Estimation of maximum water levels during spring flood on Lena river parts using artificial neural networks

G. P. Struchkova, T. A. Kapitonova, V. V. Timofeeva, D.D. Nogovitsyn

V.P. Larionov’s Institute of Physical-Technical Problems of the North, Siberian Branch of the Russian Academy of Sciences, Yakutsk, Russian Federation

E-mail: pandoramy8@list.ru

Abstract. The Republic of Sakha (Yakutia) possessing a vast territory located in various climatic zones and a developed network of water bodies is exposed to a wide range of natural emergency situations. Among them, spring-summer floods are the most frequent and bringing enormous damage, causing inundations of vast areas and national economy objects, which determines relevance of development and improvement of flood forecast methods for implementation of timely measures to prevent and reduce an inundation risk.

Artificial neural networks have proven their effectiveness in solving various forecast problems, especially when using statistical data. In particular, usage of a neural network approach based on the forecast of a time series from previous values gives good results. The artificial neural networks, unlike statistical methods of analysis, are based on parallel data processing, have an ability of self-learning and recognition of nonlinear relationships between input and output data sets. A choice of parts of the Lena river to predict maximum water levels during the floods was determined based on locations of potentially hazardous objects, the inundation of which can cause the considerable material damage. On the basis of the hydrological data obtained during 70 years, the neural network models are obtained, which make it possible to predict flood hazards from the spring floods, two variants of the transformed initial data are considered and different network structures are compared. Relative errors of the forecasts obtained during the work vary considerably (7–20%), which indicates necessity for the processing of the initial data and careful selection of the structure.

1. Introduction

The Republic of Sakha (Yakutia) has the vast territory located in the different climatic zones. A large number of rivers flowing from south to north contribute, especially during a flood period, to an occurrence of natural emergency situations (ES). The most characteristic of them are the spring-summer floods, causing the inundations of the vast areas, the objects, infrastructure and bringing the enormous damage to the national economy, destabilization processes in a technosphere, which determines the relevance of the development and improvement of the flood forecast methods to reduce the hazard levels and damage. During 13 years, from 2001 to 2013, the spring flood of the rivers in the Republic caused the damage amounting to a total of about 16 billion rubles [1]. The total damage from the flood in Yakutia in the spring and summer of 2018, as a result of which more than 5.3 thousand residents of Yakutia were harmed, exceeded 1.4 billion rubles. In the Republic during the spring and summer flood, 63 settlements in 15 districts and in the territory of Yakutsk were affected; 1482 residential houses were flooded [2]. On the basis of the statistical data obtained during 75 years and
regression modeling, the usage of neural networks to develop the model allowing one to predict the flood hazards from the spring floods is presented. The choice of the forecast part was determined based on the locations of the potentially hazardous objects, the inundation of which can cause the considerable material damage, natural factors affecting the level of the spring floods have been determined.

The time series modeling for the forecast of hydrological variables is an important step in planning and operational analysis of water resources. At present, a sufficiently large number of the methods of the forecast of the inundation hazard from the spring floods have been developed [3,4,5,6]. Traditionally, the autoregressive moving average models (ARIMA) were used to model the time series of the water resources, since such models are accepted as a standard representation of the stochastic time series. However, these models do not attempt to represent nonlinear dynamics inherent in the hydrological process and may not always work well. The time series analysis requires mapping of the complex relationships between the input (inputs) and output data (outputs), since the predicted values are displayed as a function of observed patterns in the past.

The intellectual computational methods, such as the artificial neural networks (ANN), have proven their effectiveness in solving the various problems, especially for working with the statistical data. Specifically, in the forecast the usage of the neural network approach, based on construction of the time series from the previous values, gives the good results [7, 8, 9]. The artificial neural networks, unlike the statistical methods of analysis, are based on the data parallel processing, have the self-learning ability and are able to recognize the nonlinear relationships between the input and output data sets. The theory of artificial neural networks is currently undergoing a stage of formation, which causes a variety of problem statements and basic definitions. The simplest way of application of the artificial neural networks in the forecast problems is to use a conventional perceptron with one, two or (in extreme cases) three hidden layers.

The time series analysis requires the mapping of the complex relationships between the input and output data, since the predicted values are displayed as the function of the observed patterns in the past. Due to difficulties associated with identification of the structure of the nonlinear model and estimation of parameters, the very few truly theoretical hydrological models of the nonlinear system were presented [9,10,11]. Particularly, in a prediction the use of the neural network approach, based on the construction of the time series from the previous values, gives the good results.

2. Materials and research methods

Using the data on the maximum water levels on the Lena river in the area of Tabaga during the period from 1938 to 2013, we will make the forecast till 2017 using the neural network. To increase prognostic capabilities of the neural networks, it is necessary to preliminary make a transformation of the data for the neural network [12,13]. We will consider the two variants of the transformed initial data - the series smoothed by the moving average and the time series with a subsequent seasonal adjustment. After the data transformation, we proceed to the choice of a configuration of the neural network itself. As practice shows, the multi-layer perceptron is best suited for the modeling and forecast of the time series. Using the sliding window method for the neural network, we form a block of training samples. A window size, i.e. a time series segment representing the input vector is difficult to exactly indicate due to a complex behavior of the time series. When choosing the number of the inputs of the neural network to predict the time series, quality of training should be taken into account, since an increase of the number of the inputs of the neural network accordingly reduces the size of the training sample. We adjust the model for the training: we set the value of hidden neurons, select the activation functions for the hidden and output neurons. We carry out the process of the training of the networks, then from the trained neural networks we choose the best one.
3. Results and discussions

We will analyze the data obtained to select the best model, comparing each model in terms of performance and errors. We will consider the two variants of the transformed initial data - the series smoothed by the moving average and the time series with the subsequent seasonal adjustment.

In the first case, the following models are considered with the higher performance and the smallest error:
1. MLP 12-4-1 is with an architecture of the network of a type of the multi-layer perceptron with 12 inputs, with 7 hidden neurons and with 1 output.
2. MLP 20-6-1 with the architecture of the network of the type of the multi-layer perceptron with 20 inputs, with 2 hidden neurons and with 1 output.
3. MLP 40-8-1 with the architecture of the network of the type of the multi-layer perceptron with 40 inputs, with 5 hidden neurons and with 1 output.

In the second case, the following models are selected:
1. MLP 12-7-1 with the architecture of the network of the type of the multi-layer perceptron with 12 inputs, with 7 hidden neurons and with 1 output.
2. MLP 20-2-1 with the architecture of the network of the type of the multi-layer perceptron with 20 inputs, with 2 hidden neurons and with 1 output.
3. MLP 40-5-1 with the architecture of the network of the type of the multi-layer perceptron with 40 inputs, with 5 hidden neurons and with 1 output.

Table 1. Comparative characteristic of results of retro-forecasts of water level in Lena river in area of Tabaga.

| Year | Error 2014 | Predicted | % | Error 2015 | Predicted | % | Error 2016 | Predicted | % | Error 2017 | Predicted | % |
|------|------------|-----------|---|------------|-----------|---|------------|-----------|---|------------|-----------|---|
|      | Observed   | Predicted |      | Observed   | Predicted |      | Observed   | Predicted |      | Observed   | Predicted |      |
| The series smoothed by the moving average | | | | | | | | | | | | |
| MLP 12-4-1 | 744 | 846 | 13 | 1037 | 948 | 8 | 825 | 933 | 13 | 745 | 829 | 11 |
| MLP 20-6-1 | 744 | 836 | 12 | 1037 | 959 | 7 | 825 | 890 | 8 | 745 | 822 | 10 |
| MLP 40-8-1 | 744 | 898 | 20 | 1037 | 899 | 13 | 825 | 925 | 12 | 745 | 874 | 17 |
| The series with the subsequent seasonal adjustment | | | | | | | | | | | | |
| MLP 12-7-1 | 744 | 803 | 7 | 1037 | 971 | 6 | 825 | 756 | 8 | 745 | 684 | 12 |
| MLP 20-2-1 | 744 | 789 | 6 | 1037 | 1001 | 3 | 825 | 770 | 7 | 745 | 797 | 7 |
| MLP 40-5-1 | 744 | 814 | 9 | 1037 | 923 | 10 | 825 | 924 | 12 | 745 | 909 | 22 |

The best results were shown by the MLP 20-2-1 model with the subsequent seasonal adjustment with the architecture of the network of the type of the multi-layer perceptron with 20 inputs, with 2 hidden neurons and with 1 output. From the data in Table 1 it follows that the quality of the forecast of the time series of the maximum water levels in the river during the spring flood depends on the choice of the optimal window size wherein the quality of the prognostic capabilities of the neural networks deteriorates with the excessive increase of the window size (a phenomenon of overtraining).

References
1. Burtseva E.I., Parfenova O.T. Economic damage from floods on the rivers of the Republic of Sakha (Yakutia) // Problems of modern economy, 2015, no 1 (53), pp. 256-259
2. http://ulus.media/2019/02/26/zhiteli-sela-kyuptsy-do-sih-por-ne-poluchili-kompensatsii-za-proshlogodnij-pavodok/
3. Kireeva M. B., Frolova N. L. Pozhidaeva D. S. Analysis of occurrence of extreme hydrological events on the rivers of the Russian Arctic // Changing climate and socio-economic potential of the Russian Arctic. M: League-Vent, 2015, pp. 21-35
4. Koronkevich N.I., Barabanova E.A., Zaytseva I.S. Extreme hydrological situations. M.: Media-PRESS, 2010, 464 p.
5. Struchkova G.P., Timofeeva V.V., Kapitonova T.A. The statistical model of floods during tides and ice drift on the Lena River near Yakutsk // Russian meteorology and hydrology, 2015, v 40, no 5, pp. 343-346. DOI: 10.3103/S1068373915050088
6. Krasnogorskaya N.N., Ferapontov Yu.I., Nafikova E.V. Development of methods for forecasting hydrological processes for water management tasks // Scientific notes of the Russian State Hydrometeorological University, 2013, no 28, pp. 43-50
7. Grebnev Y. V., Yarovoy A. V. Monitoring and forecasting of floods on the territory of the Krasnoyarsk region, using neural network algorithms // The Siberian fire and rescue Bulletin, 2018, v 10, no 3, pp. 13-16
8. Chen Y.H., Chang F.J. Evolutionary artificial neural networks for hydrological systems forecasting J. Hydrol., 367 (2009), pp. 125-137
9. Fazel S.A.A., Mirfenderesk H., Tomlinson R., Blumenstein M. Towards robust flood forecasts using neural networks, Neural Networks (IJCNN), 2015 International Joint Conference on, IEEE(2015), pp. 1-6
10. Volchak A., Kostyuk D., From Lavrentyev. neural forecasts for integrated flood monitoring system // proceedings of the 6th International scientific conference on the global energy and water cycle and the 2nd scientific conference on the earth’s ecosystem and atmosphere (iLEAPS), 24-28 August 2009, Melbourne, Australia. Thom. II-Pp. 713-714.
11. Holger R. Maier, Ashu Jain, Graeme C. Dandy, Sudheer K.P. Methods used for the development of neural networks for the prediction of water resource variables in river systems: Current status and future directions, Environmental Modelling & Software, Volume 25, Issue 8, August 2010, Pages 891-909
12. Osovsky S.A. Neural networks of information processing / S.A. Osovsky. - M.: Finance and Statistics, 2002, 344 p.
13. Borovikov V.P. Neural networks. Statistica Neural Networks. Methodology and technology of modern data analysis, 2nd ed., Revised and add. - M.: Hotline - Telecom, 2008, 392 p.