

**ESTIMATION AND PREDICTION
OF POLLUTANT CONCENTRATIONS INVOLVED
IN NITROGEN OXIDES CYCLES
A DOUBLY STOCHASTIC MODEL**

P. KAZIMIERCZYK*, PAWEŁ J. SZABŁOWSKI**, K. TWARDOWSKA**

The purpose of this paper is to present a new type of model that forecasts the instantaneous concentrations of several species involved in the NO_x cycle in the lower atmosphere, by treating the trajectories of pollutants' particles as trajectories of stochastic processes, and which takes all the advantages of filtering theory. The previous filtering models were found to be of limited use because of complicated chemical reactions for NO_x . Basing on the model considered here one can also identify the sources that participate in the up-crossings of the critical values of concentrations (the so-called alarm model).

1. Introduction

Processes that describe the pollutant transport in the atmosphere were systematically studied by deterministic methods (Juda, 1986; Marczuk, 1985; Nowicki, 1985; Seinfeld, 1977). The stochastic versions of these methods were intensively developed because, in general, the processes of pollutant transport have random nature that arises from random events in the atmosphere, random emission, and inaccuracy of parameters in a model. See (Korbicz and Zgurovsky, 1991) for a thorough description of these stochastic models.

Filtering theory and maximum likelihood estimation were applied to estimation of nitrogen oxide concentration in the vicinity of a roadway by Omatu *et al.* (1988). The time series analysis was used by Merz *et al.* (1972) to study the nitric oxide. The deterministic Eulerian type NO_x -model was examined by Budziński and Juda-Kuczka (compare also (Juda-Kuczka, 1986) for the Eulerian SO_2 model).

Since NO_x concentration models are more complicated than similar SO_2 models, the latter ones were studied deeper and in greater detail. Several adaptive-filtering SO_2 models were considered e.g. by Bankoff and Hanzevack (1975), Runca *et al.* (1979), Bonivento *et al.* (1980), Twardowska (1989). The computational algorithm for the instantaneous concentration prediction is based on the method of fractional steps and Crank - Nicolson method (Crank and Nicolson, 1947). For the direct numerical solution a turbulent transport partial differential equation coupled with a regressive least-squares filter are used. Such a model was tested on Cracow's (Poland) data taken from the experimental data set MONAT (1984), and in Silesian District (Poland), during continuous

* Institute of Fundamental Technological Research, Polish Academy of Sciences, 00-049 Warsaw, ul.Świętokrzyska 21, Poland

** Institute of Mathematics, Warsaw Technical University, 00-661 Warsaw, plac Politechniki 1, Poland

on-line testing experiments TEST1 (March 1988) and TEST2 (June 1988) (Twardowska, 1988; 1989).

Optimal filtering algorithms for linear distributed-time parameter systems based on the Wiener-Hopf theory were presented in the paper of Omatu and Seinfeld (1981). Two different stochastic models were studied in the papers of Finzi *et al.* (1982; 1983; 1984) both oriented towards the real-time forecast of daily SO₂ pollution. The first one is a "black-box" model with input-output transfer function identified by time-series method (Box and Jenkins, 1970). The second one is a "gray-box" model with a simple reasonable structure in accordance with the physical laws of the phenomena. These models are statistical in their main framework. A statistical model was also considered by Morawska-Horawska, (1979). Correlations between the global (for all urban area) daily average SO₂ concentrations and average meteorological conditions have been observed during a long period. The model works also on global average emission. Therefore, from such models one obtains the global average of SO₂ forecasted concentration.

Our NO_x model originates from the theory of stochastic differential equations with randomly varying structure, as well as from the adaptive filtering theory applied to minimization of modeling errors. These errors are connected primarily with the differences between the model and the real dynamics of such phenomena as random events and random emission, and result from the inaccuracy of monitoring devices and of calculations.

The first part of our model includes several input algorithms for interpolation, extrapolation etc., that process a large set of input data in order to assign proper values to every grid point of a three-dimensional network covering the tested domain. The output set of forecasted concentrations is also obtained at every grid point.

The second part consists of a system of seven parabolic partial differential equations describing the transport, advection and diffusion of NO_x, as well as the reactions with other pollutants in the atmosphere (see equations (1a)-(1f)). Processes of creation and annihilation of NO_x are also included in those equations. We apply a new mathematical theory of systems with randomly varying structure (Kazakov and Artemeev, 1980; Kazimierczyk, 1989). This theory enables us to model the chemical reactions of the NO_x cycle in a more precise way than the earlier models utilizing filtering theory. Such models, e.g. models for the SO₂ cycle (Bankoff and Hanzevack, 1975; Bonivento *et al.*, 1980; Twardowska, 1989) consist of one partial differential equation, and the only term connected with chemical evolution of pollutant concentration describes the decomposition of the pollutant, proportional to the time that elapsed from its creation (at the sources). In our model chemical creation and annihilation of several species at random moments are taken into account. The trajectories of the particles are discontinuous (they appear at the instants of creation, and disappear after removal and after chemical reactions). This is the reason for the use of varying structure models. Intensities of chemical reactions are not assumed to be constant and, in particular, depend on the concentrations of all substrata (not only on the concentration of the pollutant under consideration).

The third part of the model consists of a filter of random noises. Here, filtration theory for systems with randomly varying structure can be used.

The fourth part is formed by the output algorithms for some graphic programs. The alarm model results can also be incorporated in our model.

Let us notice that the self-learning model, which is obtained due to the use of the filter, is the best solution in the construction of the system of the, so called, on-line continuous forecasts. That is, the forecasts are calculated recursively, and a forecast for time n is made upon all available data collected up to time $n-1$. Our model learns from the differences between the forecasts and the actual measurements. Therefore, precision of the forecasts improves with filtering time. However, in order to take all advantages of the powerful filtration methodology, we have to employ chemical and mechanical models as precise as possible. An oversimplified model subjected to filtration would cause the divergence of the filter.

2. Input Data Description

2.1. Basic Set of Necessary Data

The great advantage of the model presented is that it acts on the same set of meteorological and climatic data, and on the same network as the SO_2 model. The sets of emission and immission data also have the same structure. Therefore, we can use the same input and output programs, taking into account mutual reactions between examined pollutants at every step of computations. In what follows, we present the set of input data for the computation of one twelve-hour ahead instantaneous forecast.

2.2. Meteorological and Climatic Data

The following meteorological and climatic data are considered:

- wind velocity $v_m(t)$, $m = 1, \dots, 4$, and direction measured or computed at the: $z_1 = 10$ m or 15 m (height of the ground meteorological station), $z_2 = 100$ m, $z_3 = 250$ m and $z_4 = H$ (height of the inversion layer base) at time n_0 when we start computing our forecast; these data are forecasted at times $n_0 + 1, \dots, n_0 + 12$. Between $n_0 + i$ and $n_0 + i + 1$, we use linear interpolation;
- temperature and stability class $s(t)$ of the atmosphere given as above;
- rainfall intensity and the height H forecasted for the periods $(n_0, n_0 + 6)$ and $(n_0 + 6, n_0 + 12)$;
- vertical and horizontal diffusivity coefficients K^z and K^x , K^y , respectively, solar radiation, scattered solar radiation, cloudiness for smog investigations.

The upper wind velocity $v_m(t)$, $m = 2, 3, 4$, and the diffusivity coefficient K^z may be computed, e.g., by the following formula (Runca *et al.*, 1979)

$$v_m(t) = v_1(t)(z_m/z_1)^{\alpha(s(t))}$$

$$K^z(z, s(t)) = K^D(s(t))z \exp(-\rho(s(t))z/H)$$

$$K^D(s(t)) = z_1^{-1} K^z(z_1, s(t)) \exp(\rho(s(t))z_1/H)$$

The values of $\alpha(s(t))$ and $\rho(s(t))$ as well as of $K^x(s(t))$, $K^y(s(t))$ and $K^z(z_1, s(t))$ are given in Table 1 (Runca *et al.*, 1979).

2.3. Emission Data

The set of emission data is organized as follows:

Point emission from every high source is forecasted as one value for the whole period $(n_0, n_0 + 12)$ or for every hour $(n_0 + i, n_0 + i + 1)$, separately. It can be considered as a function of time; every source is placed to an appropriate grid point (x, y, z) , by considering also plum rise H_1 (for the method of computation of H_1 see *Instruction of the Environmental Protection Ministry*, 1981).

On the level nearest to the forecasted inversion layer base we assume the following percentage of the emission amount:

$$\tilde{E} = (H_1 - h/H_1)E100\%$$

where h is the height of the considered source and E is the instantaneous emission ($[g/sec]$). The reason for this is purely numerical; otherwise, it would be uncomfortable to introduce varying height of the last level z_4 .

Surface emission is estimated for each square $[1 \text{ km} \times 1 \text{ km}]$. Afterwards, these values are assigned to neighbouring grid points of the network. Emission stemming from traffic, treated as linear sources, is added to the surface emission.

2.4. Immission Data

The values of instantaneous concentrations at the ground in $[g/m^3]$ are obtained from several monitoring sites and are represented by the values assigned to the grid points (x, y, z_1) by the well known KRIGING interpolation method (Ripley, 1981).

The values at high altitudes can be computed from a function which can be determined empirically, using, e.g., measurements from an airplane.

The immission data are used twice in our model. First, immission is the initial value of the concentration at time $t_0 = n_0$ in our parabolic partial differential equations. Secondly, immission is the base for prediction by the filter. First, we have to identify the probability distributions of the concentrations by using statistical tests, e.g. Kolmogoroff-Smirnoff, t -Student, Fisher tests. To do this, we treat concentration as a random variable X . The observable concentrations (immission) at every monitoring site separately, are considered as the values of independent random variables X_1, \dots, X_M (M is the number of monitoring sites). When the distribution is known, we can complement the missing data using statistical methods.

2.5. Chemical Data – Assumptions Concerning Coefficients Responsible for Transformation and Deposition of Pollutants from the Atmosphere

In order to compute transformations of pollutants it is necessary to measure the concentration of nitrate and sulfate aerosol as well as of formaldehyde during day and night, separately. These compounds are products of transformation of SO_2 , NO_x and hydrocarbons, respectively. Relations between primary and secondary pollutants will give estimates for the empirical transformation coefficients.

The coefficients of dry deposition velocity can be taken from (Metcalf *et al.*, 1989 – Tab. 4, p.2036) and (Derwent *et al.*, 1989 – Tab. 4, p.1900).

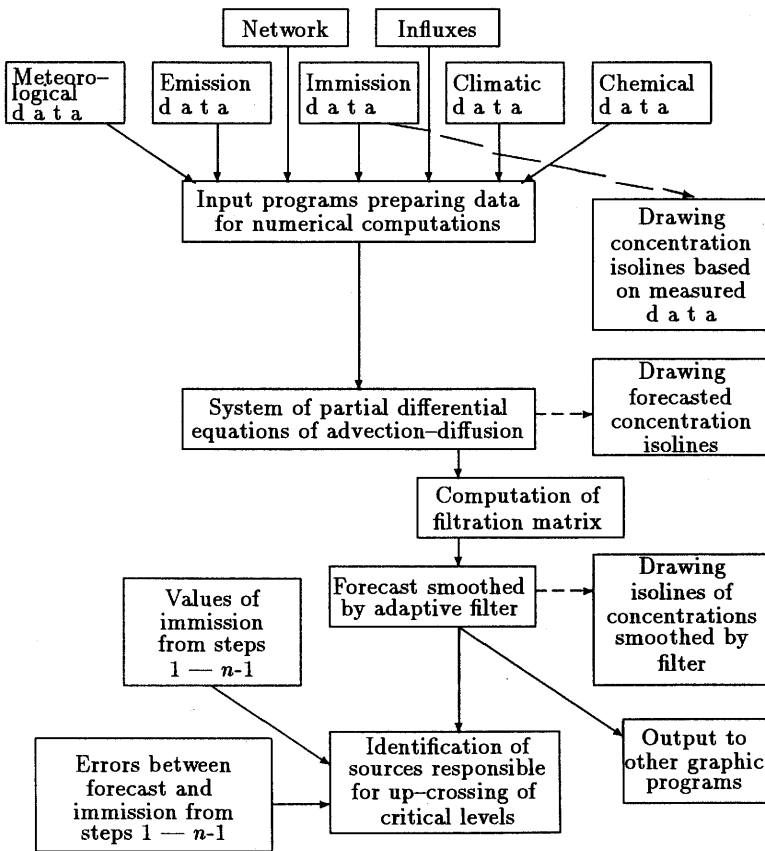
The coefficients of sulfates and nitrates wet deposition were examined in an empirical way, see e.g. (Hryniewicz, 1981; Kasina, 1981) for sulfates for Cracow area and (McMahon and Denison, 1979) for other pollutants.

2.6. Influxes

A special program for computing the forecasted influxes from beyond the investigated domain is important for preparation of input data for the model. The boundary values of concentrations treated as influx values can also be computed by the KRIGING method using measurements of immission from several monitoring sites (some of them should be located near the boundaries of the domain).

3. Block Diagram

The forecast of instantaneous concentrations made at the n -th computation step is described in the diagram below.



4. Mathematical Modeling of Evolution of the Concentration of Nitrogen Oxides and Related Air Pollutants

4.1. Chemical Reactions

The model proposed is a result of a compromise between the tendency to cover all details and the need for economy (Eykhoff, 1981). Considering the whole complicated cycle of chemical reactions that involve nitrogen oxides, we decide to use a lumped parameter model similar to that developed by Gusten (1986). It contains all basic mechanisms leading to removal of primary and secondary pollutants, as well as those responsible for build up of high diurnal concentrations, including photochemical smog build up mechanism (Kasina, 1981). Some links with the sulphur oxides cycle were taken into account. The links with the ammonium cycle, as well as with the (related) aerosol formation were not considered.

The complete list of reactions taken into account is given in Table 1, below, where the reaction rate constants are also reported from Seinfeld (1977) and Gusten (1986). The values were used in the model construction and can be employed as starting estimates. Similar sets of data can be taken from more recent publications (Chang *et al.*, 1988; Derwent *et al.*, 1989).

The concentrations of the following substances are regarded as dependent variables: O₃, NO, NO₂, NO₃, RAD, HC. RAD stands for all free radicals treated as one (lumped) group of substances (both as products and as substrata), while HC stands for all (lumped) hydrocarbons involved in the NO_x cycle. Such treatment is dictated by impossibility of gaining any detailed information about the actual hydrocarbons concentrations in the atmosphere. The concentrations of radicals will be neither traced nor measured at all. Thus, any information about their concentrations will come from the state estimation capabilities of the extended Kalman filtration procedure.

In the system (1a)–(1f) below, the concentrations of O and OH will be determined from the steady-state approximations as functions of the state variables listed above. The concentration of HNO₂ will be approximated by the Wayne and Yost method (Seinfeld, 1977). The respective formulae are as follows

$$[O] = k_1[NO_2] \{k_2 + k_4[NO_2] + k_5[NO_2] + k_7[NO] + k_{15}[HC]\}^{-1}$$

$$[OH] = k_{11}[HNO_2] \{k_{12}[NO] + k_{13}[NO_2] + k_{14}[HC]\}^{-1}$$

$$[HNO_2] = k_{11}^{-1} \{k_{12}[OH][NO] + k_{19}[NO][NO_2]\}$$

where $[x]$ denotes concentration of x while coefficients k_i are taken from Table 1. Further, the concentrations $[O_2]$, $[H]$, $[H_2O]$ and $[SO_2]$ are input data. They come with the meteorological data and as the output of the SO₂ model.

Tab. 1. Table of chemical reactions [p=parts per million, m=min]

No.	Reaction	k_i (from (Seinfeld, (1977)))	k_i (from (Gusten, (1986)))
1	$\text{NO}_2 + h\nu \rightarrow \text{NO} + \text{O}$	—	0.4 m^{-1}
2	$\text{O} (+\text{O}_2 + \text{M}) \rightarrow \text{O}_3 + \text{M}$ (M=O ₂ or N ₂)	$2.33 \times 10^{-5} \text{ p}^{-1} \text{ m}^{-1}$	$2.64 \times 10^6 \text{ m}^{-1}$ indep. of {O ₂ },
3	$\text{O}_3 + \text{NO} \rightarrow \text{NO}_2 + \text{O}_2$	$2.95 \times 10^1 \text{ p}^{-1} \text{ m}^{-1}$	$40 \text{ p}^{-1} \text{ m}^{-1}$
4	$\text{O} + \text{NO}_2 \rightarrow \text{NO} + \text{O}_2$	$1.38 \times 10^4 \text{ p}^{-1} \text{ m}^{-1}$	—
5	$\text{O} + \text{NO}_2 (+\text{M}) \rightarrow \text{NO}_3 (+\text{M})$	$4.50 \times 10^3 \text{ p}^{-1} \text{ m}^{-1}$	—
6	$\text{NO}_3 + \text{NO} \rightarrow 2\text{NO}_2$	$1.48 \times 10^4 \text{ p}^{-1} \text{ m}^{-1}$	—
7	$\text{O} + \text{NO} (+\text{M}) \rightarrow \text{NO}_2 (+\text{M})$ (M=O ₂ or N ₂)	$2.34 \times 10^3 \text{ p}^{-2} \text{ m}^{-1}$	—
8	$2\text{NO} + \text{O}_2 \rightarrow 2\text{NO}_2$	$7.62 \times 10^{-10} \text{ p}^{-2} \text{ m}^{-1}$	—
9	$\text{NO}_3 + \text{NO}_2 \rightarrow \text{N}_2\text{O}_5$	$4.43 \times 10^3 \text{ p}^{-1} \text{ m}^{-1}$	—
10	$\text{NO}_2 + \text{O}_3 \rightarrow \text{NO}_3 + \text{O}_2$	$0.46 \times 10^{-1} \text{ p}^{-1} \text{ m}^{-1}$	—
11	$\text{HNO}_2 + h\nu \rightarrow \text{NO} + \text{OH}$	—	$10^{-3} \text{ m}^{-1} \div 10^{-3} \text{ sec}^{-1}$
12	$\text{OH} \cdot + \text{NO} \rightarrow \text{HNO}_2$	$1.20 \times 10^4 \text{ p}^{-1} \text{ m}^{-1}$	$10 \text{ p}^{-1} \text{ m}^{-1}$
13	$\text{OH} \cdot + \text{NO}_2 \rightarrow \text{HNO}_3$	$1.50 \times 10^4 \text{ p}^{-1} \text{ m}^{-1}$	$30 \text{ p}^{-1} \text{ m}^{-1}$
14	$\text{OH} \cdot + \text{HC} \rightarrow 2 \left\{ \begin{array}{l} \text{RO} \left(\begin{array}{l} \text{OH} + \text{RH} \rightarrow \text{R} + \text{HO} \\ \text{R} + \text{O}_2 \rightarrow \text{RO}_2 \end{array} \right) \end{array} \right.$	—	$80 \text{ p}^{-1} \text{ m}^{-1}$
15	$\text{O} + \text{HC} \rightarrow 2 \left\{ \begin{array}{l} \text{RCO} (+\text{O}_2) \rightarrow \text{RCO}_2 \end{array} \right.$	—	$6100 \text{ p}^{-1} \text{ m}^{-1}$
16	$\text{RAD} + \text{NO} \rightarrow \text{NO}_2 + \text{radicals}$	$1.20 \times 10^4 \text{ p}^{-1} \text{ m}^{-1}$	$1500 \text{ p}^{-1} \text{ m}^{-1}$
17	RAD (ROO ₂) + NO → PAN	—	$6 \text{ p}^{-1} \text{ m}^{-1}$
18	$\text{O}_3 + \text{HC} \rightarrow \text{RO}_2$ (RAD)	—	$0.0125 \text{ p}^{-1} \text{ m}^{-1}$
19	$\text{NO} + \text{NO}_2 (+\text{H}_2\text{O}) \rightarrow 2\text{HNO}_2$	$4.30 \times 10^{-6} \text{ p}^{-2} \text{ m}^{-1}$	$0.01 \text{ p}^{-1} \text{ m}^{-1}$ (without H ₂ O)
20	$\text{SO}_2 + \text{NO}_2 \rightarrow \text{SO}_3 + \text{NO}$	$1.30 \times 10^{-4} \text{ p}^{-1} \text{ m}^{-1}$	—
21	$\text{SO}_2 + \text{O}_3 \rightarrow \text{SO}_3 + \text{O}_2$	slow in gas, not in water	—
22	$\text{SO}_2 + \text{NO}_3 \rightarrow \text{SO}_3 + \text{NO}_2$	$\leq 2.5 \times 10^{-5} \text{ p}^{-1} \text{ m}^{-1}$	—

Therefore, in the idealized analysis of the time evolution of the concentrations treated as dependent variables, we obtain the following set of differential equations (Gusten, 1986; Seinfeld, 1977):

$$\frac{d}{dt}[\text{O}_3] = k_2[\text{O}] - k_3[\text{O}_3][\text{NO}] - k_{10}[\text{O}_3][\text{NO}_2] \tag{1a}$$

$$\begin{aligned} \frac{d}{dt}[\text{NO}] = & k_1[\text{NO}_2] - k_3[\text{O}_3][\text{NO}] + k_4[\text{O}][\text{NO}_2] - k_6[\text{NO}_3][\text{NO}] - k_7[\text{NO}][\text{O}] \\ & + k_{11}[\text{HNO}_2] - k_{12}[\text{OH}][\text{NO}] - k_{16}[\text{RAD}][\text{NO}] - k_{19}[\text{NO}][\text{NO}_2] - k_{20}[\text{SO}_2][\text{NO}_2] \end{aligned} \quad (1b)$$

$$\begin{aligned} \frac{d}{dt}[\text{NO}_2] = & -k_1[\text{NO}_2] + k_3[\text{O}_3][\text{NO}] - k_4[\text{O}][\text{NO}_2] + 2k_6[\text{NO}_3][\text{NO}] + k_7[\text{O}][\text{NO}] \\ & + 2k_8[\text{NO}] - k_9[\text{NO}_2][\text{NO}_3] - k_{10}[\text{NO}_2][\text{O}_3] - k_{13}[\text{OH}][\text{NO}_2] \\ & - k_{17}[\text{RAD}][\text{NO}_2] - k_{19}[\text{NO}][\text{NO}_2] - k_{20}[\text{SO}_2][\text{NO}_2] + k_{22}[\text{SO}_2][\text{NO}_2] \end{aligned} \quad (1c)$$

$$\frac{d}{dt}[\text{NO}_3] = k_5[\text{O}][\text{NO}_2] - k_6[\text{NO}_3][\text{NO}] - k_9[\text{NO}_3][\text{NO}_2] - k_{10}[\text{NO}_2][\text{O}_3] - k_{22}[\text{SO}_2][\text{NO}_2] \quad (1d)$$

$$\frac{d}{dt}[\text{RAD}] = k_{14}[\text{OH}][\text{HC}] + k_{15}[\text{O}][\text{HC}] - k_{16}[\text{RAD}][\text{NO}] - k_{17}[\text{RAD}][\text{NO}_2] + k_{18}[\text{O}_3][\text{HC}] \quad (1e)$$

$$\frac{d}{dt}[\text{HC}] = k_{14}[\text{HC}][\text{OH}] - k_{15}[\text{O}][\text{HC}] - k_{18}[\text{O}_3][\text{HC}] \quad (1f)$$

The above equations describe an idealized air parcel under stable conditions. However, not only the conditions and rate constants do vary, but also all other phenomena such as transport, mixing, diffusion, turbulent diffusion, wet removal, deposition etc. influence those conditions. Therefore, the above equations (and similar equations concerning other substances and other reactions) are only used to make an initial selection and choice of the most important chemical reactions and groups of species.

4.2. Physical Phenomena (Equation of Particle Motion)

Let $\tilde{I} = \{0, 1, \dots, 6\}$ be the set of indices (types of particles) that are assigned to atoms observed in the atmosphere. The set of all atoms will be denoted by Ω . An atom $\omega \in \Omega$ will be assigned index 1 if it is contained in an O_3 particle, indices 2, 3, 4, 5, 6 if it is in an NO , NO_2 , NO_3 , RAD or HC particle, respectively. Index 0 is assigned to atoms in all other types of particles. The index assigned to an atom ω at time t will be denoted by $j(\omega, t)$ (the atom ω at time t is in a particle of type $j(\omega, t)$).

Let $X_t^i(\omega)$ denote the position (in Cartesian coordinates) of the atom ω at time t . Define

$$X_t^i(\omega) = \begin{cases} X_t(\omega) & \text{if } j(\omega, t) = i \\ \infty & \text{if } j(\omega, t) \neq i \end{cases} \quad i = 0, \dots, 6$$

The situation can be described in the following way: there are seven observers tracing the trajectory of an atom ω . Each of them can only see ω in one type of particle, say i -th, and he can then determine its position $X_t^i(\omega)$. If ω disappears from the i -th observer's "world" (by entering another type of particle), the observer defines $X_t^i(\omega)$ to be ∞ .

In order to describe the motion of an atom joining consecutive particles of various kinds one has to use several models. The actual model in use varies in time since the atoms

may enter different types of particles. This can be effectively treated by the theory of systems with randomly varying structure (Kazakov and Artemeev, 1980; Kazimierczyk, 1989).

Let $i \in \tilde{I}$ be any type of particles. The motion of an i -th type particle will be modeled by the following Itô stochastic differential equation

$$dX_t^i(\omega) = V(i, t, \omega, X_t^i(\omega)) dt + B(i, t, \omega, X_t^i(\omega)) dW(i, \omega, t) + dC(i, t, \omega, X_t^i(\omega)) + dA(i, t, \omega, X_t^i(\omega)) \tag{2}$$

Here $V(i, t, \omega, X_t^i(\omega))$ is a random advection velocity vector of a particle that is at the point $X_t^i(\omega)$ at time t (thus, containing the atom ω). The horizontal components of V are essentially the same for all types of particles – they are equal to the horizontal components of the wind velocity. The vertical component of V , however, is the sum of the vertical component of the wind velocity, and the velocity of dry deposition. For the initial values of dry deposition velocities for different species see §2.5. It should be stressed that these values are treated as starting estimates, and that the model is capable of improving these estimates during its operation.

In the first approximation we will assume that the velocities of dry deposition are the only nonzero parts of the vertical drift (advection). This assumption results from the lack of data concerning the vertical components of the wind velocity. The second term in the right side of equation (2) is responsible for diffusion (we will interpret this term as the eddy diffusivity term). To simplify matters we assume that the 3×3 matrix $B(i, t, \omega, X_t^i(\omega))$ is diagonal ($B = \sqrt{2} \text{diag}\{B^x, B^y, B^z\}$), where the diagonal elements are random fields whose mean values are the classically determined eddy diffusivity coefficients, assumed to be equal for all $i \in \tilde{I}$. Next, $W(i, t, \omega)$ is the standard three-dimensional Brownian motion for each $i \in \tilde{I}$. We assume that for $i \neq j$ the processes $W(i, t, \omega)$ and $W(j, t, \omega)$ are statistically independent (Ikeda and Watanabe, 1981).

The dC term in equation (2) describes the creation, and the dA term – the annihilation of atoms in the i -th model. That means that whenever the i -th observer notices the atom ω in a type- i -particle, then the state $X_t^i(\omega)$ is transferred by dC from ∞ to the point $X_t^i(\omega)$, at which the atom has been noticed. Analogously, if the atom disappears (it is drawn out of the region of interest, or it joins a particle of another type due to a chemical reaction), then the observer assigns the value ∞ to the process $X_t^i(\omega)$, by defining the annihilation process to have an infinite increment.

The processes leading to creation of particles are, firstly, immission through the boundary; secondly, emission from point, linear and area sources; thirdly, production due to chemical reactions. The process $C(i, \cdot, \cdot)$ is the sum of the respective components, assumed to be statistically independent of each other, Markovian, and progressively measurable with respect to the Brownian motion $W(i, \cdot, \cdot)$. The components are defined via their Markovian probability distributions, whose type will be identified and whose shapes will be parametrically introduced into the model on the basis of actual measurements and envisaged data and, finally, from influxes. The annihilation terms are the following: removal due to chemical reactions, by rainout and washout, out-crossing the district

boundary, sedimentation. The first two processes are W -independent, W -progressively measurable and Markovian and are defined through their intensities (Ikeda and Watanabe, 1981). The third and fourth processes are, by definition, dependent on the moment of reaching the boundary, and so they are $X_t^i(\omega)$ -measurable Markov processes with the special feature that the boundary layer is a partially absorbing, partially reflecting boundary (depending on the dry deposition velocity).

Therefore, the atoms observed can be categorized according to their macro-scale or micro-scale movements. The main assumption is that these classifications are *mutually independent*. Such random fields like wind velocity, eddy diffusivity coefficients, chemical reaction rates etc. are assumed to be *slowly* varying in space and time. On the other hand, the diffusion itself, the chemical reactions and border crossings by (Brownian-like) particles are short range processes. Thus, the set Ω of all atoms can be regarded as the direct product of two probability spaces

$$\{\Omega, \mathcal{F}, \mathcal{P}\} = \{\Omega_1 \times \Omega_2, \mathcal{F}_1 \times \mathcal{F}_2, \mathcal{P}_1 \times \mathcal{P}_2\}$$

where $\{\Omega_1, \mathcal{F}_1, \mathcal{P}_1\}$ is a macro-scale probability space, and $\{\Omega_2, \mathcal{F}_2, \mathcal{P}_2\}$ is a micro-scale probability space. We assume that the processes V , B , and the intensities defining the processes C and A depend only on ω_1 (the intensities are regarded as local averages and so they are independent of the micro-scale events). We also assume that the Brownian motion processes W are functions of ω_2 exclusively.

4.3. Equations for Concentrations (the Basic Model)

We assume that the advection field V and the eddy diffusivity coefficients B are \mathcal{P}_1 -almost surely jointly globally Lipschitzean. Then for almost all $\omega_1 \in \mathcal{P}_1$ the above assumptions allow us to consider equation (1a)–(1f) as an Itô system with creation and annihilation processes. The application of the theory of systems with variable structure leads to the following generalized Fokker–Planck–Kolmogoroff equations for conditional densities of the processes $X_t^i(\omega)$:

$$\begin{aligned} \frac{\partial}{\partial t} p(i, t, x, \omega_1) &= \mathcal{L}(i, t, x, \omega_1) p(i, t, x, \omega_1) \\ &- a(i, t, x, \omega_1, p^1, \dots, p^6) + c(i, t, x, \omega_1, p^1, \dots, p^6) \end{aligned} \quad (3)$$

with the initial conditions $p(i, 0, x, \omega_1) = p_0(i, x, \omega_1)$ and with the boundary conditions taken into account via sink and production terms a and c (for definitions see below). Here $p^i = p(i, t, x, \omega_1)$ is the density of the Lebesgue-regular part of the conditional probability that the considered atom is seen at the point x at time t , provided that it is in an i -th type particle and was subjected to the macroscopic history ω_1 . \mathcal{L} is a differential operator

$$\begin{aligned} \mathcal{L}(i, t, x, \omega_1) &= \langle \mathcal{L}(i, t, x, \omega_1, \omega_2) \rangle_2 \\ &= V^T(i, t, x, \omega_1) \nabla_x + \nabla_x^T \text{diag} \{K^x, K^y, K^z\}(i, t, x, \omega_1) \nabla_x \end{aligned}$$

where $\langle \cdot \rangle_2$ denotes averaging with respect to the measure \mathcal{P}_2 , that is, with respect to microscopic fluctuations, while A^T denotes transposition of A ; $p(0, t, \cdot, \omega_1) + \dots +$

$p(6, t, \cdot, \omega_1) = 1$ (which means that system (3) can be regarded as a system of *mass conservation* equations). Directly from the definition, the quantity $D \cdot 10^6 \cdot p(i, t, x, \omega_1)$, where D is the volume of the spatial domain above the considered area, can be understood as the instantaneous random concentration (in ppm) of particles of type i at the point (t, x) .

The terms a and c are sums of terms of the following form

$$\delta(\Gamma(i, x) - g(i))\nu(i, t, x, \omega_1)$$

(where $\delta(x) = \infty$ if $x = 0$ and 0 otherwise, δ is the Dirac's delta function)

- describing creation (with $\nu = \mu(i, t, x, \omega_1)$) or annihilation (with $\nu = \mu(i, t, x, \omega_1)p(i, t, x, \omega_1)$) on the surface $\Gamma(i, x) = g(i)$, with intensity ν ;

$$\delta(\Gamma_1(i, x) - g_1(i)) \delta(\Gamma_2(i, x) - g_2(i))$$

- describing creation at the intersection of two surfaces (along a line source);

$$\delta(\Gamma_1(i, x) - g_1(i)) \delta(\Gamma_2(i, x) - g_2(i)) \delta(\Gamma_3(i, x) - g_3(i)) \nu(i, x, \omega_1)$$

- describing creation (due to emission) at a point source;

$$p(i, t, x, \omega_1) \nu(i, t, x, \omega_1)$$

- describing removal due to rainout with ν being the intensity of rain multiplied by a factor characteristic for i -th type particles.

Here $\Gamma(i, x)$ and $\Gamma_j(i, x)$, $j = 1, 2, 3$ are the boundary surfaces in the state space, on which the state changes take place. These surfaces are given by equations $\Gamma(i, x) - g(i) = 0$ or $\Gamma_j(i, x) - g_j(i) = 0$, $j = 1, 2, 3$, where $g(i)$ and $g_j(i)$, $j = 1, 2, 3$ are known scalar values. (For the exact definitions of these surfaces see Kazakov and Artemeev, 1980).

The remaining parts of the processes a and c are just the right hand sides of system (1a)–(1f) describing the removal and reproduction of particular species. In these equations the concentrations $[\cdot]$ are just replaced by p and the rate "constants" k should be regarded as randomly fluctuating fields $k(i, t, x, \omega_1)$.

So we have the following system of coupled equations

$$\begin{aligned} \frac{\partial}{\partial t} p(i, t, x, \omega_1) &= \langle \mathcal{L}(i, t, x) \rangle_1 p(i, t, x, \omega_1) - \mathcal{A}(i, t, x, p^1, \dots, p^6) \\ &+ \mathcal{C}(i, t, x, p^1, \dots, p^6) + \mathcal{R}(i, t, x, \omega_1) \end{aligned} \tag{4}$$

with the same initial conditions as in (3), where

$$\mathcal{R}(i, t, x, \omega_1) = (\mathcal{L} - \langle \mathcal{L} \rangle_1) p - (a - \mathcal{A}) + (a - \mathcal{C}) \tag{5}$$

is regarded as the random external loading, with \mathcal{A} and \mathcal{C} formed from a and c by replacing all fluctuating intensities (dependent on ω_1) by their averages over the micro scale space Ω_2 . Here the symbol $\langle \mathcal{L} \rangle_1$ denotes averaging with respect to the measure \mathcal{P}_1 , that is, with respect to all possible macroscopic observations.

System (4) is our basic model. Residual part (5) is of stochastic nature and is an aggregate of many random factors. It is thus reasonable to assume a special form of the process \mathcal{R} to assure efficiency of the method. The main idea is to assume as much as possible about \mathcal{R} in order to simplify the method of estimation of both the state p and model parameters on the basis of observed concentrations at few points of the region.

Since for all t and x we have $p(0, t, x) \approx 1/D$, the equation for particles of type 0 can be neglected.

4.4. Discretized and Linearized Model

In order to make use of an efficient filtration algorithm the basic model (4) has to be discretized. Our discretization method is the same as for the sulphur dioxide model as far as the spatial variables are concerned. It is based on the method of fractional steps and the Crank – Nicolson (Crank and Nicolson, 1974) method. Discretization with respect to time should be performed according to Gear's method (Willoughby, 1974; Kreiss, 1978). This is because of the different reaction times in different equations of system (3). Of course, it can be done after checking whether the system is stiff.

The new feature is that we deal with a *system* of coupled equations and that the chemical production and sink terms coupling the consecutive equations are highly *non-linear*.

Therefore, we apply the method of fractional steps. The first step concerns the nonlinearities. In this way, we arrive at the system of equations for the vector of concentrations

$$\mathcal{X}(k) = [p(0, t_k, x_{vuw}), p(1, t_k, x_{lmn}), \dots, p(6, t_k, x_{qrs})]$$

of the following form

$$\mathcal{X}(k+1) = f_k(\mathcal{X}(k)) + \mathcal{R}_k(\omega_1) \quad (6)$$

The basis for the estimation of the state $\mathcal{X}(k)$ and for parameter validation is the observation of the concentrations at few nodal points

$$Z(k) = H^T \mathcal{X}(k) + v(k), \quad (7)$$

where H is a matrix of zeros and ones (zero when there is no monitoring site at the nodal point and one otherwise), while $v(k)$ is a measurement noise vector. Clearly, \mathcal{R}_k is treated as the state equation noise. \mathcal{R}_k and $v(k)$ are independent Gaussian white noises.

The extended Kalman filtering theory will serve as a tool for estimating the state vector of concentrations. The main direction of investigations will be the linearized version of system (6), (7).

5. Application of Adaptive Filter to Forecasting the Distribution of Air Pollution Concentrations

5.1. Preliminaries

Let us consider a stochastic process $\{(x_t, y_t); t \in I\}$, taking values in $\mathbb{R}^n \times \mathbb{R}^j$, defined for $t \in I$, where I is a subset of the set of integers \mathbb{Z} or real numbers \mathbb{R} .

Suppose that from the knowledge of the so called *observable process* $\{y_t; t \in I\}$ up to $s \in I$, one would like to determine, as accurately as possible, the value of the process $\{x_t; t \in I\}$ at time $s + u$. If $u > 0$, then the problem is referred to, as a *prediction problem*; if $u = 0$ then one talks about *filtration*; finally, if $u < 0$, then one talks about *smoothing* of the unobservable process $\{x_t; t \in I\}$.

In our case the observable process y_t is given by measurements of immission concentrations, and the unobservable process x_t is the concentration of pollutants.

We will assume that \hat{x}_s (the estimator of x_s) is measurable with respect to the σ -field \mathcal{Y}_s , where we have set $\mathcal{Y}_s = \sigma\{y_t; t \leq s\}$.

Our filtration problem will be considered solved, if we find the conditional distribution $P(A|\mathcal{Y}_s)$, for $A \in \mathcal{B}(\mathbb{R}^n)$. We recall that $P(A|\mathcal{Y}_s) \stackrel{\text{df}}{=} P\{x_s \in A|\mathcal{Y}_s\}$.

On the next few pages we will show how the problem of filtration is solved in the case when state and observation processes are linear, and all noises are Gaussian. This is the most important case. In more complicated cases we will refer the reader to the literature.

5.2. Filtration of Linear Systems with Discrete Time

Consider the system of recursive equations

$$\begin{aligned} x_{t+1} &= \Phi_t x_t + v_t & (8) \\ y_t &= \Theta_t x_t + e_t & (9) \end{aligned} \quad t \in [t_0, \dots, T]$$

where Φ_t and Θ_t are $n \times n$ and $j \times n$ matrices, respectively, and $\{e_{t_0}, e_{t_0+1}, \dots, e_T, v_{t_0}, v_{t_0+1}, \dots, v_T\}$ is a sequence of independent Gaussian random vectors with known covariance matrices

$$\begin{aligned} E\{v_t v_t^T\} &= R_1(t), & E(v_t) &= 0 \\ E\{e_t e_t^T\} &= R_2(t), & E(e_t) &= 0 \end{aligned}$$

The random variable x_{t_0} is also Gaussian with known mean value $E(x_{t_0}) = m$, and with covariance matrix $E\{(x_{t_0} - m)(x_{t_0} - m)^T\} = R_0$. We also assume that x_{t_0} is independent of $\{v_t, e_s; s, t \in [t_0, T]\}$. The matrices Φ_t, Θ_t depend only on t . We set $\mathcal{Y}_s = \sigma\{y_{t_0}, \dots, y_s\}$ for $s \in [t_0, T]$.

We will also need the following notation

$$\begin{aligned} \hat{x}_s &= E(x_s|\mathcal{Y}_s), & P_s &= E\{(x_s - \hat{x}_s)(x_s - \hat{x}_s)^T|\mathcal{Y}_s\} \\ S_s &= E\{(x_s - E\{x_s|\mathcal{Y}_{s-1}\})(x_s - E\{x_s|\mathcal{Y}_{s-1}\})^T|\mathcal{Y}_{s-1}\}. \end{aligned}$$

Theorem 1. *The random process $\{\hat{x}_t; t \in [t_0, T]\}$ satisfies the following system of equations (Jazwinski, 1969; Anderson and Moore, 1979)*

$$\hat{x}_{t+1} = \Phi_t \hat{x}_t + K_{t+1}(y_{t+1} - \Theta_{t+1} \Phi_t \hat{x}_t), \quad t \in [t_0, T - 1],$$

with initial condition $\hat{x}_{t_0} = E\{x_{t_0} | y_{t_0}\}$. The matrix K_t depends only on time.

Theorem 2. *The matrices P, K, S satisfy the following system of recurrent relations (Jazwinski, 1969; Anderson and Moore, 1979)*

$$P_{t+1} = S_{t+1} - K_{t+1} \Theta_{t+1} S_{t+1}, \quad T > t \geq t_0$$

$$S_{t+1} = \Phi_t P_t \Phi_t^T + R_1(t), \quad T > t \geq t_0$$

$$K_{t+1} = S_{t+1} \Theta_{t+1}^T [\Theta_{t+1} S_{t+1} \Theta_{t+1}^T + R_2(t + 1)]^{-1}, \quad T > t \geq t_0$$

with initial condition $S_{t_0} = R_0$.

It has to be underlined that this variant of filtration algorithm was applied for the SO₂ model used in Silesian District (Twardowska, 1989).

5.3. Filtration of Non-Gaussian Systems

Relaxing the assumptions of the previous section usually causes many difficulties and complications in the filtering algorithm. Nevertheless, if the disturbances $\{e_t\}$ and $\{v_t\}$ are jointly *elliptically contoured* (Szablowski, 1989), and other assumptions on systems (8) and (9) remain in force, then the resulting filter does not differ substantially from the Kalman filter. The first of the equations of Theorem 1 remains the same. The difference lies in the fact that the matrix S_s is now random. More precisely, it is of the form: P_k times some random variable depending on k . The simplicity of this case lies in the fact that this sequence of random variables can be approximated by the sequence $\{S_k^2\}$ generated by the following recurrent relationship (Szablowski, 1989)

$$S_k^2 = (1 - 1/k)S_{k-1}^2 + (1/k)(y_k - \Theta_k \hat{x}_k)^T [\Theta_k P_k \Theta_k^T + R_2(k)]^{-1} (y_k - \Theta_k \hat{x}_k)$$

5.4. Filtration of Nonlinear and Non-Gaussian Systems

For the nonlinear and non-Gaussian case, the computational difficulties become enormous. Presenting them in this short note would take too much space. That is why we refer the reader to say monograph of Anderson and Moore (1979).

5.5. Processes with Varying Structure

Processes with varying structure are described by a vector of $s + 1$ variables (s is the number of different structures); s of them take values in \mathbb{R}^n , and one is discrete, and describes the actually utilized structure.

Each of the continuous variables satisfies a system of stochastic differential equations. The discrete random variable (structure) usually forms a Markov chain with transition probabilities depending, in general, on the values of the state vector describing the current

structure. It is usually assumed that the observations are linearly related to the state vectors, and that the structure process is observable.

All these remarks lead to the conclusion that filtering of systems with varying structure is a very complicated task. It leads to a system of integro-partial differential equations which can be solved analytically only in very special cases. Hence, various methods of approximating the exact solution have been developed. The most recommended (Kazakov and Artemeev, 1980) are those based on Gaussian approximation of a posteriori (conditional) transition probability density of the state vectors, given the observations and the current structure. The idea of these methods was already described in the previous section. Presenting the precise approximation algorithm would be out of the scope of this paper, the more that we intend to resort to the varying structure filtration technique only if all the previously described techniques prove to be insufficiently powerful.

Let us remind that, although the starting point to our modeling was the varying structure equation of particle motion, it is not the particle positions which is the aim of our filtration algorithm. What we look for are the concentrations of the pollutants at the nodal points. The basic equation (4) for the concentrations, is no longer a model of varying structure. Therefore, the varying structure systems theory reenters our analysis not because our model (4) has varying structure, but because this system is highly nonlinear and the theory offers a highly sophisticated approximative method for the filtration of such systems.

Besides its nonlinearity, system (4) has one more specific feature: since the dependent functions $p_i, i = 1, \dots, 6$ are the probability densities, therefore they remain nonnegative and together with p_0 are summable over the domain D to unity. All these features can be taken care of in the following way.

As in section 4.4 we define the state variable $\mathcal{X}(t) = (\chi_i(t)) = [p_0(t, x_{vuw}), \dots, p_1(t, x_{lmn}), \dots, p_6(t, x_{qrs})]^T$. This time we will discretize model (4) with respect to the space variables only. Thus, system (6) will be replaced by

$$\dot{\mathcal{X}}(t) = f(t, \mathcal{X}(t)) + \mathcal{R}(t, \omega). \tag{10}$$

Here $\mathcal{X}(t) \in \mathcal{D} = \{\mathcal{X} = (\chi_1, \dots, \chi_n) : \chi_j \geq 0, \sum_{j=1}^n \chi_j \leq 1\}$.

Let us split $\partial\mathcal{D}$ into $\mathcal{S}_0 \cup \mathcal{S}_1$, where $\mathcal{S}_1 = \{x \in \partial\mathcal{D} : x_j > 0\}$, $\mathcal{S}_0 = \partial\mathcal{D} \setminus \mathcal{S}_1$. To assure that the boundary conditions $x_j \geq 0$ are satisfied one can embed the boundary \mathcal{S}_0 with the ε -layer $\mathcal{D}_0 \subset \mathcal{D}$ (of the thickness ε), in which model (10) is modified in such a way, that f is replaced by $F_0^\varepsilon(t)\mathcal{X}(t)$, where $F_0^\varepsilon(t)$ is a square matrix with rows F_{0j}^ε satisfying $F_{0j}^\varepsilon\mathcal{X}(t) \geq 0$ for j s.t. $x_j(t) < \varepsilon$. Further, the boundary \mathcal{S}_0 in \mathcal{D}_0 is considered reflecting. Similarly, to assure that $\sum_{j=1}^n \chi_j \leq 1$ one can embed the boundary \mathcal{S}_1 with an ε -layer $\mathcal{D}_1 \subset \mathcal{D}$ (the technicalities concerning the corners and the border between \mathcal{D}_1 and \mathcal{D}_0 are left aside here). Again, function f of (N') is replaced in \mathcal{D}_1 by $F_1^\varepsilon(t)\mathcal{X}(t)$, such that $[1, \dots, 1]^T F_1^\varepsilon(t)\mathcal{X}(t) \leq 0$ for $x \in \mathcal{D}_1$. \mathcal{S}_1 is considered a reflecting boundary.

To take care of the nonlinearities, the remaining part of the domain \mathcal{D} is divided into disjoint classes $\mathcal{D}_j, j = 2, \dots, K$ in each of which function f from (10) is linearized

with respect to \mathcal{X} and it is replaced by $F_j(t)\mathcal{X}(t)$. All these modifications are made in such a way as to preserve continuity of the piecewise linear substitute of function f , as well as to allow numerically effective and qualitatively correct approximation (since f and F_j depend on time, one has to look for a fast and robust algorithm, rather than for the one that would be sophisticated and optimal with respect to distributions). This can be achieved by defining $F_j(t)$ in consecutive cubes via equating the values of the nodal points.

In this way one arrives at $K + 1$ different linear models

$$\dot{\mathcal{X}}_i^{(j)} = F_j(t)\mathcal{X}_i^{(j)} + \mathcal{R}^{(j)}(t) \tag{11}$$

which can be treated as a varying structure system by adding in-and-out fluxes through the boundaries between the cubes (processes of creation and annihilation).

The filtration for system (11) consists of constructing the *a posteriori* estimates $\hat{\mathcal{X}}_k$ of the actual system state \mathcal{X}_i^{jact} based upon the discrete-time observation $Z_n = [Z(0), \dots, Z(n)]$, where

$$Z(n) = H\mathcal{X}_n^{jact} + V_n \tag{12}$$

Assuming that $\mathcal{R}^j(t)$ are white noises with known intensities $\mathcal{R}^j(t) dt = D^j(t) dW^j(t)$, and that the measurement noise V_n is white with $EV_n V_n^T = \delta_{nk} Q_n$, one can arrive at the following estimates (Kozakov and Artemeev, 1980):

$$\hat{\mathcal{X}}_k = E(\mathcal{X} | Z_k) = \int_{\mathcal{D}} x \hat{\omega}_k dx.$$

Here $\hat{\omega}_k(x) = \sum_{j=0}^K \hat{\omega}_K^{(j)}(x)$ and the *a posteriori* densities can be determined from the iterative equation (for $j = 0, 1, \dots, K$)

$$\begin{aligned} \hat{\omega}_{n+1}^{(j)}(x) &= \hat{\omega}_n^{(j)}(x) - \Delta t \{ \text{div} \hat{\Pi}^{(j)}(x, t_n) + \sum_{r=0}^K [\hat{v}_{jr}(x, t_n) - \hat{u}_{rj}(x, t_n)] \} \\ &\quad - \frac{1}{2} \Delta t \hat{\omega}_n^{(j)}(x) [f_n(x, z) - \sum_{r=0}^K \int_{-\infty}^{\infty} f_n(y, z) \hat{\omega}_n^{(r)}(y) dy] \end{aligned}$$

where

$$\hat{\Pi}_n^{(j)}(x, t_n) = F_j(t_n) \hat{\omega}_n^{(j)}(x) - \frac{1}{2} \sum_{\gamma=1}^M \frac{\partial}{\partial x_\gamma} [D_{k\gamma}^j(t) \hat{\omega}_n^{(j)}(x, t_n)]$$

$$\hat{v}_{jr}(x, t_n) = \sum_{k=1}^M \delta(x_k - \alpha_k^{jr}) s_{jr}^k \hat{\Pi}_k^{(j)}(x, t_n)$$

($s_{jr}^k = +1$, if by decreasing variable x_k below α_k^{jr} one leaves region r and enters region j and $s_{jr}^k = -1$ if it is obtained by increasing variable x_k ; otherwise $s_{jr}^k = 0$). Moreover, $\hat{u}_{rj}(x, t_n) = \hat{v}_{jr}(x, t_n)$.

Finally, expression $f_n(x, z)$ is defined as follows

$$f_n(x, z) = [Z(n) - H\mathcal{X}]^T Q_n^{-1} [Z(n) - H\mathcal{X}]$$

Unfortunately, the above equations are too complicated to develop a direct analytical recursive algorithm for estimates $\hat{\mathcal{X}}_k$. Therefore, the computational complexity of the above approach justifies the use of simpler methods and resorting to the varying structure filtration technique only in the last instance.

6. Parameter Identification and Model Validation

6.1. Identification of Model Parameters

Along with filtration, adaptive identification of parameters entering the model will be performed. We shall only describe the method for a linearized model. The same ideas can be applied to the nonlinear model. The difference is that the likelihood functional cannot be formed so easily in the latter case, and another criterion (based on an approximation of the likelihood functional) would have to be employed.

Let us first consider the case of identification of coefficients of the observable quantities. In this case model (6), (7) can be rewritten in the form

$$x_{k+1} = Fx_k + Gy_k + Kv_k + u_k$$

$$y_k = \theta^T x_k$$

$$z_k = y_k + v_k$$

where θ is the vector of parameters to be identified, F, G and K – some matrices, u – the external input (it can be a function of the observation z). For the above simple system, filtration and parameter identification can be performed simultaneously. The estimate of the state is based on the current value of the estimate of the parameter, and, conversely, the estimate of the parameter is calculated from the state estimate (Eykhoff, 1981)

$$\hat{x}_{k+1} = F\hat{x}_k + G\hat{y}_k + K\hat{v}_k + u_k$$

$$\hat{v}_k = z_k - \hat{y}_k$$

$$\hat{y}_k = \hat{\theta}_k^T \hat{x}_k$$

$$\hat{\theta}_k = \hat{\theta}_{k-1} - \Lambda_{k+1} \hat{x}_k (z_k^T - \hat{x}_k^T \hat{\theta}_{k-1})$$

where $\Lambda_{k+1} = \Lambda_k - \Lambda_k \hat{x}_k (\hat{x}_k^T \Lambda_k \hat{x}_k + I)^{-1} \hat{x}_k^T \Lambda_k$ and I denotes identity matrix.

A more general algorithm may be developed (if necessary) based on the extended Kalman filtration algorithm described in the previous section. Another general method leads to a prediction error algorithm resulting from the solution of an optimization problem of the form

$$g \left(\sum_k Q_k (z_k - z_{k|k-1}, \theta) \right) \rightarrow \min \tag{13}$$

where θ may be formed of all matrices F, G, H and the covariance matrices of the noises u and v . Q 's are positive definite, matrix valued functions of their arguments and g is a convex real function. In particular, for linear systems the likelihood functional takes this form (Kazakov and Artemeev, 1980).

6.2. Identification of Additional Models for Missing Meteorological Data

Several simple models for basic meteorological characteristics have been developed. Based on ten years' hourly observations of wind speed and direction, humidity, temperature and solar radiation, the basic correlations have been estimated and a coupled system of ARMA models for these quantities for each month of a reference year have been proposed (Bonivento *et al.*, 1980). Using these tools the missing hourly data can be estimated according to the local climate characteristics.

References

- Alcamo J. and Bartnicki J. (1985): *An approach to uncertainty of a long range air pollutant transport model*. — Int. Inst. for Applied Systems Analysis, A-2361 Laxemburg, Austria.
- Anderson B.D.O and Moore J.B. (1979): *Optimal Filtering*. — New Jersey: Prentice Hall.
- Bankoff S. and Hanzevack E.L. (1975): *The adaptive filtering transport model for prediction and control of pollutant concentration in an urban airshed*. — Atmosph. Environ., v.9, No.9, pp.793–808.
- Bonivento C., Fronza G., Tonielli A. and Spirito A. (1980): *The Use of Stochastic Models in Air Pollution Prediction and Control: Problems and Numerical Technics*. — Amsterdam: North Holland Publ.Co.
- Box G.E. and Jenkins G.M. (1970): *Time-Series Analysis, Forecasting and Control*. — San Francisco: Holden-Day, Inc.
- Chang Y., Carmichael G., Kurita H. and Veda H. (1989): *The transport and formation of sulfates and nitrates in central Japan*. — Atmosph. Environ., v.23, No.8, pp.1749–1773.
- Crank J. and Nicolson P. (1947): *A practical method for numerical solution of partial differential equations of heat conduction type*. — Proc. Cambridge Philos. Soc., v.43, No.1, pp.50–67.
- Derwent R.G., Hov O., Asman W.A.H., van Jaarsveld J.A. and de Leeuw F.A.A.M. (1989): *An intercomparison of long-term atmospheric transport models; the budget of acidifying species for the Netherlands*. — Atmosph. Environ., v.23, No.9, pp.1893–1909.
- Derwent R.G., Pollard G.J and Metcalfe S.E. (1988): *On the nitrogen budget for the United Kingdom and North-West Europe*. — Q.Jl R. Met.Soc., v.114, Part A, No.482, pp.1127–1152.
- Eykhoff P. (Ed.) (1981): *Trends and Progress in System Identification*. — London: Pergamon Press.
- Finzi G., Bonelli P. and Bacci G. (1984): *A stochastic model of surface wind speed for air quality control purposes*. — J. Climate and Applied Meteorology, v.23, No 9, pp.1354–1361.
- Finzi G., Garcia R. and Hernandez E. (1983): *The SO₂ pollution in Madrid, II. – A comparison between two stochastic models for real-time forecast purpose*. — II Nouvo Cimento, v.6C, No.6, pp.605–617.

- Finzi G. and Tebaldi G. (1982): *A mathematical model for air pollution forecast and alarm in an urban area*. — Atmosph. Environ., v.16, No.9, pp.2055–2059.
- Gusten H. (1986): *Formation, transport and control of photochemical smog*. — In: O. Hutzinger (Ed.), Handbook of Env. Chem., v.4, Part A, Air Pollution, S–V., pp.53–105, Berlin: Springer-Verlag.
- Hlawiczka S. (1985): *Transformations of Nitrogen Oxides in the Atmosphere*. — Inst. of Environ., Protection Report, Katowice, Poland (in Polish).
- Hryniewicz R. (1981): *Removal of some pollutants from the atmosphere with precipitation*. — Inst. of Meteorology and Water Management, Warsaw, (in Polish).
- Ikeda N. and Watanabe S. (1981): *Stochastic Differential Equations and Diffusion Processes*. — Kodansha: North-Holland Publ. Co.
- Instructions of the Environmental Protection Ministry* (1981): — No.359, Warsaw, (in Polish).
- Jazwinski A. (1969): *Adaptive filtering*. — Automatica, v.5, No.4, pp.475–485.
- Juda K. (1986): *Modeling of the air pollution in the Cracow area*. — Atmosph. Environ., v.20, No.12, pp.2449–2458.
- Juda-Kuczka K. (1986): *A Three Level Transport Model for the Dispersal of Atmospheric Pollutants in the Urban Area*. — Ph.D Thesis, Warsaw University of Technology, Poland, (in Polish).
- Kasina S. (1981): *Transport, transformations and removal of sulphur compounds from the atmosphere*. — Inst. of Environmental Formation, Warsaw, (in Polish).
- Kazakov I.E. and Artemeev V.M. (1980): *Optimization of Dynamical Systems with Randomly Varying Structure*. — Moskwa: Mir, (in Russian).
- Kazimierczyk P. (1990): *Parameter identification of diffusions with constraints*. — In: Stochastic Methods in Experimental Sciences, World Sci. Publ., Singapore, pp.249–264.
- Kreiss H.O. (1978): *Difference methods for stiff ordinary differential equations*. — SIAM J. Numer. Anal., v.15, No.1, pp.21–58.
- Korbicz J. and Zgurovsky M. Z. (1987): *State and parameter estimation of stochastic distributed parameter systems on base of combined method with application to air pollution process*. — Proc. IMACS/IFAC Int. Symp. Modelling and Simulation of Distributed Parameter Systems, Hiroshima, Japan.
- Korbicz J. and Zgurovsky M. Z. (1991): *Estimation and Stochastic Control of Systems with Distributed Parameters*. — Warszawa: Polish Scientific Publishers, (in Polish).
- Marczuk G.I. (1985): *Mathematical Modeling of Environmental Protection Problems*. — Warszawa: Polish Scientific Publishers, (in Polish).
- McMahon T.A. and Denison P.J. (1979): *Empirical atmospheric deposition parameters - A survey*. — Atmosph. Environ., v.13, No.5, pp.571–585.
- Merz P.H., Painter L.J. and Rayson P.R. (1972): *Aerometric data analysis. Time series analysis and forecast of atmospheric smog diagram*. — Atmosph. Environ., v.6, No.5, pp.319–342.
- Metcalf S.E., Atkins D.H.F. and Derwent R.G. (1989): *Acid deposition modeling and the interpretation of the United Kingdom secondary precipitation network data*. — Atmosph. Environ., v.23, No.9, pp.2033–2052.

- Morawska-Horawska M. (1979): *The influence of air termic stratification on a vertical distribution of SO₂ concentration in Cracow.* — Folia Geographica, Ser. Geogr.- Phys. v.13, No.2, pp 83-92, (in Polish).
- Nowicki M. (1985): *Monat-84; Experimental study for the needs of air pollution control in Cracow.* — Environ. Protect. Eng. v.11., No.2, pp.11-15, (in Polish).
- Omatu S. and Seinfeld J.H. (1981): *Filtering and smoothing for linear discrete-time distributed parameter systems based on a Wiener-Hopf theory with application to estimation of air pollution.* — IEEE Trans. Syst. Man and Cybernetics, v.SMC-11, No.12, pp.785-801.
- Omatu S., Seinfeld J.H., Soeda T. and Sawaragi Y. (1988): *Estimation of nitrogen dioxide concentration in the vicinity of a roadway by optimal filtering theory.* — Automatica, v.24, No.1, pp.19-28.
- Ripley B.D. (1981): *Spatial Statistics.* — Wiley Series in Probability and Mathematical Statistics. — New York: Wiley- Interscience.
- Runca E., Melli P. and Spirito A. (1979): *A Real-time forecast of air pollution episodes in the Venetian region. Part 1. The Advection -Diffusion Model.* — Appl. Math. Modelling, v.3, No.5, pp.402-408.
- Seinfeld J.H. (1977): *Air Pollution. Physical and Chemical Fundamentals.* — New York: McGraw Hill.
- Smith I.W.M. (1980): *Kinetics and Dynamics of Elementary Gas Reactions.* — Montvale, Ontario: Butterworths.
- Sorbjan Z. (1983): *Turbulence and Diffusion in the Lower Atmosphere.* — Warszawa: Polish Scientific Publishers, (in Polish).
- Szablowski P.J. (1990): *Elliptically contoured random variables and their application to the extension of the Kalman filter.* — Comput. Math. Appl., v.19, No.2, pp.61-72.
- Twardowska K. and Rożek (1988): *The stochastic processes and filtering theory in the prediction of air pollution.* — Przegląd Geofiz., v.32, No.2, pp.139-151, (in Polish).
- Twardowska K. (1989): *The stochastic processes and filtering theory in the prediction of air pollution in Silesian District.* — Proc. SICE'89 Conf., Matsuyama, Japan, pp.1111-1117.
- WHO Regional Office for EUROPE (1987) Copenhagen, Air Quality Guidelines in Europe, 1987.
- Willoughby R.A. (1974): *Stiff Differential Systems.* — New York: Plenum Press.

Received March 18, 1993

Revised December 16, 1993