Application of Nonlinear Autoregressive Neural Network with Exogen for Modeling Demographic Dynamics in Large Urban Agglomeration

DMITRY TARASOV, ALEXANDER SERGEEV, ASHAT SYDIKHOV
Department of Information Technologies and Automation
Ural Federal University
Mira 32 Ekaterinburg
RUSSIA
datarasov@yandex.ru http://www.urfu.ru

Abstract: - This work aims to study the possibility of using artificial neural networks (ANN) for the analysis of time series in special fields. The modeling of demographic dynamics of large urban agglomeration by ANN is conducted on the example of the city of Ekaterinburg, Russia. The prognosis of changes in the population and its sex-age composition, as well as the contribution of the migration factor, are carried out with help of a nonlinear autoregressive exogenous neural network from the Matlab package. The deviation of the forecast results from real data is 3%.

Key-Words: - Artificial neural networks, Time series, Forecasting, Agglomeration, Population, Russia

1 Introduction

The scope of artificial neural networks is extensive and directly related to the modeling of processes in various systems [1]. In particular, neural networks are widely used in forecasting time series. One of the well-known examples of time series is the demographic statistics, which became the subject of this study.

In 2010, Rosstat (Federal Service for State Statistics of Russian Federation) presented the results of Russia's demographic projections to 2030. Calculations of the estimated total population of the country and its age and sex groups were made on the basis of assessment of the resident population size by sex and age by January 1, 2009 [2].

Recently, a various attempts were made to predict the dynamics of the population in large urban agglomerations around the world. For example, one can select works [3] and [4].

The work [3] employed an artificial neural network for population prediction that handles incomplete and inconsistent nature of data usually experienced in the use of mathematical and demographic models. The authors emphasize the robustness of ANNs and their fault tolerance even in the face of missing data has resulted into its use in various fields of human endeavors. The experiment engaged multilayer perceptron (MLP) with the topology 19.9.1 for prediction the population dynamic for several periods of time. The neural network has also been compared with cohort component method of population projection. It is confirmed that Artificial Neural Network performs quit impressible in estimating the rates of population growth.

The work [4] focuses on population dynamics in Yangtze River delta. Authors suggests that population dynamics has major impacts on regional ecosystem and socioeconomic development. They use a back propagation MLP neural network from Matlab package for prediction of copulation dynamics in several cities in China comparing with a spatial autocorrelation model. The work shows that the prediction results using BP neural network are relatively trustworthy, due to the superiority in dealing with complex nonlinear relationships.

The purpose of this work is to build an ANN for forecasting the population size of a large urban agglomeration of the city of Ekaterinburg, Russia.

2 Methods and the experiment

In modeling the demographic dynamics, we took into account the same parameters that were used in Rosstat's forecast: birth rate, mortality, population migration and its age and sex composition [2].

The evaluation of demographic dynamics was carried out using a nonlinear autoregressive exogenous network (NARX) (Figure 1), since this architecture is often used in predicting time series [5]. As a training sample, statistical data provided by Rosstat [6] were used, according to the population, its age and sex composition, population growth due to migration from 1970 to 2010 (Table 1).

Creation, training and application of the neural network was carried out in the Matlab environment.

The training quality of the simulated neural network was tested using correlation coefficients (an example is illustrated in Figure 2).

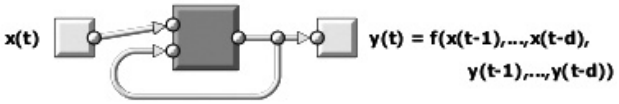

Fig. 1. Scheme of the NARX neural network

To build a more reliable forecast, a comparative analysis was made on correlation coefficients of 4 neural networks, with different number of neurons in the hidden layer: 10, 15, 20 and 25 neurons (correlation coefficients are given in Table 2).

Table 1. Input and output names of the training sample fields

Field name	Type
Date	Input
Amount of the resident population	Output
Number of women	Output
Number of men	Output
Number of births	Output
Number of deaths	Output
Grow (loss) due to migration	Output

Since the difference between the coefficients turned out to be small, an additional check was carried out: on the basis of the training set, the population forecast for 2010-2014 was based on the forecast for the population (forecast results, as well as their comparison with the actual data are given in Table 3).

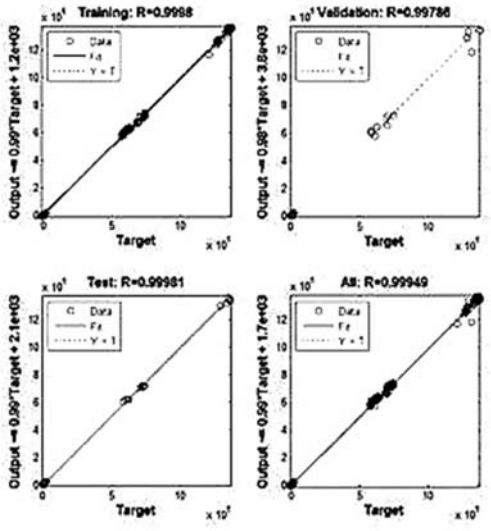

Fig. 2. Networks correlation coefficients

Table 2. Networks frameworks and correlation coefficients (C_{corr})

Neurons in hidden layer	Training C_{corr}	Vali-dation C_{corr}	Test C_{corr}	Total C_{corr}
10	0.99997	0.99980	0.99986	0.99992
15	0.99980	0.99786	0.99981	0.99949
20	0.99981	0.99836	0.99962	0.99943
25	0.99999	0.99855	0.99885	0.99954

The calculation of the average reliability was carried out according to the formula (1) (shown in Table 4), where: p_i- the predicted result, f_i -factual value.

$$d = \frac{\sum_{i\to 5}^{5} \frac{p_i}{f_i}}{5}, \quad (1)$$

For the further modeling, a nonlinear autoregressive exogenous neural network was used. The network has 1 hidden layer with 20 neurons and 1 output layer containing 1 neuron (framework 1.20.1.1). The framework of the ANN is shown in Fig. 3.

Several forecasts were obtained for the population and its age and sex composition as a result of the ANN predictions.

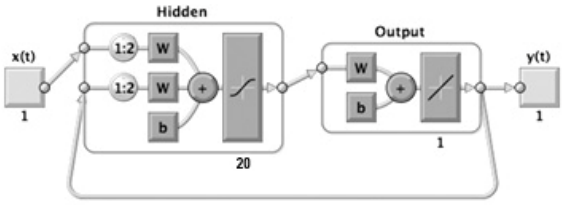

Fig. 3. A framework of the nonlinear autoregressive exogenous neural network (1.20.1.1) used for modeling

3 Results and discussion

To test the quality of network training, forecasts for 2010-2014 years were built. The results of which were compared with Rosstat data.

The average forecast error was 3%, which indicates a good predictive ability of the constructed and trained network. Then the population forecast and sex and age composition (Figures 4, 5) of the population were constructed from 2010 to 2044, as well as the level of fertility, mortality (Figure 6), and migration dynamics (Figure 7).

Table 3. The results of the forecasts by various ANN frameworks

Year	Factual data, people	Number of neurons in the hidden layer: prediction, people			
		10	15	20	25
2010	1349772	1344823	1352817	1356050	1342092
2011	1349800	1343139	1350574	1355139	1342631
2012	1377738	1341636	1348911	1352908	1341900
2013	1396074	1378630	1367804	1381312	1340452
2014	1412246	1386578	1381982	1390887	1357289

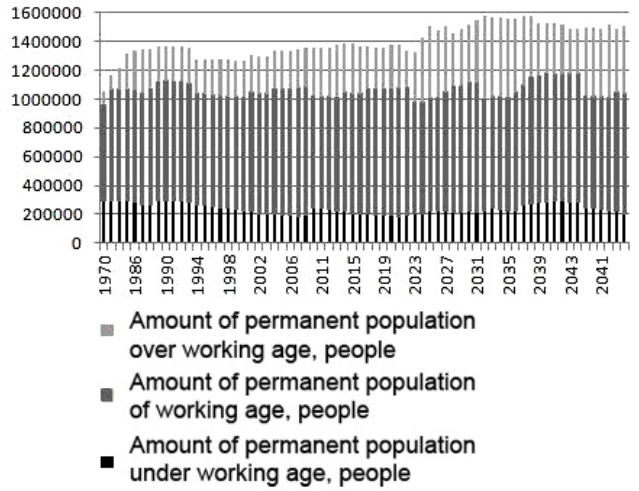

Fig. 4. Forecast of the resident population amount and its age structure

According to the constructed forecasts, it can be concluded that the number of the resident population of the city will increase. When comparing population dynamics and migration (Figures 4, 6), it was found that the level of migration plays an insignificant role in forecasting the population of the urban agglomeration.

Fig. 5. Forecast of the resident population amount and its sexual composition

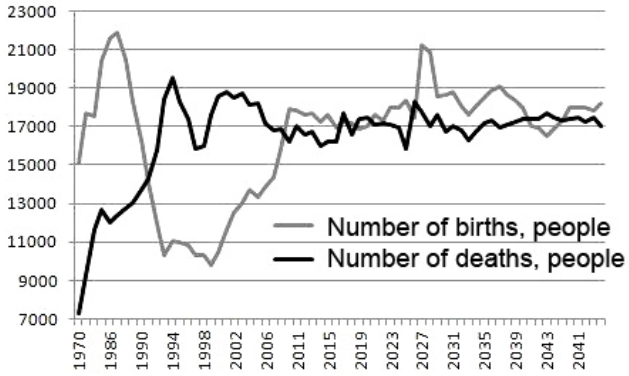

Fig. 6. Forecast of fertility and mortality

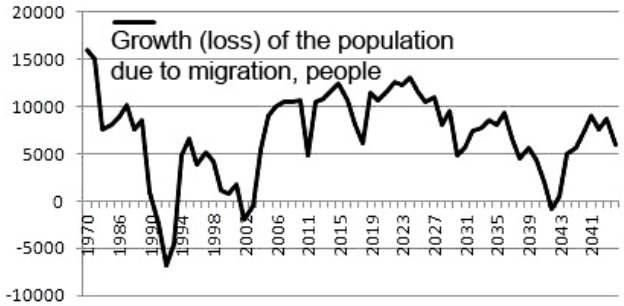

Fig 7. Forecast of migration dynamics

4 Conclusion

The study shows that ANNs can be successfully utilized to predict time series. Such ANNs can be used in decision support systems.

Application of nonlinear autoregressive exogenous model for modeling demographic dynamics in large urban agglomeration shows satisfactory results and broad prospects of application.

The best performance and reliability is achieved in the NARX network with topology 1.20.1.1.

The results obtained resemble ones from the other types of neural networks, for instance, from multilayer perceptron. However, a nonlinear autoregressive exogenous neural network demonstrates a more precise prediction accuracy. Moreover, it is fast in making prognosis and does not require high computational power.

Table 4. Comparison of the average reliability for each ANN framework

Number of neurons in the hidden layer	Average reliability
10	0.9869
15	0.9880
20	0.9929
25	0.9768

References:

[1] Gladyshev AI, Zhukov AO Advantages and disadvantages of simulation modeling using neural networks // Bulletin of the Russian New University. Series: Mathematical modeling in economics and management. 2013. № 4. P. 53-55.(In Russian. Гладышев А. И., Жуков А. О. Достоинства и недостатки имитационного моделирования с использованием нейронных сетей // Вестник Российского нового университета. Серия: Математическое моделирование в экономике и управлении. 2013. № 4. С. 53–55).

[2] Vishnevsky A. G. Population of Russia 2009: the seventeenth annual demographic report // National. Issled. University Higher School of Economics. Ed. home of the Higher School of Economics, 2011. 295-303. (In Russian. Вишневский А. Г. Население России 2009: семнадцатый ежегодный демографический доклад // Нац. исслед. ун-т «Высшая школа экономики». Изд. дом Высшей школы экономики, 2011. 295–303).

[3] O. Folorunso, A. T. Akinwale, O. E. Asiribo and T. A. Adeyemo “Population prediction using artificial neural network”, African Journal of Mathematics and Computer Science Research Vol. 3(8), 155–162.

[4] Chao Wang, Runhe Shi, Wei Wei, Chaoshun Liu, Tao Qi, Wei Gao "Population dynamics in Yangtze River delta: a neural-network and spatial statistical analysis," Proc. SPIE 8513, Remote Sensing and Modeling of Ecosystems for Sustainability IX, 85130.

[5] Learning long-term dependencies is not as difficult with NARX networks / L. Tsungan, B. G. Horne, P. Tino, C. L. Giles. Massachusetts Institute of Technology. 1996. P. 577–582.

[6] Federal State Statistics Service: methodological explanations (In Russian. Федеральная служба государственной статистики: методические пояснения). URL: http://www.gks.ru/freedoc/new_site/population/demo/metod/met-dem.htm (24.04.2014).