УДК 004.94

K.A. Horielova, V.M. Zadachyn

Simon Kuznets Kharkiv National University of Economics, Kharkiv

LONG-TERM AND MEDIUM-TERM FORECASTING OF WATER CONSUMPTION OF LARGE CITIES

The problem of long-term and medium-term forecasting of water consumption of large cities is examined. Existing time series forecasting methods are reviewed. Advantages and disadvantages of autoregressive forecasting models and neural network models are formulated. Vivid example of a long-term and medium-term forecasting of water consumption of large city using ARIMA and neural network methods is presented.

Keywords*: water consumption, water supply, forecast, forecasting, time series, model, statistical analysis.*

Introduction

Lately the problem of minimizing costs of electricity in the production process has become extremely important for large industrial facilities. Among companies for which energy savings seem to be one of the most important factors in the survival and development, are water supply companies of large cities. As the city water supply and, consequently, the volume of water pumping are determined by need, there is a direct link between the task of planning electricity consumption and the task of forecasting urban water consumption. Therefore, in the conditions of rising urbanization the problem of improving the functioning of water supply systems in the areas of energy and resource conservation is urgent. One way of its solution is implementation of water consumption forecasting.

Issues of water consumption forecasting of large cities are considered by many authors in many papers. Some works are comprehensive guides of water demand management and consider the methods and techniques of water consumption forecasting in general [2, 3, 6, 10]. Other publications are intended to identify methods and models that are useful for solving water supply problems of specific water utility [7, 8]. Most of the authors of works on the subject agree that for the modeling and forecasting of water consumption of cities such techniques of time series analysis, as artificial neural networks [1, 5, 9] and regression models [1, 4] should be used.

1. Review of existing forecasting models

Forecasting as a research with a broad coverage of analysis objects leans on many methods. By estimates of taxonomist of prognostics, there are more than 100 methods of forecasting now, therefore there is a problem of a choice of methods, which would give adequate forecasts of the studied processes or systems for experts.

1.1. Time series models

Time series models are mathematical forecasting models that seek to find dependency of future value of the process on the past and to calculate forecast using this dependency. These models are universal for different subject areas, as their general appearance does not change depending on the nature of time series. Time series models can be easily divided [5]. They can be divided into two groups: statistical and structural (fig. 1).

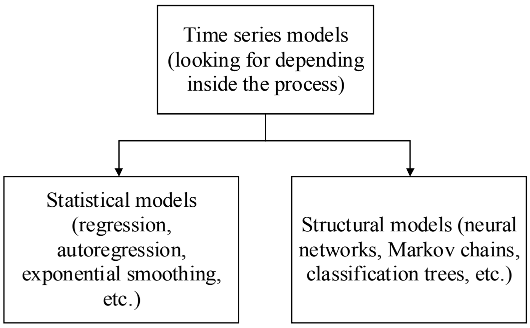


Fig. 1. Classification of time series models

In statistical models, dependency of the future value on the past is specified in the form of a given equation. These include:

– regression models (linear regression, nonlinear regression) [4];

– autoregressive models (ARIMA (Box-Jenkins model), ARIMAX, GARCH, ARDLM) [1,11];

– exponential smoothing model [13];

– maximum similarity models, etc. [13]

In structural models, future value dependency on the past is given in the form of certain pattern and its transition rules. These include:

– neural network models [1,5,9,12];

– models based on Markov chains [13];

– models based on classification and regression trees, etc. [13]

Experience shows that none of the methods, taken by itself, can provide a significant degree of reliability and precision horizon of forecast. However, in certain combinations they can be highly effective.

1.2. Autoregression forecasting models

Many tasks require the study of the relationship between two or more variables. Regression analysis is used to solve these problems [4]. The purpose of regression analysis is to determine the relationship between the original variable and many external factors (regressors).

Autoregressive models based on the assumption that the value of the process Z(t) linearly depends on a certain number of previous values of the same process Z(t-1),…,Z(t-p).

In autoregression models, the current value of the process is expressed as a linear finite set of past values of the process and impulse, called the "white noise" [11]:

$$Z(t) = C + \varphi_1 \cdot Z(t-1) + \ldots + \varphi_p \cdot Z(t-p) + \varepsilon_t. \quad (1)$$

Formula (1) describes the process of the autoregressive model order p, which is often referred in literature to AR(p) (autoregressive), C – constant, φ_1, …, φ_p – coefficients (model parameters), ε_t – model error.

Another type of model that is widely used in time series analysis and is often used in conjunction with autoregression is called the moving average model of order q and is described by the equation [13]:

$$Z(t) = (Z(t-1) + \ldots + Z(t-q)) / q) + \varepsilon_t. \quad (2)$$

In literature, the process (2) is often indicated as MA(q) (moving average); q – moving average order, ε_t – forecasting error. Moving average model is essentially a low pass filter.

To achieve greater flexibility in model fitting it is often advisable to combine in one model autoregression and moving average [11]. The general model is indicated as ARMA(p,q) and combines a filter of moving average of order q and autoregression of filtered values of the process of order p.

If as input values are used differences of times series values of order d (in practice d should be determined, but in most cases $d \leq 2$), than model is called autoregressive integrated moving average. In literature, this model is called ARIMA(p,d,q) [11].

The most part of research of autoregressive forecasting methodology was carried out by two statisticians, George E. P. Box and Gwilym Jenkins [11]. In the time series analysis autoregressive integrated moving average model (ARIMA) is one of the most widely used.

1.3. Neural network models

Currently, the most popular model among structural models is the structural model based on artificial neural networks (ANN). Neural networks are composed of neurons [12] (fig. 2).

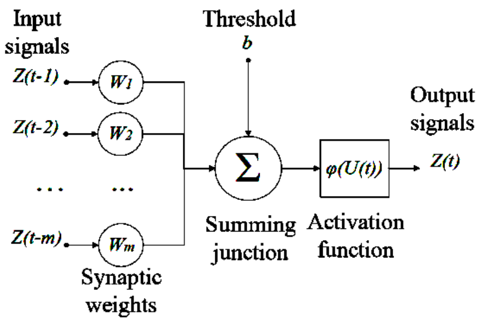


Fig. 2. Nonlinear neuron model

The neuron model can be described by a pair of equations [13]:

$$\begin{aligned} U(t) &= \sum_{i=1}^{m} \omega_i \cdot Z(t-i) + b; \\ Z(t) &= \varphi(U(t)), \end{aligned} \quad (3)$$

$Z(t-1)$, ..., $Z(t-m)$ – input signals;

ω_1, ..., ω_m – synaptic neuron weight;

b – threshold,

$\varphi(U(t))$ – activation function.

Activation functions are three main types [12]:

– binary step function;

– piecewise linear function;

– sigmoidal function.

Method of neurons communication defines the architecture of neural network. Depending on neuronal communication, networks are divided into [12]:

– single-layer models;

– multilayer models;

– recurrent networks.

Thus, using neural network modeling of nonlinear dependence of future time series values from its actual value and the external factors is possible. Nonlinear dependency is determined by the network structure and function of activation.

1.4. Advantages and disadvantages of models

Important advantages of autoregressive models are their simplicity and transparency of modeling. Another advantage is uniformity of analysis and design. Today this class of models is one of the most popular, therefore it is easy to find examples of autoregressive models for forecasting tasks for time series of different subject areas [5,13].

Disadvantages of these models include a large number of model parameters, the identification of which is ambiguous and intensive [9]; low adaptability of models, linearity and, consequently, inabilityof modeling nonlinear processes often encountered in practice.

The main advantage of neural network models is nonlinearity, which is ability to set non-linear relationship between future and actual values of processes. Other important advantages are adaptability, scalability (ANN parallel structure speeds up calculation) and uniformity of analysis and design.

At the same time disadvantages of ANN is the lack of modeling transparency; difficulty of architecture choosing, demands of consistency to the training sample; difficulty of training algorithm choosing and resource-intensive process of training.

It should be noted that none of the considered groups of models have forecasting accuracy as an advantage. This is because the accuracy of forecasting of a process depends not only on the model, but also on the experience of the researcher, the availability of data, the available hardware capacity and many other factors.

2. Example of water consumption forecasting of large cities

Input data to solve the problem is the time series of daily average values of water consumption for the month in the Moscow city in the period from January 1, 1996 to February 1, 2006 and the time series of daily average values of outside air temperature for the month in the city during the same period (fig. 3, 4). In the fig. 3 natural seasonal dependency of air temperature is clearly visible.

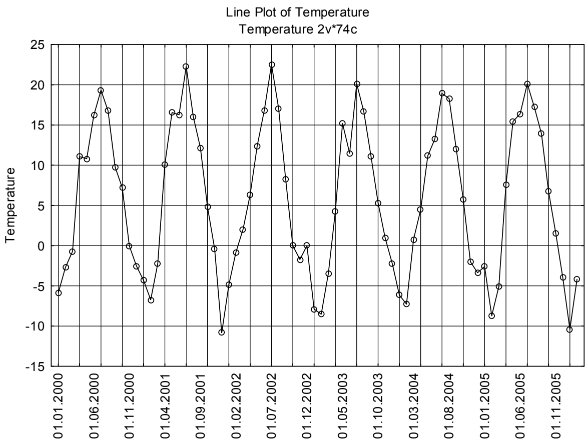


Fig. 3. Chart of air temperature

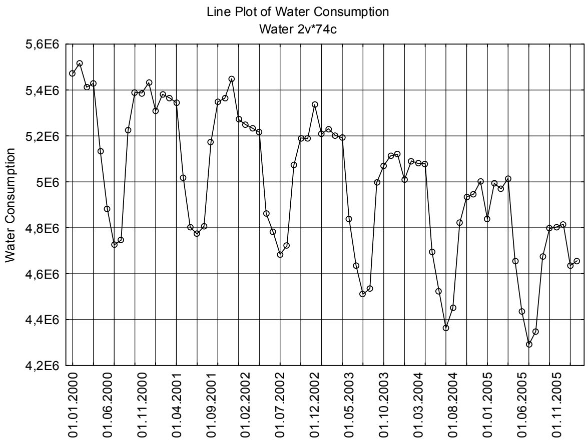


Fig. 4. Chart of water consumption

As you can see from the graph (fig. 4), water consumption decreased from year to year, but it also contains a seasonal component, which is associated with different water use in summer and winter (maximums are for December, January and February in different years). It is necessary to forecast water consumption for 2006 and test the adequacy of forecast.

2.1. Water consumption forecasting of large cities using ARIMA method

Let us determine the model parameters by method of selection: Seasonal lag = 12, p Autoregressive = 1, P Seasonal = 0, q - Mouving average = 1, Q Seasonal = = 1 (fig. 5). Chosen model gives forecast of series for 12 months a head (fig. 6).

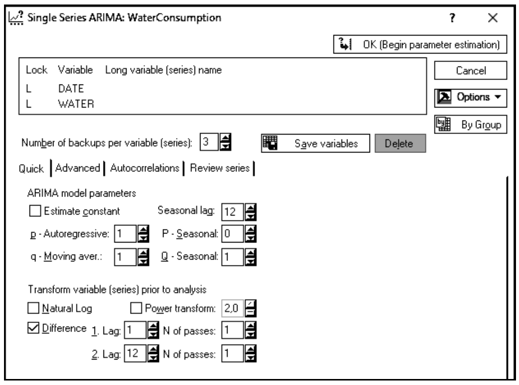


Fig. 5. Setting up ARIMA model (1,1,1) (0,1,1)

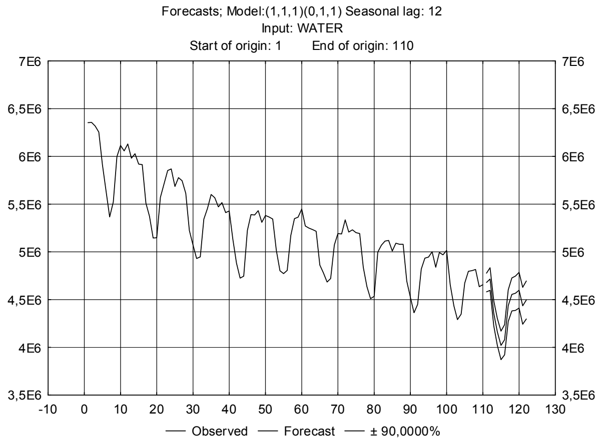


Fig. 6. Chart of forecasting for 12 months using ARIMA model (1,1,1) (0,1,1)

According to the forecast, it is intuitively clear that the model is quite adequate. To assess the built model it is necessary to use analysis of residues - a standard method of testing the adequacy of any statistical model. For this purpose, we will build auto correlation function of the residues, which is differences of the forecasted and actual values. (fig. 7).

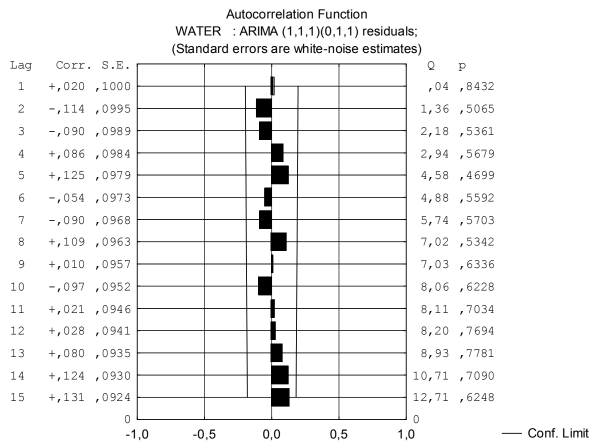


Fig. 7. Autocorrelation function of residues of ARIMA model (1,1,1) (0,1,1)

In well-selected model, correlation of residues is low, autocorrelation function and partial autocorrelation function charts do not exceed the limits.

Apparently, autocorrelation and partial autocorrelation (fig. 7) entirely lie in acceptable intervals. Another indication of good model is normal distribution of residues (fig. 8). For verifying, the quality of the forecast it is also recommended to make so-called "retro forecast" which is made for the period for which there is already a real data to compare it to the forecasted data (fig. 9).

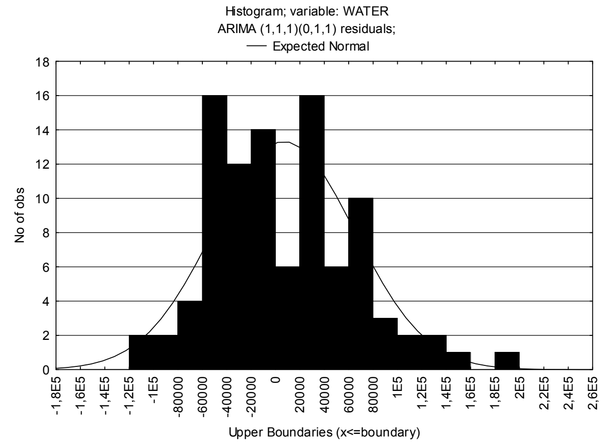


Fig. 8. Histogram of residues of ARIMA model (1,1,1) (0,1,1)

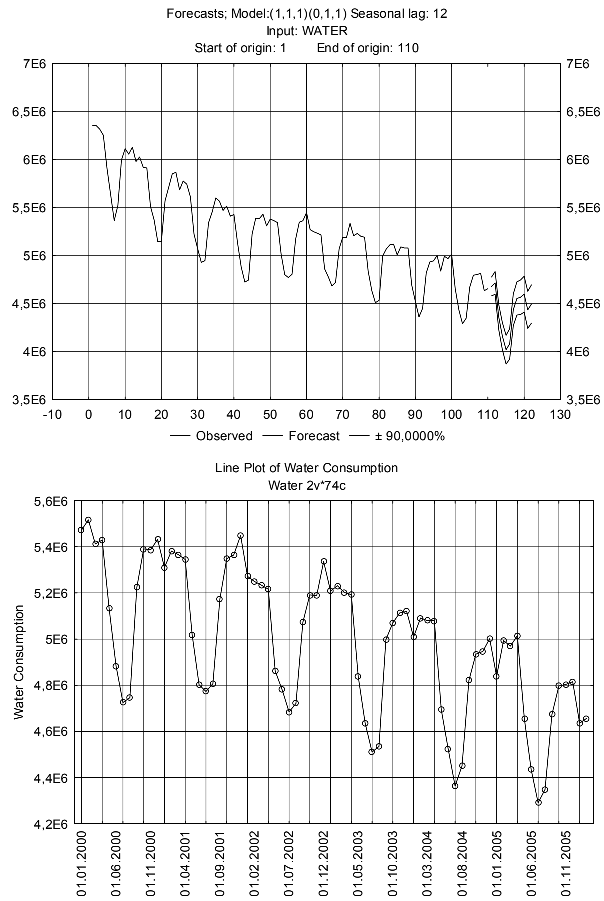


Fig. 9. Comparison of "retro forecast" and real data charts

Thus, ARIMA model allows building quite adequate city water consumption forecast through empirical selection of options.

2.2. Water consumption forecasting of large cities using neuron networks method

Let us create neural network to solve the problem of forecasting city water consumption. In the task of time series forecasting network must know how many copies of one variable it should take and how far ahead it should predict the value of a variable.

We select an option Steps equal to 12, because the data represents monthly observation and there is seasonal dependency, Lookahead parameter – 1, as the type of network – Multilayer Perceptron (fig. 10). Thus, we get a network of a three-layer perceptron type (fig. 11).

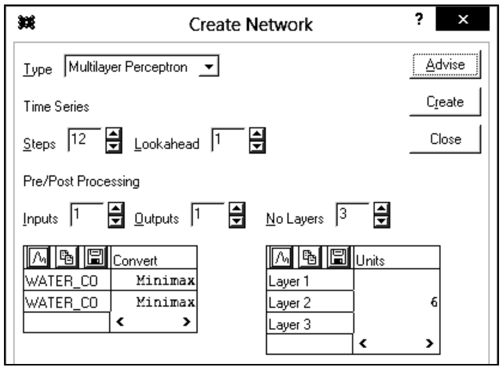


Fig. 10. Setting up neural network

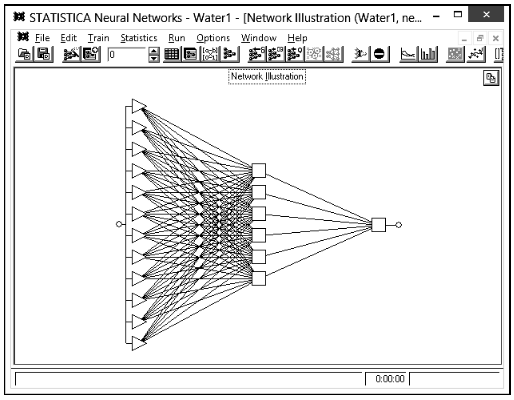


Fig. 11. Built neural network (three-layer perceptron)

Now we train the built network with Levenberg-Marquardt method (fig. 12) – one of the reliable and fast training algorithms.

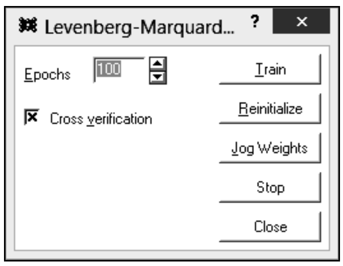


Fig. 12. Setting up Levenberg-Marquardt method

First of all network runs for the first 12 input values, resulting receiving forecast of next value. Then the predicted value together with the previous 11 input variables fed back to the input, and the network produces a forecast of the next values.

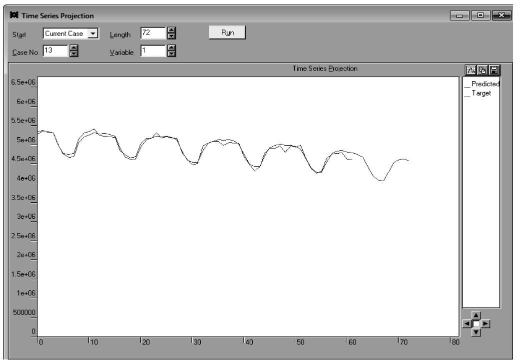
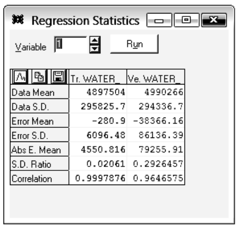


Fig. 13. The projection of time series forecast results for 12 months

In the fig. 13 it is shown that the predicted curve (blue) is very well-trained, as there is no significant deviation between the input and the forecasted series. It is possible to estimate the quality of neural network forecast, estimating the quality of the network (fig. 14).



Regression Statistics

Variable 1 Run

	Tr. WATER	Ve. WATER
Data Mean	4897504	4990266
Data S.D.	295825.7	294336.7
Error Mean	-280.9	-38366.16
Error S.D.	6096.48	86136.39
Abs E. Mean	4550.816	79255.91
S.D. Ratio	0.02061	0.2926457
Correlation	0.9997876	0.9646575

Fig. 14. Regression statistics

The correlation coefficient is nearly 1, which means that the neural network is built correctly.

Conclusions

ARIMA model allows building quite adequate medium-term city water consumption forecast through empirical selection of options. To build a long-term forecast it is necessary to use integrated approaches, such as neural networks, instead of standard statistical forecasting methods.

Among the advantages of the ARIMA method are simplicity and transparency of modeling, analysis and design uniformity and numerous examples of use, but along with it, this method is not suitable for nonlinearity modeling, it is low adaptive and not enough flexible. Neural networks instead can build nonlinear models, are scaling, highly adaptive and have many examples of use.

References

1. Collantes-Duarte J. Time Series Forecasting using ARIMA, Neural Networks and Neo Fuzzy Neurons / J. Collantes-Duarte, F. Rivas-Echeverriat // WSEAS International Conference on Neural Networks and Applications, Switzerland, 2002. – 6 p.

2. David Butler, Fayyaz Ali Memon Water Demand Management. London IWA Publishing, London, UK, 2006.– 348 p.

3. Donkor, E., Mazzuchi, T., Soyer, R., Roberson J.A. Urban Water Demand Forecasting: Review of Methods and Models. J. WaterResour. Plann. Manage., 2014.– 38 p.

4. Draper N., Smith H. Applied regression analysis. New York: Wiley, In press, 1981. – 693 p.

5. Gheyas I.A., Smith L.S. A Neural Network Approach to Time Series Forecasting // Proceedings of the World Congress on Engineering, London, 2009. – P. 1292 – 1296.

6. House-Peters, L., Chang, H. Urban water demand modeling: Review of concepts, methods, and organizing principles. [Article]. Water Resources Research, 2011. – 15 p.

7. HuienNiu. Approaches to Long-term Forecasting of Urban Water Demand in China / Department of Town and Country Planning University of Newcastle upon Tyne, 1994. – 277 p.

8. Junguo Liu. Forecast of water demand in Weinan City in China using WDF-ANN model / Junguo Liu, Hubert H.G. Savenije, JianxinXu / Physics and Chemistry of the Earth 28, 2003. – P. 219–224.

9. Morariu N. A neural network model for time series forecasting / N. Morariu, E. Iancu, S. Vlad // Romanian Journal of Economic Forecasting. – 2009. – P. 213 – 223.

10. R. Bruce Billings, Clive Vaughan Jones Forecasting Urban Water Demand, American Water Works Association, 2008. – 359 p.

11. Бокс Дж. Анализ временных рядов, прогноз и управление / Дж. Бокс, Г.М. Дженкинс. – М.: Мир, 1974. – Вып. 1. – 406 с. Вып. 2. – 197 с.

12. Хайкин С. Нейронные сети: полный курс / С. Хайкин. – М.: ООО «И. Д. Вильямс», 2006. – 1104 с

13. Чучуева И.А. Модель прогнозирования временных рядов по выборке максимального подобия: дис. канд. техн. наук: 05.13.18 / Чучуева Ирина Александровна. – М: 2012. – 153 с.

Надійшла до редколегії 19.11.2015

Рецензент: канд. техн. наук, проф. І.Г. Гусарова, Харківський національний університет радіоелектроніки, Харків.

ДОВГОСТРОКОВЕ І СЕРЕДНЬОСТРОКОВЕ ПРОГНОЗУВАННЯ ВОДОСПОЖИВАННЯ ВЕЛИКИХ МІСТ

К.А. Горєлова, В.М. Задачин

Розглянуто проблему довгострокового і середньострокового прогнозування водоспоживання великих міст. Зроблено огляд існуючих моделей прогнозування з використанням часових рядів. Сформульовані переваги і недоліки авторегресійних моделей прогнозування та моделей нейронних мереж. Здійснено наочний приклад довгострокового і середньострокового прогнозування водоспоживання міста методами ARIMA та нейронних мереж.

***Ключові слова**: водоспоживання, водопостачання, прогноз, прогнозування, часовий ряд, модель, статистичний аналіз.*

ДОЛГОСРОЧНОЕ И СРЕДНЕСРОЧНОЕ ПРОГНОЗИРОВАНИЕ ВОДОПОТРЕБЛЕНИЯ БОЛЬШИХ ГОРОДОВ

К.А. Горелова, В.М. Задачин

Рассмотрена проблема долгосрочного и среднесрочного прогнозирования водопотребления больших городов. Сделан обзор существующих моделей прогнозирования с использованием временных рядов. Сформулированы преимущества и недостатки авторегрессионых моделей прогнозирования и моделей нейронных сетей. Осуществлен наглядный пример долгосрочного и среднесрочного прогнозирования водопотребления города методами ARIMA и нейронных сетей.

***Ключевые слова:** водопотребление, водоснабжение, прогноз, прогнозирование, временной ряд, модель, статистический анализ.*