Part II: The Requirements of Data on Atmospheric Concentrations and Deposition for Models for Estimation of Emissions

Gudmund Høst and Xeni Dimakos
Norwegian Computing Center,
P.O.Box 114 Blindern, N-0314 Oslo

May 20, 1997

Summary

Information on national emissions of air pollutants in Europe may be obtained by combining the EMEP-model and EMEP monitoring data. We outline a statistical framework for extracting such information and give a discussion on possible implications on the design of monitoring networks. Our approach is illustrated by an application to emissions of nitrogen oxides in Europe during 1990. A future use of our framework could be monitoring of compliance by the European nations with their obligations to reduce air pollutant emissions.

1 Introduction

A monitoring network for air pollutants may serve several purposes. The network may be intended for direct assessment of air pollutant concentrations and depositions in space, time or space-time. The network may also provide a basis for validation of regional air pollutant models. Finally, the network may be intended to complement regional air pollutant models by providing information on spatial and temporal scales that are different from the model scales.

An important feature of a regional air pollutant model is that it provides an explanatory tool for linking depositions and concentrations with emissions. The uncertainty of the model estimates depends on model uncertainty and on uncertainty in the emission data. By combining a "good" model with monitoring data one can obtain information on the uncertainty in the emissions data. This approach was taken in Høst (1996), where a statistical method for re-estimating national emissions from EMEP monitoring data and the EMEP-model was presented.

The emissions data used in the EMEP-model are those officially supplied by the parties to the Geneva Convention on Long-Range Transboundary Air Pollution. The protocol on further reduction of sulphur emissions, adopted in Oslo 1994, puts emphasis on reviewing the information supplied by the parties. It is likely that reviewing compliance by the parties with their obligations to reduce emissions will be an important issue in future agreements. The method presented in Høst (1996) provides a framework for monitoring such compliance.

The viewpoint of compliance monitoring adds a further objective to the monitoring network. In addition to monitoring depositions and concentrations, the network should also be able to give information on emissions. Some critical research issues regarding compliance monitoring are:

- What spatial and temporal resolution is needed for the monitoring network?
- What is the optimal location of monitoring stations?
- What are the practical implications of model resolution and model quality on compliance monitoring?

It is the purpose of the present paper to elaborate some implications of the compliance monitoring perspective on the air pollutant monitoring network.

2 Method

The interested reader is referred to Høst (1996) for details of the statistical method. Our approach is based on regarding the deposition or concentrations as a realization of a second order stationary random field with partially known mean function. The mean function is taken from the EMEP-model, with the generalization that the emissions are modeled as random variables. The reported national emissions specify the prior distribution of these random variables, and the "true" emissions are estimated by fitting the statistical model to deposition or concentration data within a *Bayesian* framework (Berger 1985). The Bayesian method may be regarded as a method of regularization, giving a means of estimating a large number of parameters (emissions) from a limited number of observations. The method involves a regularization parameter, which determines how much weight should be put on the reported emissions relative to the monitoring data. At present, the regularization parameter must be selected by the user, but in the future we hope to estimate this parameter from data.

3 European Nitrogen Emissions in 1990

In this Section we present some results on the sensitivity of the uncertainty in emission estimates to the location of monitoring stations. First, the statistical model was fitted to yearly values of nitrate in precipitation from the EMEP network and predicted values from the EMEP-model. We used the national emissions of nitrogen oxides given in Barrett & Berge (1996) and the same regularization parameter as in Høst (1996). However, for the present application, it may be argued that the regularization parameter should be set to give maximum weight to the monitoring data.

Some monitoring stations near large emission sources were removed from the analysis. For each remaining monitoring station we required at least 10% data coverage in each month. Although such data screening criteria may be subject to further discussions, this is not the main focus of the current paper. The resulting 47 monitoring stations used in the analysis are shown in Figure 1. The model was fitted to an exponential spatial covariance function and our results indicated that there is very little spatial correlation for lags greater than 410 km.

We are not prticularly interested in the emission estimates in this study, but rather the uncertainty associated with these estimates. Figure 2 shows the standard deviations of the estimated national emissions relative to the assumed standard deviation of reported emissions for some nations and regions, with the monitoring network shown in Figure 1. It is seen that the largest improvements occur for Federal Republic of Germany, United Kingdom, Poland, Italy and France. Hence, emission estimates from the present network will be more precise for these countries than from other countries.

Given the parameters of the statistical model, it may be possible to find an optimal network configuration by some numerical search algorithm. Although this will require extensive analysis and fairly large computer resources, it is an interesting topic for future investigations. As an alternative, we present results for the following two exercises. The first exercise was to add a set of candidate monitoring sites were to the network, one at a time. The second exercise involved deleting existing monitoring stations from the network, one at a time. For each exercise the standard deviations of various emission estimates were calculated.

The 9 candidate sites used in the first exercise are shown in Figure 1 with symbols {ES*1,ES*2,FR*,IT*1,IT*2,PL*,RO*,UA*,BY*}. These sites were chosen randomly. Figure 3 shows the estimated reduction in the standard deviation of national emissions resulting from adding each of the candidate sites to the network. It is seen that adding any of the Spanish stations (upper left and upper middle panels) does not reduce the uncertainty of national emission estimates. This could be due to meteorology, because there are no other Spanish stations included in our configuration. However, adding a station in central Italy or on Sicily (upper right and middle left panels) will improve the estimation of Italian emissions. Likewise, adding a station in Romania will reduce the standard deviation of the Romanian emission estimate by 15 % (middle panel), and adding a station in Ukraine will reduce the standard deviation of the Ukrainian emission estimate by 10 % (lower middle panel). It is also seen that adding the Polish candidate site gives a slight improvement in the estimated emissions from the two German regions, but very little improvement in the estimated Polish emission estimate (middle right panel). Adding the French candidate site gives some reduction in the uncertainty of both the French and the Spanish emissions (lower left panel). In choosing between the French and the two Spanish candidates one would therefore be tempted to choose the French site, even if the main purpose is to estimate the Spanish emission.

Figure 4 shows the estimated increase in the standard deviation of national emissions resulting from removing some selected monitoring stations from the network. It is seen that removing the station AT4 does not affect the uncertainty of any national emission estimates (upper left panel). Furthermore, removing CS1 has some effect on the uncertainty in the emission estimate from the Czech Republic (upper right panel), while removing CS2 mainly affects the uncertainty in the Polish estimate (middle left panel). Figure 4 also indicates that removing the German station DE1 will increase the uncertainty in the United Kingdom emission by 8% (middle right panel), while removing DE4 will increase the uncertainty in the French estimate by 7% (lower right panel).

Figure 5 shows the national increases in the standard deviation of emission estimates resulting from deleting each of the monitoring stations in the network, one at a time. For Albania there is no effect, indicating that the monitoring network is not informative about Albanian emissions (upper left panel). On the other hand, the uncertainty in the French estimate is affected mainly by the stations DE4, FR11, FR3 and CH1 (lower left panel). Similarly, the uncertainty in the estimated emission from the Federal Republic of Germany is affected mainly by the stations DE5, DE2, DD2 and CS3 (lower right panel).

4 Discussion

The features shown in the previous figures result from intricate combinations of meteorology and monitoring network configuration. A unique property of the proposed mathematical framework is the capability of quantifying the effects of revising the network. Our Bayesian framework requires the specification of a regularization parameter. Ideally this parameter should be estimated from data. However, for the purpose of network design it may be argued that one should use a value that enhances the sensitivity to monitoring data.

The usual consideration in spatial sampling design would lead to minimizing prediction errors of the concentration field (Cressie 1991). The resulting optimal configuration is usually a network organized in a regular lattice. This is in contrast to the present application, where the main focus is to minimize errors of the estimated emissions. Here, an important task is to distinguish between the spatial patterns of each national contribution. An optimal monitoring network may locate monitoring stations to contrast national contribution patterns, but these patterns will depend on meteorology, and on area and geometry of each country.

In Europe, emissions from some countries, like Belgium and the Netherlands, are likely to be hard to distinguish, because their contribution patterns are very similar. Also, we would expect contributions from very small countries to be harder to estimate than contributions from larger countries due to spatial resolution of the EMEP-model. This will put specific limitations on the capability of monitoring the national emissions, given any configuration and density of monitoring stations.

In Section 3, we illustrated some implications of network design in a non-systematic manner. In the future, this should be investigated systematically. For a given optimality criterion, such as reduction of the total variance of estimated emission, the optimal way of locating (additional) monitoring stations is a search problem. This problem can be addressed by general search algorithms, such as simulated annealing (Metropolis, Rosenbluth, Rosenbluth, Teller & Teller 1953) to give an optimal combination of a set of candidate monitoring stations.

5 Concluding Remarks

Assessing the regional air pollution problem of Europe may be viewed as a decision process moving through the following three phases. The first phase is identifying the problem, which amounts to observing the negative effects. The second phase is to establish the explanatory mechanism, which involves building a model to relate sources and effects. A logical consequence of the second phase is to induce strategies for reducing the negative effects. The third phase is the control phase, which involves checking for compliance with the strategies and checking for improvement of the environment.

To reflect the relevant objectives, the optimal monitoring strategy may be different for each phase. In particular, the third phase may involve shifting focus from spatial coverage to contrasting regional contribution patterns. The framework presented in this paper may provide a tool for analyzing this problem, but a detailed and systematic study is a topic for further research.

References

Barrett, K. & Berge, E. e. (1996), Transboundary air pollution in europe. part 1: Estimated dispersion of acidifying agents and of near surface ozone, EMEP/MSC-W Report 1/96, The Norwegian Meteorological Institute, Oslo, Norway.

Berger, J. O. (1985), *Statistical Decision Theory and Bayesian Analysis*, Springer Verlag, New York.

Cressie, N. (1991), *Statistics for Spatial Data*, Wiley, New York.

Høst, G. (1996), A statistical method for estimation of European sulphur emissions using EMEP monitoring data, EMEP/MSC-W Note 3/96, The Norwegian Meteorological Institute, Oslo, Norway.

Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A. & Teller, E. (1953), 'Equations of state calculations by fast computing machines', *J. Chem. Phys.* **21**, 1087–1092.

Figure 1: *Study area and data locations.*

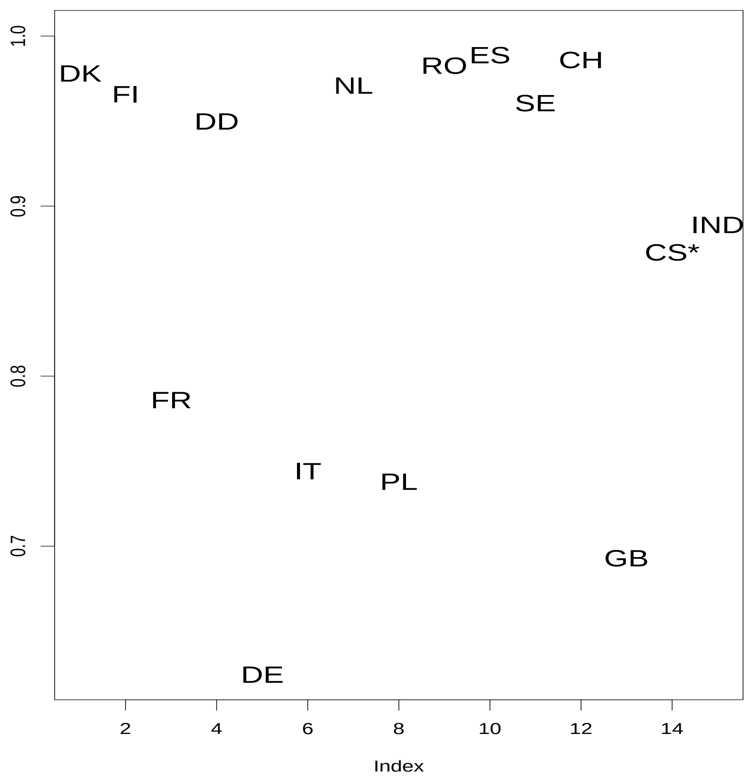

Figure 2: *Estimated reduction in the standard deviation of national emissions resulting from including monitoring data.*

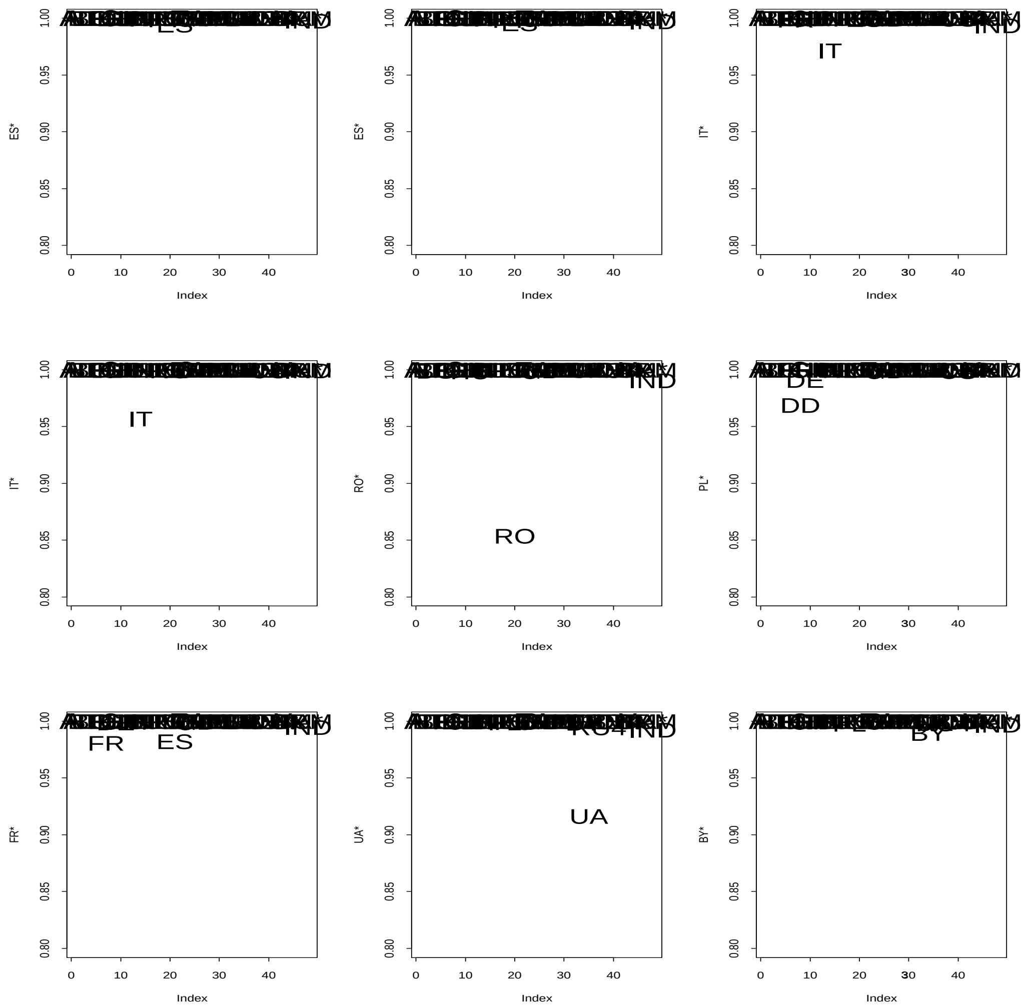

Figure 3: *Estimated reduction in the standard deviation of estimated national emissions resulting from adding pseudo-stations.*

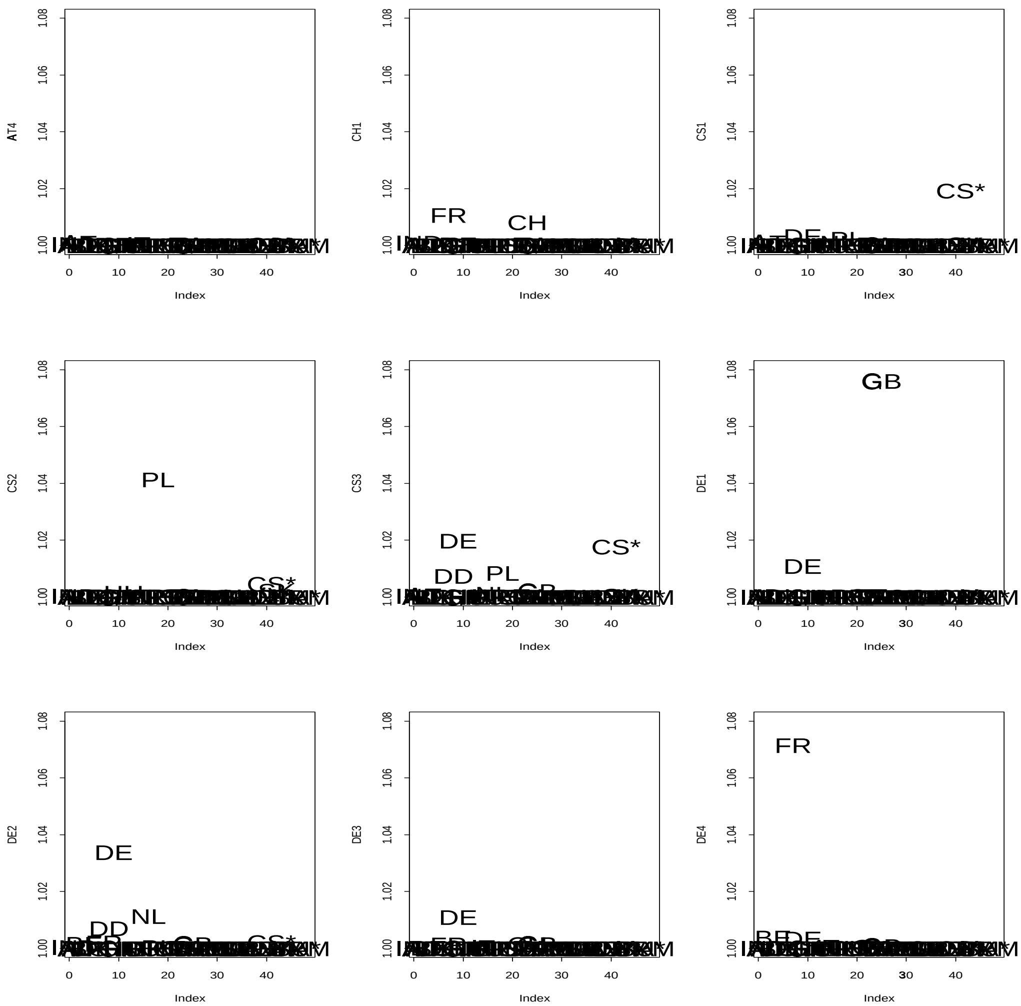

Figure 4: *Estimated increase in the standard deviation of estimated national emissions resulting from deleting selected monitoring stations.*

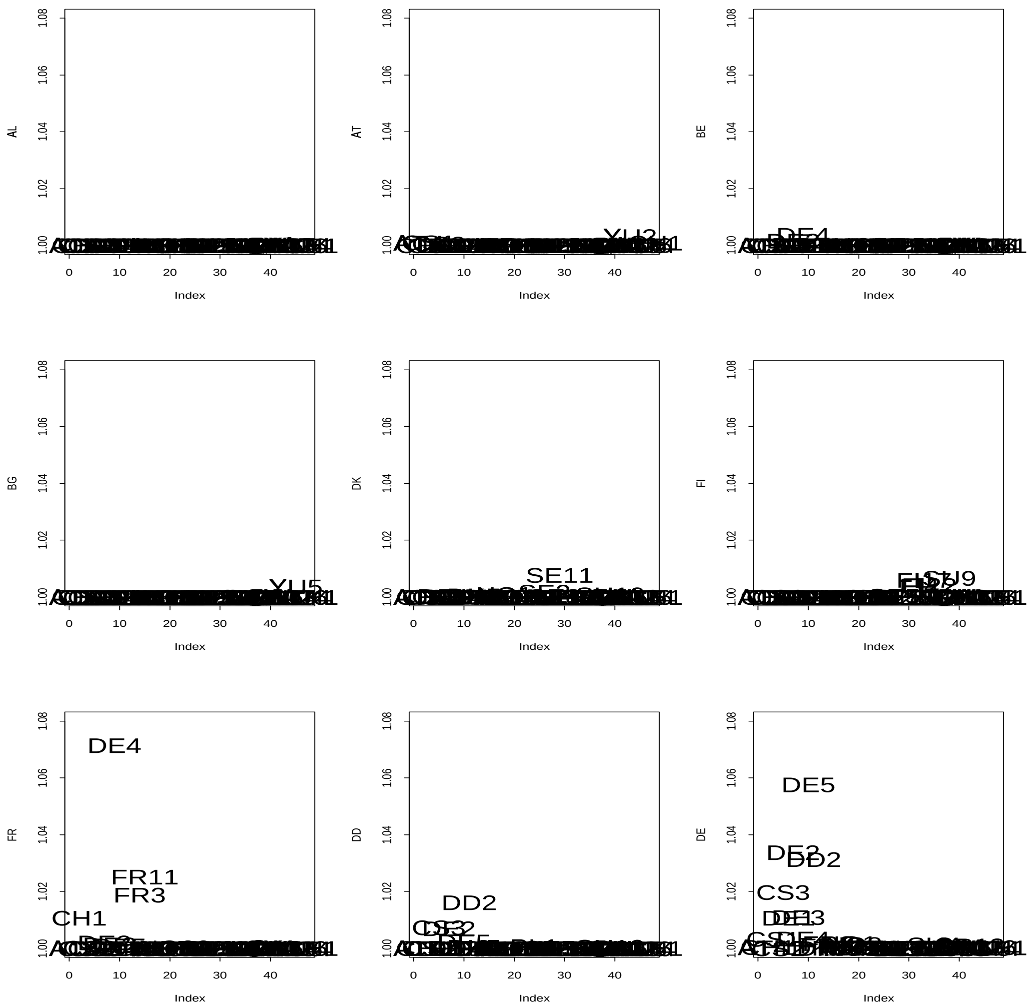

Figure 5: *Estimated increase in the standard deviation of estimated selected national emissions resulting from deleting monitoring stations.*