The reversed problem of migration streams modeling

D A Tarkhov¹, I K Shanshin¹ and D O Shakhanov¹
¹Peter The Great Saint-Petersburg Polytechnic University, Saint-Petersburg, Russian Federation, Politechnicheskaya st., 29.

Contact person: I K Shanshin, ivan.fizik92@yandex.ru

Abstract. This article discusses methods of solving the reversed problems of migration streams modeling: the analytical solution, the formula for the two-dimensional linear regression and neural networks. The cases with different amount of input data and number of neurons: 2, 10, 100, 1000 input points for linear regression method, and 5, 15, 25 neurons for network are considered. Also we have made numerical calculations for a period of time, which were increased by 10 times. The article proves the application of neural networks for constructing the models under consideration, especially for a non-linear model.

1. Introduction

Designing the complex development program of some region, for example Russia’s Arctic zone, requires consideration of a wide range of problems – economic, environmental, and social ones [1], [2], [3]. Large mining projects should be not only economically viable, but also it is necessary to take into account their impact on the environment, population, etc. Rapid changes of demographic patterns intensify the imbalance in the population of the regions which are not inhabited densely enough. For a reasonable forecast of relevant processes, it is necessary to develop an adequate medium-term model. Problems in development of such a model are caused by an insufficient amount of demographic data. For solving the problem, application of neural network is proposed.

Models of migration streams, represented as non-linear dynamic system,

\[
\begin{align*}
\frac{dx}{dt} &= sh(kx + k_1y) - x \cdot ch(kx + k_1y) \\
\frac{dy}{dt} &= sh(k_2x + ky) - y \cdot ch(k_2x + ky)
\end{align*}
\]

and linearized system at a point of zero equilibrium position:

\[
\begin{align*}
\frac{dx}{dt} &= (k - 1)x + k_1y \\
\frac{dy}{dt} &= k_2x + (k - 1)y
\end{align*}
\]
are outlined in [4]. A case of two populations interaction in two regions is considered. Here \( k_1 \) shows intensity of the first population tendency to coexist in the same region with the second population, \( k_2 \) shows intensity of the second population tendency to coexist in the same region with the first population, \( k \) shows intensity of population tendency to coexist with the people of their own population group. Variables \( x \) and \( y \) are determined as follows: 
\[
    x = \frac{(n_1^H - n_2^H)}{N}, \quad y = \frac{(n_1^V - n_2^V)}{N}
\]
where \( n_1^H \) and \( n_2^H \) - the number of people in the first population in the first and second regions respectively, \( n_1^V \) and \( n_2^V \) are the same for the second population. \( N \) – is the number of people in each population group, i.e. 
\[
    N = n_1^H + n_2^H = n_1^V + n_2^V.
\]

To compare the methods described below by constructing the migration dynamics according to equations (1)-(2) it is necessary to determine the coefficients of the given model equations in accordance with the available data. 

Here we analyze three methods of solving the given problem:

1) The first method assumes analytical solution to equation (2) with subsequent identification of coefficients by the best correspondence of the result to the data. This method has a number of shortcomings: firstly, it is impossible to apply the method to equation (1) and, secondly, in the case of models of a higher order, it is necessary to get a solution anew.

2) The second method is a discretization of equation (2) and subsequent identification of coefficients by formulas of two-dimensional linear regression. Using this method allows us to solve the problem in the case of models of higher order, but it is still difficult to apply it to equation (1).

3) The third method implies application of a neural network. It can be trained on the basis of numerical solution to the equation system (1 or 2) for quite a big set of parameters. Thus, the calculation of the required coefficients will be reduced to substitution of the data in the neural network. This method has many advantages including its universality, which makes the solution also appropriate for a non-linear system.

2. Results of calculations

2.1. Linear regression

![Figure 1. Linear regression, 2-10-100 points.](image)
When the number of points is 2, we can see that the main trend of migration is the same as in the model, but we have big divergence between graphs of theoretical and calculated models (figure 1). Having increased the number of points to 10, we can see that error reduction is significant, but still it is not big enough. By increasing the number of the input points to 100 we achieve excellent accuracy: graphs of theoretical and calculated models almost coincide. However, tenfold augmentation of the forecast period, according to which the model is constructed, leads to significant growth of error figure 2. In order to reach good approximation, we need to increase the number of points to 1000. Nevertheless, in practice collecting such an amount of input data is rarely possible.

2.2. Neural networks

First, we take a number of neurons equal to five for equation (2). A model which was plotted with this neural network by three points of an interval (0, 0, 5 and 1 moments of time) significantly differs
from a theoretical model in the migration rate, but basic tendency is preserved (figure 3). Working with

equation (2), we increase the number of neurons to twenty-five. The resulting model perfectly

approximates the theoretical one; this result is comparable to the linear regression, which was plotted

by 100 points. In case of equation (1), we get the same results. We can see that at the origin of

coordinates these models are identical. So, let’s carry out the same calculations, having multiplied the

forecast period by 10.

![Figure 4. 15-25 neurons, equation (1), forecast period multiplied by 10.](image)

Considering the work of a neural network that consists of fifteen neurons, in the case of equation (1),

we increase the forecast period tenfold. Without changing the coefficients, we can see that the accuracy

of the forecast decreases, but the main tendency of migration is preserved (figure 4). Using neural

networks that consist of twenty-five elements, in case of equation (1), we get a truly amazing

approximation in all parts of the function, even if we increase the forecast period tenfold.

![Figure 5. 25 neurons, equation (2), forecast period multiplied by 10.](image)
We analyze equation (2), using 25 neurons (figure 5). We can see that the result is better than a linear regression plotted by one hundred points, but worse than the one plotted by a thousand points, yet the divergence between the obtained result and the theoretical model is still quite big.

3. Conclusions
If the number of neurons in a neural network is greater than or equal to 25, it is possible to restore the solutions to equations (1) and (2) by data on three points in contrast to a linear regression, which needs a large amount of input data to reach the same accuracy. Neural networks can work with both linear and non-linear models; the accuracy of approximation depends more not on the volume of input data, but on the number of neurons in a neural network. In addition, neural networks have many other appreciable advantages over other methods of forecasting, such as its learning capability and universality. Many practical problems were solved or are solved by applying neural networks. This study has confirmed the universality and superiority of this method.

Our approach may be used for identification of dynamic systems’ models (e.g. socio-economic, mechanical, physical, etc.) based on measurements or observations, when a model type is known and it is necessary to determine its parameters. System dimension, equations order, its linearity or non-linearity etc. do not have a great impact on the determination procedure. For example, this technique is used in [5], [6], [7].

Acknowledgments
The paper is based on the research carried out with the financial support of the grant of the Russian Science Foundation (Project No. 14-38-00009, the program-targeted management of the Russian Arctic zone development). Peter the Great St. Petersburg Polytechnic University.

References

[1] Rudenko D, Didenko N and Skripnuk D 2015 Environmental security issues in the Russian Arctic. 15th International Multidisciplinary scientific Geconf. (SGEM 2015). Ecology, Economics, Education and legislation. Conference Proc., vol 3 (Bulgaria, 18-24 June 2015) pp 267-74

[2] Didenko N, Kunze K and Skripnuk D 2015 Russian Export Strategy and Social Sector: Consequences of Resource-Oriented Exports on Population of Russia J. Mediterr. of Soc. Sci. 6(5) 473-81

[3] Didenko N and Skripnuk D 2014 The impact of energy resources on social development in Russia J. WIT Transact. on Eco. and the Environ. 190 151-9

[4] Weidlich W 2000 Sociodynamics - A Systematic Approach to Mathematical Modelling in the Social Sciences (Amsterdam: Harwood Academic) pp 107-168

[5] Vasilyev A and Tarkhov D 2014 Mathematical Models of Complex Systems on the Basis of Artificial Neural Networks J. Nonlin. Phen. in Compl. Syst. 17 327-35

[6] Kainov N, Tarkhov D and Shemyakina T 2014 Application of neural network modeling to identification and prediction in ecology data analysis for metallurgy and welding industry J. J. Nonlin. Phen. in Compl. Syst. 17 57-63
[7] Lazovskaya T and Tarkhov D 2015 Fresh approaches to the construction of parameterized neural network solutions of a stiff differential equation *St. Peter. Polytech. Univ. J. Phys. and Math.* 1 192-8