https://doi.org/10.3176/hum.soc.sci.1987.1.03

EESTI NSV TEADUSTE AKADEEMIA TOIMETISED. ÜHISKONNATEADUSED
ИЗВЕСТИЯ АКАДЕМИИ НАУК ЭСТОНСКОЙ ССР. ОБЩЕСТВЕННЫЕ НАУКИ
PROCEEDINGS OF THE ACADEMY OF SCIENCES OF THE ESTONIAN SSR. SOCIAL SCIENCES

1987, 36, 1

Tiit KALLASTE

EXTRAPOLATION FORECASTING OF AIR POLLUTION BY MEANS OF STATISTICAL MODELLING

The author discusses some possibilities of compiling genetic forecasts of air pollution by means of correlation and regression analyses as part of long-term planning of environmental protection. Trend models and multifactor regression equations are used to forecast air pollution on the example of the five biggest Estonian towns.

1. Introduction

The reduction of air pollution below accepted standards requires timely abatement measures. National economy which is a direct or indirect source of environmental pollution is characterized by relatively high inertia. Consequently, it is impossible to reduce pollution in a short time. To avoid above-standard pollution in five or ten years, the necessary investments for technological improvement, pollution control equipment, the building of new enterprises outside residential districts instead of outdated or highly polluting ones, etc. should be planned already at present. Hence the necessity of compiling forecasts of environmental pollution for different territorial units (towns, industrial regions, recreational areas, etc.) arises. Long-term pollution forecasts (air pollution forecasts among them) could provide indispensable initial material for planning and design organizations. Such forecasts should serve as a starting point in national economy planning, they are of special importance in elaborating measures for environmental protection and planning capital investments.

Long-term environmental forecasting has been treated by the Soviet authors N. P. Fedorenko, K. G. Gofman, A. A. Gusev, R. L. Rajatskas, V. P. Sutkaitis, N. F. Reimers et al. [1—5]. As to foreign authors, G. Finzi, G. Tebaldi, T. Soeda, Y. Sawaragi, T. Yoshimura, J. M. Guldmann, A. W. C. Keddie et al. [6—9] should be mentioned.

Two kinds of air pollution forecasts can be distinguished: normative and genetic. Normative forecasts are the so-called active ones, i. e. they assume the existence of air pollution and certain management of national economic development. Normative forecasting is most often treated by means of optimum planning problems with constraints on output volume, available resources or pollution level prescribed [10, 11]. In solving such problems, peculiarities of the development of different sectors, technological innovations, changes in the composition of the fuels used, new pollution control equipment, etc. are taken into account. However, the use of these models is often limited because of the lack of initial data or normative indicators. In the Soviet Union the centre of working out normative forecasting models for air pollution is the Central Institute of Mathematical Economics of the USSR Academy of Sciences (e. g., works by K. G. Gofman and A. A. Gusev [2, 3, 12]).

Genetic forecasts will be treated in the present paper. Genetic forecasts are in essence passive, inertial. They are compiled with the assumption that the development tendencies formed during a certain retrospective time interval will not be changed before the end of the forecast period. Thus, they are estimates of air pollution provided the proportions between the growth of production and the construction of treatment plants remain unchanged, innovations accompanying technological progress are not considered. They might be called alarming or warning forecasts.

No doubt, such extrapolation of tendencies for a future period brings about certain relativity and approximation. There seem to exist a number of factors which cannot be considered in forecasting. Since we do not have at our disposal either the mathematical apparatus or the necessary initial data for more exact forecasts at present, we are inevitably facing the alternative whether to give up air pollution forecasting altogether, or compile preliminary approximate forecasts. Below some methodological problems concerning the compilation of genetic forecasts of air pollution will be treated from the standpoint of the possibilities of applying statistical methods — correlation and regression analyses.

2. On the application of statistical methods for air pollution forecasting

As is known, the forecasting process based on informational-statistical methods consists roughly of two stages. At the first stage, the so-called inductive one, observation data on the pollution of the air of a certain regional unit (city, industrial centre, etc.) are studied over some period in the past, and the obtained statistical regularities are presented in the form of a model. Two types of statistical models are constructed: an equation of the growth tendency of air pollution over time (the so-called trend models), and an equation of the dependence of air pollution on one or several factors-arguments (the so-called models of multiple regression) representing the main sources of environmental pollution. At the second, i. e. the deductive stage, the possible expected pollution levels are determined using the statistical regularities found and the forecasts of the parameters of the pollution sources.

The forecasting of the air pollution of cities and industrial centres is based on the relatively high inertia of economic processes which are the direct or indirect causes of pollution. Consequently, it can be assumed that such an inertia is characteristic of the air pollution processes as well. If a forecaster has at his disposal sufficient information on the changes in pollution levels in the past it seems that he will also be able to determine its variation tendencies for a certain future period.

The inertia of pollution processes manifests itself, first, as the preservation of the general tendencies of pollution change over time, and, secondly, as the preservation of the mechanism of interrelations, i. e. the preservation of the correlation of pollution and the set of pollution sources.

As the construction of adequate statistical models of air pollution processes is hindered by the insufficiency of initial data, it should be noted that the obtained forecasts of the air pollution of cities and industrial centres cannot be treated as final ones.

The assumption that the existing variation tendencies and rates will remain unchanged is certainly rather conditional and even inexact, but for a first estimate of the possible future pollution levels such an assumption is quite appropriate, and the forecasts have primarily a warning function. When interpreting them for planning air pollution control measures, attention should be paid to the factors and constraints not considered in compiling statistical models. To put it briefly, the extrapolated forecasts of air pollution drawn with the help of statistical methods should be critically interpreted with a consideration of the tendencies of production and technological development which are the main factors determining the quality of environment.

It should also be noted that a certain conditional character of air pollution forecasts compiled with the help of statistical methods may result from the fact that the strict requirements set to the quality of the initial data (e. g., their homogeneity, distribution pattern, etc.) have not been satisfied. As a rule, the initial information on the air pollution of cities and industrial centres is extremely nonhomogeneous and incomplete as yet. Therefore, the importance of air pollution forecasts should not be overemphasized. They should be regarded as auxiliary material for preplanning which helps to analyse a number of development alternatives.

3. Air pollution forecasting over time

Generally quite little is known about numerous factors affecting air pollution and their relations with immission.[1] Therefore, the object being forecast should be treated in a rather general way. It seems to be advisable to treat the air pollution of a certain region as a physical phenomenon affected by human impact (economic activities) with only its intensity varying over time. In so doing it is assumed that time generalizes the joint effect of all factors causing air pollution. Specific problems concerning the formation of pollution level, the intensity of individual factors, their interrelations, etc. are not directly considered.

As we mean by air pollution the content of pollutants in the air, we use dynamic immission series to characterize air pollution variation in time.

We start the forecasting with finding out the general variation tendency, i. e. the tendency of the variation of immission in time (trend) with an aim of extrapolating it on the future. A trend characterizes the basic regularity of the variation of the pollution level in time and describes the actual averaged tendency of the pollution level during the observation period. The trend model is obtained by debugging an empirical series of pollution dynamics with the help of different functions. The debugging of the dynamic series consists of the selection of the type of curve (linear, exponential, logarithmic, etc.) whose form would suit best the character of the changes in pollution level during the statistical period, and the estimation of the parameters of the functions [13].

The choice of the type of function is of vital importance in the debugging of the dynamic series since the results of the extrapolation of the trend depend namely on it. Provided that all other conditions are equal, an error in the choice of the type of the debugging curve will have greater effect on the results of the forecast than an error connected with statistical estimation of the parameters [14, p. 44].

There exist numerous approaches to the choice of the type of the curve describing the variation tendency of pollution. The form of the curve may be found on the basis of the value of one or several purely statistical criteria. Formal choice might be combined with logical reasoning. First, curves are chosen proceeding from essential deliberations with earlier research and experience taken into account. Such an approach enables to considerably limit the number of possible functions and tackle the actual solution of the problem. Finally, for the remaining functions the values of their criteria are calculated, and on their basis the ultimate choice is made. To make the choice of the best trend model easier, graphic approaches might also be applied which enable to combine visual effects and formal criteria. To this end charts are drawn of the functions which have proved to be statistically trustworthy (i. e. meaningful), e. g. according to Fisher's F-criterion.

[1] immission — the content of pollutants in the air formed as a result of their diffusion, transformation and transfer.

In the present paper the sum of the squares of the deviations of the actual values from the calculated ones obtained by means of debugging was used as the main formal criterion. Among the set of curves used for the debugging of the time series the one is selected whose mean residual squariance S is the smallest. The so-called method of least squares is also applied to estimate the parameters of the debugging curves. To characterize the closeness of the actual observations to the calculated values (obtained from the trend equation) a correlation coefficient is calculated, reflecting a linear dependence between immission and time.

After numerical estimation of the parameters of the equations characterizing the immission trends, the models are used to forecast air pollution. Using the air pollution data bases of Tallinn, the following trend models were constructed:

for dust $y_1=0.3567+0.0172t-0.1550 \ln t$;
SO_2 $y_2=0.1211+0.0067t$;
CO $y_3=2.1655+0.5728t$;
soot $y_4=0.0630+0.0288t-0.0926 \ln t$,

where $y_1, \ldots, y_4$ are the average annual immissions (mg/m^3) of the respective pollutant s and t is time (in years) as an independent variable. Respective forecasts for Tallinn were compiled up to 1990, for greater detail and numerical estimations see [15].

For the estimation of the reliability of the forecasts, they should be compared with the real immission values of the last year of the statistical period. In the case of Tallinn the values of the retrospective forecasts were quite close to the real immission values, which makes it possible to use extrapolation forecasts of this kind at the first stage of research.

4. Air pollution forecasting by means of multiple regression

Extrapolation forecasting of air pollution on the basis of multiple factors relies on the inertia of the interrelations between pollution and its sources. For modelling the interrelations are presented in the form of regression equations. The regression equations connect the changes of the dependent variable (air pollution) with the effect of a certain number of factors-arguments (pollution sources), i. e. for the data of statistical observations the most suitable equation of the type $y=f(x_1, x_2, \ldots, x_n)$ is chosen.

In most general lines the regression analysis of air pollution consists of two stages [14]. At the first stage the main pollution sources having the greatest influence on pollution level are selected, and the form of the regression equation is determined. The second stage is devoted to estimating the parameters of the regression equation.

It is advisable to choose the factors-arguments first on the basis of a qualitative analysis which enables to considerably cut the number of pollution sources included in the model. Taking into consideration the goal of the investigation as well as the availability of the respective statistical information, only the most important pollution sources can be included. The less important ones are excluded, though this results in certain inexactness of the mathematical description of the interrelations. Here it seems to be indispensable to seek help from expert estimates, logical reasoning and experience accumulated in the course of studying the processes of environmental pollution. The professional knowledge of various specialists working in the field of environmental protection may also be highly useful.

In choosing the main pollution sources some formal techniques might be applied in addition. Iterative correlation analysis, e. g. step-wise correlation analysis with the limit of partial correlation given to exclude statistically insignificant factors from the initial equation might be used. On the basis of the matrix of the coefficients of correlation of the pollution level and the characteristics of the pollution sources as well as by comparing the latter with one another, a preliminary selection is made. A strong statistical dependence between the pollution sources refers to the fact that they bring about a certain environmental pollution in a close interrelation with one another and not in isolation. Hence, the strong dependence between the factors-arguments is one of the grounds for excluding pollution sources from the initial equation [12, pp. 118—124].

An important reason for excluding (though only after the regression analysis) pollution sources is also the checking of the statistical significance of the coefficients of the multiple regression, e. g. according to Fisher's F-criterion.

When the insignificant factors-arguments are eliminated, the form of the regression can be chosen. This is a problem that cannot be solved unambiguously. Most probably a number of alternatives will be obtained as the interrelations studied may be presented by means of several types of functions rather adequately. In [16, p. 70], e. g. the following principles are suggested as the starting point for determining the form of the relation: the form of the relation must reflect the objectively existing form of interrelations between the quantities studied possibly well; the relations which can be transformed linear should be preferred; among the various types of functions those which guarantee the best approximation with regard to the selected criterion should be chosen. Bearing in mind the above-mentioned principles, out of the various different types of functions significant according to F-criterion, the one having the lowest value of the mean residual squariance S, the biggest value of the coefficient of multiple regression R and, if possible, all the regression coefficients with a positive sign, is chosen.

Thus, the selection of the best multiple regression model for the forecasting of air pollution according to the characteristics of the main pollution sources is carried out at several stages with the essential-logical solution criteria closely interwoven with formal ones. Ultimately the most appropriate equation of multiple regression is chosen.

The final stage of multiple regression analysis — numerical estimation of the parameters of regression — is conducted by means of a computer applying the least-squares method. It creates a quantitative image of the share of individual pollution sources in the formation of a certain pollution level.

The air pollution forecasts are an estimate of the mean pollution level in case of certain characteristics of pollution sources. Thus, the multifactor regression models of pollutants provide the basis for making calculation experiments when planning air pollution abatement measures. An air pollution forecast based on the characteristics of pollution sources describes the possible air quality provided the existing tendencies do not change.

4.1. Statistical multifactor models

In investigating the formation of a certain air pollution level with the help of statistical modelling the following relationships are of prime interest: (1) the operation intensity of the pollution sources — emission of pollutants; (2) the operation intensity of the pollution sources — immission of pollutants; and (3) emission of pollutants — their immission.

Quantitative research of these relationships enables to construct the respective statistical models and use them for forecasting the pollution levels. From the standpoint of planning the air pollution abatement measures and national economic growth the two first types of relationships are the most important. Their formalization into a statistical model renders it possible to start to regulate the emission of the pollution sources to the immission level not surpassing the legal standards set. The modelling of the relationships of type (1) is seriously hindered by the small statistical base, since systematic collection of data on emission started in the whole Soviet Union but a few years ago. However, in the 1980s an inventory of the pollution sources was started and the basis was laid for the respective data bank, so in the future it will be possible to study also this type of relationship. Since so far statistical reports have been made mainly about the immission, it is advisable to study the interrelations of type (2) where the content of pollutants in air is correlated with the operation intensity of the pollution sources.

For the statistical modelling of the interrelations between the operation intensity of the pollution sources and the immission of pollutants it is indispensable to solve the problem of the selection of the main pollution sources as it is evidently impossible to include all the sources. It is inevitable to analyze only the most important factors to compile the model, in so doing the resulting error will be negligible.

To find out the main air pollution sources of a city a comprehensive inventory of the environmental pollution on the example of Tallinn has been taken over several years. The basic results of this research have been published in [17, pp. 78—96]. Generalizing the experience accumulated we can draw the conclusion that in cities the air pollution sources fall into three basic groups: (1) industrial enterprises (together with the power supply systems and cargo motor transport serving them), (2) non-industrial (i. e. municipal) boiler houses, and (3) motor vehicles.

The city air is mainly polluted because of the burning of various fuels. A minor role belongs to the pollutants emitted in the course of the technological cycles by plants, but the data available on these emissions are quite inadequate and are based mainly on expert estimates. Drawing conclusions on the basis of the fuels consumed and the emission in Tallinn, it becomes evident that the total share of all non-industrial boiler houses in the air pollution is relatively small (about 3—5 per cent of the total fuel consumption) as compared to industry. Therefore, at the present stage of the statistical modelling a simplification has been made excluding non-industrial boiler houses as a separate factor among the pollution sources. The air pollution caused by them is taken into account but in an indirect way, as the so-called background pollution. Motor vehicles are not considered directly as independent pollution sources either, as we have been faced with certain difficulties in taking an inventory of this quite significant type of pollution sources. Namely, as yet we do not have trustworthy territorial surveys of the pollution caused by the burning of petrol and diesel oil.[2] Thus, transport as a separate pollution source has not been directly taken into account in the present statistical modelling. Only transport servicing certain industries can be considered conditionally together with the respective industries. The pollution caused by other types of transport is taken into account as background air pollution.

Research of emission sources allowed us to comprise only industry as the main pollution source of the city air in the present statistical modelling. The share of all other sources is taken into account as background pollution. This assumption allows us to set the problem of the modelling of

---

[2] So far only a few studies of the intensity of motor transport on city roads have been conducted by Tallinn Technical University and the Institute of Chemical Physics and Biophysics of the ESSR Academy of Sciences.

pollution processes under incomplete or partially lacking initial data in a considerably simplified way.

Because of the great number of factors involved it seems to be impossible to determine the functional relationship between the immission and the output volume of the industries. Therefore, it is advisable to study their correlational relationship and the closeness of this relationship, i. e. apply statistical research methods. For the formalization of the correlational relations the author suggests that the coefficient of specific immission (w) which characterizes statistically the immission of the pollutant studied per unit operation intensity of the pollution source can be used (the basic idea of the method is presented by the author in [18]). The operation intensity of the polluting industries may be reflected for example by their output volume.

According to the above-said, the correlational relation between the immission and the volume of industrial output may be described by means of the following statistical models.

(1) A linear model

$$y_s^{tp} = \sum_{i=1}^{m} w_i^s x_i^{tp} + G_s^{tp},$$

$$t=1, \dots, T; \; p=1, \dots, q; \; s=1, \dots, z,$$

where the coefficient of specific immission showing the share of the immission of the pollutant s in the air per unit of output of the industry i in the region p is expressed as follows:[3]

$$w_i^s = \frac{y_{si}^{tp}}{x_i^{tp}},$$

where

y_{si}^{tp} — average annual immission of the pollutant s in the air (mg/m^3; g/m^3) in the region p in the year t caused by the industry i;

x_i^{tp} — total annual output (in the year t) of the industry i (in millions or milliards roubles) within the territorial unit (region, town) p.

(2) A nonlinear model (in the present case described with the help of the exponential function)

$$y_s^{tp} = G_s^{tp} \prod_{i=1}^{m} e^{w_i^s} x_i^{tp},$$

$$t=1, \dots, T; \; p=1, \dots, q; \; s=1, \dots, z.$$

The coefficient w_i^s is determined on the basis of all the regions studied and it may either vary in time or be stable, i. e.

(a) $w_i^s = \text{const}$, or

(b) $w_i^s = f(t)$ e. g. $w_i^s = w_{0i}^s \cdot e^{-ct}$.

In both models the index t stands for the years of the statistical period,

[3] It should be noted that an indicator similar in principle to the above-suggested one, viz. the coefficient of the emission of a pollutant per unit of total production of an industry has been used by V. Leontief and D. Ford (see, e.g., [19]) in their interindustry analysis of the effect of the structure of economy on environment, and also by F. R. Førsund and S. Strøm [20] in forecasting the amount of pollutants up to the end of the present century.

s for the pollutants studied, i — the polluting industries, p — regions. G_s^{tp} denotes the background pollution.

For simplicity's sake the statistical coefficients of the specific immission w_i^s are regarded as stable in time during the initial stage of research. Formally they are regression coefficients. They show the change in the immission of the pollutant s accompanying a unit increase in the operation intensity of the pollution source i. Thus, these coefficients make it possible to quantitatively describe the share of any industry (or even plant) in the formation of a certain air pollution level.

Negative values of the coefficients of specific immission mean that such a situation has developed when an increase in the production of the polluting industry results in a decrease in its polluting effect on the environment. The underlying reason is that the effect of trapping pollutants or cleaning exhaust gases is higher when the output volume is bigger. As an example, power engineering plants might be cited. At small or out-of-date power plants the cleaning effect is extremely low or no cleaning is applied whatsoever. Large power plants, however, apply highly efficient electric filters trapping over 98 per cent or almost all pollutants. As a result, the volume of pollutants emitted per unit of output is considerably smaller at large plants.

Evidently, the coefficient of specific immission changes in the course of technological progress, process improvement, and cleaning of exhausts. A description of the possible cases of such changes has been presented by the author in [18].

If the statistical period (both for data on immission and the operation intensity of industries) were considerably longer for every district of the region, it would be possible to construct an independent model analogous to the above-presented ones for all districts studied. Such a model would simulate the relationship between the air polluting industries of the region studied and the immission caused by them considerably better than now. Thus, a specific model could be obtained for every town studied.

In the Estonian SSR a general model for the whole republic has to be used because of incomplete information. Such a model cannot take into consideration the specific character of the industrial air pollution in each town. It is constructed on the basis of 5 major towns whose total industrial production accounts for about 80 per cent of the republic's total industrial production. The general model enables to forecast the immission of pollutants studied relying on the volume of industrial production in all Estonian towns as the statistically determined coefficients of the specific immission of the towns studied are approximately equal to the general coefficients of the specific immission industry in the whole region. This means that we can estimate (though rather roughly) already now the present and future air pollution of a certain territory with such pollutants whose immissions are not presently registered. The registered immissions of the pollutants under study can be compared with the immission values prognosticated with the help of the model.

4.2. The selection of the major air polluting industries and the compilation of regression equations on the example of Estonian towns

Because of the small statistical sample it is not possible to include all the polluting industries in the models suggested. Therefore some industries have been integrated into one sector on the basis of their similar polluting effects, thus reducing the number of initial factor-arguments. On the other hand, some industries are undoubtedly heavier air polluters than others. For example, as concerns the two most important pollutants — sulphureous anhydride and nitric oxides accounting respectively for 1/2 and 1/3 of the total volume of the pollutants emitted yearly in Tallinn — the electric power and the chemical industry account for 73 and 76 per cent of the emissions. The next two branches — engineering and metal-working and the building materials industry — account for only 15—20 per cent of the emission.

As to the remaining four pollutants studied the following can be said on the basis of empirical studies of the polluting effect of industry in Tallinn. The major part of the dust emission comes from the electric power, building materials and forest, timber and paper industries. The major sources of carbon monoxide are transport and the electric power, building materials, and engineering industries as well as ferrous and non-ferrous metal industries. The same industries are the main sources of soot. The electric power, chemical and fuel-producing industries are responsible for the major part of the emission of phenol.

It is easy to see that there are but three or four industries occurring in different combinations as the most important sources of the emission of different pollutants while the total number of major air polluting industries is six. Consequently, six industries — the fuel-producing, electric power, engineering together with ferrous and non-ferrous metal, chemical, building materials, and forest, timber, pulp and paper industries — can be regarded as those determining the air quality. In the regression equations these industries are used as independent initial variables.

The share of the other industries as air polluters is, as a rule, negligible, and they can be omitted from the regression equation of factor arguments, the resulting error will be small.

Besides empirical selection of air polluting industries by experts, below also formal techniques will be applied for choosing the most important independent variables. For this purpose e. g. a programme of stepwise regression on an EC-1022 computer was applied which yielded multi-factor models with three to six factor arguments. To find out the most suitable of them a different number of industries are included in the model. To study the six main pollutants it would be necessary to analyze 66 combinations of the groups of the most important polluting industries applying a linear as well as a nonlinear model. As the input of the initial information file for an equation is a highly labour-consuming task, a computer programme[4] was compiled which forms all possible combinations of the independent variables, and then makes the standard statistical analyses for each of them. Out of 132 variants of the groups of the most heavily polluting industries the multifactoral model satisfying all the above-presented criteria in the best way is selected for each of the six pollutants. The linear and nonlinear (here exponential) models which provide basis for forecasting are the following:

for dust: $\ln y_1 = -1.5440 + 0.0001x_1 + 0.0303x_5 - 0.0552x_6$;

SO_2 $\quad y_2 = 0.0674 + 0.0003x_1 + 0.0005x_3 - 0.0026x_5 + 0.0054x_6$;

CO $\quad \ln y_3 = 1.9788 + 0.0010x_2 + 0.0005x_4 \quad (F_{0,50})$[5];

soot $\quad y_4 = 0.0268 + 0.0001x_1 + 0.00007x_4 + 0.0010x_5 - 0.0013x_6$;

NO_x $\quad y_5 = 0.0354 + 0.00003x_2 + 0.0002x_3 + 0.0002x_5$;

C_6H_5OH $\quad y_6 = 0.0052 + 0.00005x_2 + 0.000005x_3 + 0.00001x_4$,

---

[4] Both programmes have been worked out by T. Tiits at the Institute of Cybernetics of the ESSR Academy of Sciences.

[5] The reliability of the equation is very low only for carbon monoxide (50 per cent). This is due to the fact that motor transport is the main emitter of this pollutant. To forecast the immission of carbon monoxide, the respective trend equations should be preferred.

where y_s — average annual immission of the respective pollutant s (in mg/m^3);
x_1 — total output (in million roubles) of the electric power industry;
x_2 — that of the fuel-producing industry;
x_3 — the chemical industry;
x_4 — engineering and metal working;
x_5 — timber, pulp and paper industry;
x_6 — the building materials industry.

Using the above-presented models, extrapolational forecasts of the immission of the pollutants investigated were compiled for Tallinn, Kohtla-Järve, Narva, Tartu and Pärnu up to 1995. For greater detail see the research carried out by the author at the Institute of Economics.

5. Summary

The paper discusses possibilities of studying the interrelations between air pollution and the respective pollution sources by means of statistical modelling for the purposes of long-term forecasting. First, air pollution is treated as varying in time. The criteria used for choosing the functions for debugging the time series and some methodological problems concerning the selection of the most suitable trend models for forecasting are discussed. It can be said that a statistical analysis of the time series of air pollution is relatively simple and not too labour-intensive, providing at the same time satisfactory results for estimating the tendencies of air pollution changes. For this reason it could be used at the initial stage of research to obtain approximate results.

At the next stage of statistical modelling the relationships between the operation intensities of air pollution sources and the immission of various pollutants are studied. Here a new approach — the so-called method of specific immission coefficients is suggested. The method consists in combining the time series characterizing air pollution with those of the operation intensities of the main pollution sources by means of linear and nonlinear multifactor models. To select the main pollution sources the experience accumulated in the course of studying air pollution in the city of Tallinn has been used. Presently only industrial enterprises are treated as the main pollution sources while all the other sources are considered as background pollution. Using respective computer programmes, a different number of certain independent variables, i.e. the most heavily polluting industries are included into the statistical models, and standard multiple correlation and regression analyses are conducted. Accordingly, different coefficients of the multiple regression equation are obtained. As to their contents, the coefficients of specific immission show the share of the most heavily polluting industries in the immission of certain pollutants. Besides, they can serve as generalized indicators for the individual plants of the most heavily polluting industries.

The most suitable equations are chosen for every pollutant studied and used for forecasting. The suggested multifactor regression models are of a universal character. They enable to compile preliminary air pollution forecasts for any town or district located in the region (union republic) under study. This is highly important for forecasting air pollution of such areas of the region where environmental monitoring is not yet conducted though the quality of environment requires at least an approximate characterization.

Finally, general multifactor regression models for the Estonian SSR are presented. With the help of these models, the air pollution of five biggest

towns up to 1995 has been forecast. These forecasts as well as those compiled on the basis of trend models have been used in the comprehensive scheme of environmental protection and rational use of natural resources in the Estonian SSR, as well as comprehensive programmes for the development and location of productive forces and scientific and technological progress.

REFERENCES

1. *Федоренко Н. П.* Некоторые вопросы теории и практики планирования и управления. М., 1979.
2. *Гофман К. Г.* Основы долгосрочного прогнозирования состояния окружающей среды. — Вопр. географии, 1978, вып. 108, 95—104.
3. *Гусев А. А.* Проблемы совместного прогнозирования развития экономики и охраны атмосферы. — Экономика и матем. методы, 1979, XV, вып. 1, 31—44.
4. *Раяцкас Р. Л., Суткайтис В. П.* К проблеме планирования и прогнозирования состояния окружающей среды. — Изв. АН СССР. Сер. экон., 1979, № 3, 67—78.
5. *Реймерс Н. Ф.* Принципы экологического прогнозирования и планирования. — Итоги науки и техн. ВИНИТИ АН СССР. Охрана природы и воспоризводство природных ресурсов, 1978, вып. 4, 41—74.
6. *Finzi, G., Tebaldi, G.* A mathematical model for air pollution forecast and alarm in an urban area. — Atmos. Environ., 1982, 16, N 9, 2055—2059.
7. *Soeda, T., Sawaragi, Y.* a. o. Modelling for prediction of air pollution levels by nonphysical models. — In: Sci. for Better Environ. Kyoto—Oxford, et al., 1977, 566—577.
8. *Guldmann, J.-M.* Industrial location, air pollution control and meteorological variability. A dynamic optimization approach. — Socio-Econ. Plan. Sci., 1978, 12, 197—214.
9. *Кедди А. В. К.* Прогностическое моделирование загрязнения воздуха: исследование в Форт-Вэлли. — В кн.: Мониторинг состояния окружающей природной среды. Тр. I сов.-англ. симп. Л., 1977, 154—173.
10. *Калласте Т.* Об одной модели оптимального планирования качества среды региона. — В кн.: О моделях нормативного прогнозирования социально-экономического развития региона. Таллин, 1980, 31—62.
11. *Воронович Н. В., Золотаренко С. Г.* Динамическая модель управления качеством природной среды. М., 1979. Деп. в ВИНИТИ 28. 03. 1979, № 1675-79.
12. Охрана окружающей среды. Модели управления чистотой природной среды. Под ред. К. Г. Гофмана и А. А. Гусева. М., 1977.
13. *Фёрстер Э., Рёнц Б.* Методы корреляционного и регрессионного анализа. Руководство для экономистов. М., 1983.
14. *Четыркин Е. М.* Статистические методы прогнозирования. М., 1977.
15. *Калласте Т.* Экстраполяционное прогнозирование загрязненности воздуха по времени (на примере Таллина). — Изв. АН ЭССР. Обществ. н., 1980, № 4, 309—323.
16. *Лифанчиков А. Н., Воскобойник М. П.* Экономико-математические методы в планировании развития пищевой промышленности. М., 1975.
17. *Калласте Т.* Об анализе и прогнозировании загрязненности воздуха с помощью многофакторных регрессионных моделей. — В кн.: Методы планирования и управления прогнозными ресурсами (планирование и управление качеством окружающей природной среды). М., 1979, 63—65.
18. *Калласте Т.* Об одной модели прогнозирования загрязненности воздуха региона. — Изв. АН ЭССР. Обществ. н., 1979, № 3, 181—190.
19. *Leontief, W., Ford, D.* Air pollution and the economic structure: empirical results of input-output computation. — In: Input-Output Techniques. New York, 1972, 9—30.
20. *Førsund, F. R., Strøm, S.* Industrial structure, growth and residual flows. — Reprint Series N 139. University of Oslo. Oslo, 1975.

Presented by K. Habicht

Academy of Sciences of the Estonian SSR,
Institute of Economics

Received
May 12, 1986

Tiit KALLASTE

ÕHU SAASTUMISE EKSTRAPOLATSIOONPROGNOOSIMINE STATISTILISE MODELLEERIMISE ABIL

Õhu saastumise esialgseks ja ligikaudseks prognoosimiseks võib edukalt kasutada statistilist modelleerimist, rakendades selleks korrelatsioon- ja regressioonanalüüsi.

Artiklis on õhu saastumist käsitletud esmalt sõltuvana ajast kui üldistavast muutujast. Selline lahendus on otstarbekas just uurimise algetapil, kus on vaja määrata saastumise muutumise kõige üldisemad seaduspärasused. Tallinna õhu saastumise näitel on välja töötatud trendimudelid nelja saasteaine kohta.

Modelleerimise järgmisel etapil pakub huvi õhu saasteainete immissiooni korrelatiivne seostamine peamiste saasteallikate tegevusintensiivsusega. Selleks on välja töötatud uudne statistiliste eriimmissioonikoefitsientide meetod, kusjuures eeldatakse nii lineaarsete kui ka mittelineaarsete seoste kehtivust. Meetod võimaldab analüüsida regiooni suvalise linna või tööstuskeskuse õhu praegust saastatust, geneetiliselt prognoosida tulevast saastatust ja kindlaks määrata kõige enam saastavate tööstusharude osatähtsus teatud saasteaine immissiooni kujunemises. Ka on võimalik esialgselt hinnata ja prognoosida õhu saastatust seal, kus saasteainete immissioone veel ei registreerita, kuid mille saasteseisund vajab ligikaudsetki iseloomustamist.

Mitmese regressiooni mudelid Eesti NSV viie suurema linna näitel on välja töötatud kuuele enamlevinud saasteainele. Statistilise modelleerimise käigus on rakendatud mitmeid arvutiprogramme, mis võimaldavad märgatavalt ratsionaliseerida andmetöötlust. Trendimudelite ja mitmese regressiooni mudelite abil arvutatud peamiste meil registreeritavate õhu saasteainete ekstrapolatsioonprognoosid ning ka prognoosimetoodika on leidnud kasutamist «Eesti NSV keskkonnakaitse ja loodusvarade ratsionaalse kasutamise kompleksskeemis», samuti Eesti NSV tootlike jõudude arendamise ja paigutamise ning teaduse ja tehnika progressi kompleksprogrammides.

Eesti NSV Teaduste Akadeemia
Majanduse Instituut

Toimetusse saabunud
12. V 1986

Тийт КАЛЛАСТЕ

ЭКСТРАПОЛЯЦИОННОЕ ПРОГНОЗИРОВАНИЕ ЗАГРЯЗНЕННОСТИ ВОЗДУХА С ПОМОЩЬЮ СТАТИСТИЧЕСКОГО МОДЕЛИРОВАНИЯ

В целях первоначального и приблизительного прогнозирования загрязненности воздуха можно успешно использовать методы математической статистики, в частности корреляционный и регрессионный анализы.

В статье тенденции изменения загрязненности воздуха сначала рассматриваются в зависимости от времени как обобщающего фактора влияния. Такой подход целесообразен на первой стадии исследования, когда необходимо определить наиболее общие закономерности изменения загрязненности. На примере Таллина разработаны трендовые модели для четырех загрязняющих веществ.

На следующем этапе моделирования наибольший интерес представляет исследование коррелятивных связей между имиссиями загрязняющих веществ и интенсивностями деятельности основных источников загрязнения воздуха. Автором разработан оригинальный метод статистических коэффициентов удельной имиссии при допущении как линейных, так и нелинейных связей. Метод позволяет анализировать сложившуюся загрязненность воздуха любого города или промышленного центра в регионе, генетически прогнозировать загрязненность и определять вклад в нее наиболее сильных источников. Можно также первоначально прогнозировать загрязненность воздуха таких территорий региона, где наблюдения за имиссиями загрязняющих веществ пока еще не ведутся, но состояние атмосферы которых требует хотя бы приблизительной характеристики.

В результате статистического моделирования на примере пяти городов республики разработаны многофакторные регрессионные модели для шести основных регистрируемых загрязняющих веществ. Применяемые при этом вычислительные программы позволили заметно рационализировать процесс переработки информации.

Составленные на основе трендовых и многофакторных регрессионных моделей экстраполяционные прогнозы загрязняющих веществ, а также методика прогнозирования нашли применение в «Комплексной схеме охраны природы и рационального использования природных ресурсов Эстонской ССР», в схеме развития и размещения производительных сил, а также в комплексной программе научно-технического прогресса республики.

Институт экономики
Академии наук Эстонской ССР

Поступила в редакцию
12/V 1986