Long range transport of air pollution into Norway—a transfer function approach

EIVIND DAMSLETH†

The daily and monthly concentration of sulphate in the air at Birkenes in the southern part of Norway is analysed within an intervention analysis and transfer function framework. As input to the model we use the prevailing wind direction. It is shown that when the wind comes mainly from the South, that is from Central Europe and England, this gives a significant increase in the sulphate concentration, while a northern wind leads to a decrease.

1. Introduction

For quite some time now, discussions have been going on as to what extent the air pollution in Norway is caused by long range transport of pollutants from the industrial areas in Central Europe and England. This paper presents an analysis which strongly suggests that this imported pollution plays an important role in this context.

A lot of work has been carried out on the analysis of air pollution within the last ten–fifteen years. Bornstein (1980 a, b, c, d) gives a bibliography which covers approximately four hundred publications only concerning the California region. The OECD (1979) report gives some early results from a European study.

Since the basic works of Box and Tiao (1975) and Tiao, Box and Hamming (1975), time series analysis and related techniques have been widely used in the analysis of air-pollution. In this paper we use intervention analysis and transfer function modeling to describe the impact of the prevailing wind direction on the sulphate concentration.

In § 2 the data used for the analysis is presented. § 3 describes various models based upon daily observations, while in § 4 the data is compiled into monthly averages and models for these are built. The daily and monthly models are compared and found to agree reasonably well. Finally, in § 5 the findings are summarized.

2. Data

2.1. Sulphate concentration

The Norwegian Institute for Air Research collects data on air quality for a number of stations all over Norway. In this analysis we have used the observed daily mean concentration of sulphate (given as micrograms sulphate per cubic meter) at Birkenes north of Kristiansand. The location of the station is shown in Fig. 1. The measurements are taken from 0700 GMT to 0700 GMT on the following

Figure 1.   Location of Birkenes, and the eight wind sectors.

day. We have used data for the period 720701–821029, 3773 observations in all. The data are shown graphically in Fig. 2, and they give an overall impression of large random fluctuations around a rather stable mean.

Let Y_t be the observed daily mean concentration of sulphate at day t, where day no. 1 is 720701. Standard ARIMA identification and estimation suggest that the data, after log-transformation, can be well described by the AR(1)-model

$$(1 - 0.69B)(\ln Y_t + 0.29) = a_t\dagger$$
$$\quad (0.01) \qquad \quad (0.04) \tag{1}$$

with a residual standard deviation of 0.7058. Here B represents the backwards shift operator, so that $B^k \ln Y_t = \ln Y_{t-k}$. The sequence $\{a_t\}$ is a series of independent, identically distributed random variables. There may be an additional seasonal component in the series, giving rise to residual autocorrelation at lags around 365. We have not tried to incorporate this in the model.

2.2. Air trajectories

The Norwegian Meteorological Institute has the necessary data and models available to provide trajectories for any location at any time. A trajectory for a

---

† The numbers in parentheses below the estimates are the approximate standard errors.

given location and a given time is the route covered during the last 72 hours by a 'packet' of air which arrives at the location at the given time. For example, the trajectory at Birkenes at noon July 4 is the route covered in the period from noon July 1 to noon July 4 by a 'packet' of air arriving at Birkenes at noon July 4.

Four such trajectories are computed each day, with arrival times 1200, 1800, 0000 and 0600 GMT. The trajectories will have various forms. Some days they wander all the way around the horizon, other days they form more or less straight lines. If the compass is divided into eight sectors, as also is shown in Fig. 1, a day is classified as having its prevailing wind from one such sector if three or four of the trajectories falls within the sector. If not, the day is not classified. Thus, for each day we know the prevailing wind sector, or we know that the day could not be classified. Of the 3773 days in this study, 23% could not be classified. The distribution of the prevailing wind on the eight sectors is shown in Fig. 3, and it is easily seen that the dominant wind direction at Birkenes is westerly.
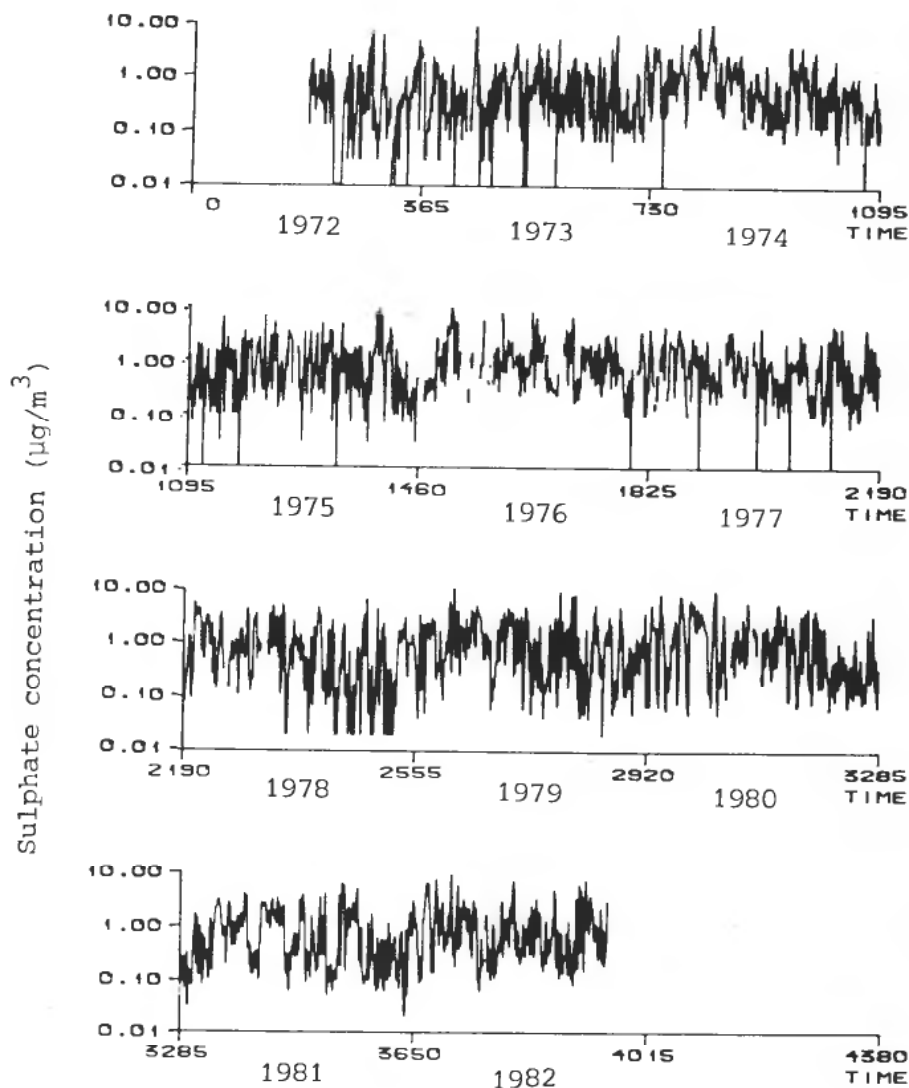

Figure 2.   Daily sulphate concentration (log-scale) 720701–821029. Day no. 1 is 720101.
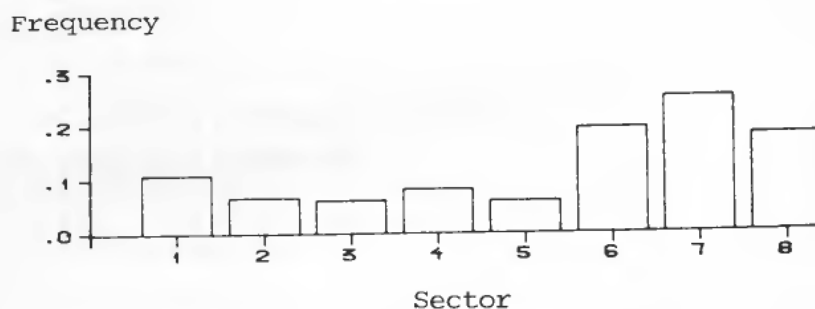
Frequency

Figure 3.   Distribution of the wind direction.

The classification, as described above, does not take the wind speed into consideration. As pointed out by a referee, more clever schemes might be applied to explain the impact of the route of the air packet on the sulphate concentration. Unfortunately, the data were presented to us as sector values only, and it would be difficult, if not impossible, to go back to the basic information to try these schemes out.

3. Models based on daily data

3.1. Multi-input intervention model

Our first approach was to introduce eight intervention series as explanatory variables. Let

$$X_{it} = \begin{cases} 1 \text{ if day } t \text{ is classified as sector } i \\ 0 \text{ elsewhere} \end{cases} \quad i = 1, \ldots, 8$$

Evidence in the data, as well as physical considerations indicate that the impulse response of the interventions is not only instantaneous, but with a delayed effect as well. The reasons are several:

The observation periods for the sulphate and for the sector calculations do not overlap completely, the sectors lagging some hours behind.

The wind calculations are made for an altitude of 1000 m, while the sulphate is measured at ground level. Meteorological experience indicates that the wind at higher altitudes is ahead of the ground wind by some hours.

There is a mixing effect. Fresh air coming in will not lead to an immediate change in the concentration, but the new air will mix with the old, leading to a gradual change.

To our surprise we also found evidence of feed-back in the system, so that today's sulphate concentration has an impact of tomorrow's wind direction! This is obviously nonsense, and there must be another explanation. We believe that the reason can be found in the precision, or rather lack of precision, in the timing of the trajectories. It is not unlikely that an air 'packet', expected to arrive at noon, arrives 8–12 hours earlier. This lack of precision in the timing of the trajectories can result in the wind direction being recorded one day too late, which in turn will give the impression of feed-back.

There is no reason to believe that this dynamic effect varies from sector to sector. As a first model we thus assumed

$$\ln Y_t - \mu = (1 - \omega_1 B - \omega_2 B^2) \sum_{i=1}^{8} \alpha_i X_{i, t+1} + n_t$$

Notice the subscript $t + 1$ for the input variables, caused by the observed feed-back in the system, as discussed above. For the noise, n_t, we initially assumed an AR(1) model similar to the one we found during the univariate analysis. The diagnostic checking, however, revealed that this noise model was not quite satisfactory, there were significant residual autocorrelations at the first lags. The noise model finally turned out to be an ARMA(1,3)-model. The complete model was estimated as:

$$\ln Y_t + 0.17 = (0.18 + 0.57B + 0.25B^2) * (-1.03X_{1, t+1} - 0.22X_{2, t+1}$$
$$(0.05) \quad (0.02) \quad (0.06) \quad (0.03) \qquad (0.11) \qquad\qquad (0.09)$$

$$+ 0.38X_{3, t+1} + 0.76X_{4, t+1} + 0.98X_{5, t+1} + 0.65X_{6, t+1}$$
$$(0.10) \qquad\qquad (0.10) \qquad\qquad (0.12) \qquad\qquad (0.11)$$

$$- 0.49X_{7, t+1} - 1.05X_{8, t+1}) + n_t,$$
$$(0.10) \qquad\qquad (0.12)$$

$$(1 - 0.91B)n_t = (1 - 0.40B - 0.19B^2 - 0.10B^3)a_t$$
$$(0.02) \qquad\qquad (0.02) \quad\; (0.02) \quad\; (0.02) \qquad\qquad\qquad (2)$$

The residual standard deviation from this model was 0.6273, which can be compared to 0.7058 for the univariate model. The residuals show satisfactory white noise behaviour, and the inclusion of the wind directions in the model has completely removed the (weak) residual seasonality we found after fitting model (1).

The MA-parameter values in (2) seem to decay exponentially with a damping factor about 0.5. If so, the model can be rewritten

$$(1 - 0.91B)n_t = (1 - 0.40B(1 - 0.5B)^{-1})a_t,$$

or

$$(1 - 0.91B)(1 - 0.5B)n_t = (1 - 0.90B)a_t.$$

However, the near cancellation of the first AR and the MA factor leads back to the AR(1)-model which initially proved unsatisfactory, and estimation of the ARMA(2,1) model gave numerically unstable results.

In (2), the model has been rewritten so that the sum of the parameters in the dynamic component, that is the operator common to all eight input series, is 1. The values can thus be regarded as an estimated probability distribution for the impact of the wind direction on the concentration on the next, same and previous day respectively. It is seen that the effect is strongest on the same day, but that the additional contributions from the neighbouring days are almost equally important.

The estimated parameters for the eight sector interventions are all significant. We can thus conclude that sector has an effect on the sulphate concentration. The sectors 1, 2, 7 and 8, which cover the compass from W to NE, have low concentrations, while the sectors 3, 4, 5 and 6, from E to SW, have high concentrations. Going back to the map in Fig. 1, there can be little doubt that the imported air pollution is important for the sulphate concentration.

3.2. Model simplification

When the model is written as in (2), with the dynamic effects totalling 1, the individual sector effects also give the gain of the eight transfer functions. These gains are shown graphically in Fig. 4. If the gain in Fig. 4 is regarded as a function of the sector, the figure suggests that this function can be well described by a simple sinusoidal. Letting

$$Z_t = \text{sector value for day no. } t$$

the simplified model was estimated

$$\ln Y_t - \mu = \delta(1 - \omega_1 B - \omega_2 B^2) \sin ((Z_{t+1} - 1)\pi/4 + \rho) + n_t,$$

or

$$\ln Y_t - \mu = (1 - \omega_1 B - \omega_2 B^2)(\delta_1 \sin ((Z_{t+1} - 1)\pi/4) + \delta_2 \cos ((Z_{t+1} - 1)\pi/4)) + n_t$$

For days with $Z_t = 9$, that is for days where the wind direction cannot be classified to a sector, we set both the sine and cosine terms in the formulae above to 0. The noise is assumed to follow the same ARMA(1,3) model as before.

Estimation of this model gave:

$$\left[\begin{array}{l} \ln Y_t + 0.16 = (0.18 + 0.56B + 0.26B^2)(0.39 \sin ((Z_{t+1} - 1)\pi/4) \\ \quad (0.04) \quad (0.02) \quad (0.07) \quad (0.04) \quad (0.05) \end{array} \right]$$

$$- 1.04 \cos ((Z_{t+1} - 1)\pi/4)) + n_t$$
$$(0.12)$$

$$(1 - 0.91B)n_t = (1 - 0.39B - 0.19B^2 - 0.09B^3)a_t \quad (3)$$
$$(0.02) \qquad\qquad (0.02) \quad (0.02) \quad\quad (0.02)$$

The residual standard deviation from this model is 0·6308, only marginally larger than the 0·6273 value from model (2). The difference, though small in value, is however significant when tested by an ordinary F-test. There are no physical reasons why the impact of the wind direction should follow a sinusoidal curve, so it is not surprising that model (3) is rejected by a significance test, especially when the large number of observations is taken into consideration.

The estimated functional relationship between sector and concentration can be re-written

$$\ln Y_t = -0.16 + 1.11(0.18 + 0.56B + 0.26B^2) \sin ((Z_{t+1} - 2.54)\pi/4)$$
$$(0.04) \quad (0.12)(0.02) \quad (0.07) \quad (0.04) \qquad\qquad (0.06)$$
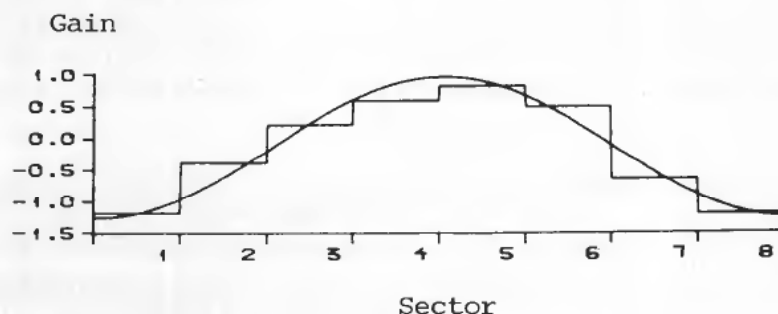

Figure 4.   Sector gains from model (2), and the sinusoidal curve from model (3).

This function is also plotted in Fig. 4, and the fit, judged visually, seems very good. The fitted curve obtains its maximum at 'sector' value 4·54, which corresponds to a geographical direction of 159·3°, pointing through Denmark to the large industrial areas in East and West Germany.

4. Models based on monthly data

We also made an analysis related to the one described in the previous chapter, but based on monthly averages. The reasons were several:

We wanted to study the possible seasonal variation of the series, which is much more pronounced in a monthly than in a daily series.

It is of interest to see how well the models for the daily and monthly series agree. From the daily model one can deduce what the monthly model should be, and this can be compared to the model actually obtained.

We wanted to see how much information we lost going from daily to monthly observations. This is of special importance since we later want to apply the same technique to the sulphate concentration in the precipitation, where daily data are not available. (It does not rain every day!)

For the sulphate we therefore calculated the monthly geometric mean of the daily observations, from July 1972 to October 1980, 124 months in all. Let W_t represent this monthly series. W_t is then given by:

$$\ln W_t = \sum_i \ln Y_i / N_t$$

where the summation is over all the N_t days in month t, and the Ys are the daily sulphate concentrations.

4.1. Multi-input intervention model

For each month we calculated the proportion of the days with prevailing wind direction within the eight sectors. The input series are thus given by:

$$V_{it} = \sum_j X_{ij} / N_t, \qquad i = 1, \ldots, 8$$

where the summation again is over all the N_t days in month t, and the Xs are the intervention variables defined in § 3.1.

The non-instantaneous effect of the sectors on the concentration which was clear from the daily data will disappear when data are aggregated on a monthly basis. After a few iterations of identification and estimation we arrived at the model:

$$\ln W_t - 0.020 = -0.90V_{1t} - 0.62V_{2t} + 0.31V_{3t} + 1.05V_{4t} + 0.53V_{5t}$$
$$ (0.20) \quad\;\; (0.38) \quad\;\; (0.53) \quad\;\; (0.39) \quad\;\; (0.40) \quad\;\; (0.59)$$

$$+ 0.33V_{6t} - 0.86V_{7t} - 1.41V_{8t} + n_t$$
$$(0.35) \quad\;\; (0.27) \quad\;\; (0.34)$$

$$(1 - 0.46B)n_t = (1 + 0.33B^{12})a_t \qquad (4)$$
$$(0.09) \qquad\qquad (0.10)$$

The residual standard deviation was 0·3080 for this model, and the residuals show no sign of model inadequacy. It is interesting to notice that the transition to

monthly data introduced a small but significant seasonal moving average component.

The only significant wind contributions in model (4) are from Sectors 1, 7 and 8, which are significantly below the mean, and Sector 4 which is above. When the actual values of the estimated gains are compared to the ones we found using model (1) and the daily data, there seems to be a very close agreement. The standard errors of the estimates are, however, much larger for model (4), so we suffer quite a loss by going from daily to monthly data.

4.2. *Sinusoidal model*

Since the simple, sinusoidal model (3) showed quite good results for the daily data, we also tried out a monthly version of this model. Thus, for a specific month t we calculated:

$$S_t = \sum_i \sin ((Z_i - 1)\pi/4)/N_t$$

and

$$C_t = \sum_i \cos ((Z_i - 1)\pi/4)/N_t$$

where the summation again is over all the days in month t, N_t is the number of days in month t, and Z_i is the sector for day i. For $Z_i = 9$, when the sector is not known, both the sine and cosine terms are set to 0 in the sums above.

Using these two series as inputs, we arrived at the simplified model:

$$\ln W_t + 0 \cdot 10 = 0 \cdot 63 S_t - 1 \cdot 05 C_t + n_t$$
$$ (0 \cdot 07) \quad (0 \cdot 11) \quad (0 \cdot 13)$$

$$(1 - 0 \cdot 41 B)n_t = (1 + 0 \cdot 35 B^{12})a_t \qquad (5)$$
$$ (0 \cdot 10) (0 \cdot 10)$$

The residual standard deviation was $0 \cdot 3023$, so that model (5), in fact, gives a better fit compared to model (4)! The seasonal MA-parameter is almost unchanged. Model (5) implies a sinusoidal relationship between the sectors and the concentration with amplitude $1 \cdot 22$ and phase $2 \cdot 22$, as compared to the values $1 \cdot 11$ and $2 \cdot 54$ we found using the daily model. When the standard error of the estimates are taken into consideration, these differences are not very significant.

The estimated gains for the various sectors from model (4) and the sinusoidal curve from model (5) are shown graphically in Fig. 5.
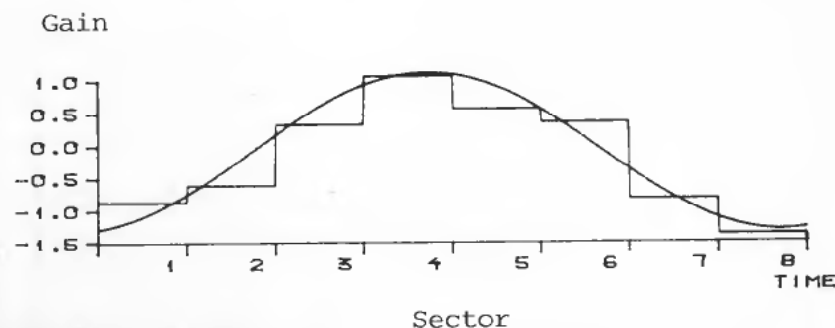

Figure 5.   Sector gains from model (4), and the sinusoidal curve from model (5).

4.3. Model consistency

There is a close agreement between the daily and the monthly models as far as the effects of the wind direction is concerned. When it comes to the noise model, the picture changes. Firstly, the transition to monthly data has shown a significant, although rather small, seasonal effect. Secondly, the univariate ARMA(1,3) model for the noise in models (2) and (3) implies that the monthly averages should be ARMA(1,2) with an AR-parameter $\approx 0{\cdot}06$ and negligible MA-parameters. With only 124 observations, this model will be indistinguishable from white noise. The fact that we have estimated an AR-parameter equal to $0{\cdot}43$ implies that the initial model for the daily data may be wrong. Going back to § 3.1, we there discussed the possibility of an ARMA(2,1)-model with near cancellation. This model, though numerically unpleasant, is much more consistent with the results from the monthly data.

5. Conclusions

We have shown that the wind direction has a significant impact on the concentration of sulphate in the air at Birkenes in southern Norway. Using daily data we have found that the (logarithmic transformed) concentration has an almost perfect sinusoidal relationship with the wind direction.

The same conclusion can also be drawn from monthly data, but the reduced number of observations leads to a considerable loss in significance.

REFERENCES

BORNSTEIN, R. D. (1980 a). Annotated Bibliography of Recent Publications on California Air Pollution Volume 1: Sources and control strategies. Technical Report no. 39, Department of Statistics, Stanford University, Stanford, California.
BORNSTEIN, R. D. (1980 b). Annotated Bibliography of Recent Publications on California Air Pollution. Volume 2: Transport, Diffusion and Transformation Processes. Technical Report no. 40, Department of Statistics, Stanford University, Stanford, California.
BORNSTEIN, R. D. (1980 c). Annotated Bibliography of Recent Publications on California Air Pollution. Volume 3: Meteorological and Concentration Climatologies. Technical Report no. 41, Department of Statistics, Stanford University, Stanford, California.
BORNSTEIN, R. D. (1980 d). Annotated Bibliography of Recent Publications on California Air Pollution, Volume 4: Removal Processes and Effects. Technical Report no. 42, Department of Statistics, Stanford University, Stanford, California.
BOX, G. E. P., and TIAO, G. C. (1975). Intervention analysis with applications to economic and environmental problems. *Journal of the American Statistical Association,* **70**, 70–79.

DAMSLETH, E. (1984). Tidsrekkeanalyse av forurensningsdata—En metodestudie. Report no. 745, Norwegian Computing Center, Oslo. (In Norwegian.)

OECD (1979) *The OECD Programme on Long Range Transports of Air Pollutants: Measurements and Findings.* Second edition. (OECD, Paris).

TIAO, G. C., BOX, G. E. P., and HAMMING, W. J. (1975). Analysis of Los Angeles photochemical smog data: a statistical overview. *Journal of Air Pollution Control Association,* **25,** 260–268.