

EMEP/MSC-W      Note 3/97  
Date:              July 97

# Estimation of Regional Emissions in Europe Using EMEP Monitoring Data and Some Implications for Network Design

by

GUDMUND HØST AND XENI K. DIMAKOS

Norwegian Computing Center,  
P.O.Box 114 Blindern, N-0314 Oslo

Det Norske Meteorologiske Institutt  
Research Report no. 53

# Preface and Acknowledgements

This report was prepared for the twentyfirst session of the Steering Body of EMEP (Co-operative Programme for Monitoring and Evaluation of the Long Range Transmission of Air Pollutants in Europe). The purpose of this report is to present results from a statistical method for estimation of sulphur and nitrogen emissions in Europe. The method uses the EMEP-model to match official emissions data from European countries and measured depositions and concentrations from the EMEP monitoring network. The application is to wet deposition of sulphate and nitrate, and to air concentrations of sulphate and sulphur dioxide. Validation results for the statistical model are also presented, and a some implications for monitoring network design are discussed.

This work was supported by the Norwegian Meteorological Institute and the Research Council of Norway, project no. 113850/720. The authors are indebted to Erik Berge of the Norwegian Meteorological Institute for helpful discussions and encouragement. Svetlana Tsyro of the Norwegian Meteorological Institute are thanked for providing model results and emissions data. The authors are also grateful to the Norwegian Institute of Air Research for providing the monitoring data and to Turid Follestad of Norwegian Computing Center for some useful software subroutines.

# Contents

<b>1</b>	<b>Introduction</b>	<b>4</b>
<b>2</b>	<b>Data</b>	<b>4</b>
2.1	Data Types . . . . .	4
2.2	Processing of Monitoring Data . . . . .	8
<b>3</b>	<b>Method</b>	<b>10</b>
3.1	General . . . . .	10
3.2	The Precision of Emission Estimates . . . . .	11
<b>4</b>	<b>Results</b>	<b>12</b>
4.1	Sulphur . . . . .	13
4.2	Nitrogen . . . . .	18
4.3	Case Study: Some Implications for Network Design . . . . .	20
<b>5</b>	<b>Validation</b>	<b>25</b>
5.1	Validation of Sulphur Components . . . . .	26
5.2	Validation of Nitrogen in Precipitation . . . . .	33
<b>6</b>	<b>Discussion</b>	<b>36</b>
<b>7</b>	<b>Concluding Remarks</b>	<b>38</b>
<b>A</b>	<b>Monitoring data</b>	<b>38</b>

# 1 Introduction

Various protocols to the Convention on Long-Range Transport of Air Pollutants have put explicit obligations on the contracting parties to reduce their emissions of various air pollutants. The Oslo Protocol of 1994 establishes an Implementation Committee to review compliance by the parties with their obligations. The technical support for reviewing national emission reductions is a responsibility of the Meteorological Synthesizing Center - West (MSC-W). This report is prepared as part of work on developing a statistical method for estimating emissions of air pollutants from the European countries.

The method used in this work is basically the same as described in Høst (1996*b*), but a new feature is the inclusion of an additional parameter for contributions of indeterminate origin. The data are again collected from 1990, using revised reported emissions and adding new chemical components to the analysis. The new components are particulate sulphate in air, sulphur dioxide in air and nitrate in precipitation. A reanalysis of the sulphate in precipitation data is also included. Our report also contains extensive results on validation of the statistical model, as well as results on sensitivity to monitoring stations. In addition, a preliminary analysis of monitoring strategies in relation to emission estimation is included.

## 2 Data

### 2.1 Data Types

*The emission data* consists of 48 official national and regional data and estimates of emissions of sulphur dioxide and nitrogen oxides within the EMEP modelling area during 1990, as described in (Barrett & Berge 1996). The data values for sulphur and nitrogen are presented in Tables 2 and 4 respectively, and we will in the following refer to these values as the *prior emissions*.

*The monitoring data* are yearly averages of various observed components from the EMEP monitoring network during 1990. For the precipitation components (sulphate and nitrate) the data value for each monitoring station, is a weighted average, with the weights taken as the proportion of daily precipitation to the total 1990 precipitation. Tables 6, 7 and 8 in Appendix A, show the number of daily values per month for sulphate in precipitation, particulate sulphate in air and nitrate in precipitation.

The EMEP-model values that we have used, describe the contribution of emissions from each country and region within the EMEP domain to each  $150\text{ km} \times 150\text{ km}$  grid block.

By dividing by the regional emissions, we obtain transfer coefficients, or contributions per unit emission. Each transfer coefficient describes the relationship between the modeled concentration at a monitoring station and a unit emission from a specific country or region. The transfer coefficients are the link between the observed concentrations and the reported and estimated emissions, and we will refer to them as the *EMEP-model coefficients*.

Figures 1–4 show the observed concentrations versus the concentrations predicted using the reported emissions after the data processing described in Section 2.2. We see that for sulphate in precipitation and sulphur dioxide the concentrations predicted by the EMEP-model are mostly smaller than the observed concentrations. For particulate sulphate in air the observed concentrations are mostly larger than the EMEP predictions, while for nitrate in precipitation there is no obvious pattern.

Of the components investigated nitrate in precipitation (Figure 4) shows the best correspondance between observed and predicted values, while the largest scatter occurs for sulphur dioxide (Figure 2). Further results for sulphur dioxide are not presented in this paper, because we were not able to fit an adequate statistical model to these data. This means that discrepancies between the EMEP-model and observations of sulphur dioxide cannot be attributed to uncertainties in the emission data. To get reasonable results for sulphur dioxide we would probably have to do more elaborate modeling.

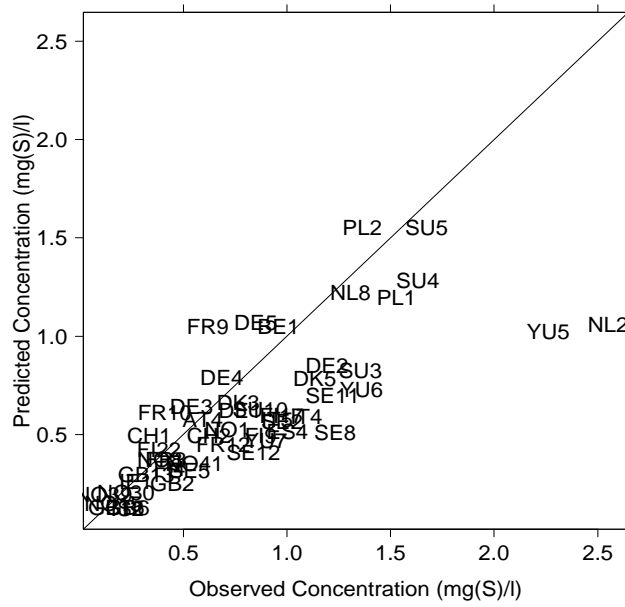


Figure 1: *Observed versus predicted concentrations for sulphate in precipitation.*

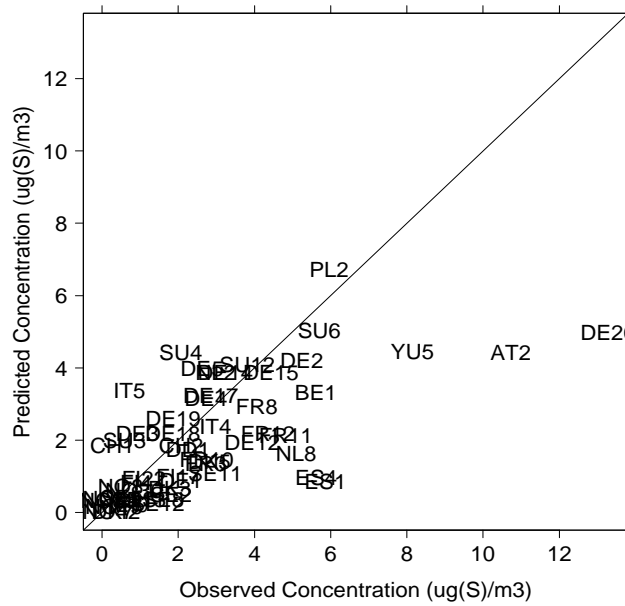


Figure 2: *Observed versus predicted concentrations for sulphur dioxide.*

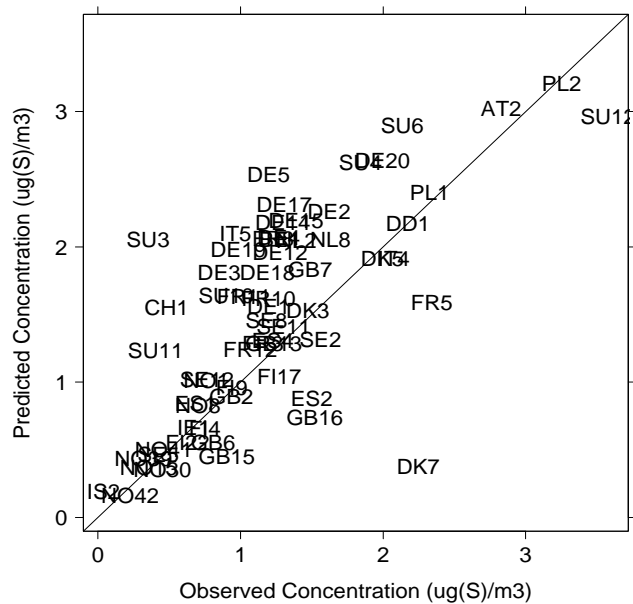


Figure 3: *Observed versus predicted concentrations for particulate sulphate in air.*

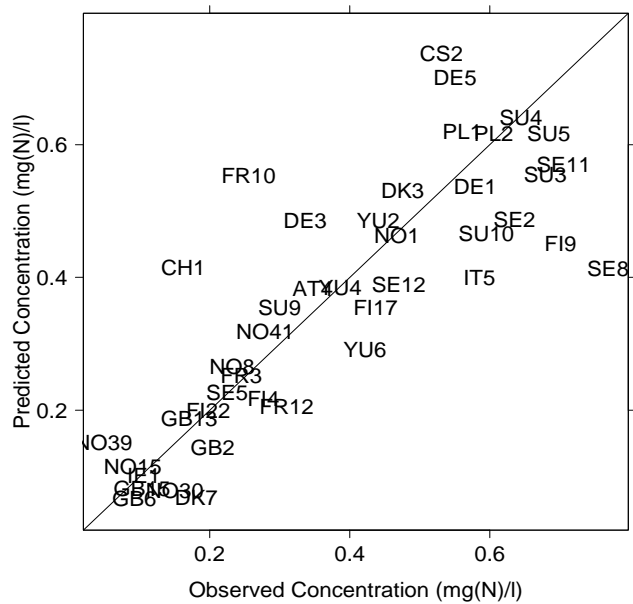


Figure 4: *Observed versus predicted concentrations for nitrate in precipitation.*

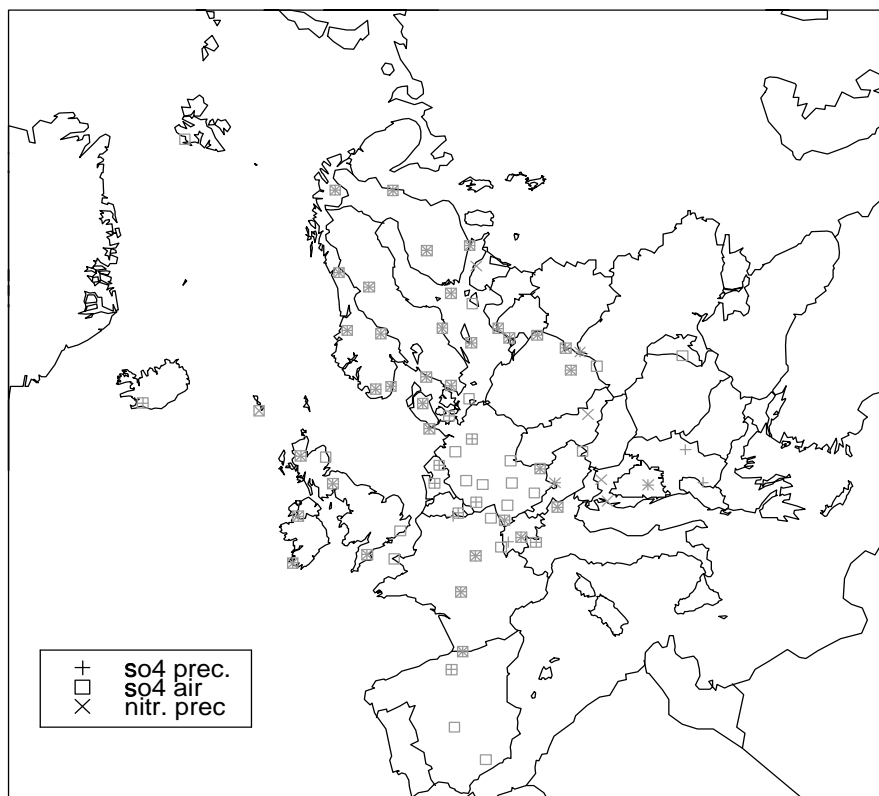


Figure 5: *The locations of the monitoring stations used in the analysis.*

## 2.2 Processing of Monitoring Data

Some screening of the monitoring data is needed to ensure a minimum of compability between predicted and measured concentrations. The procedure described here should be regarded as a preliminary tool, and it is recommended that in the future the criteria should be established in close cooperation with MSC-W and CCC. The processing of the monitoring data amounts to exclusion of monitoring stations from the analysis based on selected criterias. For all the analysed components the following rules are applied:



- (I) A station is excluded from the analysis if the monitoring data for the station includes a period of three or more consecutive months without observations. This ensures that all monitoring stations used in the analysis cover all the seasons, and that the average value is not a reflection of a seasonal factor only.
- (II) A station is excluded if the total data coverage is low. For sulphate and nitrate in precipitation the data coverage has to be larger than 20%, that is, we require that more than 20% of all days in the year are represented. For particulate sulphate in air the requirement is set to 50%. This stricter requirement is due to the fact that the data coverage for these components are much better than for sulfate and nitrate in precipitation. The 20% and 50% requirements used here, are results of a rather heuristic process of looking at the data and choosing a reasonable criterion.
- (III) A station is excluded from the analysis if it is close to large emission sources. The EMEP-model is not designed to describe local phenomena. The locations of monitoring stations are selected to measure background concentrations. As a result, a station located in a grid block with large local emissions may measure lower concentrations than the true grid block average. To avoid the systematic errors that might be induced by these effects, we exclude stations in a grid block where the total emission is larger than the 95% fractile from the distribution of the emissions from all grid squares. The 95% fractiles for sulphur dioxide and nitrogen oxide are 164134 and 117631 tonnes per grid block respectively.
- (IV) A station is removed if the quality class of the laboratories of the country is “Not satisfactory” as described in EB.AIR/GE.1/R.113, page 19, Annex III.
- (V) A station is excluded if it is an outlier in terms of an extreme deviation between the observed value and the prediction of the EMEP-model or compared to the other observations. In this context, “extreme” is taken as a factor of 10.

The locations of the monitoring stations used in the analysis are shown in Figure 5. The number of stations used in the analysis were 50, 64 and 42 for sulphate in precipitation, particulate sulphate in air and nitrate in precipitation respectively. Table 9 in Appendix A shows all the monitoring stations that are excluded from the analysis, and indicates the reason for the exclusion.

## 3 Method

### 3.1 General

This section introduces some basic notation. The interested reader is referred to Høst (1996b) for the details of the method. Consider a random field  $y(\mathbf{x})$ , where  $\mathbf{x}$  is the location within some geographic region. Let  $y(\mathbf{x})$  have the decomposition

$$y(\mathbf{x}) = \mathbf{b}'(\mathbf{x})\boldsymbol{\beta} + \epsilon(\mathbf{x}). \quad (1)$$

Here,  $\mathbf{b}(\mathbf{x})$  is an  $m$ -vector of known functions and  $\boldsymbol{\beta}$  is an  $m$ -vector of parameters to be estimated. Furthermore,  $\epsilon(\mathbf{x})$  is a Gaussian random field with zero mean and covariance

$$\text{Cov} \{y(\mathbf{x}_1), y(\mathbf{x}_2)\} = \sigma^2 \rho(\|\mathbf{x}_1 - \mathbf{x}_2\|; a). \quad (2)$$

Here,  $\sigma^2$  is the variance of the residual process,  $a$  is a “range” parameter, and  $\rho(\cdot; a)$  is a correlation function to be fitted to data. The focus of this report is on estimation of  $\boldsymbol{\beta}$ , but the unknown parameters  $\sigma^2$  and  $a$  must also be estimated. Future extensions of the statistical model may include the time dimension, but this is beyond the scope of this work.

In our application,  $y(\mathbf{x})$  represents the (true) concentration at location  $\mathbf{x}$  for a given year. Furthermore,  $b_j(\mathbf{x})$  represents the contribution to this concentration as predicted by the EMEP-model for a unit emission from country  $j$ ;  $j = 1, \dots, m$ . The unknown emission from country  $j$  is  $\beta_j$ , so  $\mathbf{b}'(\mathbf{x})\boldsymbol{\beta}$  is the concentration predicted by the EMEP model. Deviations between the true concentrations and the concentrations predicted by the EMEP model are absorbed into the residual field  $\epsilon(\mathbf{x})$ .

An important characteristic of the problem is that the number of parameters to be estimated is large. Traditional methods of estimation, such as least squares, will not work for the present application. This calls for some method of regularization.

We introduce a *Bayesian* framework (Berger 1985), and use *prior* and *posterior* to denote knowledge of the emissions  $\boldsymbol{\beta}$  without and with observations of the concentration field. The Bayesian framework gives a mathematically consistent method of regularization, with the advantage that prior information may be incorporated in the analysis. A priori, we assume independence, and take each national emission  $\beta_j$  to be a normally distributed random variable with prior mean  $\beta_{0j}$  and prior coefficient of variation  $\gamma_0$ . In this context,

$\gamma_0$  may be regarded as the regularization parameter. The model for prior variance signifies that large reported emissions are likely to be more uncertain than small emissions. Without observations of the  $y$ -field, the natural “estimate” of  $\boldsymbol{\beta}$  is  $\boldsymbol{\beta}_0$ , i.e. the reported emissions.

For the variance parameter  $\sigma^2$ , we use the commonly used improper prior  $p(\sigma^2) \propto \sigma^{-2}$ , and for the range parameter  $a$ , we use the prior  $p(a) = (1+a)^{-2}$ , as suggested by Handcock & Wallis (1994). The prior chosen for  $a$  reflects that we expect the residual field  $\epsilon(\mathbf{x})$  a priori to capture high-frequency fluctuations. The interested reader is referred to Berger (1985) for an introduction to Bayesian statistics and discussions on how to choose prior densities.

In order to avoid time-consuming integration, we estimate the parameters  $\boldsymbol{\beta}, \sigma^2, a$  by the maximum of the posterior density. Since the posterior density is just a penalized likelihood function, we may apply a modified method of numerical maximum likelihood estimation for computation of the posterior mode. We use the scoring method, which was introduced for spatial data by Mardia & Marshall (1984).

Let  $\mathbf{y} = (y(\mathbf{x}_1), \dots, y(\mathbf{x}_n))'$  be the data vector of  $n$  concentration measurements. Conditional on the residual covariance parameters  $\sigma^2$  and  $a$ , the posterior mode estimate is

$$\hat{\boldsymbol{\beta}}|(\sigma^2, a) = (\boldsymbol{\Sigma}_0^{-1} + \sigma^{-2} \mathbf{B}' \mathbf{R}_a^{-1} \mathbf{B})^{-1} (\boldsymbol{\Sigma}_0^{-1} \boldsymbol{\beta}_0 + \sigma^{-2} \mathbf{B}' \mathbf{R}_a^{-1} \mathbf{y}). \quad (3)$$

Here, the  $j$ 'th column of  $\mathbf{B}$  is  $b_j(\cdot)$ , evaluated at the data locations  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ . Furthermore, the  $(i, j)$ 'th element of the correlation matrix  $\mathbf{R}_a$  is  $\rho(\|\mathbf{x}_i - \mathbf{x}_j\|; a)$ . Also,  $\boldsymbol{\Sigma}_0 = \gamma_0^2 \text{diag}(\boldsymbol{\beta}_0 \boldsymbol{\beta}_0')$ .

### 3.2 The Precision of Emission Estimates

The uncertainty of  $\hat{\boldsymbol{\beta}}$  is described by the conditional covariance given  $(\sigma^2, a)$ . The conditional covariance of  $\hat{\boldsymbol{\beta}}$  is

$$\text{Var}\{\hat{\boldsymbol{\beta}}|(\sigma^2, a)\} = (\boldsymbol{\Sigma}_0^{-1} + \sigma^{-2} \mathbf{B}' \mathbf{R}_a^{-1} \mathbf{B})^{-1}. \quad (4)$$

It may be convenient to summarize the covariance matrix into a single number. One such measure is the total variation. The total variation of  $\hat{\boldsymbol{\beta}}|(\sigma^2, a)$  can be defined as the trace of (4):

$$\text{tr Var}\{\hat{\boldsymbol{\beta}}|(\sigma^2, a)\} = \text{tr}(\boldsymbol{\Sigma}_0^{-1} + \sigma^{-2}\mathbf{B}'\mathbf{R}_a^{-1}\mathbf{B})^{-1}.$$

This quantity does not depend on the value of  $\hat{\boldsymbol{\beta}}$ . However, it does depend on the covariance parameters  $(\sigma^2, a)$ , on the meteorology  $\mathbf{b}(\mathbf{x})$ , on the locations  $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$  and on the number of monitoring stations  $n$ . These dependencies are quite complicated, and we will restrict attention to some special cases.

1. Assume that prior information can be neglected (i.e. many monitoring stations) and that the residuals are uncorrelated. Then the conditional covariance matrix of  $\hat{\boldsymbol{\beta}}$  is  $\sigma^2(\mathbf{B}'\mathbf{B})^{-1}$ . The total variation would then be small if the monitoring stations are located to reduce co-linearities in  $\mathbf{B}$ . This means that the monitoring stations should be located so that typical patterns of air transport from distinct countries can be separated. Of course, such patterns may vary from year to year.
2. Meteorology and model resolution will put limitations on the capability of the network to monitor emissions. In particular, a spectral analysis of the conditional covariance matrix may provide information on the effective number of parameters that can be estimated.
3. Locating two monitoring stations within the same EMEP-model grid square does not reduce the total variation, although it may be useful for assessing subgrid variability.
4. Positive residual correlation is likely to increase the total variation of  $\hat{\boldsymbol{\beta}}$ , unless the inter-station distance is larger than the correlation range.
5. If the correlation range is smaller than the EMEP-model grid diameter and the monitoring network has larger inter-station distance than the EMEP-model grid diameter, then residual correlation can be neglected in the estimation of emissions.

## 4 Results

The prior coefficient of variation  $\gamma$  was set to 0.3 based on recommendations from MSC-W. This choice of  $\gamma$  implies that the reported emissions are taken to have a standard deviation of 0.3 times the reported value.

For all three components, sulphate in precipitation, particulate sulphate in air and nitrate in precipitation, the model was fitted using an exponential covariance function,  $c(h; \sigma^2, a) = \sigma^2 \exp(-3h/a)$ , where  $\sigma^2$  and  $a$  are parameters to be estimated. This covariance function implies that locations more than a distance  $a$  apart have correlation less

than 0.05. A visual inspection of fitted covariance functions and empirical correlograms showed reasonable correspondance. Hence, other parametric covariance functions were not investigated.

## 4.1 Sulphur

The estimated range and residual variance for the sulphur components are shown in Table 1.

Component	$\hat{a}$	$\widehat{\text{sd}}(\hat{a})$	$\hat{\sigma}^2$	$\widehat{\text{sd}}(\hat{\sigma}^2)$
SO <sub>4</sub> in precipitation	434	172	0.11922	0.02529
SO <sub>4</sub> in air	455	153	0.23086	0.04633

Table 1: *Estimated range,  $\hat{a}$ , residual variance,  $\hat{\sigma}^2$ , and their estimated standard deviation for sulphate in precipitation and particulate sulphate in air. The range is measured in kilometers and the residual variance in  $[\text{mg}(S)/l]^2$  for sulphate in precipitation and  $[\mu\text{g}(S)/\text{m}^3]^2$  for particulate sulphate in air.*

The estimated range parameters are very similar. These estimates indicate that the residual correlation is negligible for distances greater than 450 *km*, alternatively, that the EMEP-model resolves spatial structure at scales greater than 3 EMEP grid units. Compared to the analysis of sulphate in precipitation by Høst (1996*b*), the estimated correlation range is smaller in the present work. This may be due to contributions from inattributable sources, which was not included in Høst (1996*b*). There are also some differences in the data processing used in these two analyses. Note that the estimated standard deviation of the range parameter is smaller in the present analysis.

Figure 6 shows the reported sulphur emissions (prior emissions) and the sulphur emissions estimated from the statistical method (posterior emissions) for sulphate in precipitation. The largest difference between reported and estimated emissions seems to occur for Bulgaria, former German Democratic Republic, Italy and United Kingdom. In all these countries (regions), the estimated values are larger than the reported values. The numeric values are given in Table 2. Compared to the results of Høst (1996*b*), there are some similarities and some differences. The difference between reported and estimated emissions are in good agreement for Italy and Great Britain, but not for the German Democratic Republic and Bulgaria. The estimated sum of all emissions (42591 ktonnes) is very similar to the value obtained by Høst (1996*b*), even though the former analysis did not take

account of the inattributable sources. In the present analysis, the contribution from inattributable sources is estimated to be 34% larger than what was used in the calculations by MSC-W. The most precise estimates were obtained for Poland and for United Kingdom. The uncertainty in the results we present is large, and the difference between reported and estimated emissions are not significant. To obtain significant results we would need more monitoring stations or temporal data.

Figure 7 shows the reported sulphur emissions (prior emissions) and the sulphur emissions estimated from the statistical method (posterior emissions) for particulate sulphate in air. Here, the largest difference between reported and estimated emissions occurs for the former German Democratic Republic. The statistical method suggests that sulphur emissions from German Democratic Republic were 1622 ktonnes smaller than reported. Although surprising, this may not be totally unrealistic if one considers the political situation in German Democratic Republic during 1990.

The sulphur emissions estimated from the two sulphate data sets are not in very close agreement. This may be due to location and number of monitoring stations, as well as the measured values. Disagreements may also be due to uncertainty in the EMEP-model or in the statistical method. Since we are still using a limited amount of data, such uncertainty cannot be avoided.

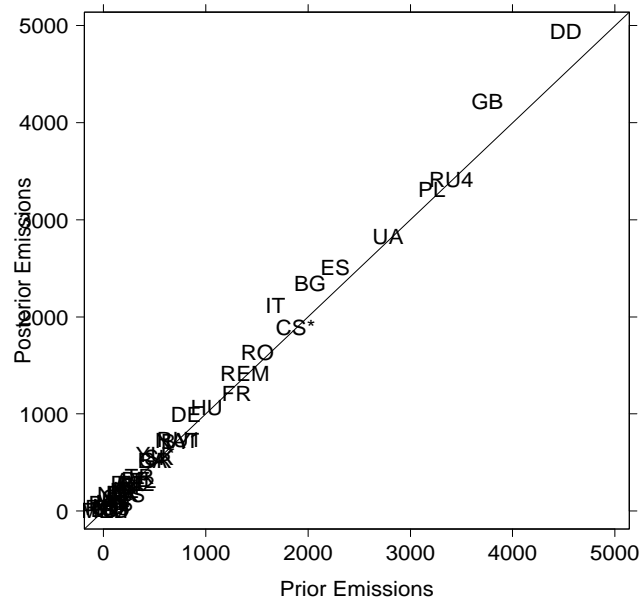


Figure 6: *Prior versus posterior emissions for sulphate in precipitation.*

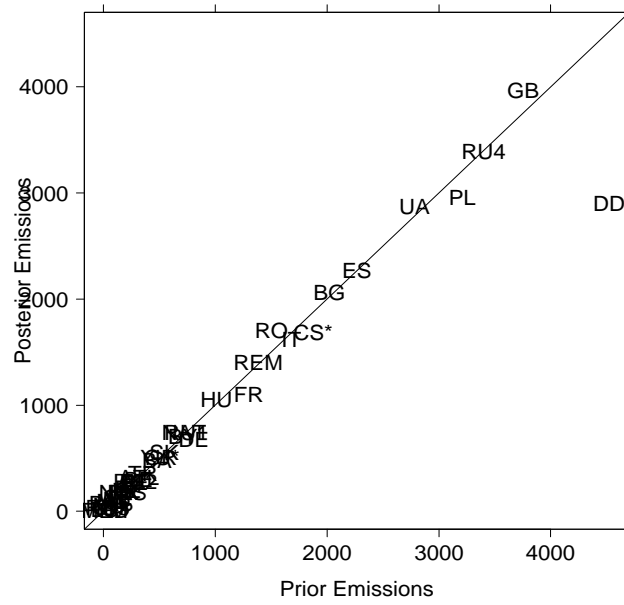


Figure 7: *Prior versus posterior emissions for particulate sulphate in air.*

Code	Region	Prior	Posterior	Increase	Rel.Change	CV
AL	Albania	120	123	3	1.03	0.3
AT	Austria	90	90	0	1	0.3
BY	Belarus	710	721	11	1.02	0.3
BE	Belgium	317	323	6	1.02	0.29
BA	Bosnia & Herzegovina	480	527	47	1.1	0.29
BG	Bulgaria	2020	2343	323	1.16	0.3
HR	Croatia	180	182	2	1.01	0.3
CS*	Czech Republic	1876	1893	17	1.01	0.29
DK	Denmark	180	187	7	1.04	0.3
EE	Estonia	275	282	7	1.03	0.3
FI	Finland	260	270	10	1.04	0.3
FR	France	1300	1213	-87	0.93	0.29
DD	Germany, Dem. Rep.	4521	4942	421	1.09	0.28
DE	Germany, Fed. Rep.	810	998	188	1.23	0.29
GR	Greece	510	518	8	1.02	0.3
HU	Hungary	1010	1072	62	1.06	0.3
IS	Iceland	24	24	0	1	0.3
IE	Ireland	178	179	1	1	0.3
IT	Italy	1681	2119	438	1.26	0.28
RU3	Kaliningrad <sup>(1)</sup>	36	37	1	1.02	0.3
RU1	Kola/Karelia <sup>(1)</sup>	740	737	-3	1	0.3
LV	Latvia	115	115	0	1	0.3
RU2	Leningrad/Novgorod-Pskov <sup>(1)</sup>	285	287	2	1.01	0.3
LT	Lithuania	222	227	5	1.02	0.3
LU	Luxembourg	14	14	-0	1	0.3
NL	Netherlands	205	214	9	1.04	0.3
NO	Norway	54	54	0	1	0.3
PL	Poland	3210	3312	102	1.03	0.24
PT	Portugal	283	284	1	1	0.3
MD	Republic of Moldova	91	91	0	1	0.3
RO	Romania	1504	1631	127	1.08	0.3
RU4	Rest of Russiav <sup>(1)</sup>	3398	3416	17	1.01	0.3
SK	Slovakia	543	554	11	1.02	0.3
SI	Slovenia	195	196	1	1.01	0.3
ES	Spain	2266	2507	241	1.11	0.28
SE	Sweden	136	140	4	1.03	0.3
CH	Switzerland	43	43	-0	1	0.3
FYM	The FYR Macedonia	10	10	0	1	0.3
TR	Turkey	354	355	1	1	0.3
UA	Ukraine	2782	2831	50	1.02	0.3
GB	United Kingdom	3752	4217	465	1.12	0.25
YU*	Yugoslavia <sup>(2)</sup>	508	588	80	1.16	0.3
BAS	The Baltic Sea	72	73	1	1.02	0.3
BLS	The Black Sea	0	0	0	1	0.3
MED	The Mediterranean	12	12	0	1	0.3
NOS	The North Sea	174	176	2	1.01	0.3
ATL	Remaining Atlantic	316	316	1	1	0.3
REM	Remaining Landbased <sup>(3)</sup>	1383	1417	34	1.02	0.3
NAT	Nat.emis. from ocean	724	731	7	1.01	0.3
IND	Indeterminate	-	-	-	1.34	0.27
	Total	39969	42591	2623	1.07	-

Table 2: *Result of emissions estimation for sulphate in precipitation. The table shows the prior emissions, posterior emissions and increase in units of 1000 tonnes of sulphur dioxide. (1) These regions form the region “Russian Federation” in EMEP MSC-W. (2) The Federal Republic of Yugoslavia. (3) This region is the sum of ‘the regions ‘Africa, North’, “Georgia”, “Kazakhstan” and “Volcanic” in (Barrett & Berge 1996).*



Code	Region	Prior	Posterior	Increase	Rel.Change	CV
AL	Albania	120	120	0	1	0.3
AT	Austria	90	90	-0	1	0.3
BY	Belarus	710	706	-4	0.99	0.3
BE	Belgium	317	301	-16	0.95	0.3
BA	Bosnia & Herzegovina	480	480	0	1	0.3
BG	Bulgaria	2020	2066	46	1.02	0.3
HR	Croatia	180	181	1	1	0.3
CS*	Czech Republic	1876	1685	-191	0.9	0.29
DK	Denmark	180	181	1	1	0.3
EE	Estonia	275	276	1	1	0.3
FI	Finland	260	263	3	1.01	0.3
FR	France	1300	1104	-196	0.85	0.29
DD	Germany, Dem. Rep.	4521	2899	-1622	0.64	0.23
DE	Germany, Fed. Rep.	810	680	-130	0.84	0.29
GR	Greece	510	511	1	1	0.3
HU	Hungary	1010	1058	48	1.05	0.3
IS	Iceland	24	24	-0	1	0.3
IE	Ireland	178	179	1	1	0.3
IT	Italy	1681	1623	-58	0.97	0.27
RU3	Kaliningrad	36	35	-1	0.98	0.3
RU1	Kola/Karelia	740	746	6	1.01	0.3
LV	Latvia	115	115	-0	1	0.3
RU2	Leningrad/Novgorod-Pskov	285	286	1	1	0.3
LT	Lithuania	222	219	-3	0.99	0.3
LU	Luxembourg	14	14	-0	1	0.3
NL	Netherlands	205	201	-4	0.98	0.3
NO	Norway	54	54	-0	1	0.3
PL	Poland	3210	2956	-254	0.92	0.26
PT	Portugal	283	285	2	1.01	0.3
MD	Republic of Moldova	91	91	0	1.01	0.3
RO	Romania	1504	1709	205	1.14	0.29
RU4	Rest of Russia	3398	3395	-3	1	0.3
SK	Slovakia	543	554	11	1.02	0.3
SI	Slovenia	195	194	-1	0.99	0.3
ES	Spain	2266	2266	1	1	0.27
SE	Sweden	136	135	-1	1	0.3
CH	Switzerland	43	42	-1	0.98	0.3
FYM	The FYR Macedonia	10	10	0	1	0.3
TR	Turkey	354	354	0	1	0.3
UA	Ukraine	2782	2877	95	1.03	0.3
GB	United Kingdom	3752	3964	212	1.06	0.24
YU*	Yugoslavia	508	511	3	1.01	0.3
BAS	The Baltic Sea	72	72	-0	1	0.3
BLS	The Black Sea	0	0	0	1	0.3
MED	The Mediterranean	12	12	0	1	0.3
NOS	The North Sea	174	173	-0	1	0.3
ATL	Remaining Atlantic	316	317	2	1.01	0.3
REM	Remaining Landbased	1383	1401	18	1.01	0.3
NAT	Nat.emis. from ocean	724	746	22	1.03	0.3
IND	Indeterminate	-	-	-	1.03	0.3
	Total	39969	38161	-1805	0.95	-

Table 3: *Result of emissions estimation for particulate sulphate in air. The table shows the prior emissions, posterior emissions and increase in units of 1000 tonnes of sulphur dioxide.*

## 4.2 Nitrogen

The estimated correlation range for nitrate in precipitation was  $444km$  (standard deviation of estimate  $190 km$ ) and the estimated residual variance was  $0.0103 [mg(N)/l]^2$  (standard deviation of estimate  $0.002 mg(N)/l$ ).

Figure 8 shows the reported emissions of nitrogen oxides (prior emissions) and the emissions estimated from the statistical method (posterior emissions) for nitrate in precipitation. The figure indicates that emissions of nitrogen oxides may be smaller than reported in France and Federal Republic of Germany, and emissions may be larger than reported in Italy and United Kingdom. The sum of estimated emissions is very close to the sum of reported emissions. The most precise estimates, in terms of posterior coefficient of variation, is obtained for Federal Republic of Germany, Italy, Poland (no significant difference in emissions) and United Kingdom.

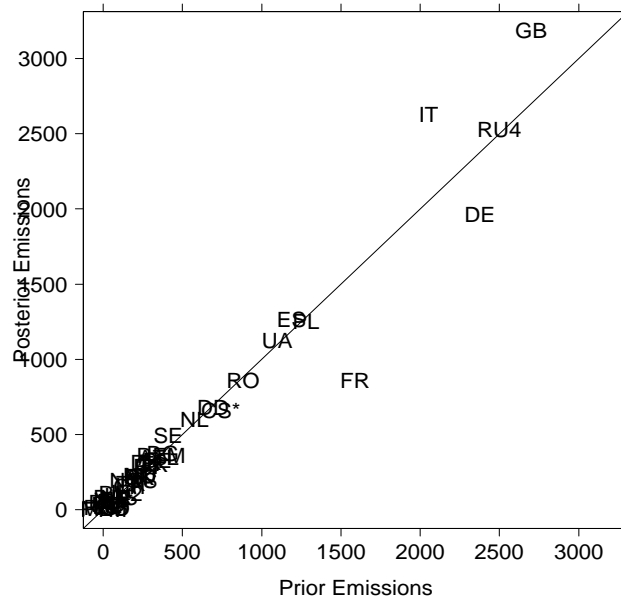


Figure 8: *Prior versus posterior emissions for nitrate in precipitation.*

Code	Region	Prior	Posterior	Increase	Rel.Change	CV
AL	Albania	30	30	0	1	0.3
AT	Austria	222	216	-6	0.97	0.3
BY	Belarus	285	291	6	1.02	0.3
BE	Belgium	343	327	-16	0.95	0.3
BA	Bosnia & Herzegovina	54	54	0	1	0.3
BG	Bulgaria	376	374	-1	1	0.3
HR	Croatia	83	81	-1	0.98	0.3
CS*	Czech Republic	742	659	-83	0.89	0.29
DK	Denmark	269	313	44	1.16	0.29
EE	Estonia	72	73	1	1.01	0.3
FI	Finland	300	309	9	1.03	0.29
FR	France	1590	861	-729	0.54	0.25
DD	Germany, Dem. Rep.	694	683	-11	0.98	0.3
DE	Germany, Fed. Rep.	2377	1966	-411	0.83	0.22
GR	Greece	306	306	-0	1	0.3
HU	Hungary	238	231	-7	0.97	0.3
IS	Iceland	20	20	0	1	0.3
IE	Ireland	115	116	1	1.01	0.3
IT	Italy	2053	2628	575	1.28	0.24
RU3	Kaliningrad	16	16	0	1.01	0.3
RU1	Kola/Karelia	48	48	0	1	0.3
LV	Latvia	93	95	2	1.02	0.3
RU2	Leningrad/Novgorod-Pskov	110	111	1	1.01	0.3
LT	Lithuania	158	165	7	1.04	0.3
LU	Luxembourg	23	23	-0	1	0.3
NL	Netherlands	575	600	25	1.04	0.3
NO	Norway	230	232	2	1.01	0.3
PL	Poland	1280	1254	-26	0.98	0.23
PT	Portugal	221	222	1	1	0.3
MD	Republic of Moldova	35	35	0	1	0.3
RO	Romania	883	862	-21	0.98	0.3
RU4	Rest of Russia	2500	2530	30	1.01	0.3
SK	Slovakia	227	220	-7	0.97	0.3
SI	Slovenia	53	53	-0	0.99	0.3
ES	Spain	1188	1264	76	1.06	0.28
SE	Sweden	411	498	87	1.21	0.29
CH	Switzerland	166	154	-12	0.93	0.3
FYM	The FYR Macedonia	2	2	0	1	0.3
TR	Turkey	175	175	0	1	0.3
UA	Ukraine	1097	1130	33	1.03	0.3
GB	United Kingdom	2702	3185	483	1.18	0.22
YU*	Yugoslavia	66	66	0	1	0.3
BAS	The Baltic Sea	80	83	3	1.04	0.3
BLS	The Black Sea	0	0	0	1	0.3
MED	The Mediterranean	13	13	0	1	0.3
NOS	The North Sea	192	195	3	1.02	0.3
ATL	Remaining Atlantic	349	348	-1	1	0.3
REM	Remaining Landbased	364	365	1	1	0.3
NAT	Nat.emis. from ocean	0	0	0	0	0
IND	Indeterminate	-	-	-	1.02	0.27
	Total	23426	23482	58	1.00	-

Table 4: *Result of emissions estimation for nitrate in precipitation. The table shows the prior emissions, posterior emissions and increase in units of 1000 tonnes of nitrogen oxides.*

### 4.3 Case Study: Some Implications for Network Design

In this Section we present some results on the sensitivity of the uncertainty in emission estimates to the location of monitoring stations using nitrate in precipitation data.

We are not particularly interested in the emission estimates in this study,

The quantity of interest in this particular study is the uncertainty associated with the emission estimates, as described by (eq:varbetahat). Figure 9 shows the standard deviations of the estimated national emissions relative to the assumed standard deviation of reported emissions for some nations and regions. 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 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.

Figure 10 shows the estimated reduction in the standard deviation of national emissions resulting from adding each of the candidate sites to the network. These arbitrary candidate sites are denoted by ES\*1,ES\*2 (Western and Southern Spain); FR\* (Southern France); IT\*1,IT\*2 (Sicily, Southern Italy); PL\* (Western Poland); RO\* (Romania); UA\* (Ukraine); BY\* (Belarus). 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 11 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 11 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 12 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).

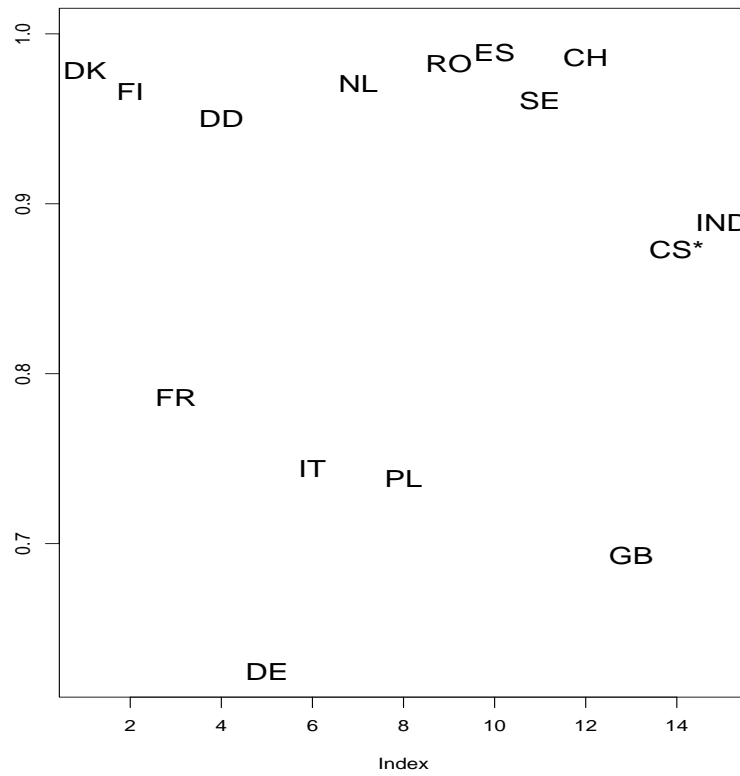


Figure 9: *Estimated reduction in the standard deviation of national emissions resulting from including monitoring data.*

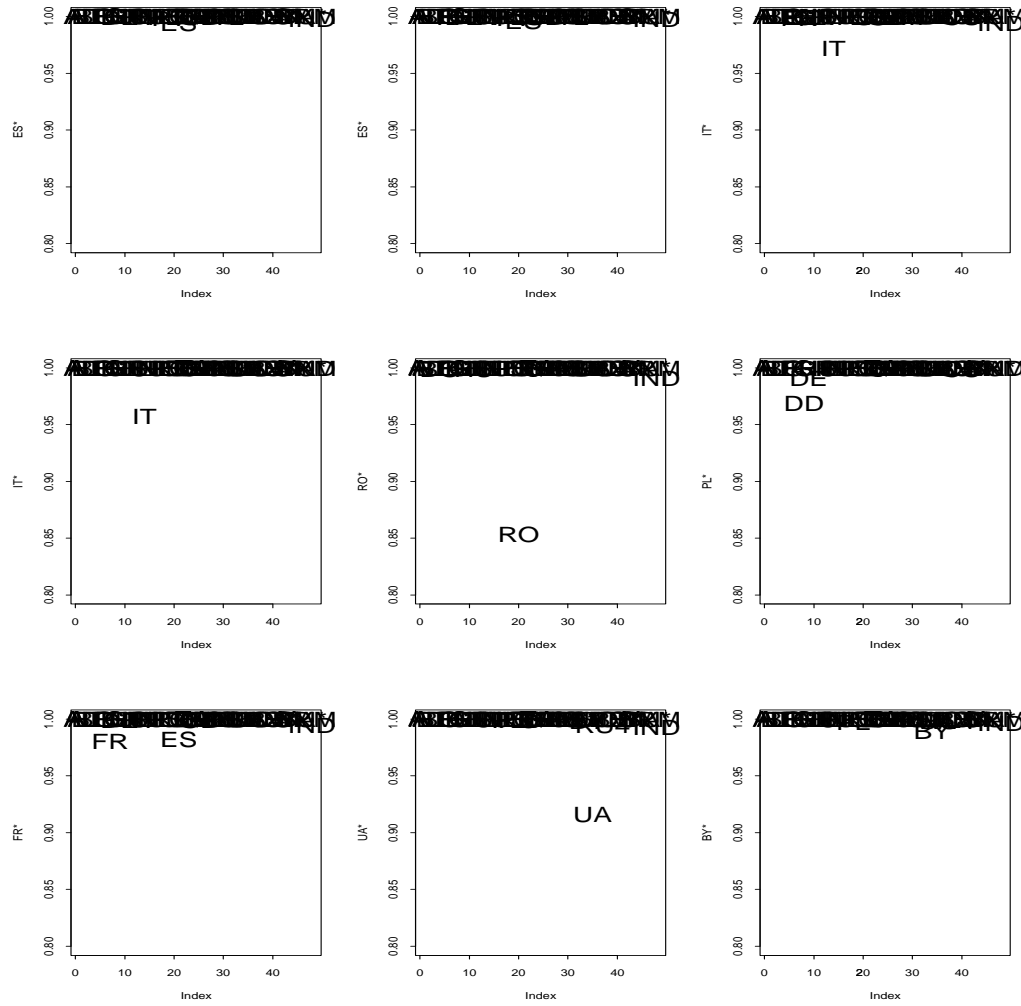


Figure 10: *Estimated reduction in the standard deviation of estimated national emissions resulting from adding pseudo-stations.*

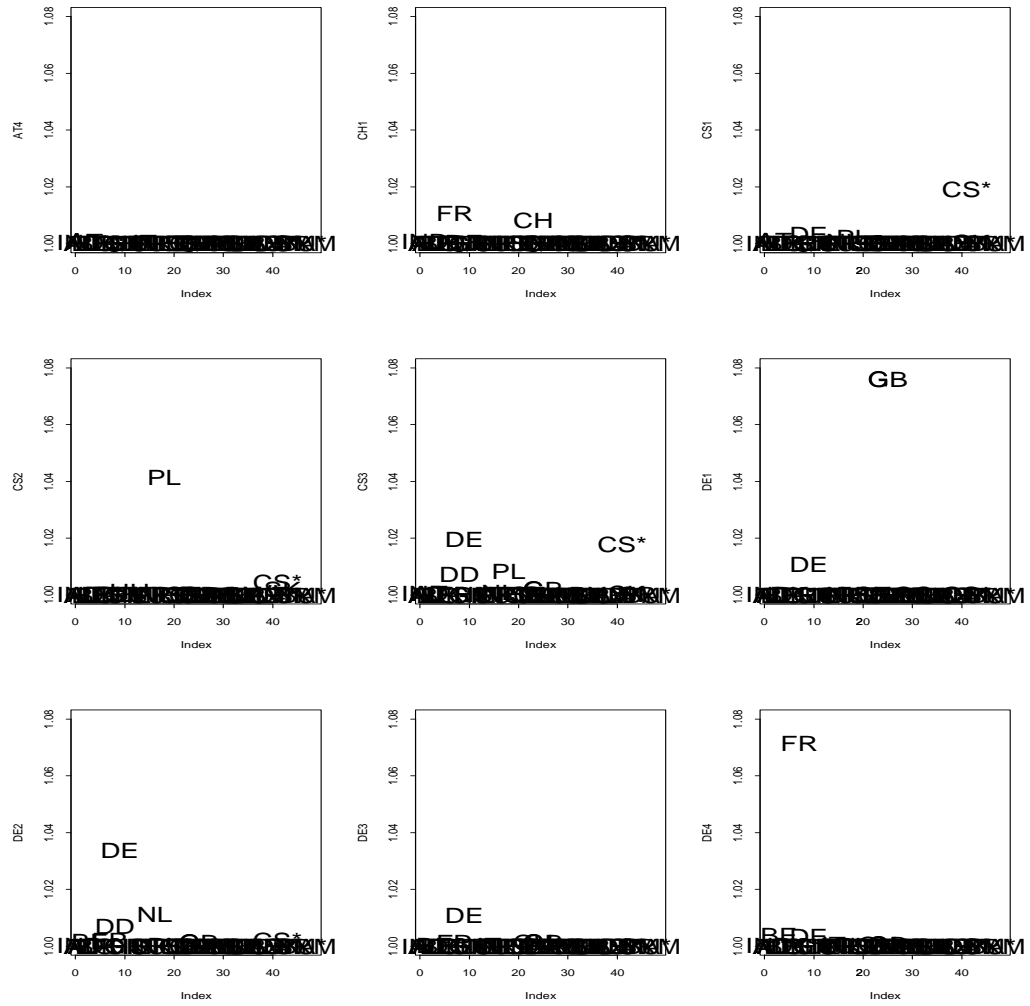


Figure 11: *Estimated increase in the standard deviation of estimated national emissions resulting from deleting selected monitoring stations.*



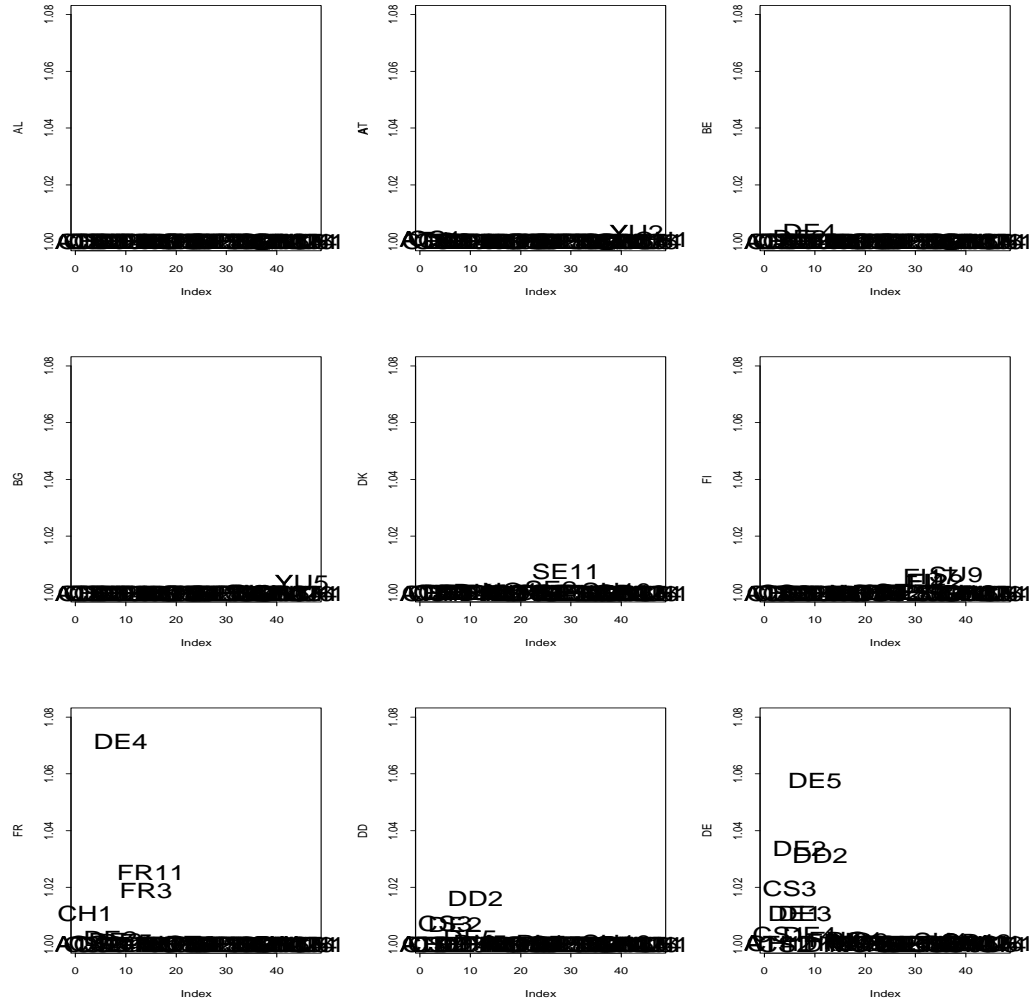


Figure 12: *Estimated increase in the standard deviation of estimated selected national emissions resulting from deleting monitoring stations.*

## 5 Validation

In this section, we investigate the statistical model and results by two different cross-validation exercises. In the first exercise, we study predicted concentrations. However, to get estimates of prediction errors we need a statistical predictor. A statistical predictor which is compatible with the EMEP-model predictions and the statistical model we have used is the Bayesian Kriging predictor described in Høst (1996b). This predictor is equal to the EMEP-model predictions at locations far from the monitoring stations (roughly more than 450 km), except that estimated emissions are used instead of reported emissions. At

the monitoring stations, the observations are reproduced. At intermediate locations, the proposed predictor will be the EMEP-prediction locally adjusted by the values of nearby monitoring stations. The first cross-validation exercise is to fix the estimated parameters and omit one monitoring station at a time. For each monitoring station the measured value is predicted using all other data. Then the root mean square (RMS) prediction error is calculated as the square root of spatially weighted average squared cross-validated prediction errors. This is compared to the EMEP-model predictions as well as the EMEP-model predictions using the reported emissions. The RMS prediction error is also compared to the Bayesian Kriging estimate of the prediction error. The results will be a check on reasonability only, since we study properties of prediction instead of estimation.

The second cross-validation exercise (full cross-validation) involves omitting one monitoring station at a time and re-estimating all the parameters (emission estimates and covariance parameters) of the statistical model. For each monitoring station omitted, the parameters of the statistical model are estimated using all other data. This exercise provides information about the uncertainty involved in fitting the statistical model. In practice, the two cross-validation exercises could be combined, but the results would be harder to interpret. In particular, it would be unfair to compare the Bayesian Kriging with the EMEP-model predictions, because EMEP-model uncertainty cannot easily be estimated.

## 5.1 Validation of Sulphur Components

Table 5 shows RMS cross-validation results for the sulphur components. For both sulphur components it is seen that the EMEP-model would give a better fit to the monitoring data if estimated sulphur emissions were used instead of reported emissions (smaller RMS errors). A further improvement would result from using the predictor proposed by Høst (1996*b*). For the present application, our main concern is regarding the quality of the prediction error estimate. The estimated prediction error is quite close to the cross-validated prediction error, indicating that our prediction error estimate has small bias. This is by no means a comprehensive investigation, but an indication that the proposed method is reasonable for assessing sulphate concentrations.

Component	Kriging Prediction		EMEP Prediction	
	True	Estimated	Prior	Posterior
SO <sub>4</sub> in precipitation	0.3103	0.3281	0.3918	0.3449
SO <sub>4</sub> in air	0.4979	0.4552	0.5858	0.5167

Table 5: *RMS prediction errors for sulphate in precipitation and particulate sulphate in air.*

Figure 13 and 14 shows the Bayesian Kriging (cross-validated) values versus the observations for sulphate in precipitation and particulate sulphate in air. These figures may be compared to Figures 1 and 3. We see that correspondance between the interpolated and observed concentrations is better than the correspondance between the predicted and observed concentrations, and the dispersion is smaller.

Figure 15 shows a density estimate of range parameter estimates from the full cross-validation exercise for sulphate in precipitation. The figure gives an impression of the uncertainty associated with estimating the range parameter. It is seen that the range parameter is sensitive to possible exclusion of the monitoring stations YU5 and NL2, but not to the other stations. Ideally, we would want our method and data set to be fairly insensitive to the inclusion/exclusion of a few stations. In Figure 17, cross-validated values of the residual variance parameter  $\sigma^2$  are shown. This parameter is not so sensitive to exclusion of monitoring stations.

Figure 19 shows density estimates of cross-validated values of national sulphur emission estimates for sulphate in precipitation, and this figure illustrates the uncertainty in the statistical model. As an example, the upper left panel shows that the estimated emission from Albania is very close to the prior value, and this estimate is very insensitive to removing any of the monitoring stations. The interpretation of other panels and similar figures are left to the interested reader.

Similar results for particulate sulphate in air are shown in Figures 16, 18 and 20. It is seen that the covariance parameter estimates are moderately sensitive to the possible exclusion of monitoring station DK7, the data value of which is located in the lower right of Figure 3.

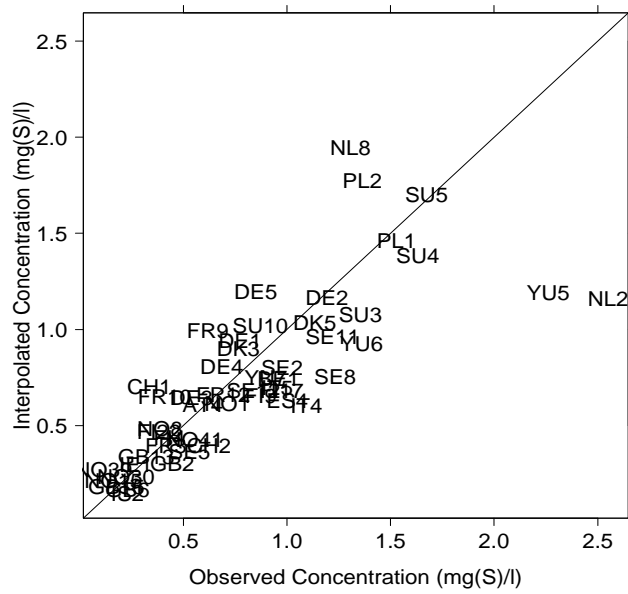


Figure 13: *Observed versus cross-validated concentrations at monitoring stations for sulphate in precipitation.*

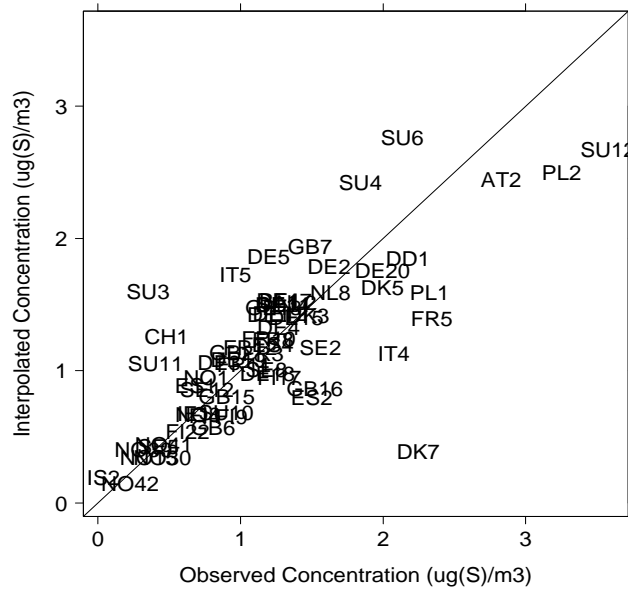


Figure 14: *Observed versus cross-validated concentrations at monitoring stations for particulate sulphate in air.*

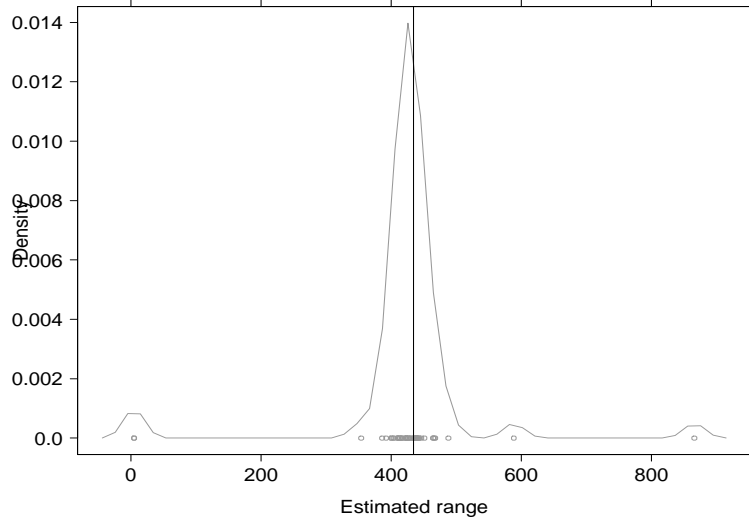


Figure 15: *Density of the estimated range in the crossvalidation for  $\gamma = 0.3$  and sulphate in precipitation. The outlier to the far left is the station YU5 and the outlier to the right NL2. The vertical line shows the estimated range  $\hat{a} = 434$  km.*

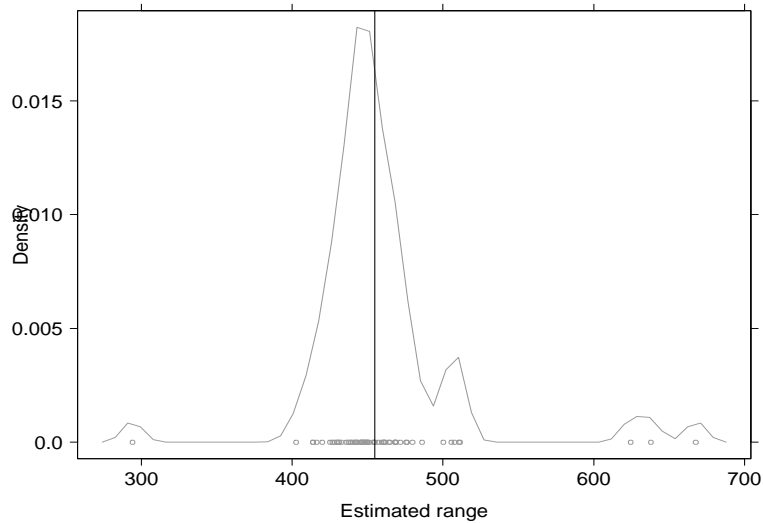


Figure 16: *Density of the estimated range in the crossvalidation for  $\gamma = 0.3$  and particulate sulphate in air. The outlier to the far left is the station DK7 and the three outliers to the right are (in increasing order) PL1, IT4 and SU3. The vertical line shows the estimated range  $\hat{a} = 455$  km.*

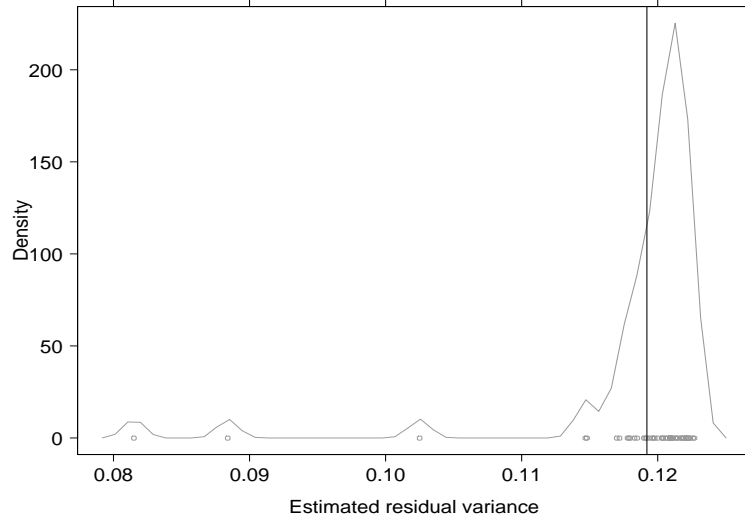


Figure 17: *Density of the estimated residual variance in the crossvalidation for  $\gamma = 0.3$  and sulphate in precipitation. The four outlier to the left are (in increasing order) NL2, YU5, SE8 and YU6. The vertical line shows the estimated residual variance  $\hat{\sigma}^2 = [0.11922 \text{ mg}(S)/l]^2$ .*

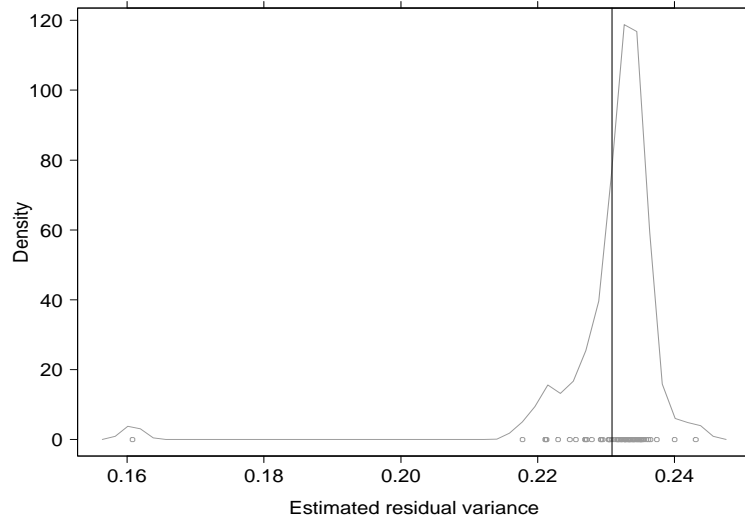
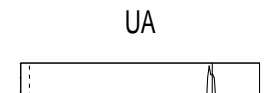
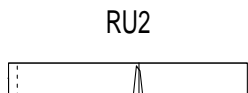
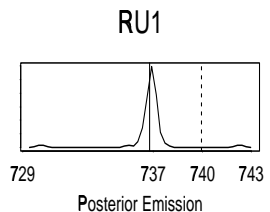
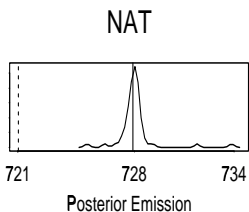
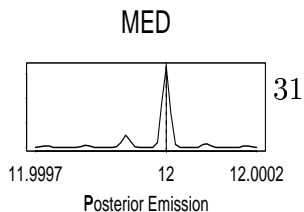
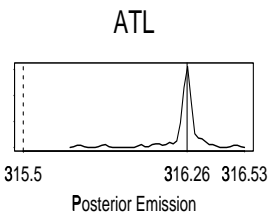
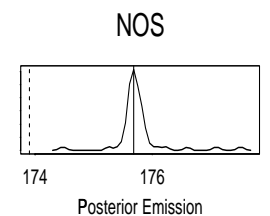
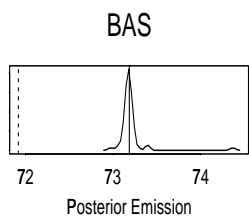
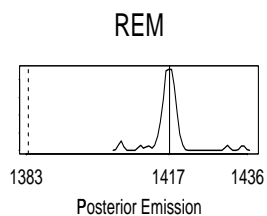
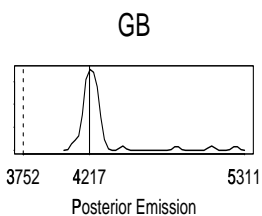
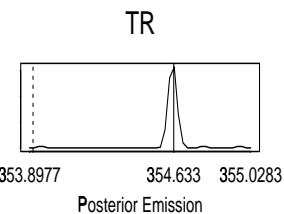
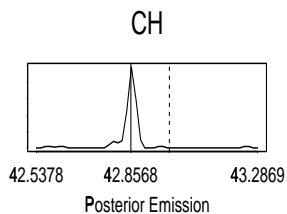
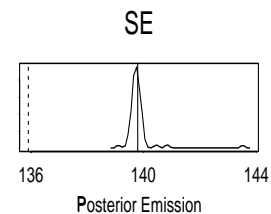
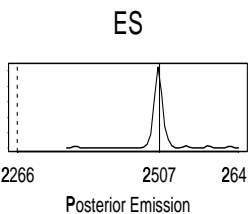
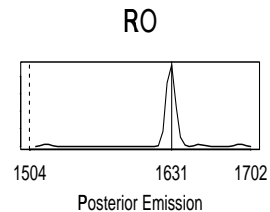
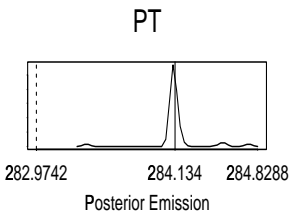
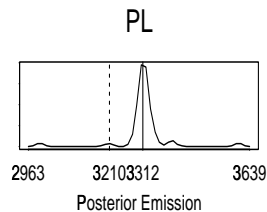
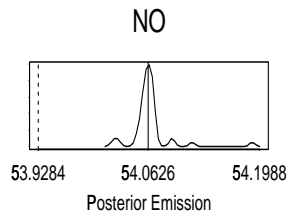
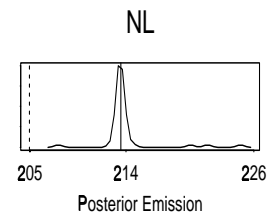
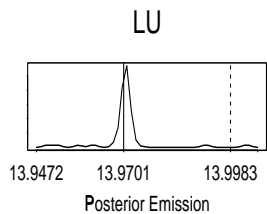
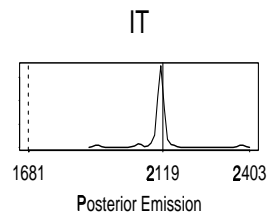
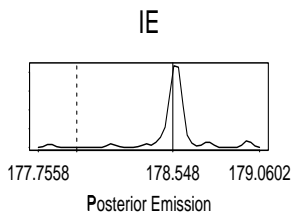
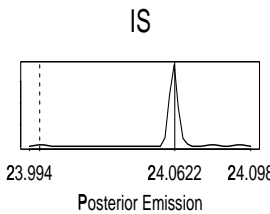
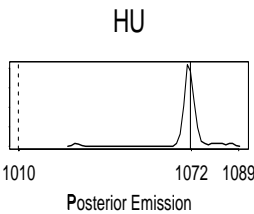
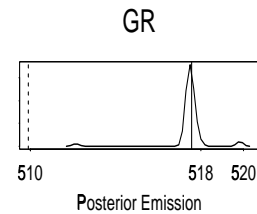
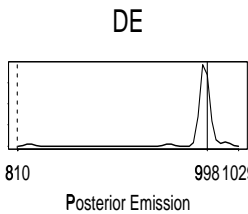
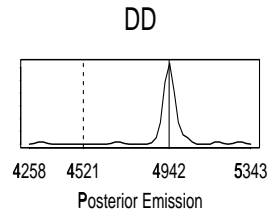
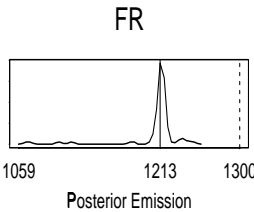
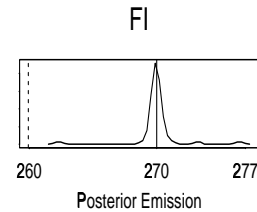
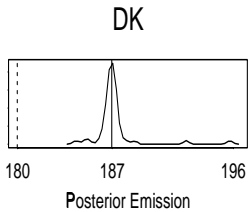
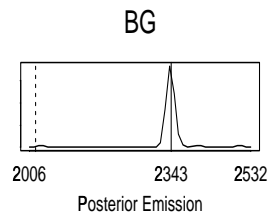
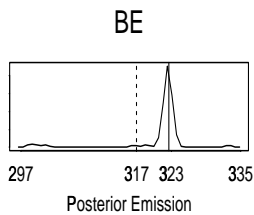
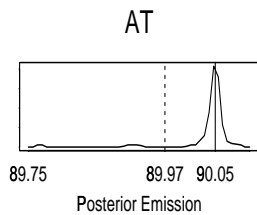
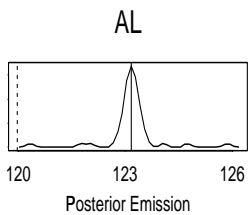
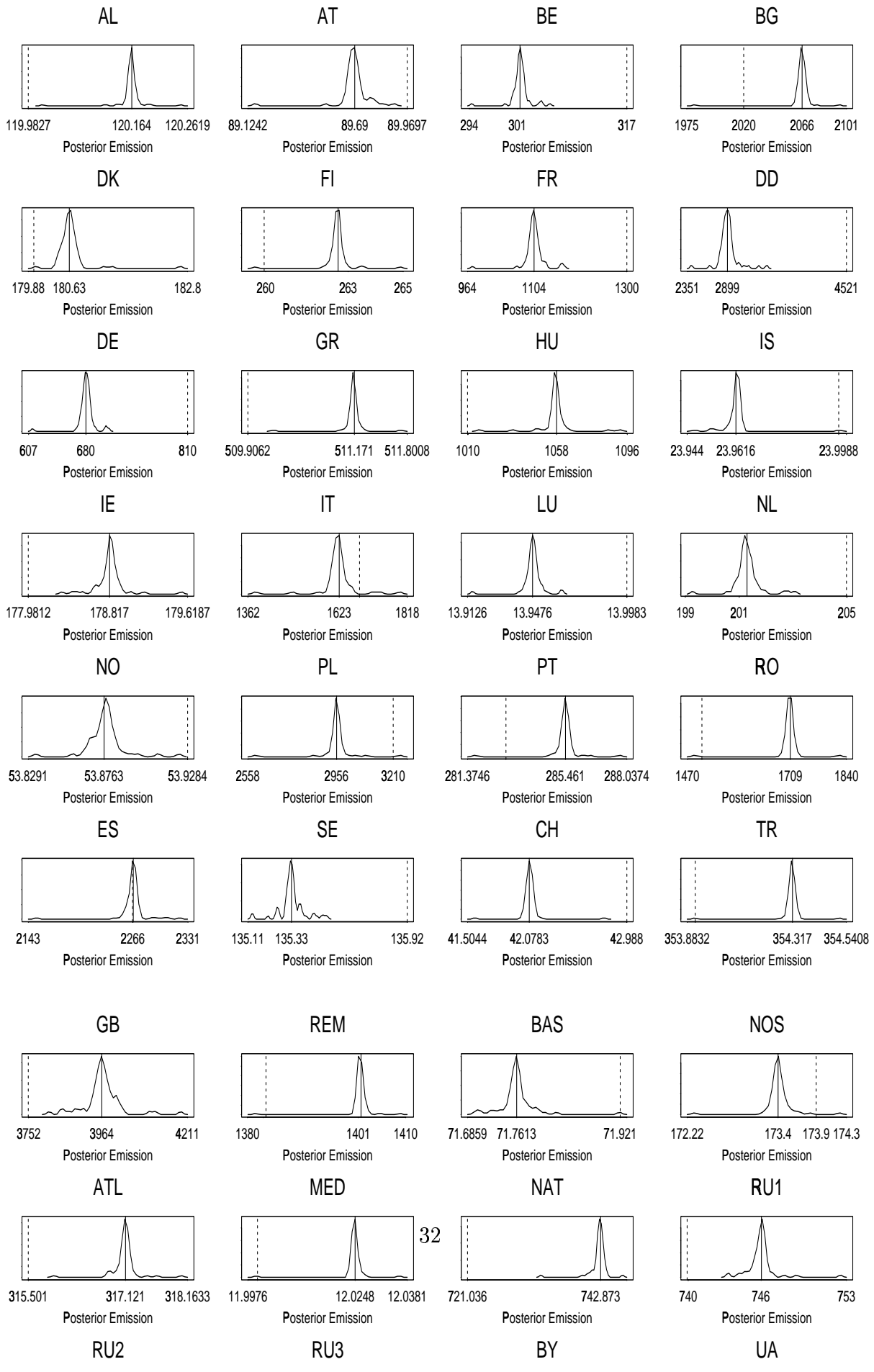


Figure 18: *Density of the estimated residual variance in the crossvalidation for  $\gamma = 0.3$  and particulate sulphate in air. The outlier to the far left is the station DK7 and the maximum value is obtained for PL1. The vertical line shows the estimated residual variance  $\hat{\sigma}^2 = [0.23086 \mu\text{g}(S)/m^3]^2$ .*







## 5.2 Validation of Nitrogen in Precipitation

Cross-validation of nitrate in precipitation data gave RMS prediction errors of  $0.12 \text{ mg(N)/l}$  for the EMEP-model with prior emissions and  $0.10 \text{ mg(N)/l}$  for the EMEP-model with estimated emissions of nitrogen oxides. In comparison, the Kriging predictor has true RMS error  $0.095 \text{ mg(N)/l}$  and  $0.099 \text{ mg(N)/l}$  estimated. Although these differences are small, the tendency is the same as for the sulphur components. Therefore, we conclude that the procedure is reasonable also for assessing nitrate concentrations.

Figure 21 shows the Bayesian Kriging (cross-validated) values versus the observations for nitrate in precipitation, and this figure may be compared to EMEP-model predictions in Figure 4. As for sulphate in precipitation and particulate sulphate in air, we see the improvement in the correspondance between the interpolated and observed concentrations compared to the predicted and observed concentrations, and the reduced dispersion.

In Figures 22 and 23, we show density estimates of the covariance parameters for nitrate in precipitation. The model is quite insensitive to the exclusion of any station except SE8. Figure 24 shows density estimates of the emission estimates.

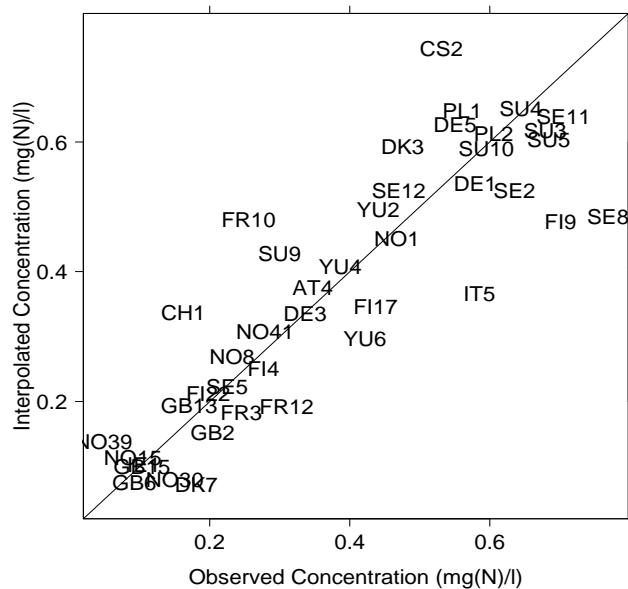


Figure 21: *Observed versus cross-validated concentrations at monitoring stations for nitrate in precipitation.*

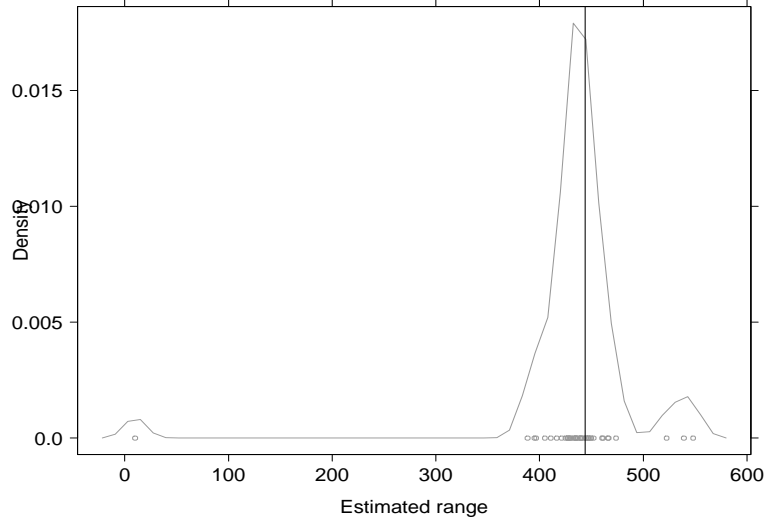


Figure 22: *Density of the estimated range in the crossvalidation for  $\gamma = 0.3$  and nitrate in precipitation. The outlier to the far left is the station SE8 and the three outliers to the right is (in increasing order) IT5, SE12 and SU9. The vertical line shows the estimated range  $\hat{a} = 444$  km.*

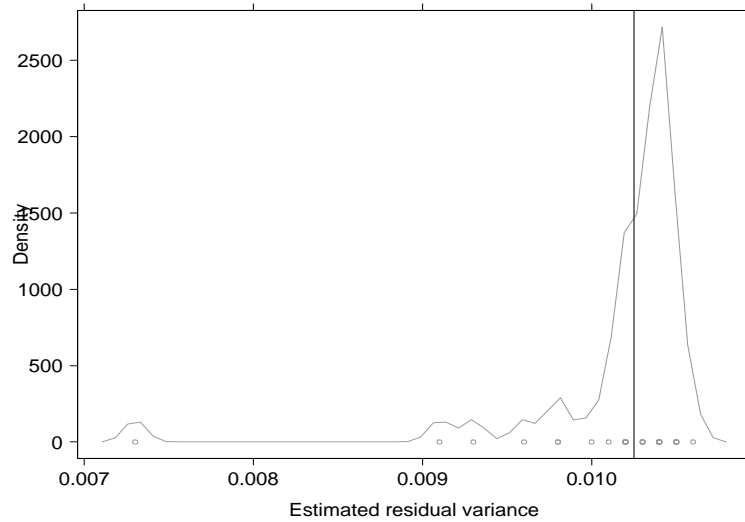
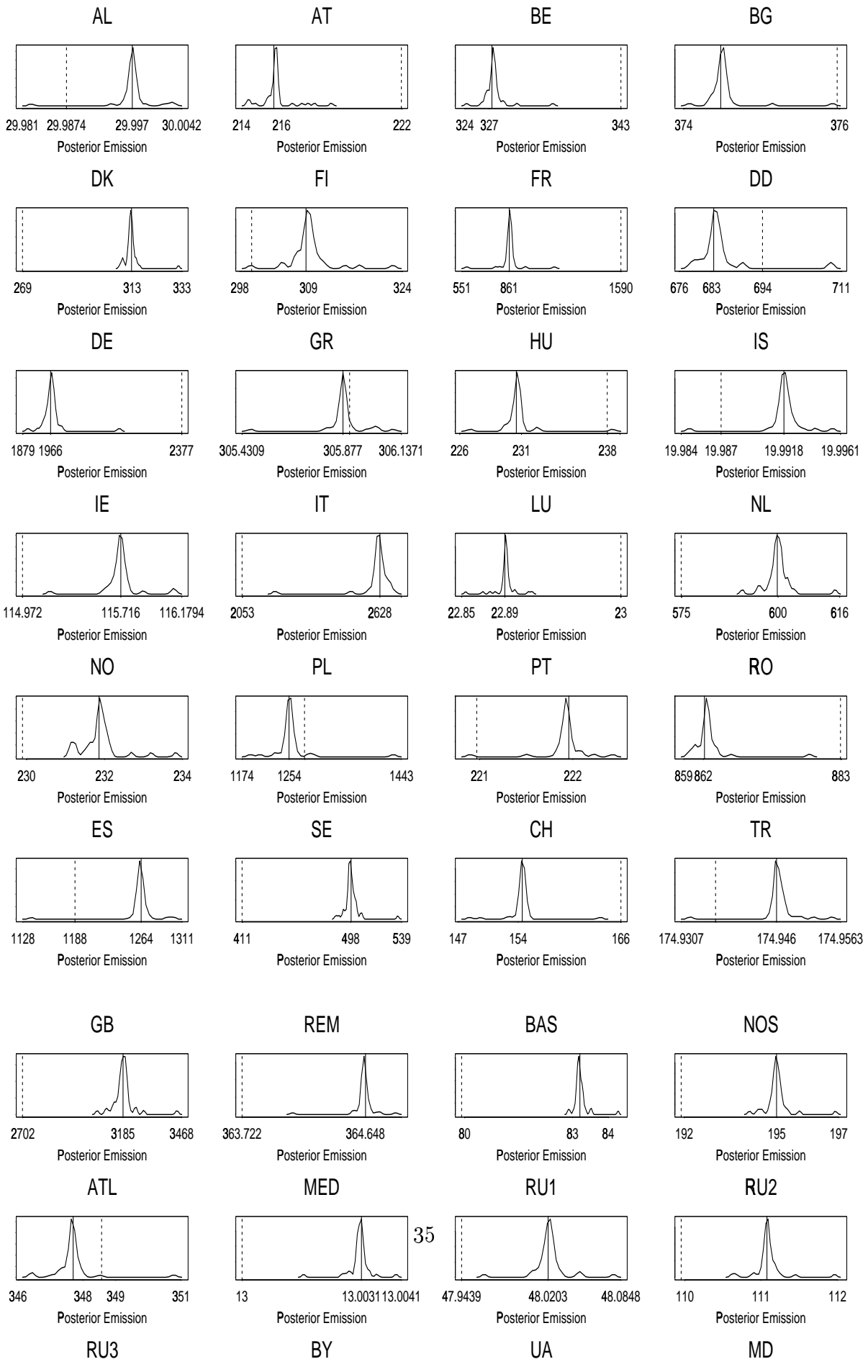


Figure 23: *Density of the estimated residual variance in the crossvalidation for  $\gamma = 0.3$  and nitrate in precipitation. The outlier to the far left is the station SE8 and the four other outliers to the left is (in increasing order) FI9, CS2, FR10, CH1 and IT5. The vertical line shows the estimated residual variance  $\hat{\sigma}^2 = [0.01025 \text{ mg}(N)/l]^2$ .*



## 6 Discussion

The estimated correlation range for sulphate in precipitation, particulate sulphate in air and nitrate in precipitation were remarkably similar. This indicates that there is little residual correlation at separating distances exceeding 3 grid units (450 km).

The results for sulphate in precipitation are roughly in correspondance with the results in Høst (1996*b*), even though a parameter for indeterminate sources was not included in the previous model. Although not significant, our results suggests that this contribution may have been larger than assumed by MSC-W in the calculations of 1990 depositions. From data of particulate sulphate in air, our results indicate that the sulphur emissions from former German Democratic Republic may have been smaller than reported during 1990. The most informative estimate from the nitrate in precipitation data is the French estimate, which suggests that the French emission of nitrogen oxides may have been smaller than reported. The reader should recognize that the methodology is still under development and that model uncertainty (of statistical model) is difficult to take proper care of. Therefore, all these results should be interpreted with care.

Some insight into model uncertainty was gained from the cross-validation exercises of Section 5. These exercises indicated reasonability of the statistical method and the fit to data. Some estimates seem sensitive to individual monitoring stations. Although other issues may be more important, this suggests that data screening and robustness issues should be taken more seriously before the method should be used as an operative tool in support of compliance reviews.

The estimated precision of the emission estimates results from intricate combinations of meteorology and monitoring network configuration. A unique property of the proposed mathematical framework is the capability of quantifying the effects on this precision from revising the network. 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. There is a need for a closer study of the properties of the EMEP-model in this regard.

In Section 4.3, 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.

Our Bayesian framework requires the specification of a regularization parameter, the prior coefficient of variation  $\gamma$ . At present, this parameter is specified a priori as  $\gamma = 0.3$  based on expert advice from MSC-W. At present, this value seems reasonable, but ideally the parameter should be estimated from data. This may be possible within a hierarchical Bayes framework, putting a hyper prior on  $\gamma$ . Principally, this is possible using a straightforward extension of the current framework. However, some preliminary testing did not give satisfactory results. This may possibly be due to skewness of the profile posterior of  $\gamma$ . In case of a very skewed profile posterior, the location of the mode may be a poor estimator.

A better approach for estimating the prior coefficient of variation would be through the use of Markov Chain Monte Carlo (MCMC) methods. A possible MCMC algorithm for our application might be a Gibbs sampler with a Metropolis-Hastings subchain (Tanner 1993, p. 140). A suggestion for such an algorithm is given in Høst (1996a). Such an algorithm will be computationally quite slow, because it involves one run of the Metropolis-Hastings algorithm for each iteration. On the other hand, an advantage of the MCMC method is that we obtain an estimate of the complete posterior density, while our current method will only give an estimate of the posterior mode and its variance.

In the future, more attention should also be given to model assumptions. This pertains to linearity and normality assumptions, as well as sensitivity to prior distributions. There is also some evidence in the data for spatially inhomogeneous second-order structure related to topography and local emission sources. It is not clear whether such inhomogeneity may be modeled adequately from available data.

The best approach to increase the precision of emission estimates would probably be to include “more data”, but this will definitely require more work. In this context, “more data” means more components, more spatial monitoring stations or more samples in the time domain (shorter averaging times or several years).

## 7 Concluding Remarks

By using the EMEP-model and the method of Høst (1996b), we have tried to relate European emissions of air pollutants to concentrations measured at the EMEP monitoring stations during 1990. Sulphur emissions were related to monitoring data of sulphate in precipitation and particulate sulphate in air, while nitrogen oxide emissions were related to nitrate in precipitation.

Our work is still at a methodological stage, but our main preliminary findings are suggestions for smaller than reported emissions of sulphur from the former German Democratic Republic, and suggestions for smaller than reported emissions of nitrogen oxides from France. However, there is still considerable uncertainty associated with these results.

Our study indicates that the proposed framework may be used for assessing monitoring strategies in relation to the purpose of improving emission estimates. A systematic analysis in this direction may prove useful.

There are also some unresolved questions regarding the properties of the method that need further attention. These are research issues that may be addressed through simulation-based statistical inference, such as Markov Chain Monte Carlo techniques.

Finally, as a practical next step, it is recommended to look closer at a time sequence of emissions and monitoring data for a single component. This could be a within-year sequence or a sequence of years, depending on practical constraints and meteorological insight.

## A Monitoring data

Some of the monitoring stations were flagged as being below the detection limit for the measurement device. For small detection limits we have taken the detection limit as the data value. The limits used were  $0.05 \text{ mg}(S)/l$ ,  $0.05 \text{ } \mu\text{g}(S)/m^3$  and  $0.03 \text{ mg}(N)/l$  for sulphate in precipitation, particulate sulphate in air and nitrate in precipitation respectively. Tables 6, 7 and 8 show the number of daily values per month for sulphate in precipitation, particulate sulphate in air and nitrate in precipitation respectively. Table 9 shows the stations excluded from the analysis based on the criteria of Section 2.2.

Station	01	02	03	04	05	06	07	08	09	10	11	12	SUM	%
AT4	6	10	10	15	8	23	13	14	17	10	15	12	153	0.42
BE1	10	13	10	14	5	15	5	9	13	14	16	15	139	0.38
CH1	4	11	9	21	15	18	7	4	10	8	17	11	135	0.37
CH2	4	13	6	8	5	13	5	4	2	10	10	3	83	0.23
DE1	19	16	12	9	3	16	8	11	22	10	11	14	151	0.41
DE2	8	12	6	10	7	13	4	12	15	4	14	10	115	0.32
DE3	9	14	9	18	11	21	12	13	9	12	19	17	164	0.45
DE4	11	14	7	15	8	15	10	9	15	9	16	15	144	0.39
DE5	8	10	5	12	7	22	11	11	14	9	16	9	134	0.37
DK3	19	18	9	11	2	7	4	10	17	11	8	11	127	0.35
DK5	9	11	9	5	4	9	6	5	13	8	12	11	102	0.28
ES4	5	5	2	12	9	7	3	5	2	7	8	10	75	0.21
FR3	9	17	5	15	13	11	5	4	7	15	11	5	117	0.32
FR9	6	8	5	8	5	11	5	5	9	7	11	6	86	0.24
FR10	1	10	10	15	10	16	7	6	10	14	14	6	119	0.33
FR12	11	8	10	9	8	9	7	2	10	14	12	4	104	0.28
IT4	4	2	2	13	7	12	2	12	2	12	6	3	77	0.21
IT5	0	4	4	11	12	19	11	9	8	15	7	3	103	0.28
IE1	28	26	18	18	8	19	11	19	10	28	17	19	221	0.61
IS2	15	8	16	14	18	11	21	18	18	14	17	15	185	0.51
NO1	20	20	15	11	3	16	9	10	14	14	4	19	155	0.42
NO8	28	26	22	15	5	16	11	16	13	16	10	19	197	0.54
NO15	22	17	23	13	6	8	15	17	8	18	14	18	179	0.49
NO30	6	8	5	1	6	4	9	9	3	11	10	6	78	0.21
NO39	14	10	23	9	10	5	17	16	11	6	15	14	150	0.41
NO41	14	11	3	13	5	12	12	11	7	8	8	12	116	0.32
NL2	24	1	11	17	8	14	12	13	24	9	20	13	166	0.45
NL8	16	17	9	14	6	11	6	10	20	14	17	15	155	0.42
PL1	18	12	16	3	10	11	12	12	18	7	14	11	144	0.39
PL2	10	10	9	11	7	7	13	13	16	5	18	13	132	0.36
SE2	21	19	12	11	8	14	10	11	17	16	10	18	167	0.46
SE5	8	6	10	8	5	11	16	9	13	7	7	10	110	0.3
SE8	18	13	6	3	6	5	14	9	14	10	11	12	121	0.33
SE11	15	12	7	7	7	8	8	8	15	10	9	7	113	0.31
SE12	15	15	6	5	5	8	21	4	21	11	10	14	135	0.37
FI4	22	21	12	9	7	10	16	13	14	7	18	18	167	0.46
FI9	14	16	7	5	3	2	8	7	16	9	13	9	109	0.3
FI17	20	17	10	8	9	10	14	8	10	8	12	15	141	0.39
FI22	21	21	14	10	11	11	17	12	12	15	17	22	183	0.5
SU3	2	0	6	3	11	5	10	7	21	8	12	3	88	0.24
SU4	14	8	8	6	6	9	16	11	17	5	20	7	127	0.35
SU5	11	7	9	11	10	7	18	10	18	3	15	5	124	0.34
SU10	16	12	12	3	7	7	13	8	16	10	20	14	138	0.38
GB2	28	26	21	18	13	22	13	26	14	21	16	19	237	0.65
GB6	27	25	12	19	11	20	13	23	16	25	17	18	226	0.62
GB13	14	7	7	5	4	14	6	5	9	16	15	15	117	0.32
GB15	23	5	2	14	10	19	5	18	16	11	16	17	156	0.43
YU5	2	7	4	12	8	7	3	3	8	9	7	16	86	0.24
YU6	3	3	6	15	6	6	7	5	9	9	10	14	93	0.25
YU7	2	4	4	15	8	6	4	4	8	9	13	21	98	0.27
SUM	820	781	621	763	552	805	683	684	870	791	896	808	9074	0.30

Table 6: *Number of daily values per month for sulphate in precipitation.*

Station	01	02	03	04	05	06	07	08	09	10	11	12	SUM	%
AT2	14	19	23	28	30	25	23	28	20	26	30	25	291	0.8
BE1	31	28	25	5	31	30	24	31	30	30	26	31	322	0.88
CH1	31	28	31	30	31	30	31	30	30	31	30	30	363	0.99
DE1	31	28	31	30	31	29	30	30	30	31	30	31	362	0.99
DE2	31	28	31	30	28	30	30	31	29	31	30	31	360	0.99
DE3	31	28	30	30	31	30	31	30	29	30	30	31	361	0.99
DE4	31	28	31	30	31	30	30	31	30	30	30	31	363	0.99
DE5	31	27	31	30	31	30	31	31	29	31	18	29	349	0.96
DE12	31	28	30	30	31	30	31	31	30	31	30	31	364	1
DE14	31	28	31	30	30	30	31	30	30	31	30	30	362	0.99
DE15	31	28	31	30	31	30	31	31	30	31	30	31	365	1
DE17	31	28	31	30	31	30	31	31	30	31	30	31	365	1
DE18	31	27	31	30	31	30	29	31	30	31	29	31	361	0.99
DE19	31	25	30	30	31	30	31	31	30	31	30	31	361	0.99
DE20	30	28	31	30	30	29	31	30	30	30	30	31	360	0.99
DD1	31	28	31	30	31	30	31	31	30	31	30	31	365	1
DK3	30	28	31	30	23	29	28	23	28	30	30	30	340	0.93
DK5	30	28	31	30	28	30	31	31	27	31	30	31	358	0.98
DK7	26	26	29	27	31	30	31	31	28	30	30	27	346	0.95
ES1	31	28	31	30	31	29	31	31	29	30	30	31	362	0.99
ES2	31	28	31	29	31	28	22	31	28	27	30	30	346	0.95
ES4	31	28	30	28	31	23	30	30	21	27	30	31	340	0.93
FR3	31	21	24	30	31	30	29	31	30	31	23	24	335	0.92
FR5	0	28	26	21	7	26	14	31	16	23	14	21	227	0.62
FR8	29	23	28	30	24	29	31	31	30	30	30	18	333	0.91
FR10	2	28	31	24	29	29	31	31	28	31	24	24	312	0.85
FR11	30	18	30	30	20	30	31	31	22	11	28	3	284	0.78
FR12	25	28	29	16	10	30	22	25	30	31	21	19	286	0.78
IT4	31	28	31	30	31	30	31	31	30	31	30	31	365	1
IT5	31	27	31	30	20	27	25	31	30	31	30	27	340	0.93
IE1	31	28	31	30	31	30	31	31	30	31	30	29	363	0.99
IS2	31	28	31	30	31	30	31	31	30	31	30	31	365	1
NO1	30	28	30	30	31	30	29	31	27	30	30	29	355	0.97
NO8	31	28	31	29	31	30	31	31	30	19	26	26	343	0.94
NO15	31	28	31	30	31	30	30	30	27	31	30	31	360	0.99
NO30	16	27	31	29	31	29	31	31	30	31	30	31	347	0.95
NO39	31	28	27	30	31	30	31	31	30	31	30	31	361	0.99
NO41	27	27	31	30	31	30	31	31	29	31	30	31	359	0.98
NO42	15	14	16	15	15	15	19	28	9	15	16	19	196	0.54
NL2	31	28	31	30	31	30	31	31	30	31	30	30	364	1
NL8	31	28	31	28	31	30	31	31	30	31	29	31	362	0.99
PL1	31	28	31	30	31	30	31	31	30	31	30	31	365	1
PL2	31	28	31	30	31	29	31	31	30	31	30	31	364	1
SE2	31	28	31	26	30	30	21	30	30	31	30	30	348	0.95
SE5	23	28	31	30	31	30	28	31	29	30	28	30	349	0.96
SE8	31	26	31	30	31	30	31	31	30	30	29	31	361	0.99
SE11	31	20	25	29	28	30	31	31	30	31	27	31	344	0.94
SE12	31	26	27	30	28	28	31	30	28	31	30	31	351	0.96
FI4	31	28	31	30	31	30	31	31	30	31	30	31	365	1
FI9	31	28	31	30	31	20	30	31	30	31	30	31	354	0.97
FI17	31	27	31	30	31	30	31	31	30	31	30	31	364	1
FI22	31	28	31	30	31	30	31	31	30	31	30	31	365	1
SU3	30	14	19	30	31	18	28	20	30	30	22	31	303	0.83
SU4	29	27	31	30	30	29	30	31	29	31	27	23	347	0.95
SU6	31	28	31	30	31	29	29	31	29	31	30	31	361	0.99
SU10	31	27	13	21	27	30	27	23	25	31	30	21	306	0.84
SU11	0	17	30	30	30	30	30	30	28	27	0	0	252	0.69
SU12	30	22	24	29	31	30	31	25	30	28	28	22	330	0.9
GB2	31	28	31	30	31	30	31	31	30	31	30	31	365	1
GB6	31	28	29	30	31	30	31	31	30	31	16	29	347	0.95
GB7	15	28	28	30	31	29	31	31	30	31	30	31	345	0.95
GB13	21	28	30	30	27	30	31	30	30	31	30	25	343	0.94
GB15	22	28	31	29	29	29	31	31	30	25	26	26	337	0.92
GB16	31	28	31	30	31	30	31	31	30	31	30	30	364	1
SUM	2208	2085	2326	2202	2270	2251	2328	2336	2210	2297	2217	2246	26976	0.87

Table 7: Number of daily values per month for particulate sulphate in air.



Station	01	02	03	04	05	06	07	08	09	10	11	12	SUM	%
AT4	6	11	10	15	8	23	13	14	17	10	15	12	154	0.42
CH1	4	11	9	21	15	18	7	4	10	8	17	11	135	0.37
CS2	10	9	12	20	16	16	15	10	20	10	18	22	178	0.49
DE1	19	16	13	9	3	16	8	10	22	10	11	14	151	0.41
DE3	9	14	9	18	11	21	12	13	9	12	19	17	164	0.45
DE5	8	10	5	12	7	22	11	11	14	9	16	9	134	0.37
DK3	19	18	10	11	2	8	4	10	17	11	9	14	133	0.36
DK7	25	23	28	17	11	13	14	21	16	17	14	17	216	0.59
FR3	9	17	5	15	13	11	5	4	7	15	11	5	117	0.32
FR10	1	10	10	15	10	16	7	6	10	14	14	6	119	0.33
FR12	11	8	10	9	8	9	7	2	10	14	12	4	104	0.28
IT5	0	4	4	11	12	19	11	9	8	15	7	3	103	0.28
IE1	25	23	14	14	7	16	10	15	8	24	16	17	189	0.52
NO1	20	20	15	11	3	16	9	10	14	14	4	19	155	0.42
NO8	28	25	22	15	5	16	11	16	13	16	10	19	196	0.54
NO15	22	13	23	13	6	8	15	17	8	18	14	18	175	0.48
NO30	6	8	5	1	6	4	9	9	3	11	10	6	78	0.21
NO39	14	10	23	9	10	5	17	16	11	6	15	14	150	0.41
NO41	14	11	3	13	5	12	12	11	7	8	8	12	116	0.32
PL1	18	11	16	3	10	11	12	12	18	7	14	11	143	0.39
PL2	10	10	11	11	7	7	13	13	16	5	18	13	134	0.37
SE2	21	19	12	11	8	14	10	11	17	16	10	18	167	0.46
SE5	8	6	10	8	5	11	16	9	13	7	7	10	110	0.3
SE8	18	13	6	3	6	5	14	9	14	10	11	12	121	0.33
SE11	15	12	7	7	7	8	8	8	15	10	9	7	113	0.31
SE12	15	15	6	5	5	8	21	4	21	11	10	14	135	0.37
FI4	22	21	12	9	7	10	16	13	14	7	18	18	167	0.46
FI9	15	16	7	5	3	2	8	7	16	9	13	9	110	0.3
FI17	20	17	10	8	9	10	14	8	10	8	12	15	141	0.39
FI22	21	21	14	10	11	11	17	12	12	15	17	22	183	0.5
SU3	2	0	6	2	10	5	7	7	21	9	11	3	83	0.23
SU4	12	6	8	5	5	9	15	11	15	4	15	7	112	0.31
SU5	10	6	7	11	10	7	16	10	17	3	15	5	117	0.32
SU9	17	14	12	3	6	7	8	9	6	7	12	8	109	0.3
SU10	15	11	9	3	7	6	13	10	16	9	19	13	131	0.36
GB2	28	26	21	18	13	22	13	26	14	21	16	19	237	0.65
GB6	27	25	12	19	11	20	13	23	16	25	17	18	226	0.62
GB13	14	7	7	5	4	14	6	5	9	16	15	15	117	0.32
GB15	23	5	2	14	10	19	5	18	16	11	16	17	156	0.43
YU2	4	7	6	18	9	20	7	5	15	13	11	16	131	0.36
YU4	5	5	5	15	13	12	9	4	11	14	13	23	129	0.35
YU6	3	4	8	17	7	8	7	6	10	10	11	15	106	0.29
SUM	805	769	618	765	547	813	669	672	872	792	891	829	9042	0.3

Table 8: *Number of daily values per month for nitrate in precipitation.*

Component	Criterion	Stations
nitrate in precipitation	(I)	FR11, GR1, PT4, SE3, SE13, SU6, SU11, SU13, SU14
	(II)	AT2, AT3, DD1, DD2, ES1, ES2, ES3, ES4, FR8, FR5, GR1, HU2, PT1, PT3, PT4, SE3, SE13, SU1, SU11, SU12, SU13, YU1
	(III)	NL8, BE1, FR9, DE4, DK5, FR8, DE2, DD2, NL2, IT4, CS1, GB14, CS3
	(IV)	YU5, YU7
	(V)	CH2
SO <sub>4</sub> in precipitation	(I)	FR11, GR1, PT4, SE3, SE13, SU6, SU11, SU13, SU14
	(II)	AT2, AT3, DD1, DD2, ES1, ES2, ES3, FR8, FR5, GR1, HU2, PT1, PT3, PT4, SE3, SE13, SU1, SU11, SU12, SU13, YU1
	(III)	GB14, SU1, GB4, ES3, DD2, CS3, CS1, CS2, HU2
	(IV)	YU2, YU4, SU9, SU11
	(V)	DK7
SO <sub>4</sub> in air	(I)	DE16, GR1, PT3, PT5, SE3, SE13, SU5, SU9, SU14
	(II)	DE16, GR1, HU2, PT3, PT5, SE3, SE13, SU5, SU9
	(III)	GB14, SU1, GB4, ES3, DD2, CS3, CS1, CS2, HU2
	(IV)	GR1, HU2, PT1, PT3, PT4, PT5
	(V)	CH2

Table 9: *Monitoring stations excluded from the analysis. The criteria are (I) Periods of three or more consecutive months without observations. (II) Data coverage not exceeding 20% for nitrate and sulphate in precipitation and 50% for sulphur dioxide and particulate sulphate in air. (III) Closeness to large emission sources. (IV) Unsatisfactory laboratory quality. (V) Extreme outliers.*

## References

- Barrett, K. & Berge, 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.
- Handcock, M. S. & Wallis, J. R. (1994), ‘An approach to statistical spatial-temporal modeling of meteorological fields’, *Journal of the American Statistical Association* **89**(426), 368–378.
- Høst, G. (1996a), Contributions to the analysis of spatial and spatial-temporal data, Dr.scient. thesis, Department of Mathematics, Statistics Division, University of Oslo.

- Høst, G. (1996*b*), 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.
- Mardia, K. V. & Marshall, R. J. (1984), 'Maximum likelihood estimation of models for residual covariance in spatial regression', *Biometrika* **71**, 135–146.
- 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.
- Tanner, M. A. (1993), *Tools for Statistical Inference*, Springer series in statistics, 2. edn, Springer Verlag Inc., New York.