp. 207–220.

10. **Lopes V. A., Fan F. M., Pontes P. R., Siqueira V. A., Col-lischonn W., Marques D. M.** A first integrated modelling of a river-lagoon large-scale hydrological system for forecasting purposes, *Journal of Hydrology*, 2048, vol. 565, pp. 177–196.

11. **Velasco L., Serqunia R., Zamad M., Juanico B., Lomoso J.** Week-ahead Rainfall Forecasting Using Multilayer Perceptron Neural Network, *Procedia Computer Science*, 2019, vol. 161, pp. 386–397.

12. **TensorFlow**, available at: <https://www.tensorflow.org>

13. **Kratzert F., Klotz D., Shalev G., Klambauer G., Hochreiter S., Nearing G.** Benchmarking a Catchment-Aware Long Short-Term Memory Network (LSTM) for Large-Scale Hydrological Modeling, *Hydrology and Earth System Sciences Discussions*, 2019, vol. 22, pp. 1–36.

14. **Palchevsky E. V., Khristodulo O. I., Pavlov S. V.** Threat prediction in complex distributed systems using artificial neural network technology, *CEUR Workshop Proceedings*, 2020, vol. 2763, pp. 289–294 (in Russian).

15. **Mokhov V. G., Tsimbol V. I.** Electrical Energy Consumption Prediction of the Federal District of Russia on the Based of the Recurrent Neural Network, *Journal of Computational and Engineering Mathematics*, 2018, vol. 5, no. 2, pp. 3–15 (in Russian).

ООО "Издательство "Новые технологии". 107076, Москва, ул. Матросская Тишина, д. 23, стр. 2
Технический редактор *Е. М. Патрушева*. Корректор *А. В. Чугунова*.

Сдано в набор 27.04.2022 г. Подписано в печать 30.06.2022 г. Формат 60×88 1/8. Заказ PI622
Цена свободная.

Оригинал-макет ООО "Авансд солюшнз". Отпечатано в ООО "Авансд солюшнз".
119071, г. Москва, Ленинский пр-т, д. 19, стр. 1. Сайт: www.aov.ru

«STABILITY DECISION SUPPORT SYSTEM BASED ON APPLICATION
OF THE SECOND GENERATION NEURAL NETWORK»

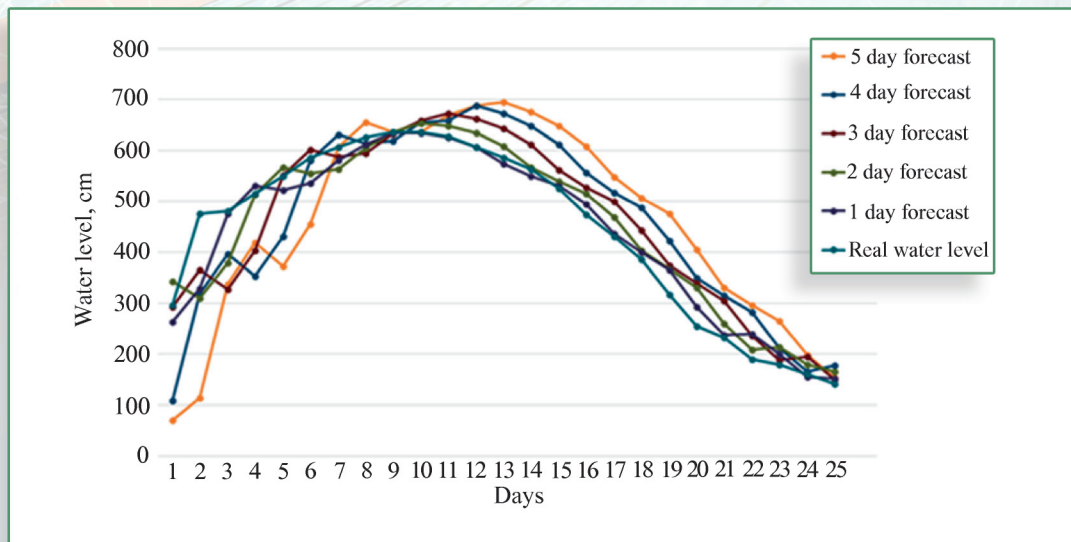


Fig. 4. The result of forecasting water levels for 5 days
at the stationary hydrological station 76289 in Ufa