ISSN: 0975-8585

Research Journal of Pharmaceutical, Biological and Chemical Sciences

# Comparative Assessment of Methods for Forecasting River Runoff with Different Conditions of Organization.

**Fedor N Lisetskii*, Vitaliy I Pichura, Yaroslava V Pavlyuk, and Olga A Marinina.**

Belgorod State University, Russia, 308015, Belgorod, Pobedy str., 85.

**ABSTRACT**

The article presents the results of comparative ansalysis of application of traditional statistical methods and non-linear multilayer neural networks for processing multi-year data of the river runoff with the aim of its forecasting. It has been shown that universal possibilities are possessed by neuromodels based on multilayer perceptron (MLPN), that have provided test sampling reliability of modeling of 82–93%. The obtained in the work optimal parameters for learning predictive neuronetworks are recommended for learning transformation and forecast of river runoff in the situation of permenant anthropogenic influence on basin.

**Key words:** neuron networks , Box-Jenkins method, Winters method, river runoff, modeling, forecasting.

**Corresponding author*

ISSN: 0975-8585

## INTRODUCTION

The actual problem of using water resources is the improvement of method of analysis of river runoff assessment under fluctuated natural-economic conditions. The methods of statistical modeling and forecasting allow to get the most probable scenarios of random processes. The major part of series, including series with trend-cycling, seasonal (non-regular) components is possible to be described by autoregressive integrated moving average model (ARIMA). It is flexible enough and covers a wide range of stationary and non-stationary processes [1]. Seasonal components make it possible to take Winters method into account (Three-PES – three-parameter exponential soothing). But noisy and distorted information puts difficulties in the way of getting reliable forecasts. Artificial Neural Networks (ANN) allow to successfully solve problems which one cannot do via traditional methods [2]. Various approaches and results of possibilities of ANN applications have been studied by ASCE Task Committee [3], including flood forecast depending on precipitation total and catch water basin [4, 5], the modeling of within-year, annual river runoff for different climatic conditions, analysis of hydrochemical regime of the rivers [6-9].

## METHODS

Similarity of multi-year fluctuation in the river runoff has been studied using non-parameter methods: Kendall coordination and Spearman correlation. The rivers have been identified and grouped according to similar conditions of runoff organization using the method of difference integral curves. The methods of Box-Jenkins, Winters and multilayer ANN (RBFTN – radial basis, MLPN – multilayer perceptions) have been used for comparative forecasting of the river runoff movement. Instruction of ANN with input data and correction of weight factors is by means of back propagation of error algorithm (BFGS):

$$w_{ni}(t+1) = \eta\delta_i x_n(t) + \alpha(w_{ni}(t) - w_{ni}(t-1)),$$

where $w_{ni}(t)$ – weight from neuron $n$ or input signal element $n$ to neuron $i$ at the moment of time $t$; $x_n$ – neuron yield $n$ or $n$'th element of the signal; $\eta$ – learning rate coefficient; $\alpha$ – inertial coefficient; $\delta_i$ – error of neuron $i$.

The error function is difference of current output (approximate values) and ideal output (empirical values) of the network. The mathematical expectation of absolute percentage error is used to cross-check forecasting models (in physical measurements and percentage – MAPE). STATISTICA and Automated Neural Networks for Windows are used for processing.

## MAIN BODY

It has been studied the annual runoff of eight river stations (fig. 1) of the rank V with catch water basin ($F$) that is less than 800 km$^2$ and of the rank VI (1200 km$^2$ < $F$ < 9000 km$^2$) (see table 1). Duration of series sensing covers 1930–2012 years, standard error value is not >10% (3 % < $\sigma_{w_0}$ < 7%), variation coefficient is not > 0.5 (0.22 ≤ $C_v$ ≤ 0.5), variation coefficient error according to moments method is within 8-10.5%. The results of calculated hydrologic characteristic indicate the representativeness of the studied periods of runoff formation.

**Table 1: Statistics of average multi-year water expenditure for the rivers of the region of Belgorod**

| Statistical indicators | | Rivers[a] | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|
| | | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 |
| Observations, years | N | 81 | 62 | 50 | 63 | 66 | 57 | 60 | 31 |
| Average consumption, m$^3$/s | $Q_0$ | 5.81 | 1.00 | 2.57 | 1.92 | 6.25 | 14.36 | 3.26 | 5.80 |
| Variation coefficient | $C_v$ | 0.36 | 0.50 | 0.32 | 0.40 | 0.27 | 0.22 | 0.46 | 0.25 |
| Coefficient of skewness | $C_s$ | 0.67 | 0.51 | 0.42 | 0.57 | 0.55 | 0.69 | 1.40 | 1.35 |

[a] Numbering of hydroposts in fig. 1.

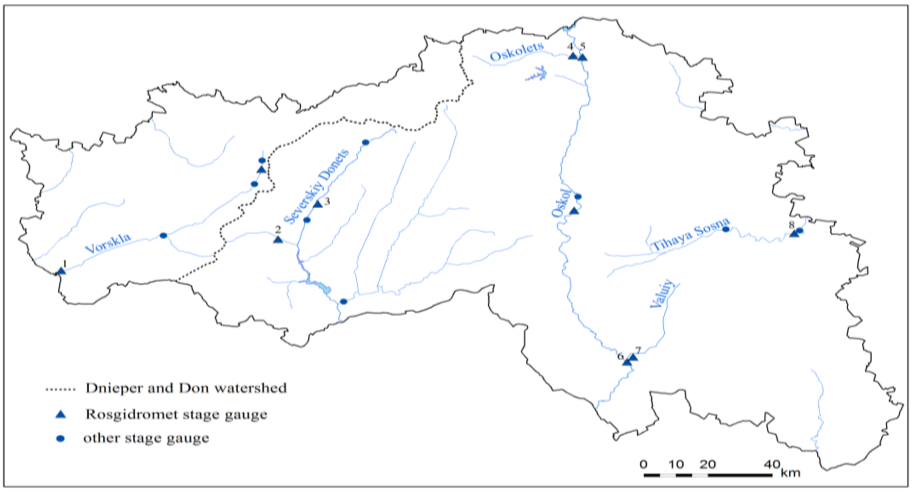


**Figure 1: Rivers runoff observational hydrostations in the region of Belgorod: 1 – Kozinka (Vorskla); 2 – Belgorod (Vezelka); 3 – Kiselevo (Severskiy Donets); 4 – Staryy Oskol (Oskolets); 5 – Staryy Oskol (Oskol); 6 – Razdolye (Oskol); 7 – Valuyki (Valuiy); 8 – Alekseyevka (Tihaya Sosna).**

The analysis of annual expenditure dynamics has shown a certain synchronism in changes of water content of the rivers of the Region of Belgorod, high values of Kendall correlation and ranking order correlation - 0,95 point at this, and it is indicative of runoff isotropism in the explored territory.

Statistical methods of the forecasting have been approbated and compared in terms of characteristic types of annual changes of the river runoffs [10], located on the southwestern slopes of Sredne-Russkaya Vozvyshennost' in the basins of the Dnieper and the Don (see fig. 1). In terms of the analysis of differential integral curves of the annual runoff three groups of rivers (fig. 2), for which characteristic are individual seasonal, cyclic and trend changes and also uneven emissions, depending on natural conditions and hydroeconomic situation of the basin, are differentiated. The first group comprises the rivers of the V rank with indigenous transformation of hydroregimes and water-collecting area of $F$ < 500 km$^2$; the second one – the rivers of the VI rank with disturbed type of hydrofunctioning and water-collecting area of 1000 km$^2$ < $F$ < 1800 km$^2$, the rivers of the third group are the rivers with relatively stable type of hydrofunctioning of the Vrank (500 km$^2$ < $F$ < 1000 km$^2$) and of the VI rank (1800 km$^2$ < $F$ < 9000 km$^2$).

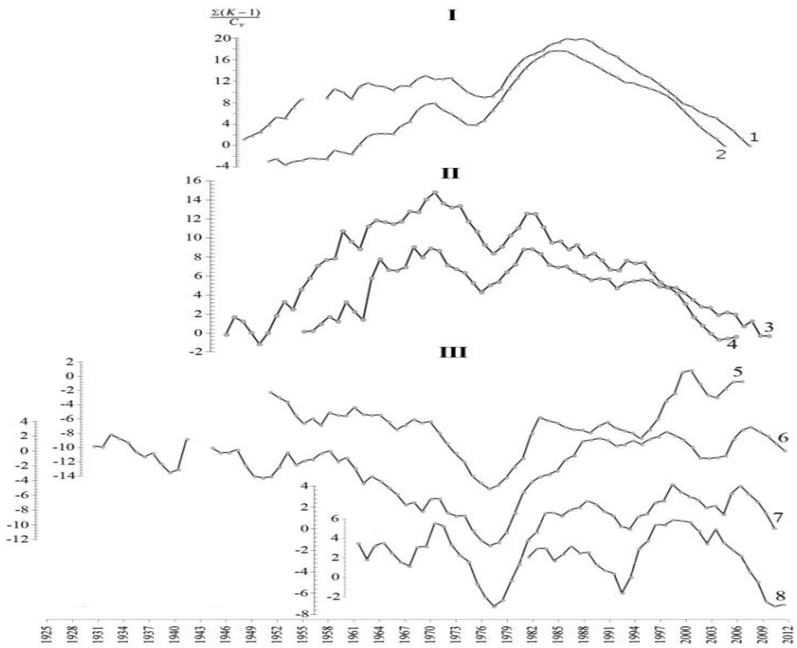


**Figure 2: Mixed differential integral curves of the river runoff:**

(Σ(K-1)/Cv) – values of integral curves, where K – ratio of actual condition of the river runoff in a certain period t to mean annual value; Cv – coefficient of annual river runoff variation.

River groups I–III: 1 – Vezelka (rank V, F = 394 $km^2$), 2 – Ooskolets (rank V, F = 494 $km^2$), 3 – Valuiy (rank VI, F = 1290 $km^2$), 4 – Oskol (river sources, rank VI, F = 1540 $km^2$), 5 – Oskol (Razdolye – lower current, rank VI, F = 8640 $km^2$), 6 – Vorskla (rank VI, F = 1870 $km^2$), 7 – Severskiy Donets (rank V, F = 740 $km^2$), 8 – Tihaya Sosna (rank VI, F = 2060 $km^2$).

Hydrologic regime using selected methods of forecasting has been modeled by using multiple combining of the parameters and model types. In optimal models (table 2) there is the obtained reflection of the best qualities, that have been gained in terms of learning (80%) and testing (20%) sets.

**Table 2: Comparative assessment of the time analysis methods with the aim of forecasting the average annual rivers runoff**

| Method | Model parameters | MAPE, % | |
|---|---|---|---|
| | | in terms of learning sample | in terms of testing sample |
| the I group of the rivers | | | |
| ARIMA | p(1), d(1), q(1), Ps(1), Ds(0), Qs(0) | 16–25 | 25–32 |
| Three-PES | α = 0.5; 0.7, β = 0.1, γ = 0.1 | 20–32 | 24–43 |
| MLPN | Learning algorithm – BFGS<br>Error processing function – SSE<br>Hidden and output neurons activation function– Exponential (range: 0;+∞), Hyperbolic tangent (range: –1;+1), Logistic sigmoid (range: 0;1) | 12–18 | 18–22 |
| the II group of the rivers | | | |
| ARIMA | p(1), d(1), q(0), Ps(0), Ds(0), Qs(1) | 20–25 | 26–28 |
| Three-PES | α = 0.1, β = 0.1, γ = 0.1 | 20–32 | 23–38 |
| MLPN | Learning algorithm – BFGS<br>Error processing function – SSE<br>Hidden and output neurons activation function – Hyperbolic tangent (range: –1;+1), Hyperbolic tangent (range: –1;+1), Exponential (range: 0;+∞) | 7–14 | 16–18 |
| The III group of the rivers | | | |
| ARIMA | p(1), d(1), q(1), Ps(1), Ds(0), Qs(0) | 25–27 | 30–35 |
| Three-PES | α = 0.5; β = 0.1, γ = 0.1 | 30–32 | 34–38 |
| MLPN | Learning algorithm – BFGS<br>Error processing function – SSE<br>Hidden and output neurons activation function – Logistic sigmoid (range: 0;1), Hyperbolic tangent (range: –1;+1), Exponential (range: 0;+∞), Identity (range: –∞;+∞) | 7–10 | 7–18 |
| The I–III groups of the rivers | | | |
| RBFTN | Learning algorithm – RBFTN<br>Error processing function – SSE<br>Hidden neurons activation function – Gaussian<br>Output neuron signal function – Identity (range: –∞;+∞) | Group I (19–21)<br>Group II (30–32)<br>Group III (20–34) | Group I (22–30)<br>Group II (35–40)<br>Group III (20–40) |

Abbreviations: MAPE – mathematical expectation of absolute error; MLPN – Multilayer Perceptron Network; RBFTN – Radial Basis Function Training Network; SSE – sum-squares error function, BFGS – Broyden-Fletcher-Goldfarb-Shanno.

The analysis of table 2 shows the advantage of ANN with architecture of MLPN. The realization of ANN for a definite aim (quantity of layers, hidden neurons, neuron activation functions, size of input set) depends on hydrologic characteristics of the rivers and the degree of their disturbance. The main feature of ANN-modelling is in its being based only on initial data without efforts to attract a priori reasons.

The most reliable runoff forecasting can be provided for the rivers with defined natural cyclicity in terms of water basin > 1000 $km^2$. For the rivers of group I the reliability of forecasting of MLPN based on testing sample is 81.9%. Runoff fluctuation under the influence of anthropogenic loads determines the optimal period of water content of the rivers. For group I it should not exceed the period of 8–9 years. For the river content forecasting of group II the initial interval should exceed the period of 12–18 years, then the reliability of forecasting of MLPN in terms of testing sample reaches 84.3%. To create effective neuromodels of the forecasting of the river runoff of group III with stable type of hydrofunctioning (1800 km$^2$ < $F$ < 9000 km$^2$) input time interval should be equal to or exceed three minor cycles (> 27 years), for rivers of 500 km$^2$ < $F$ < 1000 km$^2$ it should be equal to or exceed one minor cycle (8–9 years and more), maximum reliability of forecasting of MLPN in terms of testing sample is 92.7%. For the rivers of group III the forecasting is carried out for long-term period (> 15 years). The reference period of time series should be 2–3 times greater than predictive period.

Non-linear multilayer ANN have been created to forecast changes in dynamics of the river water content. Forecasting error in terms of testing sample is 7-18% that proves a high degree of reliability of the obtained data of the river runoff modelling within the time period.

## CONCLUSION

Comparative analysis of application of traditional statistic methods (ARIMA, Three-PES) and non-linear multilayer ANN (RBFTN, MLPN) has shown that ANN with the architecture of multilayer perception have the advantage in solving the problem of river runoff. Specifics of aged dynamics of the annual change of the river runoff on hydrological posts in the Dnieper and Don basins are conditioned by differences in hierarchy (rank) of the rivers, their water catchment area and degree of hydrological regime disturbance. On the basis of methods of non-parameter statistics it has been determined non-linear connection of synchronism of river water content with different types of hydrological functioning. By the results of differential integral curves of annual runoff the three types of rivers with specific seasonal, cyclic and trend regularities have been determined. For the river runoff forecast as optimal multilayer perception neuromodels (MLPN) have been determined. The reliability of the neuromodels in terms of testing samples is 82–93%. The present optimal variants of input set use for learning expected neuronetworks can be applied for learning transformation and high accuracy forecasting of the river runoff with different types of hydrological functioning in undiminishing anthropogenic influence on catch basins.

## REFERENCES

[1] Box, G.E.P. and Jenkins, C.M, 1970. Time series analysis: Forecasting and control. San Francisco: Holden-Day.

[2] Nigam R., S. Nigam and S. K. Mittal, 2014. The river runoff forecast based on the modeling of time series. Russian Meteorology and Hydrology, 39: 750–761.

[3] ASCE Task committee on application of artificial neural networks in hydrology. Artificial neural networks in hydrology II: Hydrologic applications, 2000. J. Hydrol. Eng. ASCE, pp: 213.

[4] Aziz K., et al., 2014. Application of artificial neural networks in regional flood frequency analysis: a case study for Australia. Stoch Environ Res Risk Assess., 28: 541–554.

[5] Dawson, C. W. and R. L. Wilby, 2001. Hydrological modelling using artificial neural networks. Hydrological modelling using artificial neural networks. Progress in Physical Geography, 25(1): 80–108. DOI: 10.1177/030913330102500104.

[6] Lisetskii F.N., Pavlyuk Ya.V., Kirilenko Zh.A., et al., 2014. Basin organization of nature management for solving hydroecological problems. Russian Meteorology and Hydrology, 39(8): 550–557. DOI: 10.3103/S106837391408007X.

[7] Sivakumar B., M.Persson, R.Berndtsson and C. Bertacchi, 2002. Is correlation dimension a reliable indicator of low-dimensional chaos in short hydrological time series. Water Resources Research, 2: 31–38.

[8] Ji Z., et al, 2013. Comprehensive assessment of flood risk using the classification and regression tree method. Stoch. Environ. Res. Risk Assess. 27: 1815–1828. DOI 10.1007/s00477-013-0716-z.

[9] Kim S., et al., 2014. Assessment of future water resources and water scarcity considering the factors of climate change and social–environmental change in Han River basin, Korea. Stoch. Environ. Res. Risk Assess. 28: 1999–2014.

[10] Lisetskii F. N., et al, 2014. New opportunities of geoplanning in the rural area with the implementing of geoinformational technologies and remote sensing. Advances in Environmental Biology, 8(10): 536–539.