Regional Life Quality Management: Methodological Approaches to its Forecasting using Neural Networks

Elena Petrova
Volgograd State University
Institute of management and regional economics
Volgograd, Russia
ea_petrova@mail.ru

Abstract— **This paper proposes theoretical and methodical approaches to forecasting of quality of life in regions using neural networks. Assessment of quality of life main indices is a subject of study of both federal statistical agencies and several expert organizations. However, an adequate measurement of these indicators is confronted with information and methodological problems, which include: the incompatibility of methods of measurements in Soviet Union and modern Russian statistics; incomplete pools of data associated with various negative processes, administrative reforms and institutional changes in the Russian economy; partial loss of historical statistical data. This determines the inability to use time series models of economic indicators to identify the main development patterns, and the resulting trends are limited and not reliable enough. Furthermore, most methods have applicability limits and do not work with incomplete or "noisy" data. Therefore, it is proposed to use neural network technologies and intelligent information systems developed on their basis as alternative approach for solving semistructured and unformalized problems of analyzing the quality of life, as well as constructing forecasting models. A comparative analysis of the researches of Russian and foreign authors on the problems of assessment and constructing a integral indicator of the quality of life of the population is carried out. The author proposes an assessment model, as well as the results of a neural network construction, and an economic interpretation of the obtained results.**

Keywords— regional economy, quality of life, forecasting models, neural networks

I. INTRODUCTION

The problems of regional development management include, first and foremost, quantitative measurements of the main basic indicators of quality of life, as well as analytical methods for assessment of the integral indicator of the quality of life and building of forecasting models on their basis.

Models of indicator systems reflect hidden (not measurable directly) factors that form this category, and can also be viewed as a result of the activity and interaction of authorities of different levels (federal, regional and municipal). The efficiency of public administration is determined by the possibility of a positive impact of relevant management structures on the economic and organizational processes taking place in the region, which should be reflected in the relevant indicators of economic and social development. In

this regard, the development of a composite model of metrics that define the standard of living and quality of life of the population will allow forming scientifically-based economic and social policies of the region.

Assessment of quality of life main indices is a subject of study of both federal statistical agencies and several expert organizations. In this connection it is necessary to notice the statistical yearbook published by Rosstat "The Social Situation and the Quality of Life of the Population" [1], which is the most composite edition that accumulates data obtained from sample observations, as well as data published by the ministries of the Russian Federation. Another analytical study in this field is the annual rating of the magazine "Expert" and the rating agency "RIA rating", that is published since 2012. [2] However, an adequate measurement of these indicators is confronted with information and methodological problems, which include: the incompatibility of methods of measurements in Soviet Union and modern Russian statistics; incomplete pools of data associated with various negative processes, administrative reforms and institutional changes in the Russian economy; partial loss of historical statistical data.

This determines the inability to use time series models of economic indicators to identify the main development patterns, and the resulting trends are limited and not reliable enough.

Similar problems arise in the building of forecasting models, constructed as systems of equations connecting individual characteristics through weight coefficients, or balance sheet identities that determine certain parameters through the system of others.

II. MATERIALS AND METHODS (MODEL)

At present, there are many statistical methods and models for analyzing indicators of the quality of life and constructing a integral indicator. However, most methods have applicability limits and do not work with incomplete or "noisy" data. Therefore, it is proposed to use neural network technologies and intelligent information systems developed on their basis as alternative approach for solving semistructured and unformalized problems of analyzing the quality of life, as well as building forecasting models.

The description of quality of life and standard of living requires consideration of many interrelated factors that are added to the research system in a nonlinear way. To solve this problem, it is proposed to use neural network technologies. Neural networks are used in the case when relationship between the input and output data is a priori unknown. In such situations, both traditional mathematical and expert systems fail. It is the ability of artificial neural networks for generalization and data mining within network elements that allows them to successfully cope with the tasks of forecasting and the classification of objects in economic analysis.

As a rule, measurements of the standard of living and the quality of life of the population are conducted in the direction of assessing the income of population groups, demand and consumption that have developed at the regional level. Several studies are dedicated to the research of savings, housing, social characteristics of the population, the dynamics and differentiation of these indicators in the context of population groups and regional economies. [3].

In these studies, systems of absolute and relative indices are constructed to characterize the degree of provision and satisfaction of the needs of different categories of the population. [4]

In many models of the standard of living, authors apply methods where they identify the concepts of "the cost of living" and "the quality of life". Classical economics determines the cost of living as the value of a set of commodities which is determined, first of all, by the level of prices for food and services, as well as by changes in types of income, consumption patterns, tax legislation, and demographic indicators of population groups. The main calculated rates are consumer basket and commodity bundle, as well as living wage. [5] The quality of life is another category that requires taking in consideration not only cost indicators but also other factors.

Quality of life, unlike the category of "standard of living", is a more complex compound category that reflects all the conditions of life and its development that are essential for a person. It should not go without mentioning the research of Aivazian S., who showed that the integral indicator (integral indicator / integrated index) that define the quality of life should be constructed on the basis of 5 groups of indicators - the quality of the population, the welfare of the population, the quality of the social sector, the quality of the environment, the natural and climatic conditions. [6]. However, there remain problems of the selection of particular criteria for each block and the problems of folding the indicators into an integrated one. The construction of the forecasting models based on the proposed system of indicators, faces the problems that have already been mentioned above.

Over time, the term "standard of living of the population" replaced in Russian scientific literature such concepts as "people's welfare", "degree of satisfaction of material and nonmaterial needs", and was used more often than the broader term "quality of life". This is due to, first of all, the possibility of using quantitative statistically monitored indicators, as well as its wider spreading in foreign studies. [7]

The main methodological approach in the study of poverty and social stratification are various kinds of sociological surveys and observations. However, the micro-level of these studies does not allow to apply their results to the regional economy and to develop on their basis a suitable socio-economic policy.

Among the foreign practices of assessment of the standard of living (or the standard of well-being), the most famous and widely recognized index is the Human Development Index (HDI). The index was created in 1990 by a group of economists led by the Pakistani economist Mahbub ul Haq. However, the conceptual structure of the index was developed on the basis of the work of Amartya Sen. The index is published since 1990 by the United Nations in the annual Human Development Report.

The HDI measures the country's achievements in terms of state of health, education and actual income of its citizens, in three main areas for which the indices are valued:

- Life expectancy index: health and longevity, measurable indicators of the average life expectancy at birth.

- Education index: access to education, measured by the average expected length of schooling for school-age children and the average length of adult education.

- The index of gross national income: a adequate standard of living, measured by the gross national income (GNI) per capita in US dollars at purchasing power parity (PPP). "The Human Development Report 2016. Human Development for Everyone." is available at the UN website http://hdr.undp.org/.

In 2010, the family of indicators that measure the HDI was expanded; three new indicators were introduced in addition to the indicators used before: The Inequality-adjusted Human Development Index (IHDI), the Gender Inequality Index (GII) and the Multidimensional Poverty Index (MPI).

Different approaches to measuring well-being have been undertaken in different countries at different levels. In some countries, was built a unified index of well-being, in others - a preference was given to the "dashboard" (monitor panel). Each of these approaches has its strengths and weaknesses. [8]

As for the possibilities of predicting the standard of living, there is also no unequivocal scientific approach to these issues.

The most promising application of neural networks is the analysis of numerical data in a certain limited range. This creates problems in cases where the data has a non-standard scale, in case of missing values and non-numeric data. [9] Many real tasks have to deal with not completely reliable data; the values of individual variables can be distorted by noise or partially absent. Neural networks are generally resistant to noisy data, which allows applying them in prediction tasks. [10]

Among all the properties of artificial neural networks, the most interesting one is their ability to learn. The learning process identifies the statistical properties of the learning set and classifies similar vectors into classes. The input of a vector from this class will give a certain output vector, but it is

impossible to predict before learning which output will be produced by this class of input vectors. Consequently, the outputs of such a network must be transformed into some understandable form, conditioned by the learning process, which, as a rule, is not a difficult task. [11]

Currently, foreign academic journals publish results of applied research in the field of economics and sociology, based on the use of neural network technologies and machine learning, including methodological and practical aspects of forecasting economic growth [12].

III. RESULTS AND DISCUSSION

Building a neural network requires careful selection of input data that affects the expected result. It is necessary to exclude from the input data, all information that is not related to the problem under study in order to avoid problems with the network setup. Therefore, at the first stage, a cognitive analysis of the problems and factors of the social and economic development of the region was carried out.

Thus, when constructing forecasting models, it is proposed:

- To rely on the methodological approach proposed by Aivazian S (in the construction of indicator systems);

- To use neural networks as a tool to build such a model.

The neural network implies the formation of two indicator sets that represent input and output parameters. Indicators that describe the standard of living are used as output parameters. The indicators that characterize the factors of internal and external environment of regional development, affecting the standard of living, are used as input parameters.

Attempts to build a single integral indicator (integral indicator /integrated index), accumulating all the listed components of the standard of living have not given positive results. Difficulties arose already at the stage of choosing the convolution scheme - additive or multiplicative. In some regions, negative results were obtained, which did not allow making an adequate economic interpretation of them. Furthermore the construction of such indicators reduces their information value and leads to the information loss.

Therefore, at this stage of research, it was proposed to use the Human Development Index (HDI) calculated and published, since recently, for the regions of the Russian Federation as the output indicator.

The results of cognitive analysis were used to describe the impact factors. However, when creating the data set, the author encountered difficulties in collecting primary data. Some of them were practically absent in Rosstat's statistical compilations, some of them used different methods of collection and assessment. Therefore, as a result, the second group included 7 indicators, which allowed building a neural network. Among these are: the incidence rate per 1000 population; emission of pollutants into the atmosphere; the number of recorded crimes per 100,000 population; total fertility rate; number of professional educational organizations; number of higher education institutions; households that have access to the Internet.

255 observations were collected over the period of 3 years for 85 regions of the Russian Federation to train the network, later data for some regions were excluded (Republic of Crimea, Sevastopol, Nenets Autonomous District). The network was trained on data collected between 2010 and 2015. Verification of the predictive capabilities of the network was carried out on the data of 2016 to compare the real values and the obtained theoretical data (network forecast).

Analytics platform Deductor was used for the analysis. As a result of building a neural network, the Deductor Studio program generates a mapping of the neural network in the form of a graph represented on the slide. The graph has vertices that correspond to the network modules, as well as edges that describe the connections between the modules. Under the graph is placed a measuring scale corresponding to the strength of the connection between the modules.

The scatter plot (see Figure 1) shows that the predicted values of five out of 255 observations exceed the error of 0.05, i.e. 2% of observations are not included in the given limitations.

Particular attention should be paid to the "what if" tool, because it allows to apply the generated model to data for 2014. To do this, we collected data for 2014 for 7 variables.

For example, the data for the Volgograd Region for 2014 were entered. At the output, the model calculated a value equal to 0.829. The actual Y value for the region for 2014 is 0.859, i.e. the predicted value differs from the actual value by 3.5%. For the Moscow region, the forecasted value was 0.935, and the actual value was 0.946, the difference was 1.2%
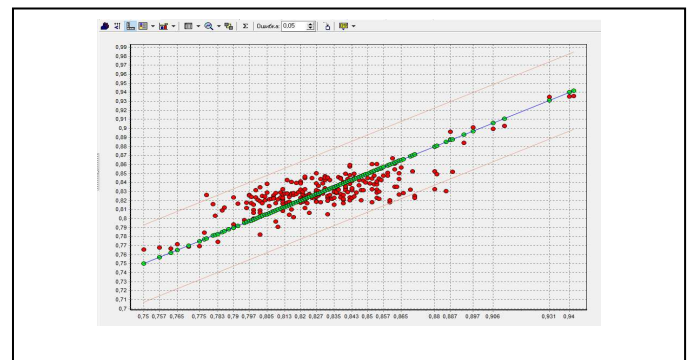

Fig. 1. The scatter plot with a given error value of 0.05.

Based on the results of the data evaluation, the average error value of the predicted data was 2.77%, the maximum error value was 6.93%, and the minimum error was 0.07%. 90% of the observations have an error of less than 5%. Based on these data, it can be concluded that the network has been fully trained and the model has been successfully built.

Furthermore, the "what if" tool has graphs showing dependence of the indicators on the output parameter under otherwise equal conditions. The diagrams are mainly used to select the value of one of the input fields to obtain the desired output. In this study, diagrams can show the type of dependence of input variables on the output parameter.

The graph in Fig. 2 shows that in regions with a higher standard of living, there are more pollutants are being emitted into the atmosphere, which is due to the presence in industrialized regions of manufacturing enterprises. In this case, the graph is not linear. However, it is interesting that the form of such dependence is somewhat different than in the previous case.
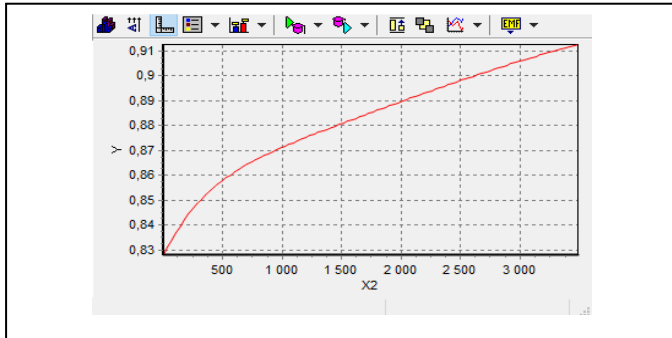

Fig.2 dependency graph that plots Y (HDI) versus X2 (emissions of pollutants into the atmosphere from stationary sources)

Fig. 3 shows the inverse proportion dependence of the standard of living and the indicator "The number of recorded crimes per 100,000 population". This dependence is coherent; there are fewer crimes in regions with a higher value of the standard of living.
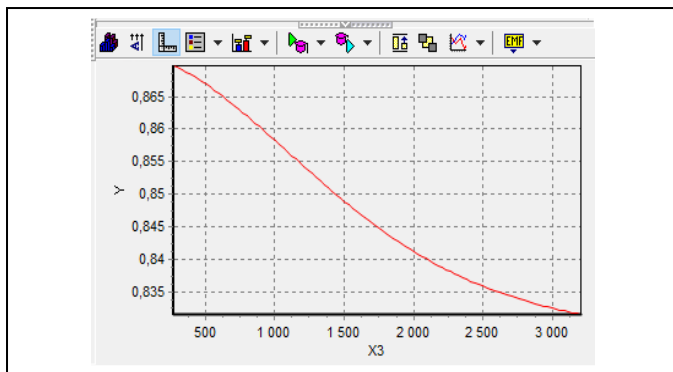

Fig. 3 dependency graph that plots Y (HDI) versus X3 (The number of recorded crimes per 100,000 population)

Interesting dependencies are obtained in terms of the number of households with access to the Internet.

At the Fig.4 the dependence of the standard of living on the variable looks like a parabola, up to a certain point (in this example up to 60), the dependence is direct at first, and then inverse. Thus, it can be concluded that in regions with a high standard of living, the number of households with access to the Internet, after reaching a certain level (in this case - 60), no longer affects the quality of life in the future.
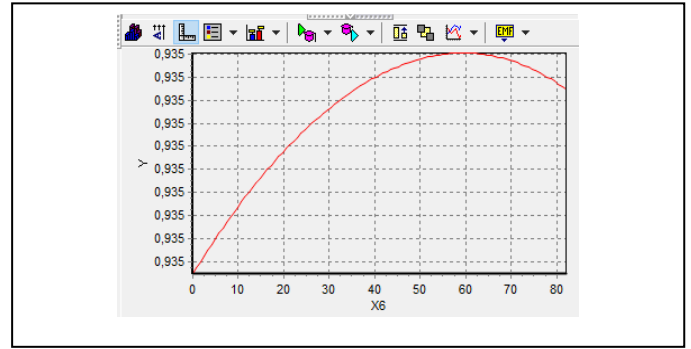

Fig.4 dependency graph that plots Y (HDI) versus X4 (the ratio of households that have access to the Internet)

IV. CONCLUSION

Thus, the obtained neural network allows not only to obtain predicted values, but also to research on the obtained dependences between the output parameter and the parameters influencing the level of the output parameter.

However, the presented results require further research and raise new questions. For example, the choice of one output parameter - HDI as a parameter describing the standard of living of the population is not entirely justified.

Author proposes in the further research to extend beyond only one output parameter, and to build integral indicator for the proposed 5 blocks.

The model requires refinement and a larger set of input parameters. It is proposed to adjust the systems of indicators is to be carried out using the canonical correlation approach, that allows to estimate the strength of relationships between groups of indicators.

However, even in a limited form, the obtained model showed quite good predictive properties.

References

[1] Social status and standard of living of the population of Russia. 2017: Statistical book. [Sotsial'noe polozhenie i uroven' zhizni naseleniia Rossii . 2017: Stat.sb] Moscow: Rosstat., 2017, 332 p., URL: http://www.gks.ru/wps/wcm/connect/rosstat_main/rosstat/ru/statistics/publications/catalog/doc_1138698314188

[2] Rating of regions by living standards of families - 2017. [Reiting regionov po urovniu zhizni semei – 2017], URL: http://riarating.ru/regions/20170515/630062559.html

[3] Suvorov A.V. Methodological problems of forecasting the standard of living of the population [Suvorov A.V. Metodologicheskie problemy prognozirovaniia urovnia zhizni naseleniia]. Problems of forecasting [Problemy prognozirovaniia], 2000, vol 1, pp. 22-38

[4] Eliseeva I.I. Social Statistics [Eliseeva I. I. Sotsial'naia statistika]. Mocow: Finance and Statistics, 2012, 558 p.

[5] Gorelov N.A. Policy of incomes and quality of life of the population.[Gorelov, N. A. Politika dokhodov i kachestvo zhizni naseleniia]. St. Petersburg.: Piter, 2013, 352 p.

[6] Ayvazyan S.A. Comparative analysis of integral characteristics of the quality of life of the population of the subjects of the Russian Federation [Aivazian S. A. Sravnitel'nyi analiz integral'nykh kharakteristik kachestva zhizni naseleniia sub"ektov Rossiiskoi Federatsii]. Moscow: Central Economics and Mathematics Institute of the Russian Academy of Sciences, 2001, p.7

[7] Shakin V.V. Global Environmental Assessment, Sustainable Development Criteria and Mathematical Modeling of Living Systems under Extreme Conditions. Encyclopedia of Life Support Systems Publisher, Ltd. – Proceedings of the International Conference on Environmental Indices/System Analysis Approach (INDEX-97). London, U.K.,1998, pp. 202–220; Stiglitz J.E., Sen A., Fitoussi J.-P. The Measurement of Economic Performance and Social Progress Revisited // Documents de Travail de l'OFCE / Observatorie Français des Conjonctures Economiques (OFCE). 2009. Dec. No. 2009-33; Dolan, P., Layard, R., Metcalfe, R. (2011). Measuring Subjective Wellbeing for Public Policy: Recommendations on Measures. London:CEP Special Reports. Centre for Economic Performance. London School of Economics; Eger R.J., Maridal J.H. (2015). A statistical meta–analysis of the wellbeing literature // International Journal of Wellbeing. 5(2). P. 45–74. и др.

[8] Kroll C. (2011). Measuring Progress and Well-Being Achievements and Challenges of a New Global Movement. London: Friedrich–Ebert–Stiftung

[9] Kachanovskiy Yu. P. Preprocessing of data for learning neural network [Kachanovskii Iu. P. Predobrabotka dannykh dlia obucheniia neironnoi seti]. Basic researc [Fundamental'nye issledovaniia], 2012, vol. 12, p. 34

[10] llarionov M.G. Prediction based on the apparatus of neural networks [Illarionov M. G. Prognozirovanie na osnove apparata neironnykh setei]. Bulletin of Kazan Technological University [Vestnik Kazanskogo tekhnologicheskogo universiteta], 2012, vol. 1, p. 15

[11] Dmitrienko V.D. Neumann networks of Hamming and Hebb, able to learn [Dmitrienko V. D. Neironnye seti Khemminga i Khebba, sposobnye doobuchat'sia] Bulletin of the National Technical University Kharkov Polytechnic Institute. Series: Informatics and modeling [Vestnik Natsional'nogo tekhnicheskogo universiteta Khar'kovskii politekhnicheskii institut. Seriia: Informatika i modelirovanie]. 2013, vol. 19, p. 57

[12] Bang, James & Sen, Tinni & Basuchoudhary, Atin. (2015). New Tools for Predicting Economic Growth Using Machine Learning: A Guide for Theory and Policy), распространения бедности (Jean, Neal & Burke, Marshall & Xie, Michael & Matthew Davis, W & B. Lobell, David & Ermon, Stefano. (2016). Combining satellite imagery and machine learning to predict poverty. Science. 353. 790-794. 10.1126/science.aaf7894) и "самочувствия" населения (Agarwal, Ankur & Baechle, Christopher & Behara, Ravi Multi-method approach to wellness predictive modeling (2016). Journal of Big Data. https://doi.org/10.1186/s40537-016-0049-0