УДК 004.94

K.A. Horielova, V.M. Zadachyn

Simon Kuznets Kharkiv National University of Economics, Kharkiv

PLANNING OF CITY WATER SUPPLY SYSTEM MODERNIZATION BASED ON WATER CONSUMPTION FORECAST

The problem of planning of large cities water supply system modernization based on long-term water consumption forecast is reviewed. A review of existing models of long-term forecasting on the basis of time series is made.

Advantages and disadvantages of forecasting models based on autoregressive models ARIMA, neural networks and exponential smoothing are formulated. Vivid example of a long-term forecasting of water consumption of large city using these methods is presented to identify the most efficient and adequate model.

Keywords*: water supply, water consumption, forecast, forecasting, time series, model, statistical analysis, exponential smoothing, neural network.*

Introduction

Problems of water supply needs of the growing economy and world population are aggravated every day involving more and more attention. Currently one of the most important factors for the survival and development of large industrial enterprises, which include water supply companies in large cities, there are energy savings. Thus extremely important task is to minimize its costs in the production process.

The current stage of society economic development requires the implementation of water supply system that must provide the necessary technological parameters of water consumption, reliability and energy efficiency. The water system in many cities of Ukraine in the industrial and technically extremely worn, and their financial position becomes unprofitable. Low solvency of the population and the critical condition of industrial enterprises does not allow to pay for water in time. At the same time, the water system is a major consumer of electricity, gas, petrochemical and other resources. Considering this, the water supply companies of Ukraine suffer from significant costs caused by the need to spend a lot of electricity in order to pump the water by pumping stations.

With growing urbanization actual problem is upgrading the water supply system in the areas of energy and resource saving. City water supply with corresponding amounts of water pumping are determined by direct request, so there is a direct link between the objective of water system improving planning and power consumption planning, which in turn depend on the volume of future city consumption. One solution to this problem is the implementation of long-term water consumption forecasting.

The literature on forecasting of consumption large cities water consumption is not well lighted and consists mainly of small foreign publications.

For example, scientific article of L.A House Peters and R. Chang [6] is an overview of concepts, methods and organizational principles of city water consumption modeling. The methodological research in urban water consumption modeling over the past three decades was made. In Huien Niu’s scientific work [5] the municipal water system of China was systematically analyzed and he dynamic model for long-term forecasting of urban water consumption was built on the results detected by the analysis. Article of Robert Palmer from University of Washington [7] reviews the newly developed by Seattle public water utility model of long-term city water consumption forecasting. The work of Dutch scientists M. Bakker, J.H.G. Vreeburg, L.C. Rietveld, T. Blom, M. Van der Roer [11] is reviewing building adaptive models of water consumption forecasting in the Netherlands. Actual calculations and charts, detailed composition of forecasting model are given and its accuracy and results of implementation are analyzed. Nazario D. Ramirez-Beltran's work devoted to forecasting water consumption in Puerto Rico [9], considers regression model and time series analysis in general.

Some papers are comprehensive guides to managing water consumption and consider the methods and techniques of consumption forecasting in general. For example, the book of R. Bruce Billings and Clive W. Jones [10] provides all the necessary tools for accurate forecasting needs of drinking water for the city in the short, medium and long term. It considers the full range of spheres of influence on urban water consumption, including weather, climate, water prices, wage rates, and so on, and describes in details all the methods of water consumption forecasting used by water utilities USA. David Butler’s and Faiz Ali Memon’s book [2] is a comprehensive guide to water consumption managing. The concept of this work was to gather a complete picture of water demand management specifics, from technical to social and legal aspects.

Thus, the problem of water consumption forecasting is not widely covered in the Ukrainian literature, and is mainly represented by the works of foreign authors who have focused on the purpose of detection methods and models useful for solving water supply problems of specific water utility.

The purpose of this article is to identify the most effective and adequate long-term forecasting models for planning the modernization of the water supply of large cities in Ukraine.

1. Review of forecasting methods

Forecasting is a process of predicting the future state of an object or phenomenon by analyzing its past and present or systematic information on the qualitative and quantitative characteristics of the subject or phenomenon in the future [13].

Forecasts vary according to forecasting horizon. Depending on the period of time for which the forecast is made, there are operational, short, medium and long-term forecasts. Each time series has its own classification ranges. Long-term forecast on which modernization of the water utility should be based is used to develop strategic plans. It is characterized by the combination of qualitative and quantitative forecasting methods. It is made for the future perspective, for which significant qualitative changes are expected .

Overall, the forecasting method is a sequence of actions that need to be done to obtain a forecasting model, which, in turn, is a functional notion that adequately describes the studied process and is the basis for receiving its future values.

Forecasting as research with a wide coverage of analysis objects is based on a variety of methods. According to foreign and native systematics of prognostics forecasting already has more than 100 methods and because of this there is a problem of selection methods that would provide adequate forecasts for the studied processes or systems [15].

Term "forecasting method" is much broader than term "forecasting model". In this regard, on the first stage of classification methods are usually divided into two groups: intuitive and formalized [15].

Intuitive forecasting methods deal with judgments and experts' estimates. They are used in cases where it is impossible to take into account many factors influence due to the considerable complexity of the object of forecasting

Formal methods are described in the literature forecasting methods resulting in the building forecasting models, i.e. determining the mathematical relationship, which allows to calculate the future value of the process, that is to make a forecast.

Next step is the classification of forecasting models. In the first phase models should be divided into two groups: subject area models and time series models (fig. 1).

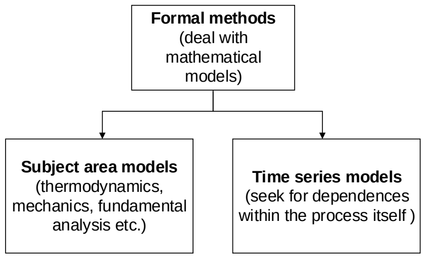

Figure 1. Classification of formal forecasting methods

Subject area models are such mathematical forecasting models, which are used for building subject area laws. For example, the model, which make weather forecasts, contains fluid dynamics equations and thermodynamics equations.

Time series models are mathematical forecasting models that seek addiction of future value of the process in the past inside the process and on these dependences calculate the forecast. These models are universal for different subject areas, i.e. their general appearance does not change depending on the nature of time series.

Time series models can be divided into two groups: statistical and structural (fig. 2).

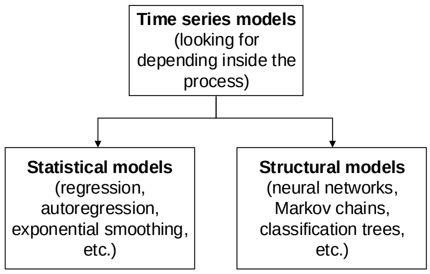

Figure 2. Classification of time series models

As we can see from the classification, significant group of forecasting methods are statistical methods.

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

– regression models (linear regression, nonlinear regression) [3];

– autoregressive models (ARIMA (Box-Jenkins model), ARIMAX, GARCH, ARDLM) [1,12];

– exponential smoothing model [15];

– maximum similarity models [15], etc.

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

They include:

– neural network models [1,4,8,14];

– models based on Markov chains [15];

– models based on classification and regression trees [15], etc.

1.1. Autoregressive ARIMA models

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" [12]:

$$Z(t) = C + \phi_1 Z(t-1) + \phi_2 Z(t-2) + \ldots + \phi_p Z(t-p) + \varepsilon_t. \quad (1)$$

Formula (1) describes the process of the auto-regressive model order p, which is often referred in literature to AR(p) (autoregressive), C – constant, $\phi_1, \ldots, \phi_p$ – 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 [15]:

$$Z(t) = \frac{1}{q}\big(Z(t-1) + Z(t-2) + \cdots + Z(t-q)\big) + \varepsilon_t. \quad (2)$$

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≤2), than model is called autoregressive integrated moving average. In literature, this model is called ARIMA(p,d,q) [12].

Important advantages of autoregressive models are their simplicity, transparency of modeling, analysis and design consistency.

The disadvantages of this class of models are: a large number of model parameters, the identification of which is ambiguous and intensive [8]; low adaptability of models, linearity and, consequently, the lack of ability modeling nonlinear processes that often found in practice.

1.2. Neural network models

Currently, the model based on artificial neural networks (ANN) is the most popular among structural models. Neural networks are composed of neurons [14].

The neuron model can be described by a pair of equations [15]:

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

where $Z(t-1), \ldots, Z(t-m)$ – input signals; $\omega_1, \ldots, \omega_m$ – 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.

The main advantage of neural network models is nonlinearity, which is ability to set non-linear relationship between future and actual values of processes, 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.

1.3. Exponential smoothing models

Exponential smoothing models were developed in the middle of XX century and today are widespread due to their simplicity and clarity [15].

Exponential smoothing (ES) is based on the idea of constant review of the predictive values as of actual receipt. ES model assigns exponentially declining weights of observations as they age. Thus, the last available observation has a greater impact on the forecasted value than the older observations.

Function of ES model looks like:

$$Z(t) = S(t) + \varepsilon_t, \quad (4)$$
$$S(t) = \alpha \cdot Z(t-1) + (1-\alpha) \cdot S(t-1),$$

where α – smoothing factor, 0<α<1; initial conditions are determined, such as S(1)=Z(0). In this model, each following smoothed value S(t) is a weighted average between the previous value of the time series Z(t) and the previous smoothed value S(t-1).

Holt model or double exponential smoothing is applied to modelling processes with trend [13]. In this case, the model should consider two components: level and trend. The level and trend are smoothed separately:

$$Z(t) = S(t) + \varepsilon_t; \quad (5)$$
$$S(t) = \alpha \cdot Z(t-1) + (1-\alpha) \cdot \big(S(t-1) - B(t-1)\big);$$
$$B(t) = \gamma \cdot (S(t-1) - S(t-2)) + (1+\gamma) \cdot B(t-1).$$

Here α – level smoothing factor, as like in model (4), γ – trend smoothing factor.

Holt-Winters or triple exponential smoothing is applied to modelling processes with trend and seasonal component:

$$Z(t) = (R(t) + G(t)) \cdot S(t). \tag{6}$$

Here $R(t)$ – smoothed level without seasonal component:

$$R(t) = \frac{\alpha \cdot Z(t-1)}{S(t-L)} + (1+\alpha) \cdot \big(R(t-1) + G(t-1)\big), \tag{7}$$

$G(t)$ – smoothed trend:

$$G(t) = \beta \cdot \big(S(t-1) - S(t-2)\big) + (1-\beta) \cdot G(t-1), \tag{8}$$

and $S(t)$ – seasonal component:

$$S(t) = \frac{\gamma \cdot Z(t-1)}{S(t-L)} + (1-\gamma) \cdot S(t-L). \tag{9}$$

Size of lag L is determined by the length of the season of the studied process.

The advantages of exponential smoothing models are simplicity and uniformity of analysis and design. Exponential smoothing models are most popular for long-term forecasting [15].

Disadvantage of this class of models is the lack of flexibility.

2.1. Long-term water consumption forecasting for planning of city water supply system modernization

As an example of long-term water consumption forecasting of large city, we consider a time series of average year values of water consumption in Moscow city in the period from 1 January 1972 to 1 January 2009 and the time series of population in the city during the same period (fig. 3, 4). As the software we use well-known package STATISTICA.

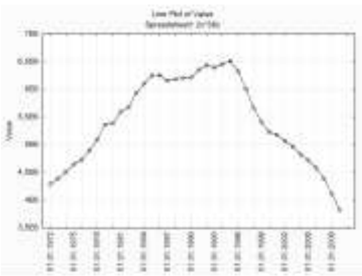

Figure 3. Line graph of water consumption by years

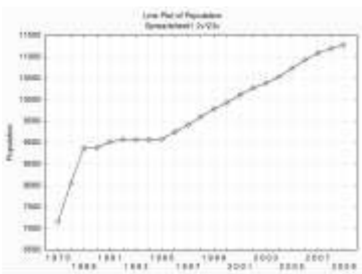

Figure 4. Line graph of the population by year

As you can see from the graph (fig. 3), water consumption increased until 1995 and decreased rapidly from year to year after, and the population has increased steadily (fig. 4).

It is necessary to forecast the consumption of water for 10 years ahead and test the adequacy of the forecast.

Let us forecast water consumption using three models: model ARIMA, neural networks and exponential smoothing, and then compare the results.

Determine the ARIMA (p, d, q) (P, D, Q) model parameters by selection as: p Autoregressive = 1, q - Moving average = 1, d =1. As there is no seasonality in annual water consumption time series, P Seasonal = 0, Q Seasonal = 0. The results of the forecast are presented in fig. 5.

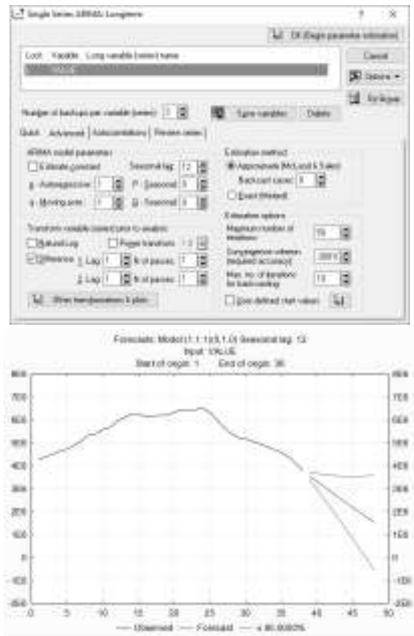

Figure 5. Configuration and forecast chart for 10 years using ARIMA model (1,1,1) (0,1,0)

To evaluate the constructed model, it is suitable to use analysis of residues. For this purpose, we build autocorrelation function balances, i.e. differences of forecasted and actual values (fig. 6).

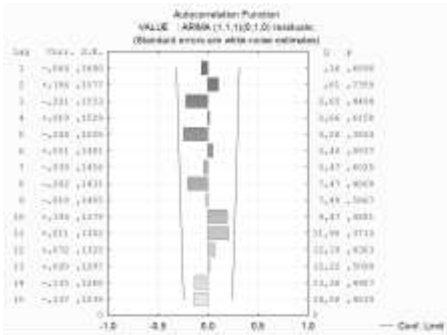

Figure 6. Autocorrelation function of residues of ARIMA model (1,1,1) (0,1,0)

To create a city water consumption forecast based on artificial neural network we select volume of water consumption and population as the input values of the time series. Then create Multilayer Perceptron as type of neural network and teach it with conjugate gradient method (fig. 7).

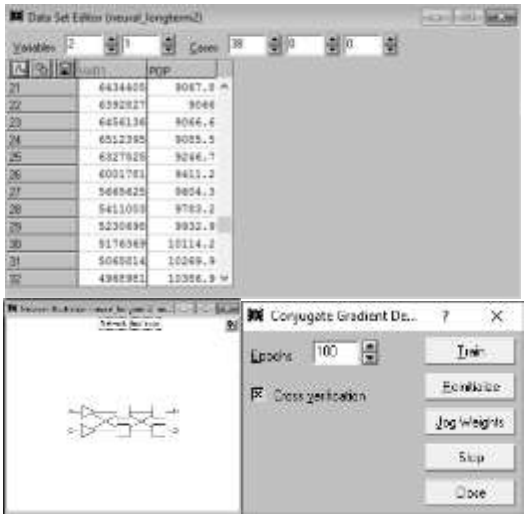

Figure 7. Selecting the input values and creating a three-layer perceptron

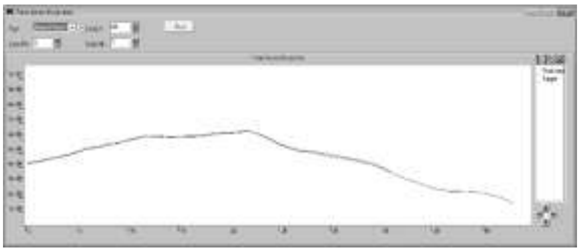

Figure 8. Forecast results projection of time series for 10 years

It is possible to rate the quality of forecast based on neural network, considering the quality of the network (fig. 9).

Figure 9. Regression statistics

Choose the following options for exponential smoothing model: α=0,9, γ=0,9, φ=0,9, Damped trend without seasonal component. Set Forecast = 10, defining the forecast for 10 years ahead (fig. 10).

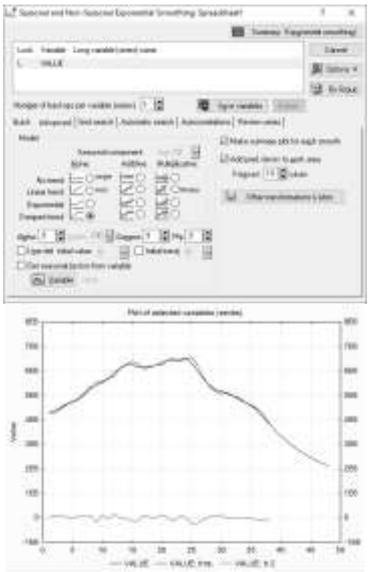

Figure 10. Model setting up and forecast chart for 10 years

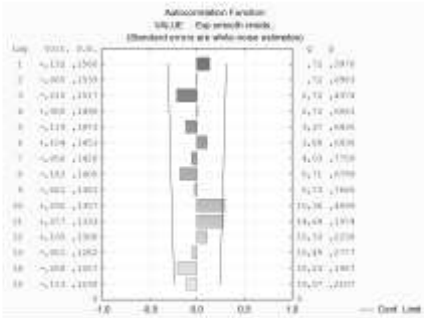

Figure 11. Autocorrelation function of residues of exponential smoothing model

It is intuitively clear that all the forecasting models give quite believable water consumption forecasts.

On neural network forecast chart (рис. 8) 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. The correlation coefficient is close to 1 (fig. 9), that means the neural network is built correctly.

On exponential smoothing model forecast chart (fig. 10) original series and its smoothed version match, that's why the forecast is adequate

In well-selected model, correlation of residues is low, autocorrelation function and partial autocorrelation function charts do not exceed the limits.Apparently, autocorrelation function of ARIMA (1,1,1) (0,1,0) model (fig. 6) entirely lies in acceptable intervals, instead exponential smoothing model autocorrelation slightly higher than permissible limits (fig. 11).

3. Conclusion

Thus, the results of a forecasting models comparison revealed that ARIMA models and models based on neural networks are most effective and adequate models of long-term forecasting of large cities water consumption.

Considering the fact that according to the results of forecasting the next 10 years consumption will be reduced, we can conclude that the water supply company is advisable to upgrade the water supply system so as to reduce the amount of cleaning and pumping facilities to thereby minimize economic costs.

The represented analysis of large cities water consumption forecasting methods allows to determine which methods and models are the best solution of the urban problems of public water supply company management decisions, which depend on future levels of demand.

Modern water utilities functioning system differs by a wide range of integration of all kinds of resources, a large number of infrastructure, internal and external communications. Providing significant economic autonomy of water utilities and the need for formation of social-market elements of communal policy allow to build an adequate regulatory system of water consumption and its resource intensity based on a long-term forecast.

References

1. *Collantes-Duarte J., Rivas-Echeverriat F. Time Series Forecasting using ARIMA, Neural Networks and Neo Fuzzy Neurons // 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. *Draper N., Smith H. Applied regression analysis. New York: Wiley, In press, 1981. – 693 p.*
4. *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.*
5. *Huien Niu. 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.*
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. *Long-Term Water Demand Forecasting Methods [Electronic resource]. – Access mode to the resource: http://www.ecs.umass.edu/waterresources/projects/KingCounty/KCReviewSeattleDemand_9-8-06.doc.*
8. *Morariu N., Iancu E., Vlad S. A neural network model for time series forecasting // Romanian Journal of Economic Forecasting, 2009, P. 213 – 223.*
9. *Prediction of Water Use in Pueto Rico [Electronic resource]. – Access mode to the resource: http://prwreri.uprm.edu/publications/Prediction%20of%20Water%20Use%20in%20Pueto%20Rico.pdf*
10. *R. Bruce Billings, Clive Vaughan Jones Forecasting Urban Water Demand, American Water Works Association, 2008. – 359 p.*
11. *The use of an adaptive water demand prediction model [Electronic resource]. – Access mode to the resource: http://repository.tudelft.nl/assets/uuid:19ad57bd-054e-4c88-8f26-598676853f1a/283886.pdf*
12. *Бокс Дж., Дженкинс Г.М. Анализ временных рядов, прогноз и управление. – М.: Мир, 1974. вып.1 — 406 с.; вып.2 — 197 с.*
13. *Прогнозування [Електронний ресурс]. – Режим доступу до ресурсу: https://uk.wikipedia.org/wiki/Прогнозування*
14. *Хайкин С. Нейронные сети: полный курс. – М.: ООО «И. Д. Вильямс», 2006. – 1104 с*
15. *Чучуева И.А. Модель прогнозирования временных рядов по выборке максимального подобия: дис. канд. тех. наук: 05.13.18 / Чучуева Ирина Александровна. – М: 2012. – 153 с.*

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

Authors: *HORIELOVA Kseniia Anatoliivna*
Simon Kuznets Kharkiv National University of Economics, Kharkiv, student, Tel. – (067)572-86-22, E-mail - kseniia.horielova@hneu.net.
ZADACHYN Victor Mykhailovych
Simon Kuznets Kharkiv National University of Economics, Kharkiv, PhD in Physics and Mathematics, Associate Prof. Tel. -(067)397-57-47 , E-mail – zadachinvm@gmail.com.

Планування модернізації системи водопостачання міста на базі прогнозу водоспоживання

К.А. Горєлова, В.М. Задачин

Розглянуто проблему планування модернізації системи водопостачання великих міст на базі довгострокового прогнозу водоспоживання. Зроблено огляд існуючих моделей довгострокового прогнозування з використанням часових рядів.

Сформульовані переваги і недоліки моделей прогнозування на основі авторегресійних моделей ARIMA, нейронних мереж та експоненціального згладжування. Здійснено наочний приклад довгострокового прогнозування зазначеними методами водоспоживання великого міста для виявлення найбільш ефективної та адекватної моделі..

***Ключові слова:** водопостачання, водоспоживання, прогноз, прогнозування, часовий ряд, модель, статистичний аналіз, експоненціальне згладжування, нейронна мережа.*

Планирование модернизации системы водоснабжения города на базе прогноза водопотребления

К.А. Горелова, В.М. Задачин

Рассмотрена проблема планирования модернизации системы водоснабжения крупных городов на базе долгосрочного прогноза водопотребления. Сделан обзор существующих моделей долгосрочного прогнозирования с использованием временных рядов.

Сформулированы преимущества и недостатки моделей прогнозирования на основе авторегрессионых моделей ARIMA, нейронных сетей и экспоненциального сглаживания. Осуществлен наглядный пример долгосрочного прогнозирования указанными методами водопотребления большого города для выявления наиболее эффективной и адекватной модели.

***Ключевые слова:** водоснабжение, водопотребление, прогноз, прогнозирование, временной ряд, модель, статистический анализ, экспоненциальное сглаживание, нейронная сеть.*