Adjustment of the Model of the Agent-Determinant Type in the Forecasting of Pollution on the Section of the City Road

Mykhailo Susla[1], Roman Pasichnyk[2], Natalia Pasichnyk[3], Andriy Melnyk[4]

Department of Computer Science, Ternopil National Economic University, UKRAINE, Ternopil, 8 Cheлhova str., email: suslamisha91@gmail.com[1], roman.pasichnyk@gmail.com[2], natalia.pasichnyk@gmail.com[3], melnyk.andriy@gmail.com[4].

Abstract: **The problem of constructing a dynamic model of pollution is considered. The model of pollution in the form of a set of differential equations is proposed, which is identified by means of difference expression with subsequent refinement using gradient methods. Numerical experiments allow you to choose the model that best approximates the experimental data.**

Keywords: **air pollution; mathematical model; set of ordinary differential equations; difference expressions; Levenberg–Marquardt method.**

I. Introduction

The rapid growth of the number of vehicles brings out the problem of control over air pollution with motor vehicles. The latter affects the level of respiratory and other diseases for people living in the areas of high traffic density. Therefore, there is a need for sufficiently accurate monitoring of information on the level of air pollution to make managerial decisions regarding the configuration of residential quarters, design and reconstruction of roads.

The distribution of pollution in the city is characterized by significant spatial heterogeneity, since emissions to the atmosphere are carried out from the network of roads, and the pollution level declines rapidly as we move away from the pollution source [1, 2]. These features take into account the Land use regression (LUR) method, which combines measurements of air pollution in a relatively small number of locations characterized by qualitatively different types of pollution, and the construction of statistical models based on measurements taking into account the features of the points of observation.

Numerous researchers use LUR to estimate the concentration of contaminants in a number of cities in Canada, the United States and Europe [3]. However, this method allows you to build only stationary models. At the same time, it is necessary to build dynamic models for deeper understanding of the pollution effects using apparatus differential equations.

Such a model, which takes into account the influence of several key factors, should be as simple as possible and at the same time sufficiently precise. If we allow interaction of factors with each other, in the simplest way it is modeled using the product of the corresponding indicators.

The generalization of this approach is a model-type agent-determinant that reflects the evolution of the agent under the influence of the determinant so that a significant concentration of the determinant does not compensate for the low concentration of the agent. A representative of models of this type is the Monod model. In a number of simulated situations, such a generalized model is extremely effective [8]. This work is devoted to the study of the possibility of using models of the specified type in the simulation of pollution concentration.

II. Model of Pollution on the Local Section of an Urban Road

In order to construct a model of the dynamics of pollution in an area where it is potentially high, we have to identify the main variables that affect it. No significant influence of humidity and temperature on the dynamics of pollution levels X has been found after the analysis of the measurements obtained with the help of special sensors. Instead, a significant effect of the traffic intensity R and wind speed V has been established.

The apparatus of differential equations is chosen to simulate the dynamics of pollution, since it's much more flexible than regression relations. Numerical differentiation is very sensitive to random perturbations in the measurement results, so it is subjected to multiple smoothing by the method of moving average.

The criterion of multiplicity of smoothing served The minimization of the correlation between the remnants of measurements after their elimination from the smoothed values served as a criterion of multiplicity of smoothing.

When constructing the differential equation of the dynamics of pollution it is taken into account that the growth of pollution is associated with an increase in the intensity of pollutants, ie, the movement of vehicles. Contamination reduce occurs as a result of their dispersal, which is associated with the speed of the wind.

However, the speed of diffusion of contaminants depends not only on the mentioned factor but also on the product of it and of the concentration of contaminants themselves. Since pollution itself decomposes over time, the contaminants concentration decreases in proportion to the pollution itself. The above statements are established on the basis of data analysis and confirmed during the preceding procedures of parametric identification. As a result, we come to the following differential equation of the dynamics of pollution on the road section

$$\frac{dX(t)}{dt} = p_1 R(t) - (p_2 + p_3 V(t)) X(t), \qquad (1)$$

$$X(t_0) = X_0 . \qquad (2)$$

where X is pollution concentration; R is traffic intensity; V is wind speed; $p_1,\ldots,p_4$ are model parameters.

In an empirically constructed model, the interaction of contaminants is described by their product with a constant relative intensity of interaction. In some cases, the actual intensity of the interaction may vary with the change in the determinant characteristic, in this case, the wind speed. Often there is a variable intensity of interaction, which is lower at small values of the determinant and is obtained at the maximum value with saturation at large values of the determinant. This intensity is fed by a multiplicand of the Monod type

$$\frac{V(t)}{p_4 + V(t)} \qquad (3)$$

With its application, the model equation (1) takes the form

$$\frac{dX(t)}{dt} = p_1 R(t) - p_2 X(t) - p_2 V(t) X(t) \frac{V(t)}{p_4 + V(t)} \qquad (4)$$

As a result, we obtain a more complex differential equation containing an additional parameter, which is included nonlinearly. To compare the effectiveness of the proposed models, they need to be identified.

III. The Identification Methods of Pollution Model

It is necessary to establish a method of parametric identification of the pollution model after we have built it. Usually, identification is carried out by minimizing the appropriate quality functional. One of the simplest quality functional is the square root, which is used in the least squares method.

Since the differential equation is nonlinear, its quality functional has a large number of local extrema. The general approach to building a global extremum of this kind is the use of methods of random search, the method of the directing cone of Rastrigin, in particular. However, this method requires great amount of computing resources and the development of special procedures for locating local extremes to find a global one.

At the same time, taking into account the peculiarities of certain classes of tasks, it is possible to set up search domains containing a single global extremum. In particular, a whole class of methods of this kind is proposed for the identification of models of systems with limiting factors.

In these methods, the initial approximation of the values of the models' parameters is based on the difference ratios and scanning the values of one of the key parameters on the grid. The parameters of this grid are also pre-evaluated. The initial approximation is further specified by the gradient method.

These methods have shown their high efficiency and therefore the corresponding method of identification of systems of nonlinear differential equations modeling the dynamics of processes with limiting factors is chosen as the basis for the method of model (1) - (2) identification [4].

Since this differential equation (1) is much simpler, the identification method itself is also simplified. At the initial stage, we construct a system of linear equations with respect to the parameters basing on the difference relations:

$$\frac{X_{i+1} - X_{i-1}}{t_{i+1} - t_{i-1}} = p_1 R(t_i) - (p_2 + p_3 V(t_i)) X(t_i) \quad i = i_1, i_2, i_3 . \qquad (5)$$

Where X_i, R_i, V_i are the values of corresponding functions at the moment of time t_i

The ratio for constructing the initial approximations of the coefficients of the differential equation must reflect the most significant features of the resulting function of the process, that is, in this case, the dynamics of pollution. The numbers of the identification points are chosen in case of the maximum absolute values of the derivatives of the pollution concentration function. Further, the initial values of the parameters of the model are specified by the method of least squares.

The procedure for identifying a differential equation (4) is somewhat more complicated. It includes a method for checking the values of a nonlinear parameter p_4 on a grid, whose parameters are selected experimentally. After selecting a specific parameter value p_4 for choosing the initial values of other parameters, an analogue of the system of linear equations (5) is used.:

$$\frac{X_{i+1} - X_{i-1}}{t_{i+1} - t_{i-1}} = p_1 R(t_i) - p_2 X(t_i) -$$

$$- p_3 V(t_i) X(t_i) \frac{V(t_i)}{p_4 + V(t_i)} i = i_1, i_2, i_3 \qquad (6)$$

IV. Numeric Experiments

Let us demonstrate the possibilities of the proposed methodology on the example of modeling the daily dynamics of pollution on one of the streets of the city of Ternopil. Discrete observations are interpolated using piecewise Hermite interpolation. In particular, the dynamics of wind speed during the day of observation is given on Figure 1.

The following figures show the smoothed results of observations of pollution and traffic. As you can see, the dynamics of the concentration of pollution is quite complicated. To verify the reality of the identification of the proposed model, we analyze the dynamics of the left and elements of the right-hand side of the differential equation (1), given in Figure 4.

It is worth noting that the components of the left parts of the differential equation are brought to comparable values with the pollution derivative by multiplication on corresponding scaling multipliers keeping "+" or "-" sign, how they are included in the equation. The comparison of these functions reveals some similarity in their behavior, as well as the complexity of the task of bringing their sum to zero with the help of just three constant coefficients.

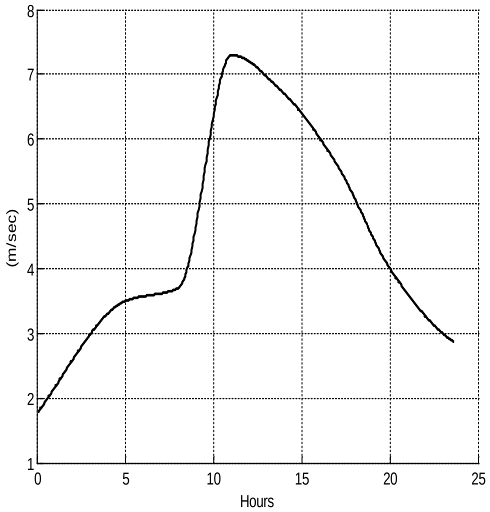

Fig 1. Observation of wind speed during the day

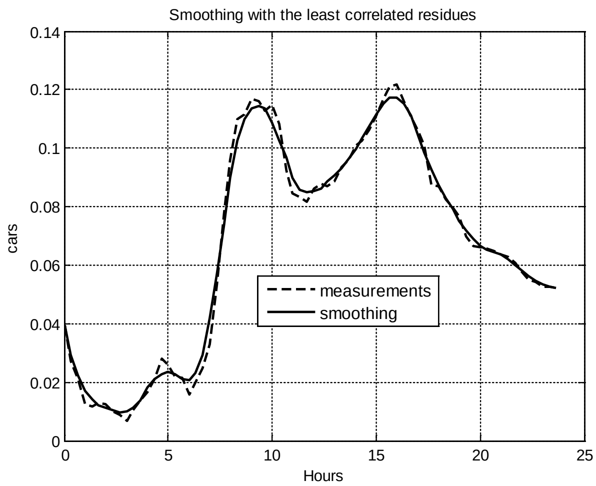

Fig 2. Observation of pollution during the day

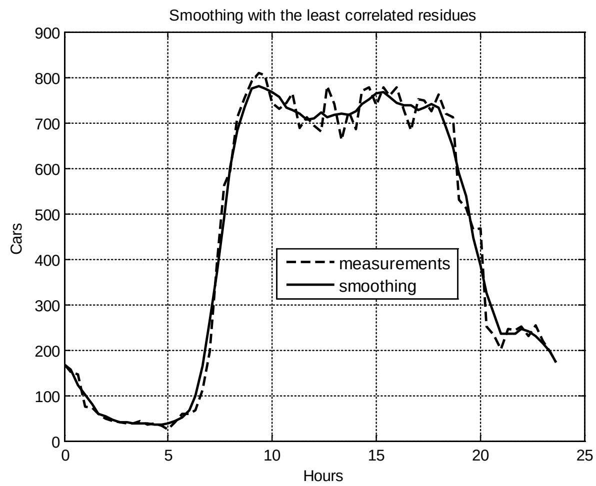

Fig 3. Observation of traffic during the day

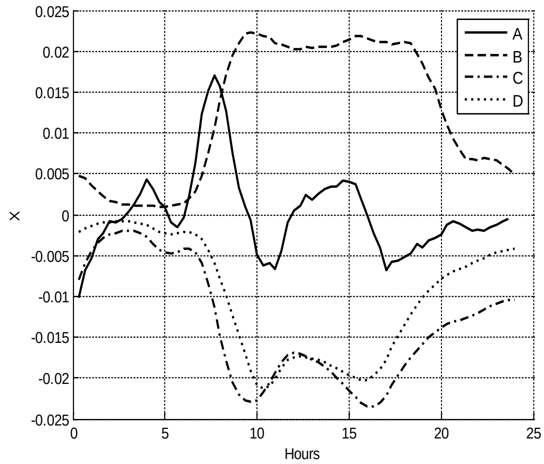

Fig 4. The comparison of the left hand and elements of the right-hand parts of the equation (1), where a) pollution derivative; b) traffic; c) pollution; d) wind speed and pollution product.

By the criterion of the maximum of the derivative module the points 7.3, 10.7, 14.3 have been selected among the internal points of the time interval. As a result of the solution of the system of linear equations (3) the following values of the parameters of the differential equation (1) are obtained: $p_1 = 7.0958e-4,\ p_2 = 3.0172,\ p_3 = 0.3175$.

After optimization of the initial approximation of coefficients using the gradient method of Levenberg-Marquardt, only the first coefficient was somewhat specified to the value $p_1 = 7.0942e-4,$ which allowed to somewhat decrease the average identification error.

The identification results for the equation (1)-(2) are presented on the figure 5, average identification error is 8.6%. Details of the distribution of errors in points of observation can be found using Figure 6.

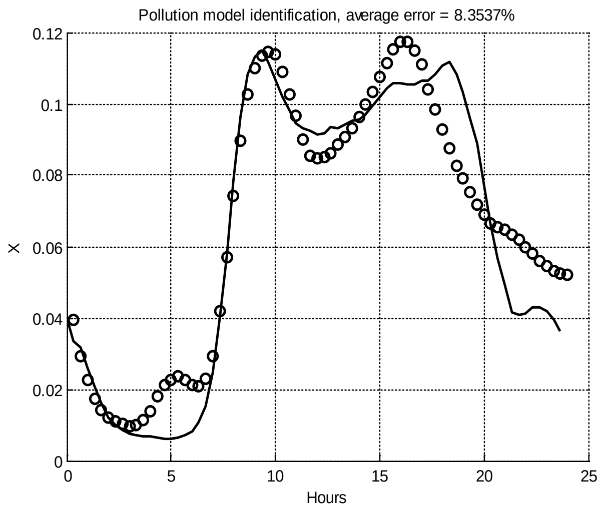

Fig 5. Approximation of observed pollution using identified model (1) -(2)

The question arises whether we can significantly improve the accuracy of identification by using a more complex model (4) - (1)?

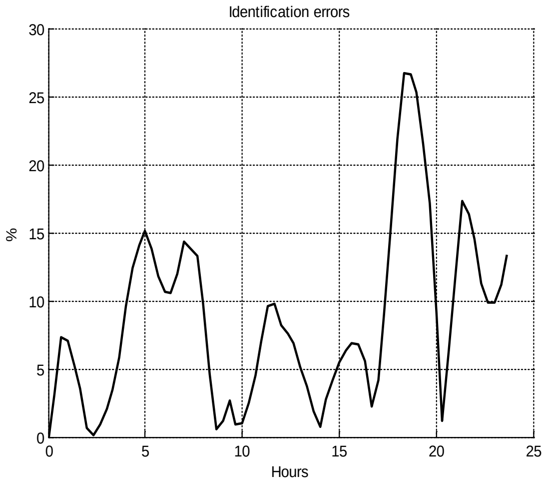

Fig 6. Distribution of model (1)-(2) identification errors

We will investigate this question experimentally. The identification of the model (4) - (1), with the results of which is shown in Figure 7, was carried out by checking the values of the parameter p_4 on the specially selected grids and the identification procedure given by the relations (6). As a result of solving the system of linear levels (6), refining the parameters by the Levenberg-Marquardt method and selecting the parameter p_4 as the basis of the performed calculations, the criterion for minimizing average ratios is the following values of the parameters of the differential equation (1):

$$p_1 = 6.9316_{e-4}, p_2 = 2.9586, p_3 = 0.3683, p_4 = 0.6810$$

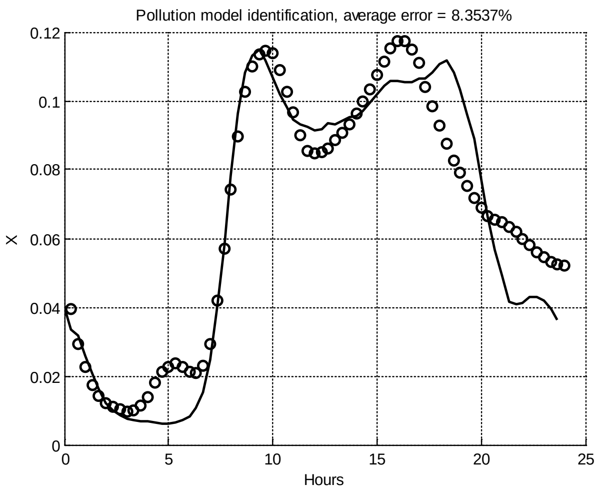

Fig 7. Approximation of observed pollution using identified model (4) -(2)

The average error of identification was 8.35%, which improves insignificantly significantly the result of the approximation of experimental data using the model (1) - (2). The analysis of the distribution of errors of approximation by model (4) - (2) also did not reveal any significant differences with the distributions of errors in the model (1) - (2).

III. Conclusion

As a result of the conducted research, a model of pollution on the local section of the city road was proposed in a form of an ordinary differential equation, which includes observation of typical daily traffic and daily forecast of wind speed. The methods of parametric identification of the constructed models were proposed. The methods are based on the use of difference approximation of the differential equation in specially selected points for construction of initial approximations of the model coefficients with their further refinement by the Levenberg-Marquardt method. Also, based on the use of selection of the p_4 parameter value included in the equation (4) nonlinearly in the sequence of special grids. As a result of the experimental study, the proposed models established their practical identity with the accuracy of the approximation of experimental data. Therefore, the simplest of them, namely the model (1) - (2), should be preferred.

The chosen model can serve as the basis for constructing a dynamic map of pollution of the city.

References

[1] K.M.Antropov, J.I.Kazmer, A.N. Varaksin. “Assesing spatial variability of air pollution in industial city with land use regression (review)”. Ecological systems and devices. Moscow, vol.1, pp. 28-41, 2010.

[2] T. Bush, S. Smith, K.Stevenson, S. Moorcroft, “Validation of nitrogen dioxide diffusion tube methodology in the UK”, Atmospheric Environment, vol. 35, 2, 2001.

[3] I. Voytyuk, M. Dyvak, V. Spilchuk. “Research of quality characteristics of models structure in kind of interval difference operator”. Proceedings of International Conference CADSM’2011, Polyana-Svalyava, pp. 87, 2011.

[4] N. Porplytsya, M. Dyvak,, I. Spivak,, I. Voytyuk, “Mathematical and algorithmic foundations for implementation of the method for structure identification of interval difference operator based on functioning of bee colony”, 13th International Conference: The Experience of Designing and Application of CAD Systems in Microelectronics, CADSM 2015, pp. 196-199, 2015.

[5] N. Ocheretnyuk, I. Voytyuk, M. Dyvak, Ye. Martsenyuk, “Features of structure identification the macromodels for nonstationary fields of air pollutions from vehicles”, Modern Problems of Radio Engineering, Telecommunications and Computer Science - Proceedings of the 11th International Conference, TCSET'2012, pp. 444.

[6] T. Sahsuvaroglu, A. Arain, B. Beckerman, P. Kanaroglou, J.R. Brook, N. Finkelstein, M.M. Finkelstein, N.L.Gilbert, B. Newbold, M.Jerrett. “A LUR model for predicting ambient concentrations of nitrogen dioxide in Hamilton, Canada”, J. Air Waste Manage. Assoc., vol. 56, 8, 2006.

[7] R. Pasichnyk, M. Dyvak, N. Pasichnyk. “Identification and modeling of limiting factors systems”. Data Stream Mining & Processing (DSMP), IEEE First International Conference, pp. 336-34, 2016.

[8] R. Pasichnyk. “Modeling of resources accumulation and operational management in biotechnology, biomedical and web information systems”. Computational Problems of Electrical Engineering, vol. 4, 2, pp. 37-46, 2014.