International Institute for
Applied Systems Analysis
IIASA www.iiasa.ac.at

Statistical Analysis of Winter Sulphur Dioxide Concentration Data in Vienna

Bolzern, P., Fronza, G., Runca, E. and Ueberhuber, C.

IIASA Collaborative Paper
September 1981

Bolzern, P., Fronza, G., Runca, E. and Ueberhuber, C. (1981) Statistical Analysis of Winter Sulphur Dioxide Concentration Data in Vienna. IIASA Collaborative Paper. Copyright © September 1981 by the author(s).

STATISTICAL ANALYSIS OF WINTER SULPHUR DIOXIDE CONCENTRATION DATA IN VIENNA

P. Bolzern
G. Fronza
E. Runca
C. Überhuber

September 1981
CP-81-25

Research supported by MA 22, Umweltschutz des Magistrates der Stadt Wien, Vienna, Austria and by IIASA, Laxenburg, Austria

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

THE AUTHORS

P. BOLZERN and G. FRONZA are from the Centro Teoria Sistemi, Instituto Elettrotecnica ed Elettronica, Politecnico, Milan, Italy.

E. RUNCA is a research scholar in charge of air pollution research at the International Institute for Applied Systems Analysis, Schloss Laxenburg, A-2361 Laxenburg, Austria. Formerly, he was project leader at IBM Italy Scientific Center in Rome.

C. UEBERHUBER is with the Institut für Numerische Mathematik, Technische Universität, Vienna, Austria.

PREFACE

The modelling of the time evolution of air pollutants concentration in urban industrial areas allows administrative bodies responsible for the protection of the atmospheric environment to identify those situations which may lead to a serious increase in atmospheric pollution.

This paper discusses the modelling of the daily average winter sulphur dioxide concentration in the city of Vienna, Austria. A statistical approach was adopted, that is, the modelling was not achieved through the solution of the physical laws governing the dynamics of atmospheric diffusion, but through regression equations whose parameters were estimated from the available measurements of sulphur dioxide concentrations and meteorological parameters.

This study is part of the activity of the Resources and Environment Area (Task 2 Environment Quality Control and Management) on development and application of air pollution models to real situations. The study was conducted jointly with the Institut für Numerische Mathematik, Technische Universität of Vienna and Centro Teoria Sistemi, Politecnico of Milano.

ABSTRACT

The paper describes two nonlinear regression models, applied to winter daily SO_2 concentration data and to the corresponding meteorological data from the metropolitan area of Vienna. The first model accounts for the role of wind speed and temperature (a proxy for emissions due to residential heating) on average SO_2 concentration in the area. The second regression has an additional wind direction input and tries to point out the contribution by the industrial emissions (located primarily near the south-eastern border of the area) to concentration in the most polluted subarea.

Both models offer a satisfactory fitting performance (e.g., correlations around 0.85 between observed and regression values). However, since model validation is a critical point for regressions, sensitivity tests of model fitting performance are carried out by using various data sets for the estimation of regression coefficients. One of such tests points out that there is an "optimal length" of the data set to be used, namely neither a too short set nor a set including "too past" data offer a satisfactory fitting quality.

CONTENTS

1. INTRODUCTION 1

2. MAIN CHARACTERISTICS OF SULPHUR DIOXIDE POLLUTION IN VIENNA 3

3. THE MODEL ARX1 7

4. THE MODEL ARX2 10

5. DESCRIPTION OF MODEL SIMULATION RESULTS 12

Fitting Performance of ARX1 and ARX2 on Winter 1978/79 Data 12

Sensitivity of Model Performance to the Estimation Data Set 17

REFERENCES 20

STATISTICAL ANALYSIS OF WINTER SULPHUR DIOXIDE CONCENTRATION DATA IN VIENNA

P. Bolzern, G. Fronza, E. Runca, and C. Überhuber

1. INTRODUCTION

In general terms, mathematical representations of pollutant dispersion in an airshed belong to one of the following two classes (see also Seinfeld (1975))

a) Representations ("deterministic models") derived from the pollutant continuity equation, usually under suitable simplifying assumptions. The coefficients of such mathematical relations have a physical meaning and are determined on the basis of experimental evidence.

b) Representations ("stochastic models"), which consist of statistical regressions of one or more "ambient pollution variables" on their past values (see for instance Merz et al. (1972), Chock et al. (1975), McCollister and Wilson (1975), Tiao et al. (1975)) and on factors, like emissions and meteorological variables, which affect the dispersion phenomenon (see for instance Finzi et al. (1980)). The mathematical form of the regressions is usually derived not on physical grounds but from a statistical point of view. Namely, provided that regression fitting is satisfactory, the model is considered acceptable, at least if its relations somehow "resemble" the true cause-effect mechanism of the phenomenon (see also Benarie (1981), Smith and Jeffrey (1972)). In particular, although stochastic models are often gross simplifications of the physical mechanism, their structure turns out relatively complex in some cases (see for instance Bacci et al. (1981)). As for the coefficients of the regression, they are estimated by statistical data fitting techniques (usually according to the least squares principle or its modifications) and have generally no direct physical meaning.

By their very nature, precisely by the way their coefficients are evaluated, stochastic models sometimes give a fitting of concentration records better than deterministic ones, at least when only an aggregate description of the pollution phenomenon is required. However, the relevant lack of physical interpretation makes the reliability of the fitting performance by stochastic models a critical point. In other words it may happen that a model shows a satisfactory fitting performance on a particular data set but does not turn out to be a success for a different data set at the same site. Moreover, statistical confidence theory is not very helpful, since it is generally based on assumptions (e.g. normality), which are unrealistic in many applications or, most of all, can hardly be checked (e.g. stationarity). Therefore, a systematic analysis of model fitting performance on different data sets must be carried out.

The present paper illustrates two stochastic models of winter sulphur dioxide pollution in the metropolitan area of Vienna. The work has been committed by the municipality authorities for screening purposes, namely for preliminary understanding of relevant aspects like the dilution by moderate-to-strong winds, the transport from outside industrial sources to the city center in particular wind conditions, the dependence of residential heating emissions upon temperature variations.

Specifically, two regressions are described in the third and fourth section respectively:

i) A regression accounting for the role of meteorological factors on daily average SO_2 concentration in the area. The two factors considered are wind speed and temperature, which is a proxy for emission due to residential heating (a datum not available at daily time scaling).

ii) A regression supplying a rough measure of the contribution by the industrial sources located in the south-eastern part of the area to the concentration in the most polluted subarea (the city center).

The results are described in the last section. Specifically, both models have been first set up by using winter 1977/78 data for estimating regression coefficients (by an ad-hoc least squares algorithm), while winter 1978/79 data have been used for checking the model fitting performance, which has turned out rather satisfactory (e.g. correlations between regression values and observations around .85, also in "episode" situations).

Moreover, the following systematic test of the sensitivity of model performance to the data set, used for estimating model coefficients, has been carried out. The coefficients have been reestimated after every five days, by using the concentration and meteorological data of the last M days. The overall fitting performance has turned out poor for low values of M (twenty, say), corresponding to low statistical confidence (too few past data are used for model coefficient estimation, thus fitting in the next five days is unsatisfactory). Then, the performance increases with M up to $M \simeq 40$ and subsequently decreases. Hence, for instance, using the last eighty days of data for coefficient estimation gives a worse fitting of the next five days than using the last sixty days of data. In other terms, it is better to forget "too past" data, a clear sign of the non-stationarity of the process.

Finally, as for the role of industrial emissions in the pollution of the city center it turns out modest as a whole, although significant in particular situations characterized by moderate-to-strong winds.

2. MAIN CHARACTERISTICS OF SULPHUR DIOXIDE POLLUTION IN VIENNA

The contour of Vienna metropolitan area is shown in Fig.1, together with the network of sixteen SO_2 sensors, the meteorological station (20 m above the ground) and the three main industrial sources S_1, S_2 and S_3. In particular S_1 is a power plant with two 150 m stacks, S_2 is a power plant with four

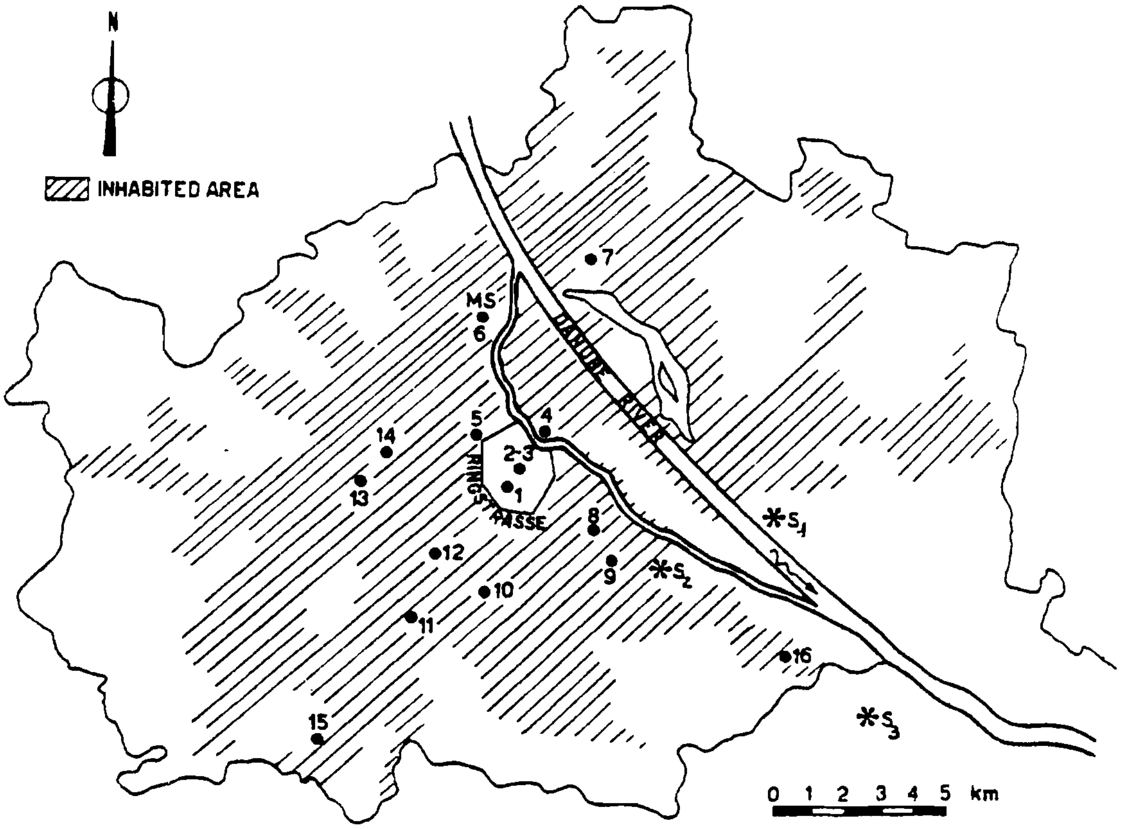


Figure 1. Vienna metropolitan area, SO_2 monitoring network, meteorological station MS (coincident with sensor 6) and industrial sources S_1, S_2 and S_3.

stacks (80 m, 120 m, 120 m, 200 m) while source S_3 corresponds to a refinery, where stacks do not exceed 80 m.

Each sulphur dioxide sensor records 30-min. concentration, while the meteorological station supplies hourly wind speed, wind direction and 3-hour temperature data (see for instance Löffler (1980) for a description of the monitoring and data recording system).

Vienna is located in a relatively flat area of the Danube Valley, although the hills of Wienerwald on the western border create a local orographic effect on air circulation. A picture of wind speed and direction distributions is given by Fig. 2, showing the hourly wind rose in winter 1977/78. Stagnation phenomena are practically absent. Moreover, wind speed is scarcely persistent at daily time scaling, as pointed out by Fig. 3, which illustrates the time pattern of daily speed in winter 1977/78

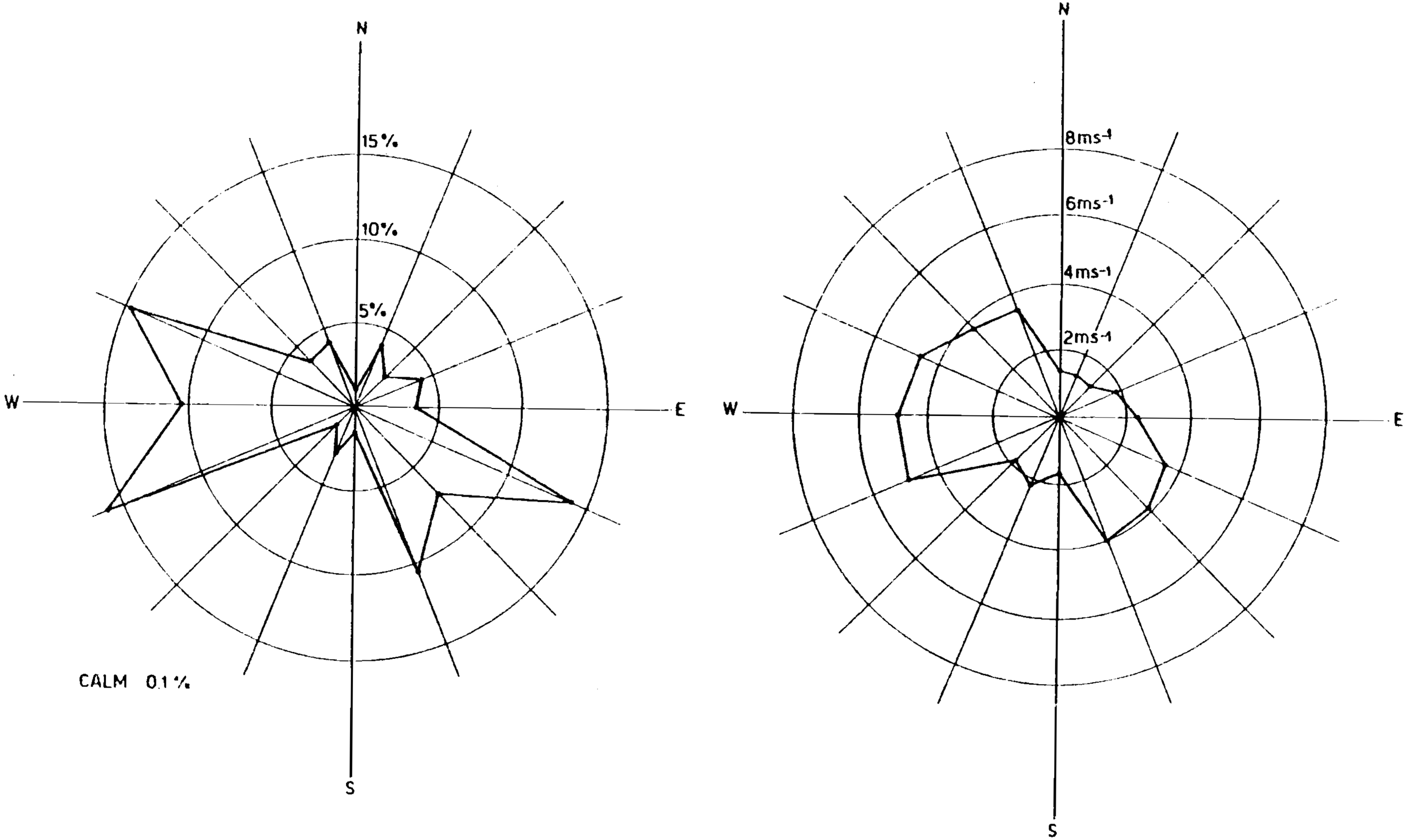


fig 2a

fig 2b

Figure 2. Winter 1977/78 (a) hourly wind rose; (b) hourly wind speed rose.

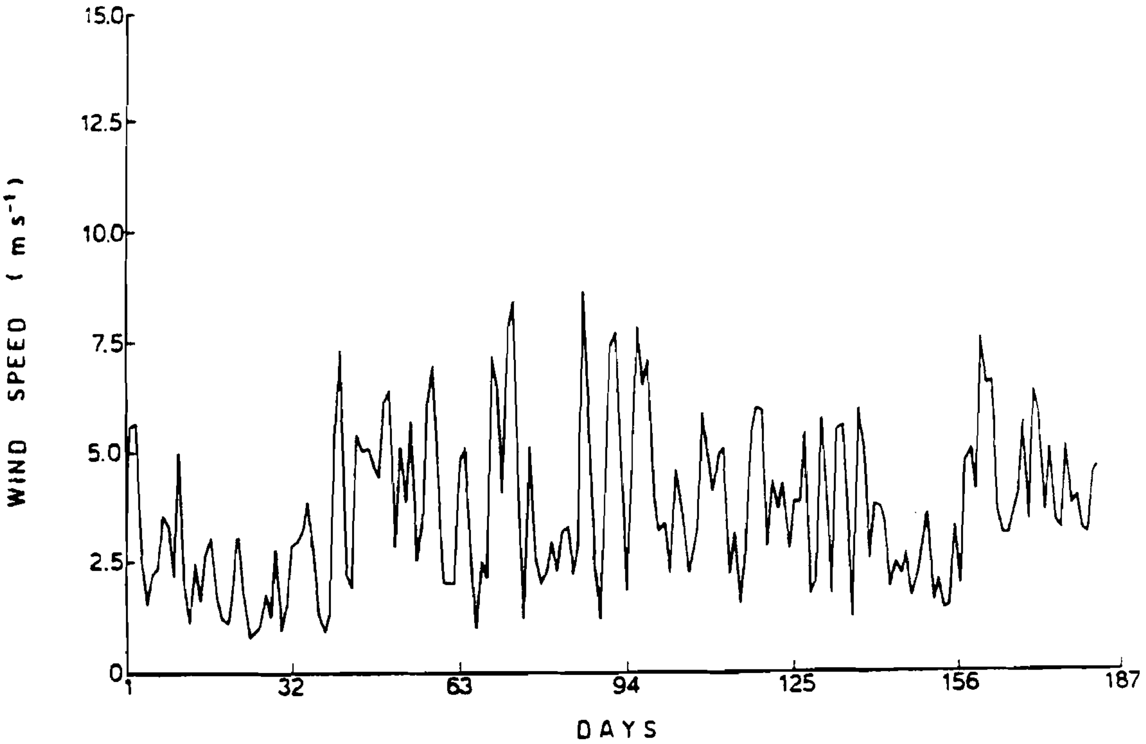


Figure 3. Time pattern of daily wind speed in winter 1977/78.

and gives an idea of the conspicuous weather variability in the region.

Sulphur dioxide pollution in winter is mainly due to residential heating, which is approximately uniformly distributed in the dashed area of Fig. 1. SO_2 concentration reaches its highest levels in the city center, the area inside the Ringstrasse (see Fig. 1). When moderate-to-strong winds (in the range 3 ms^{-1} - 6 ms^{-1}, say) blow from the eastern/south-eastern sectors, a contribution by the industrial sources to pollution in the central area can reasonably be assumed. As a matter of fact, such contribution is partially revealed by the comparison between the wind speed rose of Fig. 2b and the SO_2 concentration rose of Fig. 4. Precisely, Fig. 4 points out the dilution effect by the moderate-to-strong winds blowing from the western sectors but does not indicate a similar effect by winds blowing from the eastern/south-eastern sectors, characterized by an average wind speed only slightly lower. Hence, a reasonable explanation (validated in

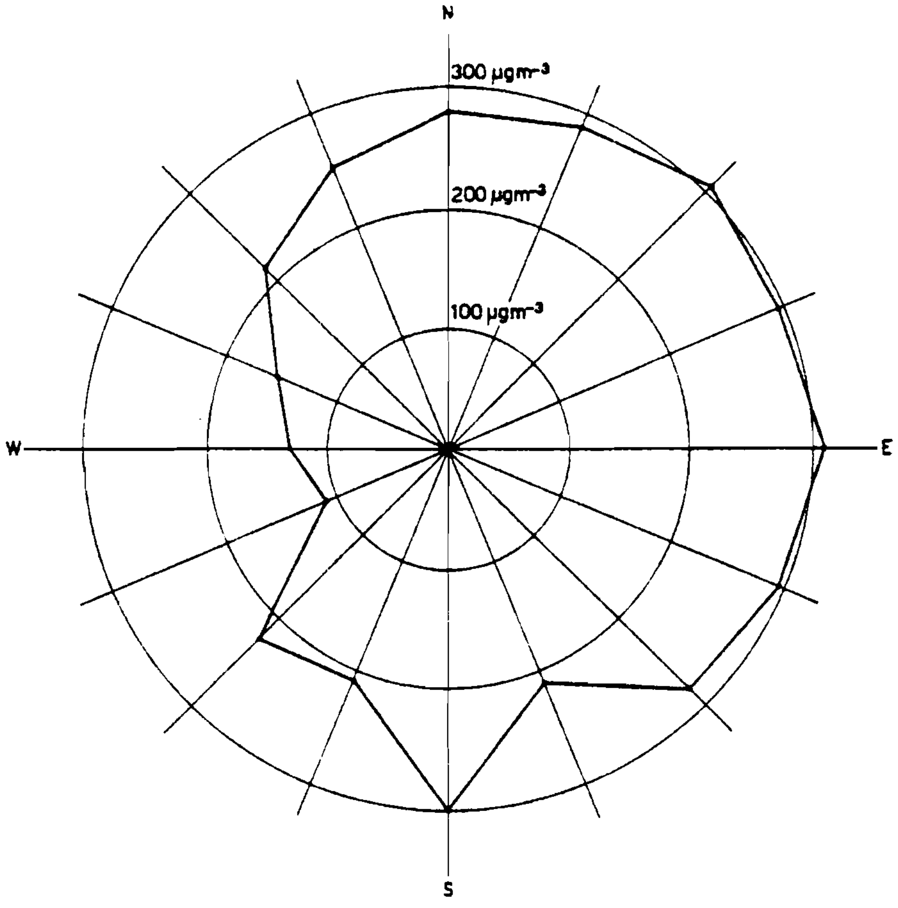


Figure 4. Winter 1977/78 hourly SO_2 concentration rose in station 1, in the city center.

the paper under the limitations and approximations of the statistical approach) is that the smaller contribution by residential heating emissions in such moderate-to-strong wind conditions is balanced by the transport of pollutant from industrial sources.

3. THE MODEL ARX1

The two regression models described in this section and in the next one are classified as ARX (AutoRegressive with eXogenous inputs) in the stochastic modelling literature, therefore they will respectively be labeled as ARX1 and ARX2.

The objective of ARX1 is to fit observed average daily concentration in the area only by using concentration and available meteorological data. Let $x_i(k)$ $(k=1,2,\ldots)$ denote the average SO_2 concentration ($\mu g.m^{-3}$) measured by the i-th sensor $(i=1,2,\ldots 16)$ in the k-th winter day. Moreover, let $x(k) \doteq \sum_{i=1}^{16} x_i(k)/16$ be the average of concentration measures in the area, in the k-th day. Then, ARX1 consists of the following nonlinear regression:

$$x(k+1)=a_1x(k)+a_2\left[T(k+1)+a_3\right]^{-\alpha}+ a_4\left[v(k+1)\ +a_5\right]^{-\beta} \qquad (1)$$

where:

$T(k) \doteq$ average temperature (°C) in the k-th day;

$v(k) \doteq$ average wind speed (ms^{-1}) in the k-th day;

$\{a_j\}_{j=1}^{5} \doteq$ regression coefficients, determined by an ad-hoc estimation procedure (see below);

$\alpha,\beta \doteq$ other (positive) coefficients, assigned by trial (see also below).

A gross "physical" justification of ARX1, namely of the form of the right-hand side of eq. (1) is the following. By eq. (1), concentration in the (k+1)-st day is made to depend upon concentration in the previous day (x(k)), temperature (T(k+1)) and wind speed (v(k+1)). Concentration in the k-th day is somehow correlated to the pollutant mass existing in the area at the end of the k-th day. Thus, the term $a_1x(k)$ in eq. (1) can somehow be regarded as the contribution to x(k+1) by the pollutant mass existing in the area at the beginning of the (k+1)-st day. As for the other two addenda of the regression, the first one aims at taking into account the increase of residential heating emissions when temperature decreases (a concept similar to that of degree-days), while the other one aims at representing pollutant dilution by wind speed on emissions from the residential heating sources.

In conclusion, eq. (1) attempts to account for three "phenomena" affecting x(k+1), namely the existence of previous pollutant, the temperature-dependent emissions by residential heating, the wind dilution effect.

The special nonlinear form of the last two terms of the right-hand side of eq. (1) differs from previously used ARX models. In particular, the "inversely proportional" wind speed term in eq. (1) replaces the corresponding negative linear term used by Finzi et al. (1980) in the Milan case. In fact, an addendum linearly decreasing with wind speed has proved insufficient in the Vienna area, characterized by a range and distribution of wind speed completely different from the Milan region.

In detail, the coefficients α, β and $\{a_j\}_{j=1}^{5}$ of eq. (1) have been determined as follows.

Assume to have selected a certain data set from the available concentration and meteorological records (which data have actually been used for estimating the coefficients of eq. (1) in the Vienna application is explained in the last section).

A) Try a certain pair (α, β).

B) For such values of α and β, the coefficients $a_1, a_2, \ldots, a_5$ are estimated from the data set by a computer algorithm. Specifically, each step of such algorithm consists of the following operations.

B_1) Values of a_3 and a_5 are taken from the previous step. Introducing these values into eq. (1) makes ARX1 formally a linear regression $x(k+1) = a_1 x(k) + a_2 T'(k+1) + a_4 V'(k+1)$, where $T'(k+1) \doteq [T(k+1)+a_3]^{-\alpha}$ and $v'(k+1) \doteq [v(k+1)+a_5]^{-\beta}$. Hence, coefficients a_1, a_2, a_4 can be estimated from the data set by the ordinary least squares formula.

B_2) Fitting of regression (1) over the data set is improved (=the overall sum of the error squares is reduced) by determining new a_3 and a_5 values according to optimal search in the plane (a_3, a_5) (e.g. Hooke-Jeeves search, see for instance Himmelblau (1972)).

C) When the iterative procedure B) converges to some estimates of $\{a_j\}_{j=1}^{5}$, a new pair (α, β) is tried, the iterative estimation B) of $\{a_j\}_{j=1}^{5}$ is repeated and so on.

Obviously, that pair (α, β) and the corresponding set of estimates of $\{a_j\}_{j=1}^{5}$ is selected, which gives the best absolute fitting of the data set. Naturally, trials on (α, β) could be replaced by optimal search also in the (α, β) plane. In practice, it has turned out reasonable and sufficient to test a few pairs (α,β) in the ranges $1 \leq \alpha \leq 3$, $1 \leq \beta \leq 3$.

A final remark concerns the convergence of the iterative procedure B) for estimating $a_1, a_2, \dots a_5$. Convergence is not guaranteed a priori, however it has always been obtained in a reasonable number of steps in the Vienna application. Moreover, the estimates resulting from algorithm B) have been validated by assuming different starting values for the procedure.

4. THE MODEL ARX2

Model ARX2, which is described in this section, attempts to give some information about the contribution of industrial emissions to SO_2 concentration in the city center.

First, the wind speed range has been subdivided into a certain number of intervals, from now onwards called wind speed classes. A posteriori, the most suitable subdivision in the Vienna area has turned out to be: $I = [0,\ 1.3\ ms^{-1})$, $II = [1.3\ ms^{-1},\ 3\ ms^{-1})$, $III = [3\ ms^{-1},\ 4\ ms^{-1})$, $IV = [4\ ms^{-1},\ v_{max}]$ (maximum wind speed v_{max} recorded in winters 1977/78 and 1978/79 has been $8.7\ ms^{-1}$).

For each wind speed class w = I, II, III, IV, ARX2 consists of the following regression:

$$y(k+1) = b_1^w\, y(k) + b_2^w\, [T(k+1)+b_3^w]^{-\gamma^w} + b_4^w\, N(k+1) \qquad (2)$$

where

$y(k)$ ≑ average daily concentration in the central subarea in the k-th day (precisely, average of daily concentrations recorded in stations 1, 2, 3 and 4);

$N(k)$ ≑ number of hours in the k-th day characterized by winds blowing from the eastern/south-eastern sectors (112°30', 135°);

$\{b_j^w\}_{j=1}^4$ ≑ regression coefficients;

γ^w ≑ (positive) coefficient found by trial;

w ≑ index of wind speed class (w = I, II, III or IV) in the (k+1)-st day.

The justification of model (2) and, in particular, the reason why wind speed has been introduced in a way different from ARX1 are now given.

The first addendum of the right-hand side of eq. (2) accounts for the contribution to y(k+1) by the "initially existing" pollutant in the area (see the comment in the previous section). The second and third addendum can respectively be considered as representatives of the contributions to y(k+1) by temperature-dependent residential heating and industrial emissions. As in ARX1, emissions are introduced in the model through the related inputs, temperature and wind direction. Moreover, the impact of these inputs on y(k+1) is made to depend upon wind speed, through assuming dependence of the b and γ regression coefficients upon the speed class w.

In conclusion, from eq. (2), concentration y(k+1) is the sum of contributions by initially existing pollutant, by temperature-dependent residential heating sources and by industrial sources, each contribution depending upon the wind speed class in the (k+1)-st day.

So, while in eq. (1) the wind speed term $a_4[v(k+1) + a_5]^{-\beta}$ could not be interpreted as the contribution of existing or newly emitted pollutant, here the above interpretation of eq. (2) allows to draw a gross comparison of the relative impact of residential heating and industrial emissions on the most polluted subarea. Precisely, such comparison can be obtained by taking day-by-day the ratio between the third and the second addendum in the right-hand side of eq. (2).

Naturally, since the above "physical" interpretation of ARX2 is gross, this ratio can be regarded as a too rough measure, if one is reluctant to accept the statistical viewpoint. However, it is the only ratio computable through the presently available measures and can, at least, be considered as an indicative value.

Finally, as for the estimation of ARX2 coefficients, a procedure quite similar to the one described for ARX1 has been used (in correspondence with each wind speed class). Obviously, since there is only one nonlinear term in eq. (2), in this case only a one-dimensional optimal search on the b_3^w-axis is required by algorithm B) (e.g. Fibonacci search, see for instance Wilde and Beightler(1967)).

5. DESCRIPTION OF MODEL SIMULATION RESULTS

Fitting performance of ARX1 and ARX2 on winter 1978/79 data

First, the representativity of the two models ARX1 and ARX2 has been checked as follows.

- Model coefficients have been assigned through the A) - C) estimation algorithm (see sect. 3) by using the winter 1977/78 concentration and meteorological data set (here "winter" means the heating season Oct. 1st - March 31st). In particular, the b_4^w-coefficient of ARX2 (the one multiplying the wind direction input N(k+1) in eq. (2)) has turned out to be very near to zero (in fact slightly negative, due to the unconstrained optimization procedure B)) for w = I and w = II. This is not surprising, because it is reasonable to assume that there are no contributions by industrial sources to pollution in the city center for wind speed lower than 3 ms^{-1}. Furthermore, setting a priori $b_4^I = b_4^{II} = 0$ has given a slight performance improvement, thus the zero value has been definitely accepted for those two coefficients. With this specification, the overall set of ARX1 and ARX2 coefficient estimates is given in Table 1.

As for the residuals of the regressions (= time series of regression errors), their means have come out negligible. Moreover, the residuals have turned out uncorrelated (i.e. white noise) under the cumulative periodogram test (see for instance Box and Jenkins (1970)).

- Then, both models ARX1 and ARX2 have been tested on winter 1978/79 data. The performance has been satisfactory, as shown by the direct comparison (Figs. 5 and 6) between day-by-day observations and regression values. In particular, both models fit the conspicuous measure fluctuations with acceptable accuracy.

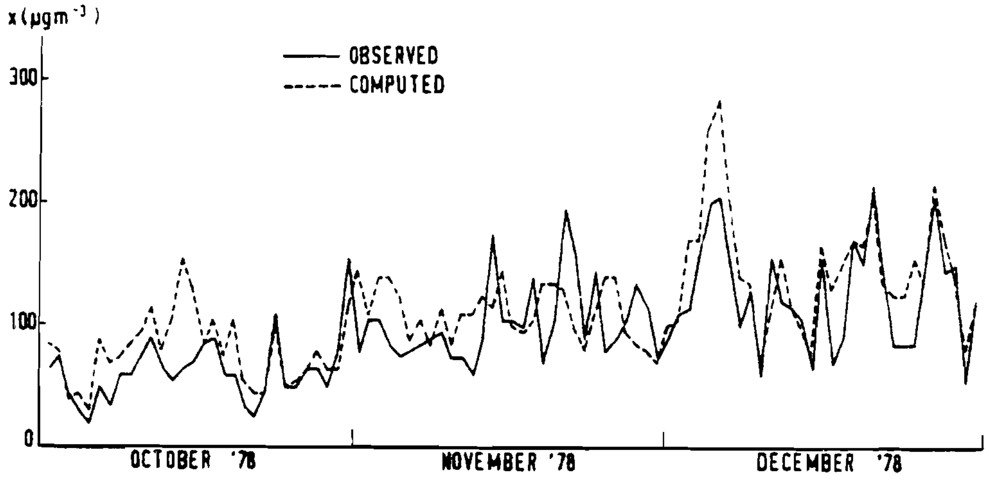
Table 1. (a) ARX1 and (b) ARX2 coefficient estimates from winter 1977/78 data.

α	β	a_1	a_2	a_3	a_4	a_5
2.0	1.0	0.26	$3.89\text{x}10^4$	23.1	443	1.16

(a)

Coefficients / Wind speed class w	γ^w	b_1^w	b_2^w	b_3^w	b_4^w
I	1.6	0.31	$9.81\text{x}10^4$	34.9	0
II	1.6	0.30	$6.18\text{x}10^4$	33.6	0
III	1.0	0.39	$0.83\text{x}10^3$	10.7	0.64
IV	1.5	0.25	$1.53\text{x}10^4$	22.1	2.5

(b)



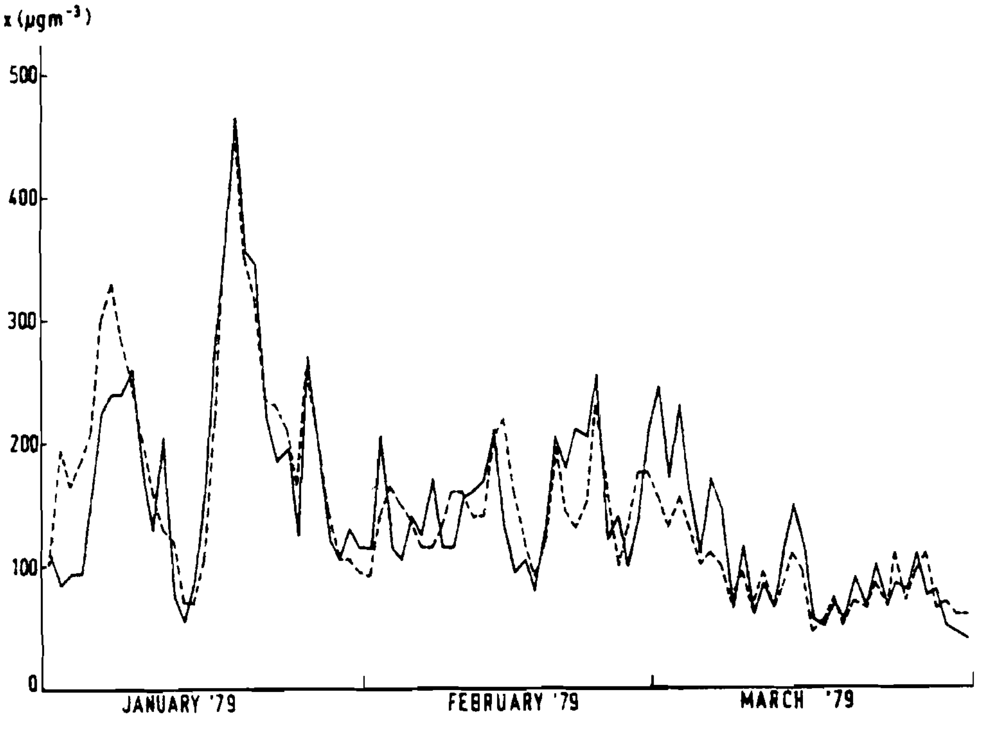
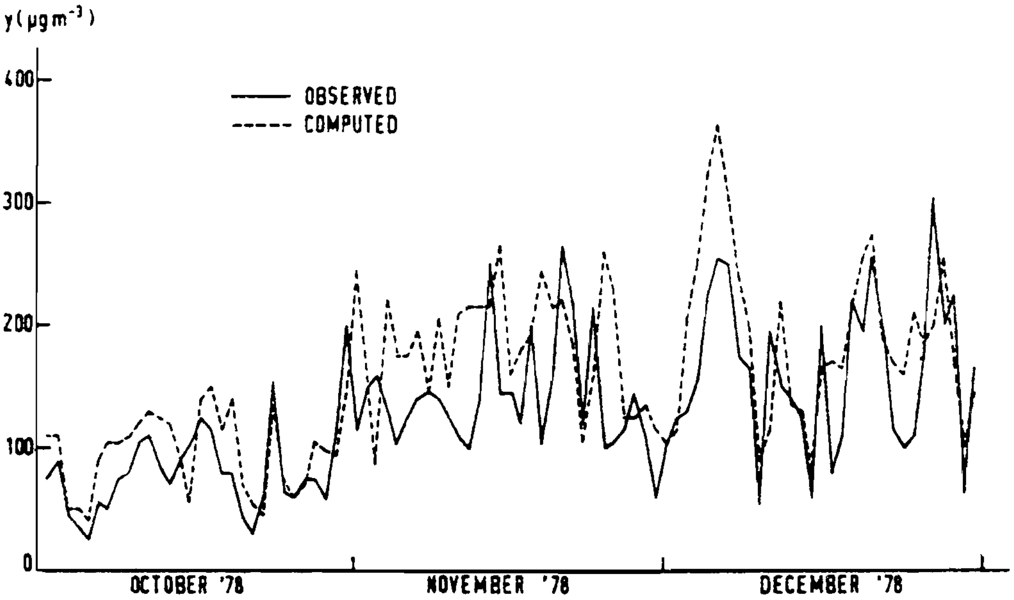


Figure 5. Winter 1978/79 observed versus regression concentration values (model ARX1).



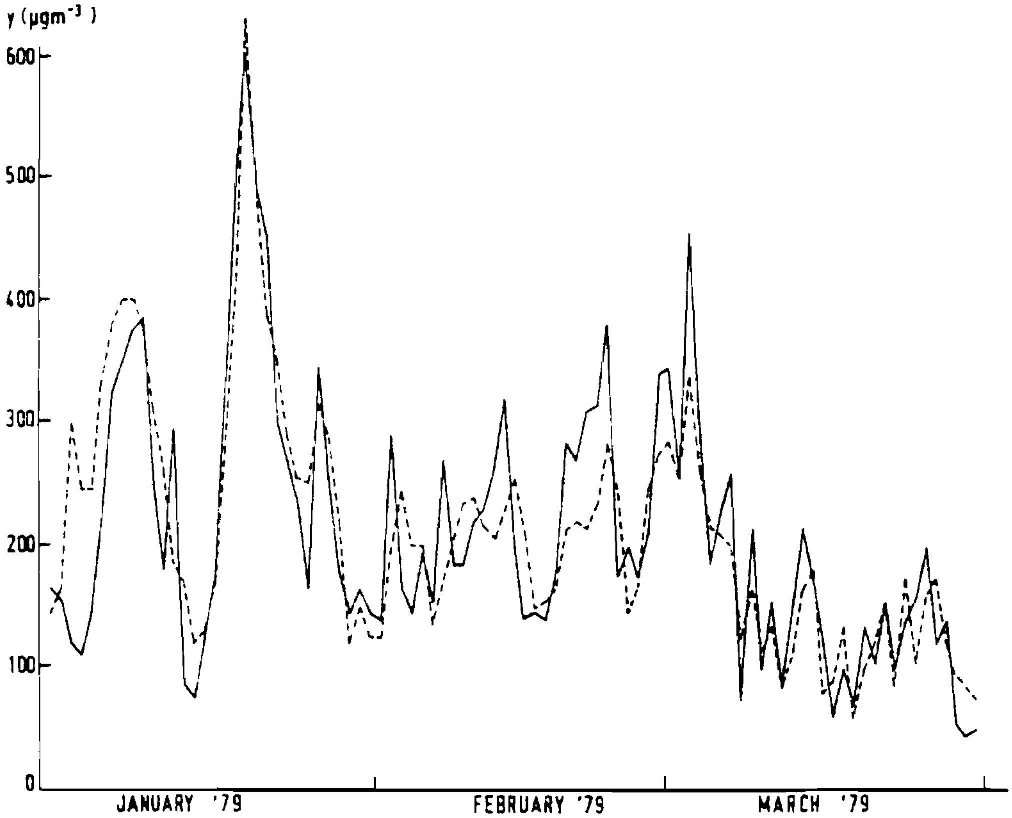


Figure 6. Winter 1978/79 observed versus regression concentration values (model ARX2).

The overall winter 1978/79 model performance has been evaluated by the following indexes, reported in Table 2 :

$\rho \doteq$ correlation between observations and regression values;
$\rho^E \doteq$ correlation between observations and regression values in "episode" situations; here an episode is defined as a situation where the measured concentration $x(k)$ (or $y(k)$ for ARX2) exceeds its mean plus standard deviation;
$\sigma_{err}/\mu \doteq$ ratio between the standard deviation of the regression error and the concentration mean;
$\sigma^E_{err}/\mu^E \doteq$ same as σ_{err}/μ, but in episode situations.

- The day-by-day ratio $R(k+1) = b_4^w\ N(k+1)/b_2^w[T(k+1)+b_3^w]^{-\gamma^w}$, w = III, IV, between the third and second addendum in ARX2 has been evaluated and reported in Fig. 7. Such ratio, under the limitations pointed out in the previous section, is an indication of the weight of the contribution by industrial sources to pollution in the city center. The overall average of $R(k+1)$ in winter 1978/79(i.e. by including also situations under wind speed classes I and II, when a zero contribution by the industrial stacks is assumed) is modest (~ 3%). However, as pointed out by Fig. 7,the

Table 2. Fitting performance of ARX1 and ARX2 on winter 1978/79 data.

Model / Perf.Index	ARX1	ARX2
ρ	.84	.81
ρ^E	.80	.83
σ_{err}/μ	.31	.35
σ^E_{err}/μ^E	.20	.17

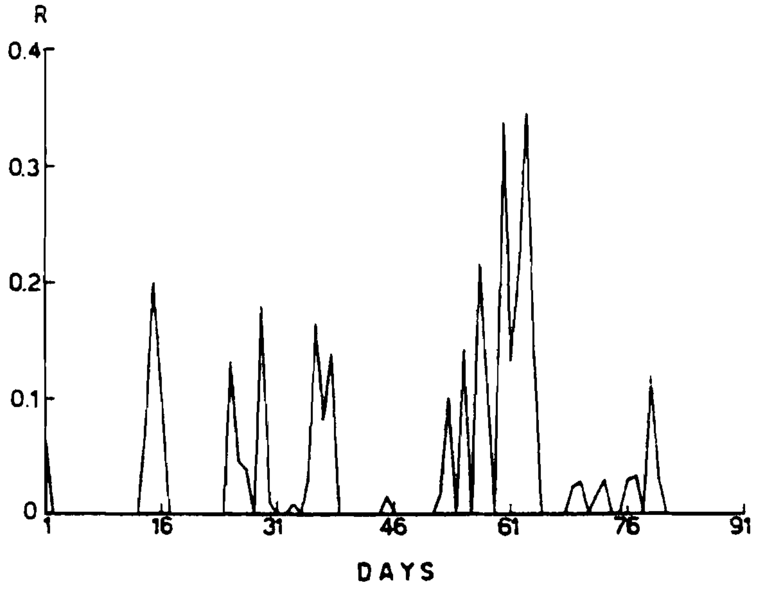


Figure 7. Time pattern of ratio R in wind speed classes III and IV (non-consecutive days in the abscissa).

ratio $R(k+1)$ reaches nearly 40% in certain days, characterized by wind speed in class III or IV and by wind blowing from the eastern/south -eastern sectors during a relevant part of the day.

Sensitivity of model performance to the estimation data set.

Sensitivity of model fitting quality to the data set used for applying the coefficient estimation procedure A) - C) has also been analyzed. The sensitivity tests have been performed only on ARX1. In fact, ARX2 has a higher number of coefficients and the statistical significance of their estimates would become poor near the lower bound of the size of the estimation data set considered by the sensitivity tests.

- First, it has been ascertained whether it is useful or not to distinguish within the heating season between the months of October and March and the "winter core" (November-February) In particular, the behaviour of residential heating polluters is likely to be partially different in October and March, when some of the oil burners are switched off or work for a smaller number of hours. Thus, two estimates of ARX1 coefficients have been carried out on 1977/78 data: one by the October and March data and one by the data of the "winter core". There has actually turned out a slight improvement when fitting the concentration of the following heating season 1978/79, as pointed out by Table 3.

Table 3. Performance of modified ARX1 (different coefficients between October/March and "winter core").

Performance Index	Modified ARX1
ρ	.85
ρ^E	.82
σ_{err}/μ	.29
σ^E_{err}/μ^E	.20

- A more systematic sensitivity test has also been performed, based on adaptive coefficient estimation. Specifically, ARX1 coefficients have been reestimated after every five days, by using the data of the last M days. In explicit words, ARX1 coefficients are first estimated at the beginning of the second winter, namely on Oct.1st 1978 (by using the data of the last M days of March 1978). Such coefficient estimates are used for fitting by ARX1 the concentration record Oct.1st-Oct. 5th, 1978. On Oct. 6th, the coefficients are reestimated (by using the data set Oct. 1st- Oct.5th, 1978 and the last (M-5) days of March 1978). Such new estimates are used for fitting the concentration record Oct. 6th - Oct. 10th, on Oct. 11th the coefficients are reestimated and so on, up to March 31st 1979.

The overall performance indexes mentioned in Table 2 have been evaluated for different values of M. For instance, the plot of ρ and ρ^E versus M is shown in Fig. 8. The main conclusions

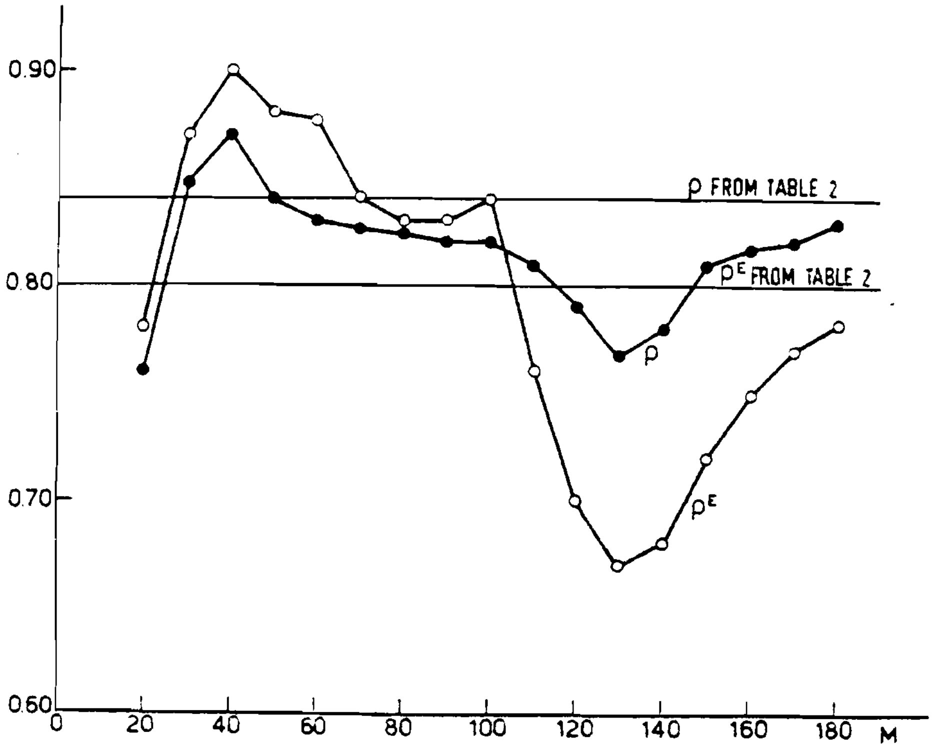


Figure 8. Performance indexes ρ and ρ^E versus "memory" M.

which can be drawn from Fig. 8 are the following (quite similar considerations would be suggested by the analysis of the other two overall indexes of Table 2).

For low values of M the reliability of coefficient estimates is poor and thus model fitting performance is low. The quality increases up to M = 40, which is therefore approximately the "optimal" length of the data set for the adaptive coefficient estimation (actually M has been varied with steps of 10). For higher M values, there is an even conspicuous decrease of performance. Hence, keeping memory of "too past" data worsens the quality of the fitting by adaptive regression. This is clearly due to the non-stationarity of the winter concentration process, otherwise increasing the length of the data set should give better coefficient estimates and consequently a better performance. Note that here non-stationarity turns out more conspicuous than in the previous sensitivity test.

In correspondence with the optimal M, the ARX1 fitting performances ρ and ρ^E reach .87 and .90 respectively, which are significantly higher than the values reported in Table 2. Therefore, for suitable "memory" M, adaptive regression, namely frequent coefficient reestimation, gives an improvement with respect to the "batch" regression described above.

REFERENCES

Bacci, P., P. Bolzern, and G. Fronza. 1981. A stochastic predictor of air pollution based on short-term meteorological forecasts. J. Appl. Met. 20.

Benarie, M.M. 1981. Air pollution modeling operations and their limits. In: Mathematical Models for Planning and Control of Air Quality, edited by G. Fronza and P. Melli. Oxford/ New York: Pergamon Press (in press).

Box, G.E.P., and G.M. Jenkins. 1970. Time-series Analysis, Forecasting and Control. San Francisco: Holden-Day.

Chock, D.P., T.R. Terrel, and S.B. Levitt. 1975. Time-series analysis of Riverside, California air quality data. Atmospheric Environment 9:978-989.

Finzi, G., G. Fronza, and A. Spirito. 1980. Multivariate stochastic models of sulphur dioxide pollution in an urban area. J. Air Pollut. Control Ass. 30:1212-1215.

Himmelblau, D.L. 1972. Applied Nonlinear Programming. New York: McGraw-Hill.

Löffler, H. 1980. Umweltschutz and Luftreinhaltung. Der Aufbau 34:314-320 (in German).

McCollister, G.M., and K.R. Wilson. 1975. Linear stochastic models for forecasting daily maxima and hourly concentrations of air pollutants. Atmospheric Environment 9:417-423.

Merz, P.H., L.J. Painter, and P.R. Ryason. 1972. Aerometric data analysis: time series analysis and forecast and an atmospheric smog diagram. Atmospheric Environment 6:319-342.

Seinfeld, J.H. 1975. Air Pollution. New York: McGraw-Hill.

Smith, F.B., and G.H. Jeffrey. 1972. The prediction of high concentrations of sulphur dioxide in London and Manchester air. Proc. 3rd NATO-CCMS Int. Tech. Meeting on Air Poll. Model and its Application.

Tiao, G.C., G.E.P. Box, and W.J. Hamming. 1975. A statistical analysis of the Los Angeles ambient carbon monoxide data 1955-1972. J. Air Poll. Contr. Ass. 25:1129-1136.

Wilde, D.G., and C.S. Beightler. 1967. Foundations of Optimization. Englewood Cliffs: Prentice Hall.