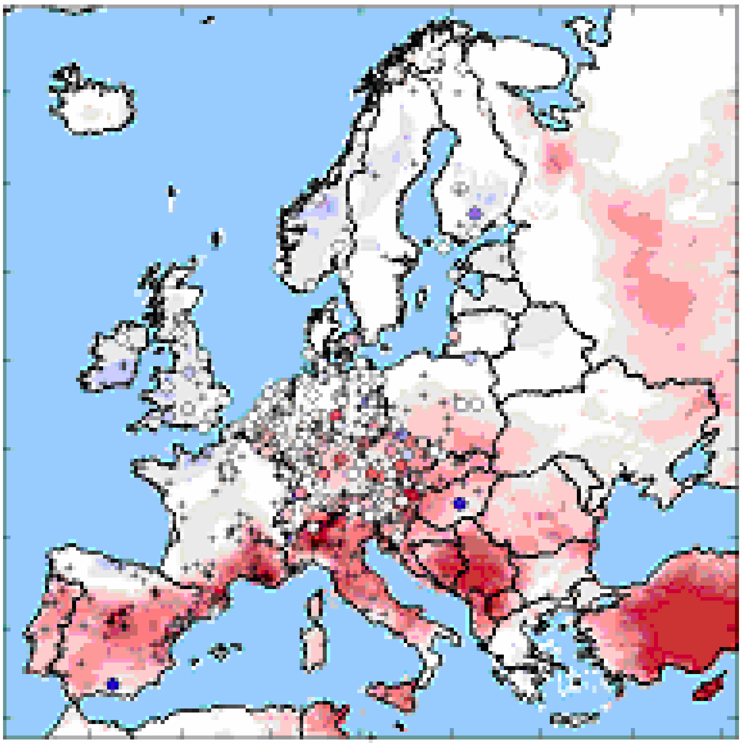
Preliminary assessment report on the spatial mapping of air quality trends for Europe



ETC/ACC Technical Paper 2008/3
November 2008

Bruce Denby, Ingrid Sundvor, Peter de Smet, Frank de Leeuw

European Topic Centre on Air and Climate Change

The European Topic Centre on Air and Climate Change (ETC/ACC)
is a consortium of European institutes under contract of the European Environmental Agency
MNP UBA-D UBA-V NILU AEAT AUTh CHMI MET.NO ÖKO TNO REC

Front page map: *This paper, Figure 5.5, left map.*

DISCLAIMER

Preliminary assessment report on the spatial mapping of air quality trends for Europe

ETC/ACC Technical Paper 2008/3
Final version

Authors
Bruce Denby [1], Ingrid Sundvor[1], Peter de Smet[2] and Frank de Leeuw[2]

[1] Norwegian Institute for Air Research (NILU), Norway
[2] The Netherlands Environmental Assessment Agency (PBL), The Netherlands

Contents

1 Introduction

The European Environmental Agency (EEA) and the European Topic Centre for Air Quality and Climate Change (ETC/ACC) have in the past few years developed methodologies for the spatial assessment of air quality on a European wide basis (see Horalék et al., 2005; 2007; 2008 and Denby et al., 2008). The work is intended to provide the best quality spatial assessment of a number of directive related pollutants, with focus on ozone and PM_{10}, that will provide both policy support as well as information to the public. The maps are used to assess European wide exposure and the resulting health effects at a resolution that provides 'background' pollutant concentrations, i.e. 10 x 10 km grids.

The main thrust of the development to date has been to establish a robust mapping method that can be implemented operationally. As a result a wide range of mapping methodologies have be assessed for the years 2003 – 2005. These mapping methods include the use of monitoring data (provided by the AirBase and EMEP databases), atmospheric chemistry model calculations (provided by the Unified EMEP model and the LOTOS-EUROS model) and other spatially resolved supplementary data such as altitude and a selection of meteorological parameters. These are combined using multiple linear regression of all the data with the observations to provide a base spatially resolved map. Ordinary kriging is applied to the residuals, i.e. differences between the base map and the observations, to obtain the final maps.

The resulting maps provide an assessment for an individual year. The question arises as to whether longer term trends ('trends' refers in this work exclusively to temporal changes and 'long term' over periods of 10 years or longer) can be seen in the maps and whether this can provide useful information for policy assessment. Though trend analysis has been carried out for individual monitoring, or sets of monitoring, stations in the past a spatially resolved trend analysis has not been carried out. This report is intended to provide a preliminary assessment of the feasibility of such an assessment using the mapping methodologies already developed by ETC/ACC.

1.1 Aim

This report will assess the feasibility for assessing the European wide mapping of air quality trends. It will identify the requirements, both in data and methodology, and will test some basic applications. It will provide evidence, arguments and a work plan for a more detailed assessment to be carried out.

1.2 Scope

Since the study is intended as a feasibility study it must be limited in scope and application. On the other hand it must also test some methodologies in order to truly assess the feasibility of the approach. Two indicators have been chosen for the assessment. The first of these is AOT40, an ozone indicator for which target values have been set in the European air quality directive (EC, 2008) for the protection of vegetation, and the second of these is the annual mean SO_2. AOT40 does not show a clear trend in the observations whilst models calculations indicate a negative trend (Solberg et al., 2008). A much clearer trend signal is observed and modelled for SO_2 where both models and observations show significant decreasing trends.

The following aspects will be addressed in this report

- Chapter 2: Brief literature review of relevant trend analysis data for Europe and trend assessment methods generally. Focus is on EEA, EMEP and ETC reporting.
- Chapter 3: Review and assessment of data availability for the maps in the period 1996 - 2005
- Chapter 4: Methodological description of the mapping technique, station selection, trend analysis and uncertainty estimates.
- Chapters 5 and 6: Results of a number of selected tests on the data in regard to interpolation method, the trend analysis method and the selection of stations for the two pollutants AOT40 and SO_2.

- Chapter 7: Discussion and recommendations resulting from the study. Plans for future work.

2 Literature review on trend analysis

Trend analysis is used in a wide range of scientific disciplines including the environment, economy, biology, demography, etc. In this literature review we will concentrate on applications from the environmental sciences and wherever possible applications within air quality or meteorology.

2.1 Trend analysis methodologies in environmental applications

In the environmental literature much work in trend analysis has been carried out in detecting trends in climate, e.g. trend analysis of precipitation and temperature, but less focus has been given to air quality trends over similar time scales, i.e. decades. This is chiefly due to the shorter time scales of available and reliable monitoring data for air compounds compared to traditional weather parameters such as wind, temperature and precipitation.

Several statistical approaches are available for detecting and estimating trends in environmental data, such as regression analyses, time-series analyses and methods of non-parametric statistics. A good over all text on the use of trend analysis in environmental applications is provided in Gilbert (1987). Seven different linear methods have been compared in Hess et al. (2001) and Porter at al. (2002) with the use of both simulated and real data. Some useful insight into trend testing and its challenges can be found in Schär and Frei (2001), who look for trends in rare events. A brief overview lecture of analytical methods and a review of trend analysis is given by Chandler (2002).

In this study annual data for the period 1996 to 2005 are used. This means that, at the most, 10 data points are available for the trend analysis. A common analysis method for such small data sets is the Mann Kendall test (Mann, 1945; Kendall, 1975), which is a simple and robust nonparametric method intended for detecting monotonic trends. It has been further developed to include seasonality and serial correlation (Hirsch and Slack, 1984), but this is not required for the current application. The Sen's slope estimate (Theil ,1950 and Sen ,1968) is applied for quantifying the trend. A short and simple resume of the method can be found in Salmi et al. (2002) and this is outlined in section 4.2 of this report. The method has recently been applied with success for precipitation data in Turkey (Partal and Kahya, 2006) and Spain (Mosmann et al. 2003). The latter also implements mapping of areas showing significant trends.

2.2 Trend analysis for air quality applications in Europe

There are a variety of trends analyses carried out on monitoring data, mostly at individual monitoring stations or at groups of stations. Some examples include the application of the Mann Kendall test to air quality data in Europe in OSPAR (2005), by Salmi et al. (2002) and by De Leeuw (2000). The last of these has been applied for detecting trends in ground level ozone concentrations. Baldsano et al. (2005) look at trends in EMEP data since 1977 to assess the effectiveness of policy information. Solberg et al. (2008) has also recently studied trends in ozone using both modelling and monitoring data. Ibarra-Berastegi (2001) has applied the KZ filter to ozone data in Bilbao. Bronnimann et al. (2002) apply linear regression to assess trends in ground level ozone in Switzerland.

An overview of the different trend and uncertainty assessment methods is provided in table 1.

Table 1: Overview of trend analysis methods

Method	Key feature	Description	References
Linear Regression	Provides an estimate of slope (trend)	Least squares fitting technique for a linear regression model. A range of tests and confidence and prediction interval assessments are available.	Bronnimann et al. (2002)
Box-Jenkins test	Forecasts near future values of a time series.	An auto regressive moving average method. It is quite flexible, needs at least moderately long time series and regularly spaced data.	Box and Jenkins (1976), Gröger and Rumohr (2006), Hu et al. (2006)
Kolmogorov –Zurbenko method	Removes "noise" first to make trend appear.	A low pass (KZ) filter, with iterative moving average. The data is smoothed and then regressed in time to obtain the trend slope. Useful for large datasets.	Zurbenko (1986), Rao and Zurbenko (1994), Ibarra-Berastegi et al. (2001)
Mann Kendall test	Gives a Yes/No for existence of a significant trend.	The test is suitable when the trend can be considered monotonic and no cycles are present in the data. Missing values are allowed. Can be applied to small data sets. Is not sensitive to outliers	Mann (1945) Kendall (1975), De Leeuw (2000), Mosmann et al (2003)
Sen's method	Estimate the magnitude of a linear trend.	A nonparametric method to find the magnitude of a linear trend. Not affected by outliers and allows missing data.	Sen (1968), Thiel (1950), Partal and Kahya (2006)
Monte Carlo	Ensemble technique for assessing uncertainty based on prior distributions	Calculates uncertainty using an ensemble of possible realisations. Can be applied to any trend analysis method if uncertainty of the data used in the trend is known	

3 Data availability

In this scoping study the data availability of ozone, PM_{10} and SO_2 is provided. Though PM_{10} is not assessed in this report, due to its lack of available observational and modelling data, it is included in this survey of data availability to assess any further use of PM_{10} in future trend analysis. This report makes use of rural background stations only but includes urban and suburban stations in the data survey as these may also be included in further studies.

3.1 Monitoring data

Monitoring data is exclusively taken from the AirBase database (AirBase, 2008). Though more air quality data, especially for the years prior to 2000, may be available from national or local networks, AirBase is currently the most comprehensive database available of air quality monitoring in Europe. In figures 3.1 – 3.3 the availability over the period 1995-2005 of regional and (sub)urban stations are shown for the compounds of ozone, SO_2 and PM_{10}. In appendix III data on the total number of stations for the individual countries, is also provided.

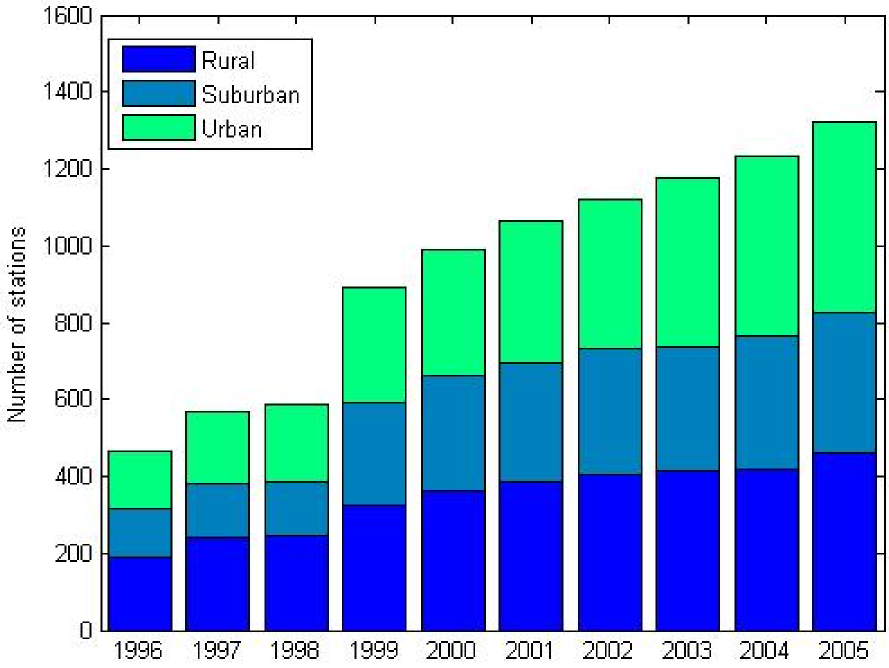


Figure 3.1. The number of available ozone stations in AirBase with hourly temporal coverage greater than 75% for the calculation of AOT40 crops.

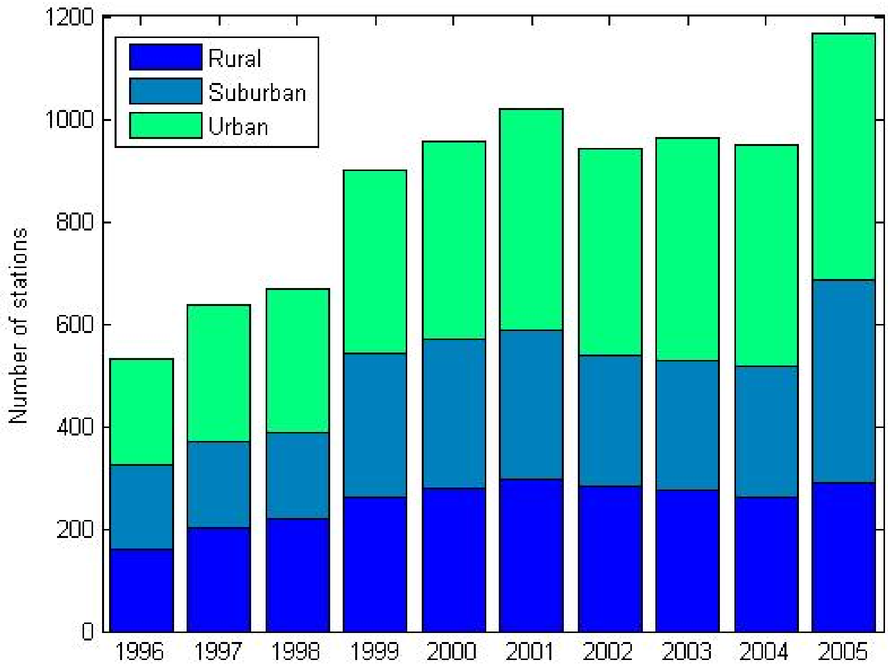


Figure 3.2. The number of available SO_2 stations in AirBase with daily mean temporal coverage greater than 75% for the calculation of annual mean concentrations.

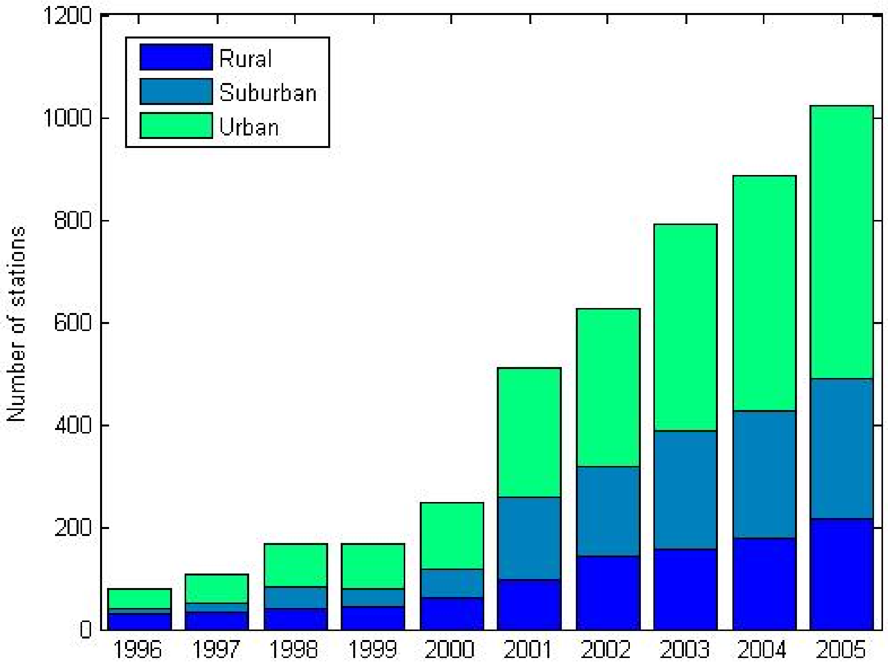


Figure 3.3. The number of available PM_{10} stations in AirBase with daily mean temporal coverage greater than 75% for the calculation of annual mean concentrations.

To indicate the spatial distribution of the available monitoring data, figure 3.4 shows the number of countries reporting data to AirBase for the pollutants PM_{10}, SO_2 and ozone. Up until 2001 less than 15 countries have reported PM_{10} data to AirBase. For SO_2 and ozone the proportion is higher, with the majority reporting since 1997. Only since 2003 has there been a good country representation for all pollutants.

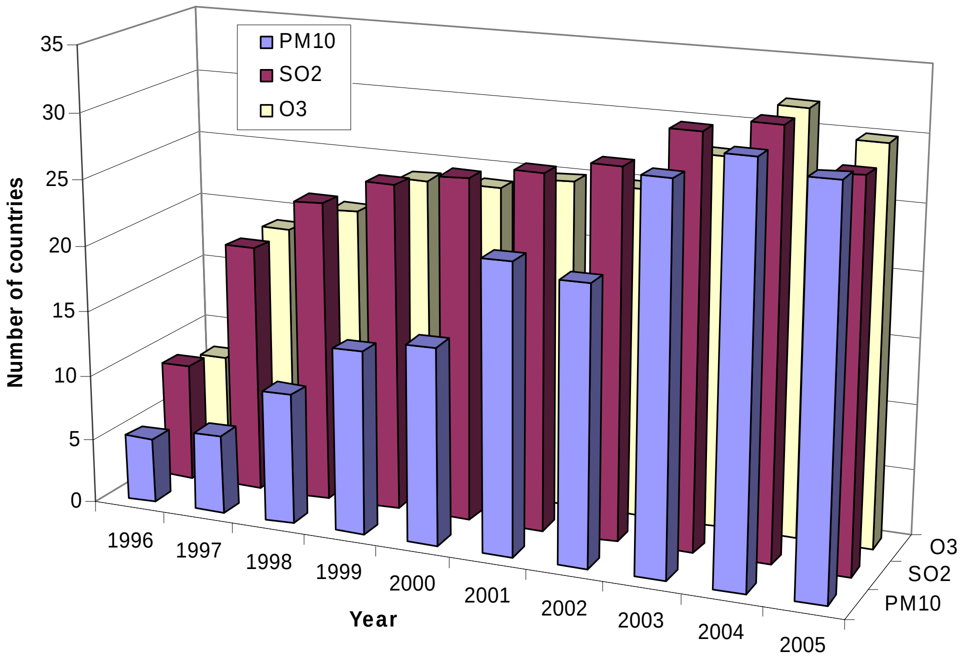


Figure 3.4. The number of countries providing data to AirBase for the three pollutants PM_{10}, SO_2 and ozone. Only countries providing station data (urban, suburban and rural) with a mean coverage > 75% are included.

3.2 Modelling data

The unified EMEP Eulerian Photochemistry model is used in this trend assessment. The model has a polar stereographic projection with a horizontal resolution of approximately 50×50 km^2 and applies 20 vertical layers below 100 hPa. The model domain is centred over Europe and also includes most of the North Atlantic and the polar region. The EMEP model uses 3-hourly resolution meteorological input data from a dedicated version of the HIRLAM model. The model is public domain and can be downloaded from www.emep.int. A comprehensive description of the model is also available via this site.

In connection to the long term trend report, 'Assessment of ground-level ozone within the EEA Member Countries with focus on long-term trends' (Solberg et al., 2008), 10 year runs of the Unified EMEP model have been carried out. These runs include all standard compounds including ozone, PM and SO_2. A complete description of the EMEP model runs can be found in that report.

In regard to the availability of data, it is worth noting that emission data was not available for the primary emissions of PM_{10} for the years prior to 2000.

3.3 Supplementary data

In addition to the EMEP modelling data for the years 1996-2005 other supplementary data is also available for the interpolation. For the methodologies applied here altitude is the only supplementary data currently used. This has been derived from the GTOPO30 (Global Digital Elevation Model) at a resolution of 30 x 30 arcsec. (source: ESRI, Redlands, California, USA, 2005) dataset and averaged for the 25 km grids used for the interpolation.

Meteorological data can also be used in the multiple regression, but this has not been included in the current analysis. Such meteorological data, derived from ECMWF reanalysis data, is available form the Meteorological Archival and Retrieval System (MARS) of the ECMWF (European Centre for Medium-range Weather Forecasts; http://www.ecmwf.int). The spatial resolution of the data used is 0.25 x 0.25 degrees. This data is available for all the assessment period, back to 1990, if required.

Conclusion: As a result of the data availability assessment trend maps of AOT40 and SO_2 only will be made. The poor coverage of PM_{10} data in AirBase prior to 2001 and the lack of primary emission data in the EMEP model prior to 2000 make trend analysis of the PM_{10} data unrealistic prior to 2001.

4 Methodology

In this paper a selected number of trend assessments and methodologies will be applied and these are described in the following sub-sections. Common to all the trend assessments is that trends of annual concentration maps will be made. This is in contrast to the other possible method for mapping trends when combining model and monitoring data, i.e. by firstly determining the observed and modelled trends and then combining the trends to create a map. In all maps rural stations only are applied and the maps are made on a 25 x 25 km grid.

The following aspects will be investigated:

- Trend maps of two pollutants, annual mean SO_2 and AOT40, will be made. The first since it has a clearly decreasing trend signal and the second since it is one of the relevant indicators of interest and has relatively good data coverage.
- Four versions of these trend maps will be assessed based on: EMEP modelling, kriging interpolation of observations, regression of the EMEP model with observations and residual kriging of the regression model (section 4.1).
- Two trend assessment methods will be applied: the trend (linear change in concentration per year) is estimated using linear regression and using Sen's method (section 4.2).

- Two methods for ascertaining the uncertainty in the trend will be applied. Standard deviation of the trend residuals and Monte Carlo simulation of the trend using the mapped uncertainty (section 4.3).
- Two types of station selection will be made. Using all available stations for every year or using only stations with a temporal coverage of at least 8 of the 10 years (section 4.4).

4.1 Mapping method and spatial uncertainty assessment

There are a number of interpolation methods available and these have been thoroughly investigated in previous ETC/ACC technical reports (Horalék et al., 2005; 2007; 2008). In this report we will present results from four of these methods. These are:

- Pure modelling data
- Ordinary log-normal kriging of the observations. In this study only rural background station s are used
- Multiple linear regression of the observations with the EMEP model (and in the case of AOT40 also with altitude)
- Residual log-normal kriging after multiple linear regression

These four results are intended to show the trend of the model, the trend of the observations and the trend derived from the interpolation and combination of model and observations. A brief description of the above methods, and relevant elements of these, is given here.

4.1.1 Log transformation

Investigation of the frequency distributions of the observational and model data indicates that the concentrations are close to log-normally distributed. To ensure that data used in the interpolations are normally distributed for the interpolation, and to avoid the undesirable result of negative concentrations, the model and observational concentrations are transformed using the natural logarithm prior to interpolation and then back transformed afterwards. The back transformed expectation value is determined using

$$E[C] = \exp\left(\mu + \frac{\sigma^2}{2}\right) \tag{1}$$

where μ and σ are the mean and standard deviation of the log-normal distribution and C indicates either an observed, interpolated or modelled concentration. σ^2 is determined using the log-normal kriging variance. The back-transformed variance of a log-normal distribution is similarly given by

$$\mathrm{var}[C] = \left(\exp\left(\sigma^2\right) - 1\right)\exp\left(2\mu + \sigma^2\right) \tag{2}$$

4.1.2 Multiple linear regression

Regression of the annual mean EMEP model concentrations with the observed annual mean concentrations provides one of the simplest methods for correcting bias in the model. Linear regression, after logarithmic transformation, is applied each year. The regression coefficients are determined and these are applied to adjust the model concentration field. It has been shown (e.g. Horálek et al., 2005) that multiple linear regression with other spatially distributed data, such as topography, can improve the regression statistics for annual fields. For AOT40 the altitude is also used in the regression.

4.1.3 Ordinary kriging and residual kriging

Ordinary kriging is a widely used interpolation technique that determines the statistically most likely concentration at a point in space by weighting the available observational data so that the interpolated spatial variance is minimised. The spatial variance is generally defined using a variance model known as the variogram (Cressie, 1993), whose parameters of sill, nugget and range need to be determined.

In this report ordinary kriging is applied to the observational data only as well as the residual of the linear regression model for each year. The kriging parameters are determined by fitting the variogram with a spherical variance function for each year. No optimisation is carried out of the variogram parameters, as was done in Denby et al. (2008). A maximum allowable range of 1000 km is specified and only the 50 nearest stations to the interpolation point are used for the kriging interpolation.

4.1.4 Spatial uncertainty assessment

Spatial uncertainty is estimated from the kriging or residual kriging variance, expressed in terms of the standard deviation. When using the logarithmic transformation the uncertainty becomes proportional to the concentrations rather than an absolute concentration. This may underestimate the uncertainty in areas with low concentrations but it provides more realistic estimates of uncertainty than the use of normal kriging.

4.2 Trend analysis

Though there is a range of methodologies available for trend analysis (section 2) we have selected two common methods for assessing trends. These are linear regression and Sen's trend analysis.

4.2.1 Linear regression

Linear regression is a straight forward and well known methodology that fits a linear model in the form of

$$y = a + bx \tag{3}$$

Fitting is achieved by minimising the sum of the square residual errors. The resultant trend in the data, given by the slope *b*, can be determined algebraically. Since it minimises the sum of the squares of the errors, regression tends to give higher weighting to larger values. This makes it more sensitive to outliers than other methods that do not rely on the squared error for fitting.

Associated with linear regression are a number of parameters and tests that can be used to indicate the uncertainty or test various hypothesise. These have not be applied in the current assessment, see section 4.3.

4.2.2 Sen's method

To find the trend one can use Sen's slope estimate if the trend is linear. First all the slopes of all data pairs are calculated

$$Q_i = (x_j - x_k)/(j - k)\,. \tag{4}$$

where x_j is the data measurement at time *j*, x_k is the data measurement at time *k*. Q_i is the slope between data points x_j and x_k

With *n* data points one gets

$$N = n(n-1)/2 \tag{5}$$

values for Q_i and the Sen's estimator is then the median of these *N* slope values. Estimates of uncertainty can be based on the variance of Q_i but this methodology has not been applied here.

4.3 Trend significance and uncertainty

We will present a limited selection of methodologies for assessing the significance and the uncertainty of the trends. The methods outlined are the standard error of the regression residual, the Mann Kendall significance test and Monte Carlo simulations using data uncertainty.

4.3.1 Standard error of the regression residual

This method provides indicative information on the variability of the data around the regression trend. The standard deviation of the regression residuals (standard error) is calculated and normalised with the number of years. This gives a simple first impression of the variability of the results. When the normalised standard error is larger than the trend, then the trend is not considered to be well defined. However, large standard deviations may also exist if the trend is not linear. In addition this method does not take into account uncertainty in the individual concentrations for each of the years.

4.3.2 Mann Kendall

The Mann Kendall (MK) statistic is used to ascertain the significance of a trend, without specific reference to the actual value of the trend

The Mann Kendall (MK) statistic, *S*, is defined as follows:

$$S = \sum_{k=1}^{n-1} \sum_{j=k+1}^{n} \operatorname{sgn}\left(x_j - x_k\right) \tag{6}$$

where *x* are the data values, *n* is the length of the dataset and the *sgn* function is defined as

$$\operatorname{sgn}\left(x_j - x_k\right) = \begin{cases} 1 & if \quad \left(x_j - x_k\right) > 0 \\ 0 & if \quad \left(x_j - x_k\right) = 0 \\ -1 & if \quad \left(x_j - x_k\right) < 0 \end{cases} \tag{7}$$

It is assumed that if there is no trend, S is normally distributed with zero mean and with variance

$$VAR(S) = n(n-1)(2n+5)/18 \tag{8}$$

The Z test statistic of the MK test is given by

$$Z = \begin{cases} \dfrac{S-1}{\sqrt{VAR(S)}} & if \quad S > 0 \\ 0 & if \quad S = 0 \\ \dfrac{S-1}{\sqrt{VAR(s)}} & if \quad S < 0 \end{cases} \tag{9}$$

The p-value, a measure indicating if the trend is likely the result of random processes, is found as the integral of the standard normal distribution *f(y)* for *y >Z* if *Z >0* or *y<Z* if *Z<0* . If the p-value is small it is a measure that the trend is quite unlikely to be the result of random sampling. If one chooses a significance level of α= 0.05 and find $p \le 0.05$, then the existing trend is considered to be statistically significant and not caused by random processes. This then tells us, with a 95 % confidence level, that the hypothesis of no trend can be rejected. However, this does not say anything about the size of the trend or the uncertainty of the value of the trend.

4.3.3 Monte Carlo methods

As a more general method for determining the uncertainty in the trend, Monte Carlo methods can be used. For each year and at each grid square an uncertainty is determined using the kriging variance. This uncertainty is randomly sampled, assuming a normal distribution, and for each set of samples a slope is determined by linear regression. Random sampling is carried out 100 times and the standard deviation of the resulting slope is used to indicate the uncertainty in that slope.

This method requires information concerning the statistical distribution of the data, which is in principle available through the kriging variance. However, if this information is not available then it is also possible to use the standard error, section 4.3.1, to indicate the general statistical distribution of the data and apply Monte Carlo methods to that. This method is not applied here but can be considered for further development. Monte Carlo methods can also be applied to Sen's trend analysis but this was not carried out extensively in this study.

4.4 Testing the effect of station selection

As shown in section 3.1 the AirBase database is not homogenously spatially distributed for the 10 year period under investigation. There are thus regions of Europe where 10 year observational trends are not available. This will lead to a bias in the information for the 10 year trend analysis if only stations with 10 years of data are used. To test the effect of station selection two selection criteria are used.

1. No selection criteria are used when making the interpolation. All stations available for each of the years are used.
2. A pre-selection of stations is made where only stations with data for 8 or more years, irrespective of timing, are used in the interpolation.

5 Results of the tests for AOT40

Given the number of different combinations of possible methodologies, as described in section 4, it is not possible to provide results on all permutations. The following results will be presented:

1. Assessment of annual statistics of the mapping methodologies
2. Average trends based on modelling and observations
3. Comparison of the two trend analysis methodologies (i.e. linear regression and Sen's method) when applied to the residual kriging interpolation
4. Comparison of the trends based on different interpolation mapping methodologies (i.e. kriging of the observations and kriging of the residual after regression with the EMEP model)
5. Comparison of the trends using two different station selection criteria (i.e. all stations and stations with $\geq$ 8 data years)
6. Comparison of the trend uncertainty methods (i.e. standard deviation of regression residuals, Monte Carlo and Mann Kendall)

5.1 Annual statistics of the interpolation maps

Before looking at the individual maps, the statistical results of the different interpolation methods will be shown for each year. In figure 5.1 the number of available rural stations is shown for the two different station selections. The number of ozone stations in AirBase has increased steadily over the past decade (see figure 3.1). The spatial distribution of the stations with availability of 8 years or more can be seen in the maps provided in section 5.3.

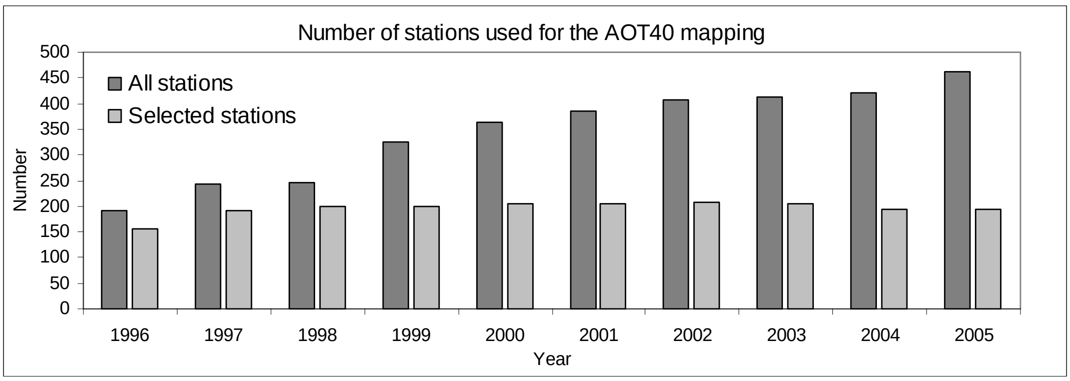


Figure 5.1. Number of stations available for the AOT40 trend assessments. Shown are the two station selection cases. 1) All available stations are used (dark grey) and 2) where stations with at least 8 years of data are used (light grey). Only rural stations are used.

To provide an overview of the statistical results of the different interpolations for each year the normalised cross-validation RMSE and correlation coefficients for AOT40 are shown for the two station selections in figures 5.2 and 5.3. The results reflect previous tests where the RMSE has been shown to be lowest and the correlation highest for the residual kriging methods (e.g. Horálek, 2005). However, there are exceptions to this rule. In 2002 and 2004 for the stations with a coverage ≥ 8 years the RMSE is higher for the residual kriging method than for the ordinary kriging. For 2002 the correlation is also lower for the residual kriging. This may be associated with poor performance from the EMEP model, and also regression model, in that year or it may be due to outliers in the observational dataset. In either case, the regression model is not providing the spatial distribution required for production of 'improved' residuals.

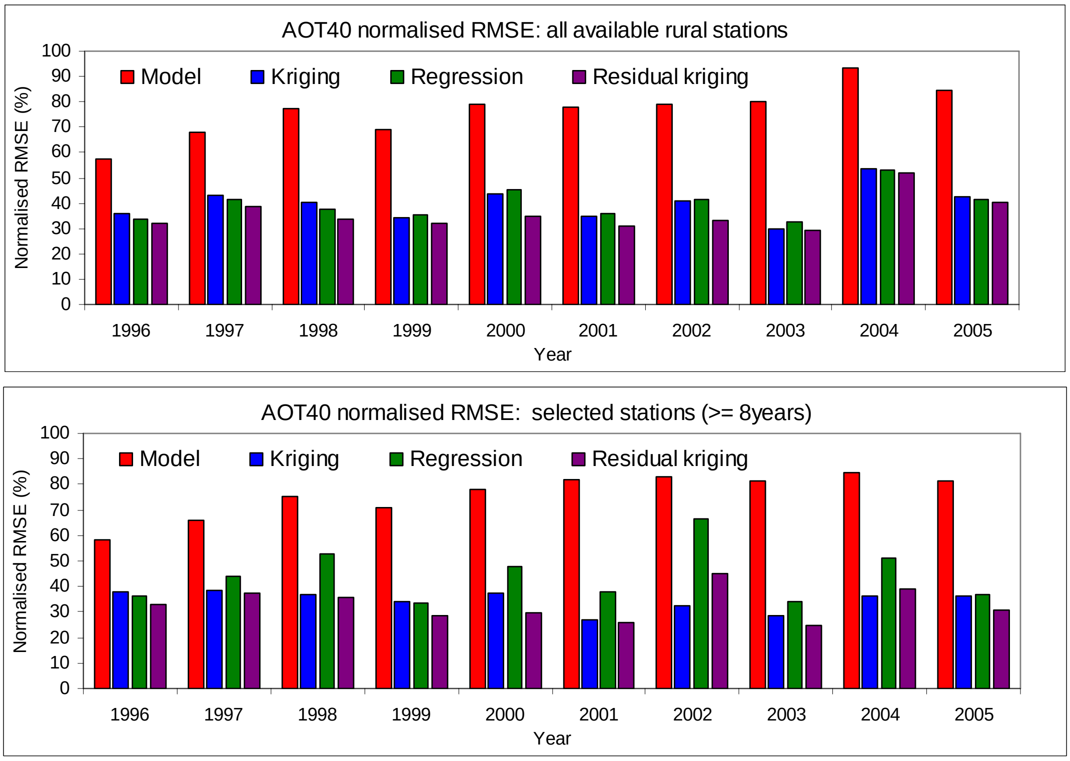


Figure 5.2. The normalised cross-validation RMSE of the different mapping methods used for AOT40. All available stations (top) and selected stations (bottom). Shown are the results from the EMEP model (red), ordinary kriging (blue), regression with EMEP and altitude (green) and the residual kriging after regression with EMEP and altitude (purple).

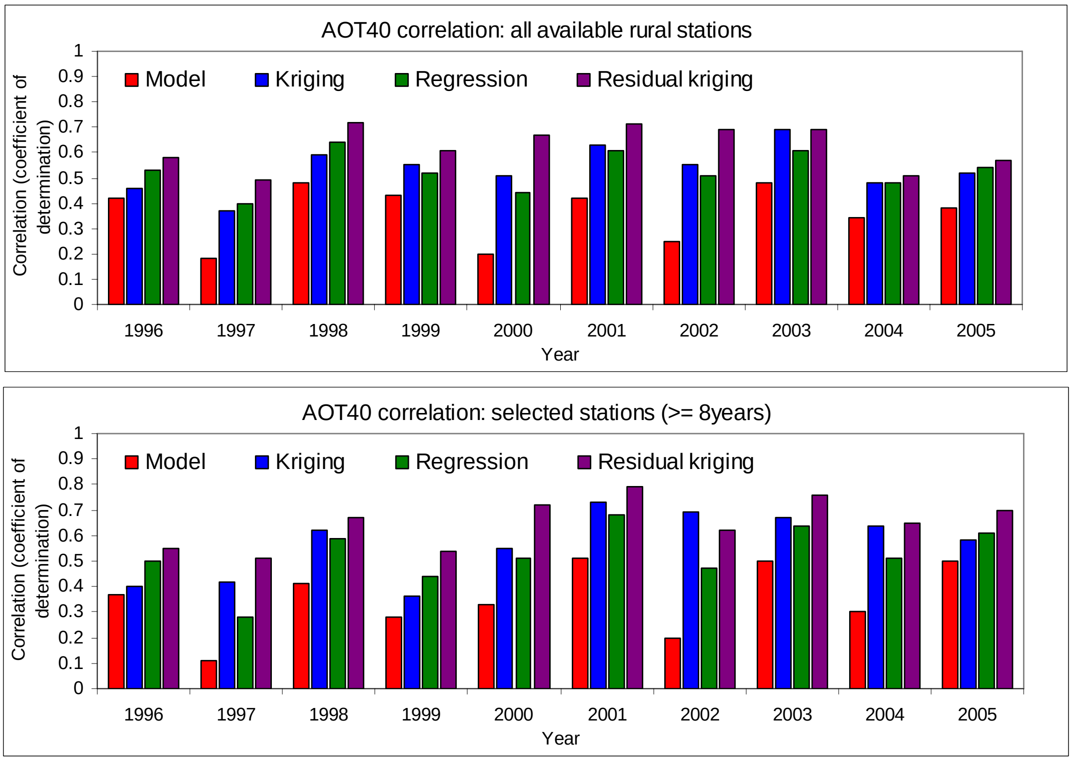
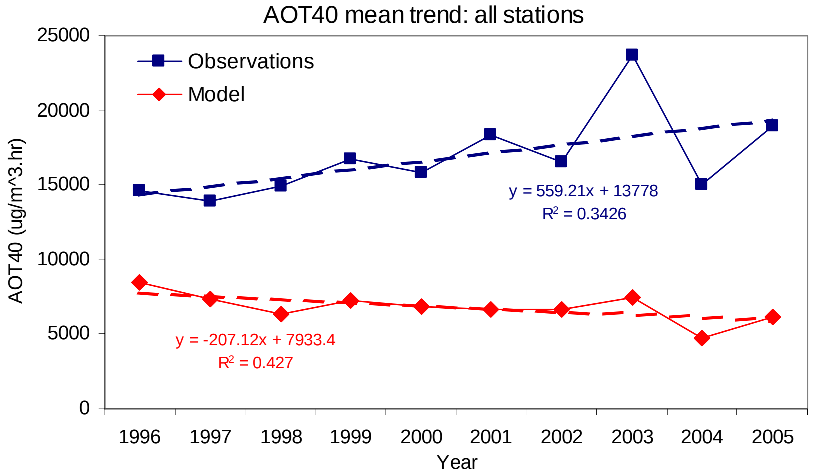


Figure 5.3. The correlation coefficient of the different mapping methods used for AOT40. All available stations (top) and selected stations (bottom). Shown are the results from the EMEP model (red), ordinary kriging (blue), regression with EMEP and altitude (green) and the residual kriging after regression with EMEP and altitude (purple).

5.2 Average trends

It is useful, before moving onto the spatial trends, to also show the mean trends as determined from the observations and the model when using the two different station selections (figure 5.4). Mean trends of the regression and kriging are not shown as these follow very closely the mean observed trend. The following points can be noted:

- The model underestimates the observed AOT40 by more than a factor of two.
- The model trend in AOT40 is decreasing whilst the observational trend in AOT40 is increasing, for both station selections.
- The observed trend is a factor of 3 higher when all the available stations are used, compared to the selected 8 year stations.
- The modelled trend is less negative when all the available stations are used, compared to the selected 8 year stations.
- The inter-annual variability of the observations is larger than the inter-annual variability of the EMEP model (attributed to variability in meteorology).



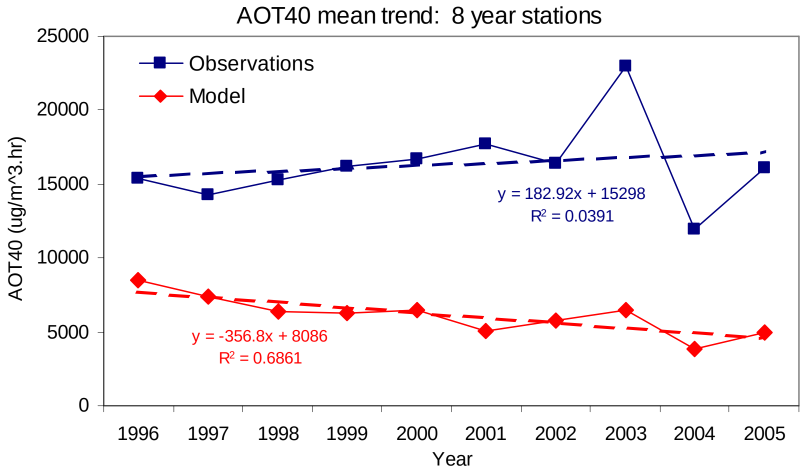


Figure 5.4. Trends of the mean observed and modelled AOT40 for the two types of station selection. All available stations (top) and selected stations with 8 or more years of coverage (bottom). Mean model concentrations are calculated at the same points in space as the available stations used for the interpolation.

5.3 Comparison of two methodologies for determining trend

We have selected linear regression and Sen's method as the two methods for determining the trend. Maps made of the trend using these two methods are shown in figure 5.5. The annual maps used for the trend analysis were created using all available stations and the interpolation methodology was residual kriging with regression of the EMEP model and altitude.

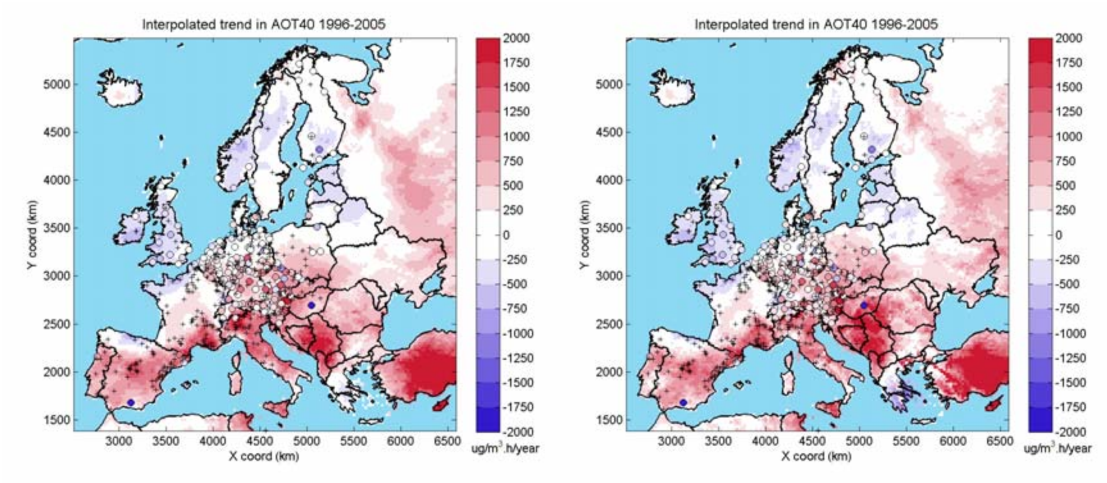


Figure 5.5. Trend maps showing the determined trend for AOT40 (1996-2005) using linear regression (left) and Sen's (right) methodology. The maps are made using residual kriging after regression with the EMEP model and altitude. All available stations are used in the annual interpolations. Also shown for comparison are the station trends. Circles with colour indicate stations with at least 8 years of data and crosses indicate other stations used for the annual interpolations but with less than 8 years of data.

Discussion: The two methods for determining trends give very similar results. There is only one area where the trend differs significantly. This is in the Balkan region where a single station shows a large negative trend. The Sen's methodology shows more positive trends in this region. This is likely due to the fact that the Sen's method uses the median of the station trend pairs. The influence of outliers on the median is less than the effect of outliers when using linear regression.

Conclusion: Due to the small differences between the two methods we will make use of linear regression only to investigate other aspects of the trend analysis. This has been selected since it is a more transparent method for assessing trends and is also computationally more efficient.

5.4 Comparison of trends for different interpolation methods

In addition to the residual kriging maps, shown in figure 5.5, it is worth comparing these to maps made using pure model calculations and maps made using pure observational interpolation. These two are shown in figure 5.6 below, using linear regression as the trend analysis technique and using all station data. These are directly comparable to figure 5.5 left which uses the same trend assessment method and station selection.

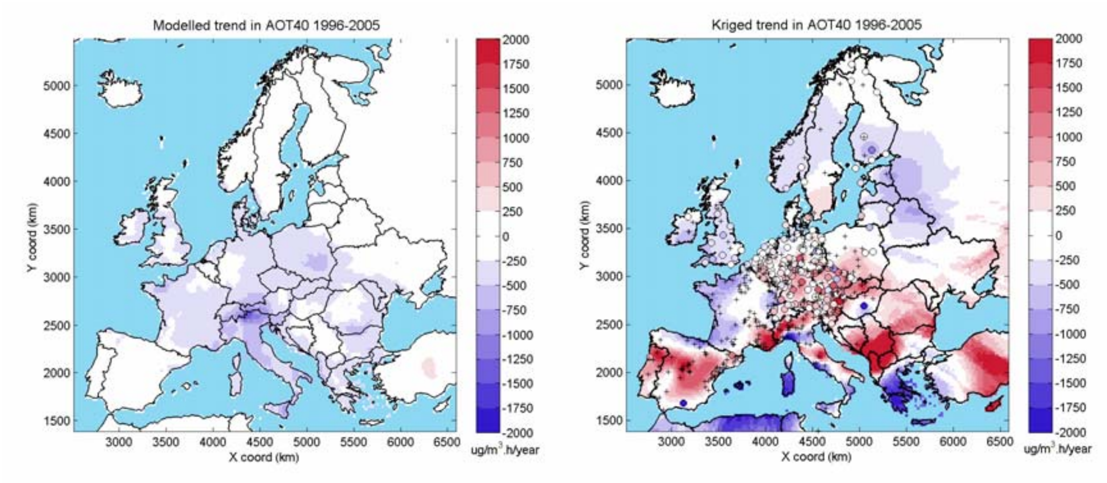


Figure 5.6. Trend maps showing the linear regression trend for AOT40 (1996-2005) of the EMEP model (left) and ordinary kriging of the observations (right), using all station data. Also shown for comparison in the kriging map are the station trends. Circles with colour indicate stations with at least 8 years of data and crosses indicate other stations used for the annual interpolations with less than 8 years of data.

Discussion: The model trend for AOT40 is slightly negative and reflects the average model trend shown in figure 5.4. Only in Turkey is there a slight positive trend in the model. In contrast the pure kriging trend map shows both strong positive and negative trends. This is an interesting map because it indicates the effect of introducing new stations into the interpolations. In areas such as Greece where station data is only available for the last 3 years strong negative trends appear. This is because the kriging interpolation overestimated the AOT40 levels in this area for the first 7 year period for which no observations were available. This can be seen in the individual maps provided in Annex I. The observed AOT40 data in this region for the last 3 years shows lower levels of AOT40 than earlier interpolated. This leads to a sudden drop in the interpolation AOT40 and a resulting negative trend. The same can occur in other areas but with positive trends as the result of the introduction of new data. A similar effect, but not as pronounced, can be seen in the residual kriging maps shown in figure 5.5.

Conclusions: The trend maps can be very sensitive to the introduction of new stations, particularly when the new data differs strongly from the previously interpolated data and when the maps are based on observational data only.

5.5 Comparison of two station selection methods

To assess the influence of station selection on the results two selection criteria have been used. The first includes all available data for all years and the second only stations with 8 or more years of data. The results are shown for both residual kriging (figure 5.7) and ordinary kriging of the observations only (figure 5.8).

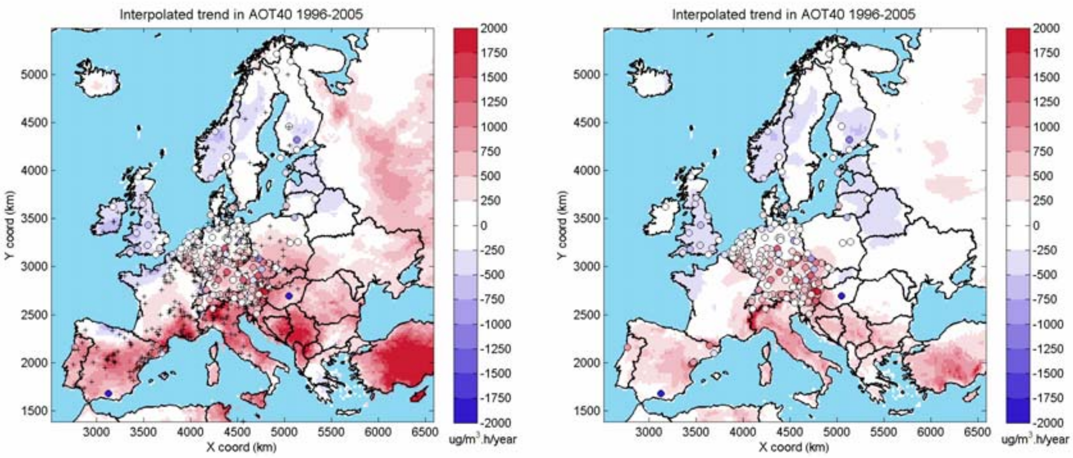


Figure 5.7. Trend maps showing the linear regression trend for AOT40 (1996-2005) all stations (left) and 8 year stations (right). The maps are made using residual kriging after regression of the EMEP model and altitude. Also shown for comparison are the station trends. Circles with colour indicate stations with at least 8 years of data (left and right) and crosses indicate other stations used for the annual interpolations with less than 8 years of data (left only).

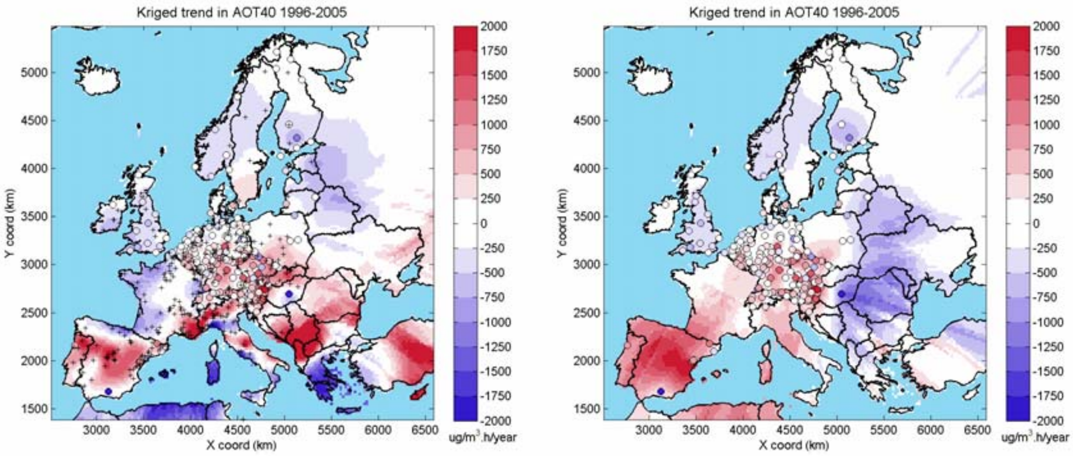


Figure 5.8. Trend maps showing the linear regression trend for AOT40 (1996-2005) all stations (left) and 8 year stations (right). The maps are made using ordinary kriging of the observation data only. Also shown for comparison are the station trends. Circles with colour indicate stations with at least 8 years of data (left and right) and crosses indicate other stations used for the annual interpolations with less than 8 years of data (left only).

Discussion: There are distinct differences between the maps in figures 5.7 and 5.8 for the two station selection methods and these differences exist mostly in the areas where observational data is not available for the required 8 years. When residual kriging is used then areas where stations are not available depend mainly on the model regression to interpolate the concentrations. The regression tends to correct the bias of the model and so the model is essentially rescaled according to the mean of the observations after regression. The result is that in areas far from stations the trend tends towards the average observational trend. When using all stations this trend is around 560 $\mu g.m^{-3}.h.year^{-1}$, whereas this trend is around 180 $\mu g.m^{-3}.h.year^{-1}$ when only 8 year stations are used (figure 5.4). This is reflected in figure 5.7 where similar, but lower, spatial distributions of the trend are found far from stations when using the two station selection methods. The absolute value of the trend is less for the 8 year stations than for all the stations.

When kriging of observations is used (figure 5.8) then the interpolated trend reflects the trends in the nearest 50 stations, the number of nearest stations used for the kriging interpolation. Since the range value used in the interpolations is less than 1000 km this means that when the interpolation is more than 1000 km away from any station then the interpolation weights the 50 nearest stations evenly, producing the average of the 50 nearest stations. In South-western Europe positive observational trends dominate whilst in Eastern Europe negative trends tend to dominate. This effect would not be present if all the stations were used for the kriging interpolation.

Conclusion: The interpolations are sensitive to the station selection, particularly when only kriging of the observations is used. For this reason it is not recommended to use pure kriging for trend analysis. When residual kriging is applied the spatial distribution of the trends is similar but the absolute value of the trends far from stations reflects the mean observational trends. In areas where the station coverage is good for the at least 8 years there is little difference between the two station selection methods.

5.6 Comparison of two methodologies for determining trend uncertainty

Determining the uncertainty or significance of a trend is essential information whenever a trend is presented. For the application here we concentrate on the uncertainty, rather than the significance, of the trend. This is because we are interested in knowing the possible range of the trend, rather than simply ascertaining if the calculated trend is due to random processes or not.

Two methods are applied here, see section 4.2, for determining the uncertainty in the trend. Firstly the standard deviation of the regression residuals is calculated and divided by the number of years over which the trend is determined, providing an estimate of the sample variability over the trend period. Secondly, use is made of the estimated uncertainty for each of the concentration estimates (based on the kriging or residual kriging variance) and Monte Carlo methods are applied, see section 4.3.3, to determine the uncertainty of the regression slope.

The second of these is demonstrated in figure 5.9 for a randomly selected point in space. In this figure the black squares and error bars represent the annual AOT40 value and its standard deviation based on the residual kriging interpolation and its variance. The pink lines represent 100 realisations of the regression trend, where each realisation is achieved by randomly sampling from the normal distribution associated with each of the years. The thick blue line shows the mean of the 100 realisations. The standard deviation of the regression slope from all the realisations is used to determine the uncertainty of the calculated trend.

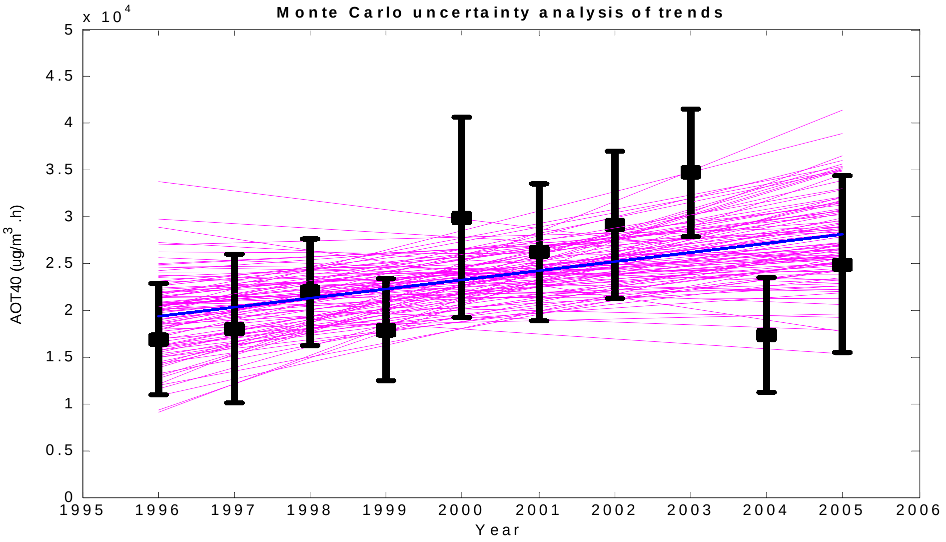


Figure 5.9. Example of a Monte Carlo simulation for deriving the uncertainty in trends given an uncertainty in the input data.

In figure 5.10 and 5.11 the two trend uncertainties are presented for the two station selection cases. These represent the standard deviation of 1) the Monte Carlo ensemble and 2) of the regression residual standard deviation.

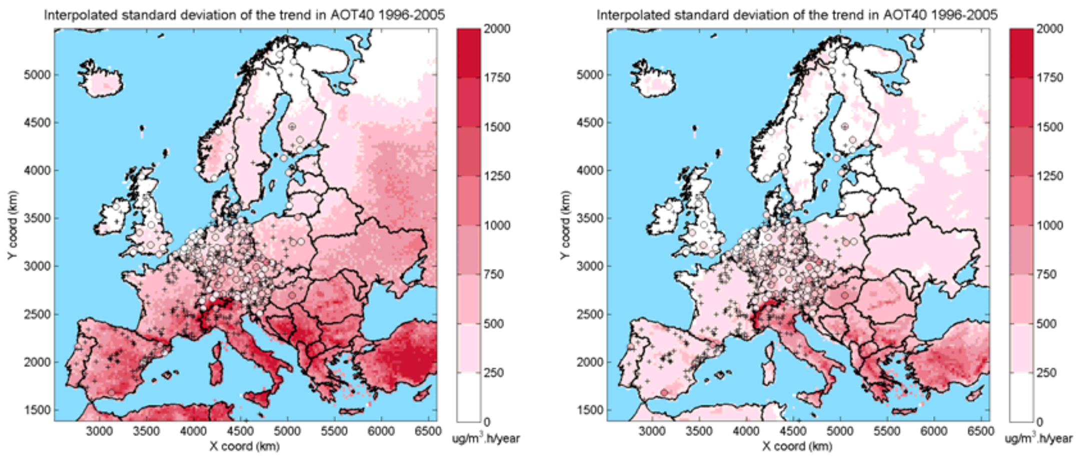


Figure 5.10. Trend uncertainty maps showing the uncertainty (standard deviation) in the linear regression trend for AOT40 (1996-2005) using the Monte Carlo method (left) and the residual standard deviation method (right). The maps are made using all available station data. Residual kriging after regression with the EMEP model and altitude is the interpolation method. Also shown for comparison are the station trend uncertainties. Circles with colour indicate stations with at least 8 years of data and crosses indicate other stations used for the interpolations but with less than 8 years of data.

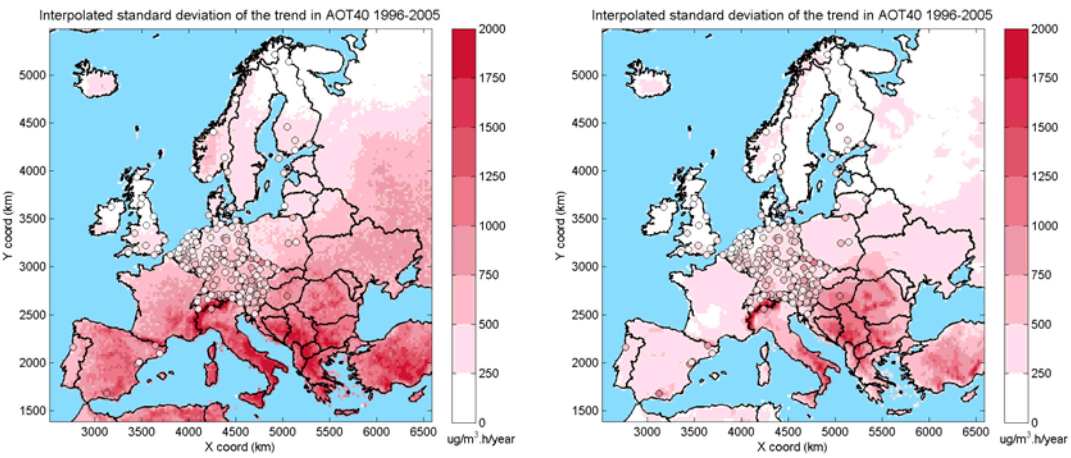


Figure 5.11. Trend uncertainty maps showing the uncertainty (standard deviation) in the linear regression trend for AOT40 (1996-2005) using the Monte Carlo methods (left) and the residual standard deviation method (right). The maps are made using stations with at least 8 years of data. Residual kriging after regression with the EMEP model and altitude is the interpolation method. Also shown for comparison are the station trend uncertainties, indicated by circles.

Discussion: The uncertainty calculated using the residual of the regression trend does not directly indicate the full uncertainty in the trend itself. It is intended more to provide information on the variability of the data with the expectation that where the variability is low then the trend is also better defined. It does not take into account the uncertainty in the data caused by interpolation. As previously mentioned this method can also give high values when the trend is non-linear. The Monte Carlo method is expected to provide more realistic uncertainties of the trend itself as it is based on the statistical distributions of the data used, in this case the residual kriging variance. Though not shown here, trend analysis using Sen's method shows a similar uncertainty when the Monte Carlo method is applied.

The magnitude of the uncertainty is, in many regions of the maps shown in figures 5.10 and 5.11, larger than the estimated trend, shown in figure 5.7. This is demonstrated in figure 5.12 where the ratio of the magnitude of the trend to the standard deviation of the trend is presented for the two uncertainty methods. Areas where the trend is larger than the standard deviations have a ratio larger than unity. In the case of the regression residual standard deviation, figure 5.12 right, there are large areas where the trend is larger than the standard deviation of the trend itself. For the Monte Carlo analysis (figure 5.12 left) the only region with a trend significantly larger than the estimated uncertainty is the United Kingdom and Ireland. This is a region with a reasonable coverage of stations that shows a negative trend in AOT40.

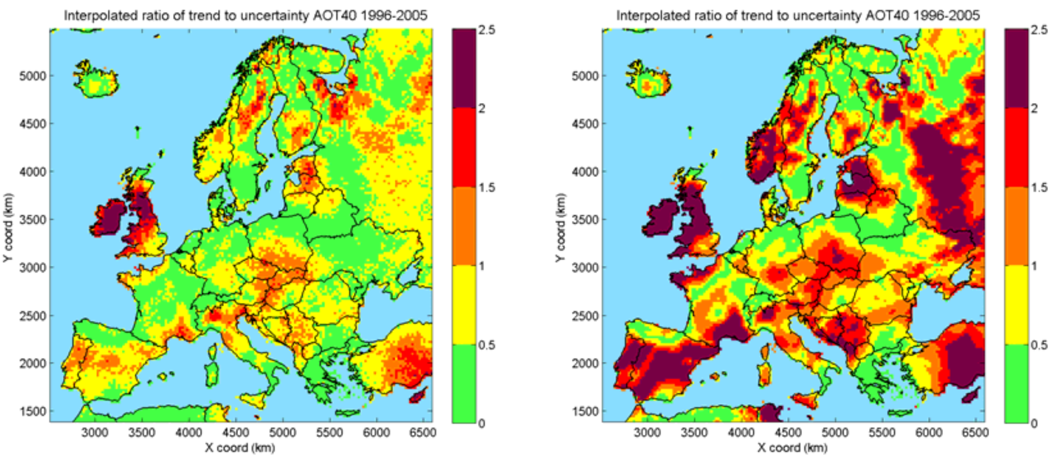


Figure 5.12. Maps showing the ratio of the absolute trend to the standard deviation of the trend for AOT40 (1996-2005) using the Monte Carlo methods (left) and the residual error method (right). The

maps are made using all station data. Residual kriging after regression with the EMEP model and altitude is the interpolation method.

It is possible to compare the results presented in figure 5.12 with the Mann Kendall significance test, section 4.3.2. In this case the alpha parameter is set at 0.05, meaning that there is at maximum a 5% chance that the trend is the result of a random process. Figure 5.13 shows this result for the same case as in figure 5.12. The significance map is very similar to the residual error ratio map, figure 5.12 right, at the contour level value of 2, but covering a slightly reduced area. This is not surprising since they both represent a 95% percentile based on the same available data distribution.

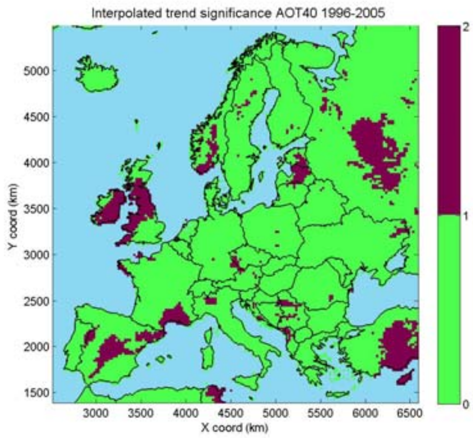


Figure 5.13. Map showing the Mann Kendall significance test for AOT40 (1996-2005). Significance is based on the assumption of a maximum 5% chance that the trend is the result of random processes. Areas in dark purple are considered significant, i.e. the trend is considered not to be the result of random sampling processes. The map is made using all station data. Residual kriging after regression with the EMEP model and altitude is the interpolation method.

Conclusion: Monte Carlo is not the only method for determining uncertainty in the trends but it provides a more realistic expression of the uncertainty than just using the regression residual standard deviation as it includes more information concerning the statistical distribution of the data used for the trend. The usefulness and application of significance tests needs to be further considered, specifically in regard to the information conveyed in such presentations.

6 Results of the tests for annual mean SO_2

In this section the results of the trend analysis of annual mean SO_2 are presented. This follows a similar structure to the AOT40 assessment but the analysis is reduced so as not to duplicate general aspects highlighted in the AOT40 analysis. SO_2 has been chosen as it has a clearly observed negative trend in Europe, in contrast to AOT40 where the trend is not so well defined. The following results will be presented:

1. Assessment statistics of the mapping methodologies
2. Average trends based on modelling and observations
3. Comparison of the two trend analysis methodologies when applied to the residual kriging interpolation
4. Comparison of the trends based on different mapping methodologies
5. Comparison of the trends using two different station selection criteria
6. Comparison of the trend uncertainty methods

6.1 Annual statistics of the interpolation maps

In figure 6.1 the number of available rural stations is shown for the two different station selections. The number of SO_2 stations in AirBase has not increased over the past 6 years (see figure 3.2). The spatial distribution of the stations with availability of 8 years or more can be seen in the maps provided in section 6.3.

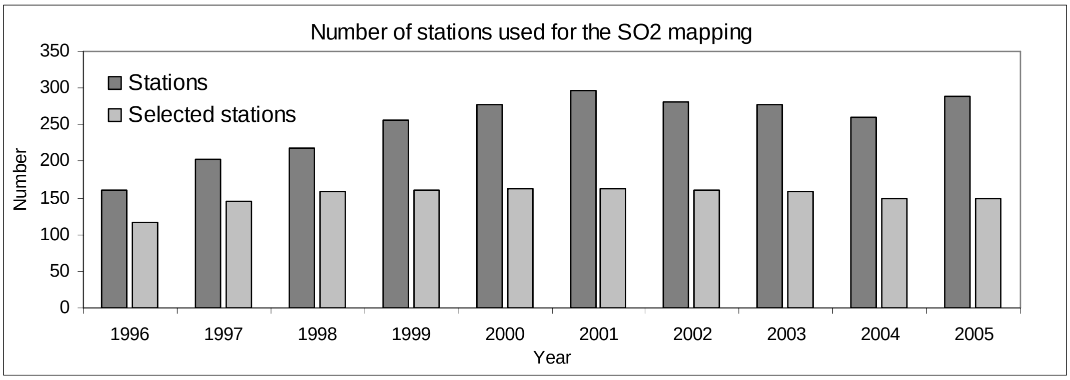
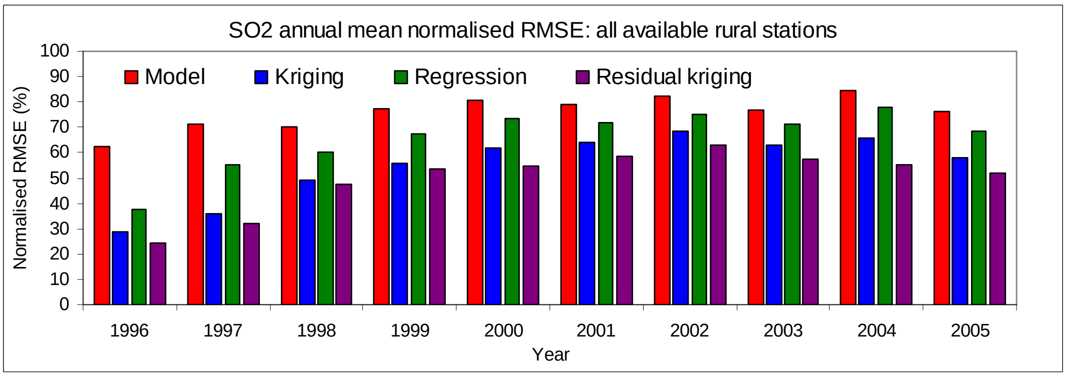


Figure 6.1. Number of stations available for the annual mean SO_2 trend assessments. Shown are the two station selection cases. 1) All available stations are used (dark grey) and 2) where stations with at least 8 years of data are used (light grey). Only rural stations are used.

To provide an overview of the statistical results of the different interpolations for each year the normalised RMSE and correlation coefficients for annual mean SO_2 are shown for the two station selections in figures 6.2 and 6.3. The results reflect previous tests where the RMSE has been shown to be lowest and the correlation highest for the residual kriging methods.



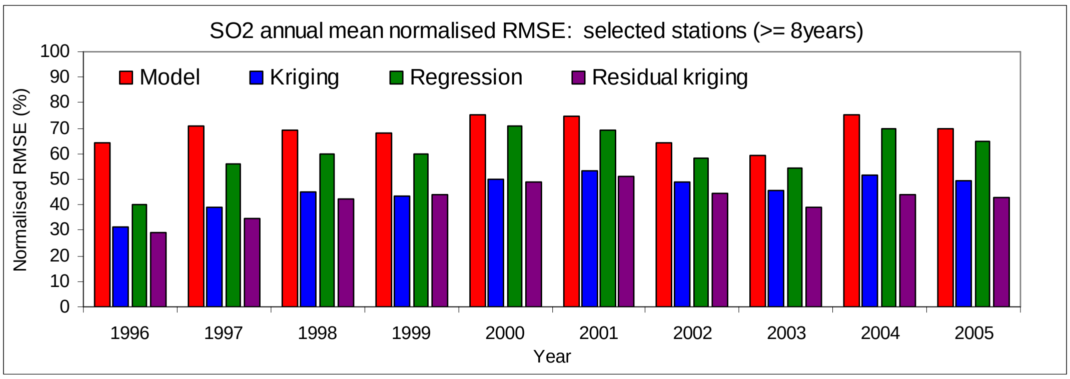


Figure 6.2. The normalised cross-validation RMSE of the different mapping methods used for annual mean SO_2. All available stations (top) and selected stations (bottom). Shown are the results from the EMEP model (red), ordinary kriging (blue), regression with EMEP (green) and the residual kriging after regression with EMEP (purple).

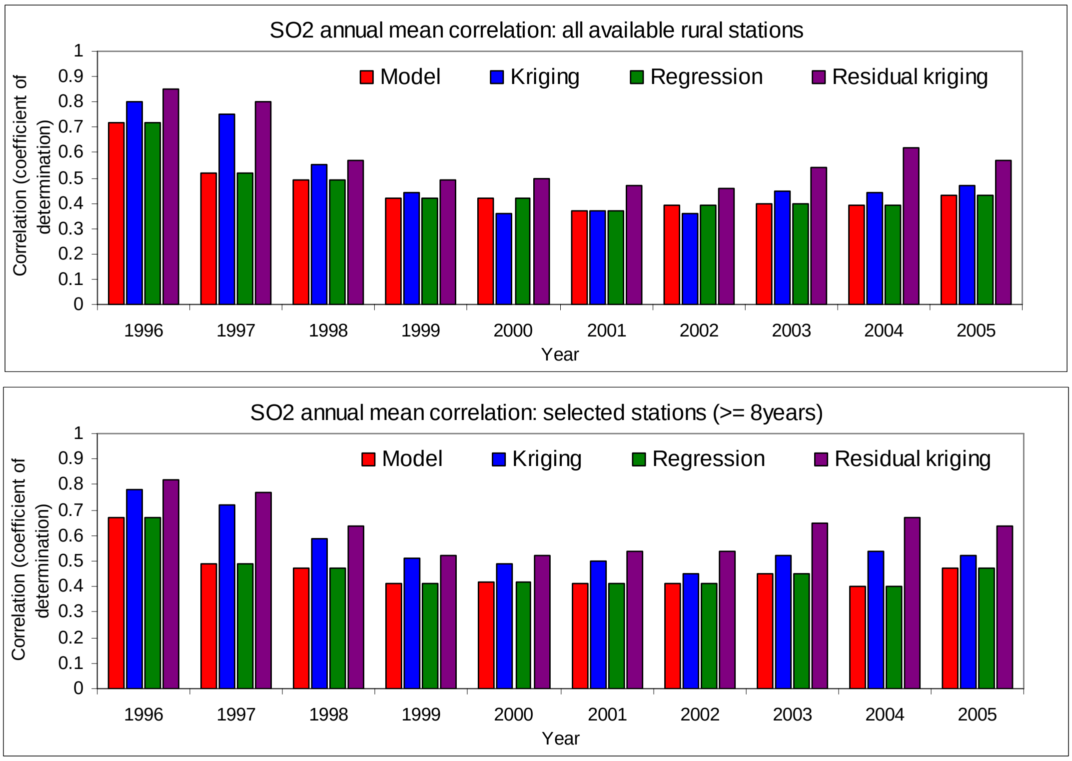
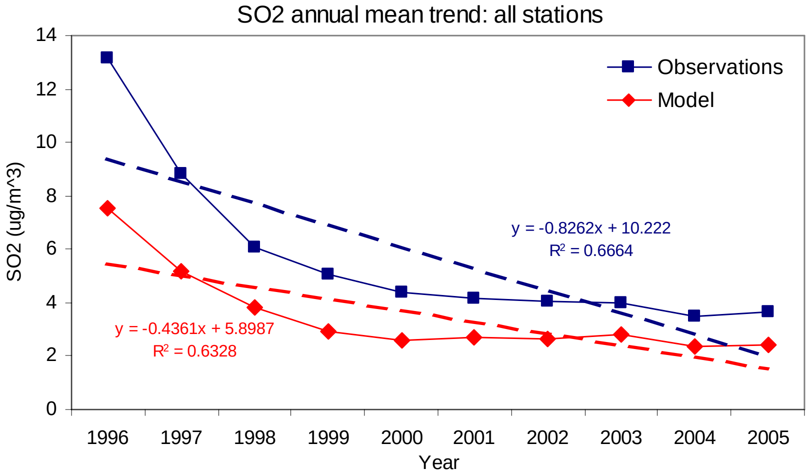


Figure 6.3. The correlation coefficient of the different mapping methods used for annual mean SO_2. All available stations (top) and selected stations (bottom). Shown are the results from the EMEP model (red), ordinary kriging (blue), regression with EMEP (green) and the residual kriging after regression with EMEP (purple).

6.2 Average trends

It is useful, before moving onto the spatial trends, to also show the mean trends as determined from the observations and the model when using the two different station selections (figure 6.4). Mean trends of the regression and kriging are not shown as these follow very closely the mean observed trend. The following points can be noted:

- The model underestimates the observed SO_2 by1/3. This in agreement with EMEP reports, e.g. Simpson et al. (2005) which reports a bias of 32% (2003) and 24% (2002).
- The model and observational trends for both station selections have a similar character but differ in magnitude by a factor of two, when regression is used to estimate a linear trend.
- The trend over the 10 year period is non-linear. Before 2000 the trend was large but after this period the trend reduces significantly in magnitude



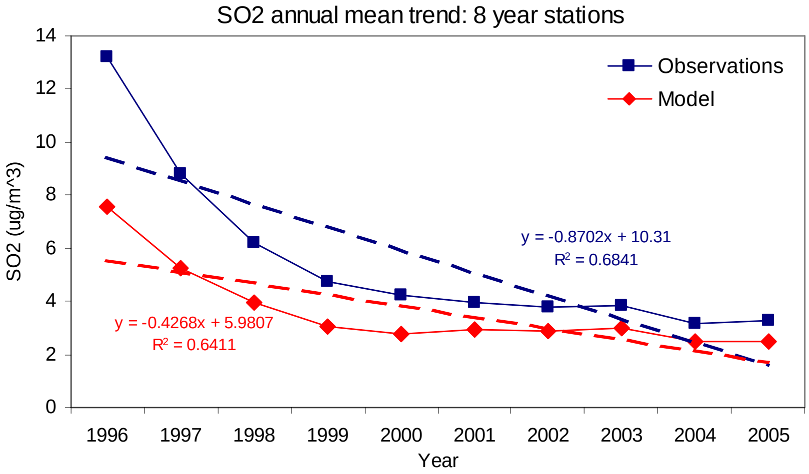


Figure 6.4. Trends of the mean observed and modelled annual mean SO_2 for the two types of station selection. All available stations (top) and selected stations with 8 or more years of coverage (bottom). Mean model concentrations are calculated at the same points in space as the available stations used for the interpolation.

6.3 Comparison of two methodologies for determining trend

Maps made of the trend using the two trend assessment methods, linear regression and Sen's method, are shown in figure 6.5 below. The annual maps used for the trend analysis were created using all available stations and the interpolation methodology used was residual kriging with regression of the EMEP model.

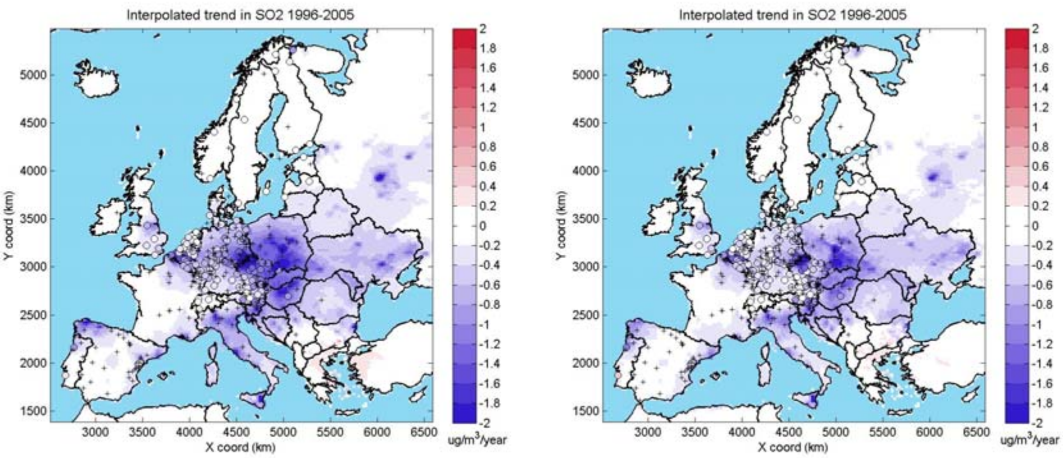


Figure 6.5. Trend maps showing the determined trend for annual mean SO_2 (1996-2005) using linear regression (left) and Sen's (right) methodologies. The maps are made using residual kriging after regression with the EMEP model. All available stations are used in the annual interpolations. Also shown for comparison are the station trends. Circles with colour indicate stations with at least 8 years of data and crosses indicate other stations used for the annual interpolations but with less than 8 years of data.

Discussion: The two methods for determining trends give very similar results, with a negative trend over almost all of Europe. The exceptions are the countries of Greece and Turkey where a slight positive trend is seen. Note that no monitoring data for Turkey is available in AirBase. The magnitude of the negative trend is slightly less when using Sen's method. This may be due, once again, to the use of the median of the station trend pairs in Sen's method, which tends to reduce the influence of outliers in the dataset, in comparison to the use of the mean in the linear regression. For this reason the initial strong trend in SO_2 has less of an influence on the Sen trend analysis compared to linear regression.

Conclusion: In this case the Sen trend analysis provides an overall reduced magnitude in the negative trend. This is likely due to the non-linear nature of the trend. More attention should be given to these differences in any further assessment. However, the differences are not large so we will make use of linear regression to investigate other aspects of the trend analysis here.

6.4 Comparison of trends for different interpolation methods

In addition to the residual kriging maps, shown in figure 6.5, it is worth comparing these to maps made using pure model calculations and maps made using pure observational interpolation to assess any differences that result with the use of different data sources, i.e. model or observations. These are shown in figure 6.6 below, using linear regression as the trend analysis technique and using all station data.

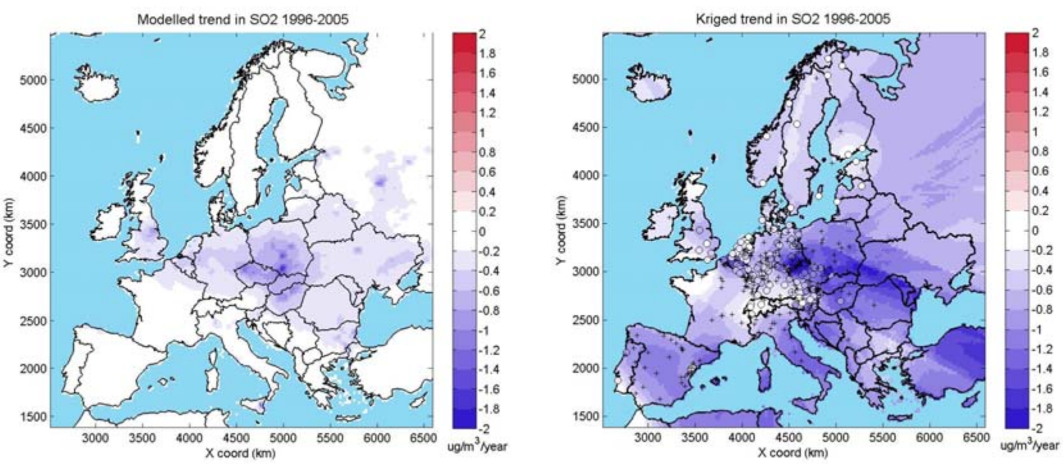


Figure 6.6. Trend maps showing the linear regression trend for annual mean SO_2 (1996-2005) of the EMEP model (left) and ordinary kriging of observations (right). The right map is made using all station data. Also shown for comparison in the kriging map are the station trends. Circles with colour indicate stations with at least 8 years of data and crosses indicate other stations used for the annual interpolations with less than 8 years of data.

Discussion: The model trend for annual mean SO_2 is everywhere very small or negative and reflects the average model trend of -0.44 $\mu g.m^{-3}.year^{-1}$, shown in figure 6.4. The pure observationally based kriging trend map also shows negative trends but these are higher, reflecting the average trend of -0.84 $\mu g.m^{-3}.year^{-1}$ also shown in figure 6.4. As previously described in section 5.4 the kriging map will tend to show the mean trend of the 50 nearest stations when the interpolation is far from stations.

Conclusions: The kriging trend maps do not provide good spatial information on the trend in areas where there are few stations for a good interpolation.

6.5 Comparison of two station selection methods

To assess the influence of station selection on the results two selection criteria have been used. The first includes all available data for all years and the second only stations with 8 or more years of data. The results are shown for the residual kriging (figure 6.7).

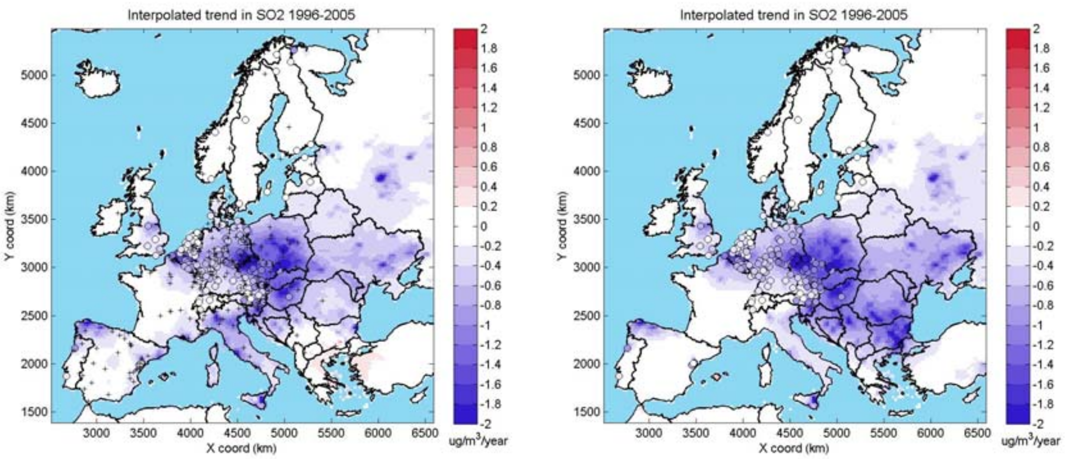


Figure 6.7. Trend maps showing the linear regression trend for annual mean SO_2 (1996-2005) all stations (left) and 8 year stations (right). The maps are made using residual kriging after regression of the EMEP model. Also shown for comparison are the station trends. Circles with colour indicate stations with at least 8 years of data (left and right) and crosses indicate other stations used for the annual interpolations with less than 8 years of data (left only).

Discussion: In contrast to the results found for AOT40 trend, section 5.5, there is little difference between the two station selection methods for annual mean SO_2. This reflects the average trends using the two selection criteria, figure 6.4, and also reflects on the general agreement between observations and model in regard to the trends.

Conclusion: The inference here is that the results will be similar if the selected stations are representative of the general trends found at all stations. This was not the case for the AOT40 calculations where the trend was smaller than the inter-annual variability, and varied substantially spatially.

6.6 Comparison of two methodologies for determining trend uncertainty

Two methods are applied here, see section 4.3, for determining the uncertainty in the trend. Firstly the standard deviation of the regression residuals is calculated and divided by the number of years over which the trend is determined, providing an estimate of the sample variability. Secondly use is made of the estimated uncertainty for each of the concentration estimates (based on the kriging or residual kriging variance) and Monte Carlo methods are applied to determine the uncertainty of the regression slope.

The second of these is demonstrated below (figure 6.8) for a randomly selected point in space. This figure, in the same way as in figure 5.9, shows 100 realisations of the Monte Carlo runs. The thick blue line shows the mean of the 100 realisations. The standard deviation of the regression slope from all the realisations is used to determine the uncertainty of the calculated trend.

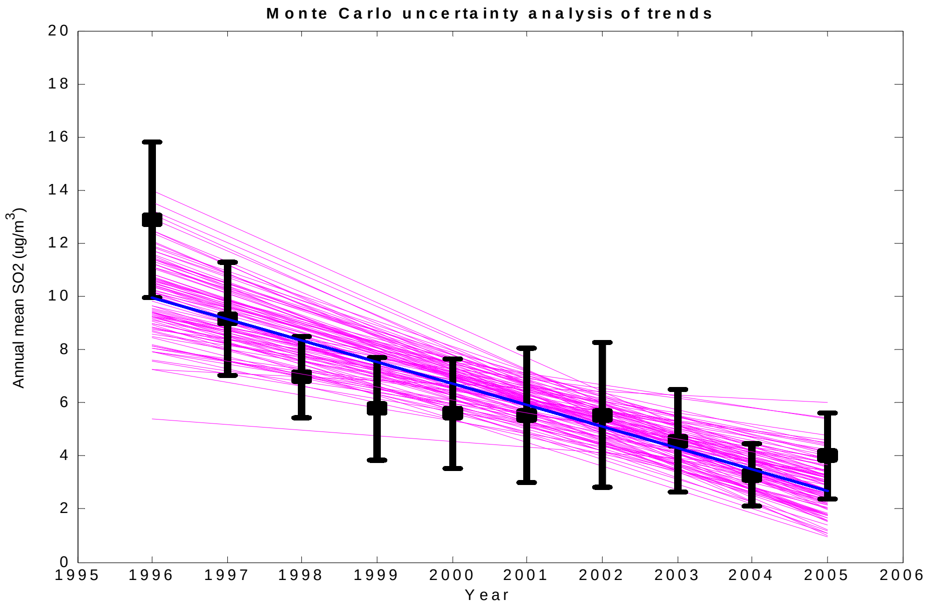


Figure 6.8. Example Monte Carlo simulation for deriving the uncertainty in trends given an uncertainty in the input data.

Due to the similarities in the two station selection trend maps, figure 6.7, we demonstrate the different uncertainty methods for the case where all station data is used, figure 6.9. These represent the standard deviation of 1) the Monte Carlo ensemble and 2) of the regression residual standard deviation.

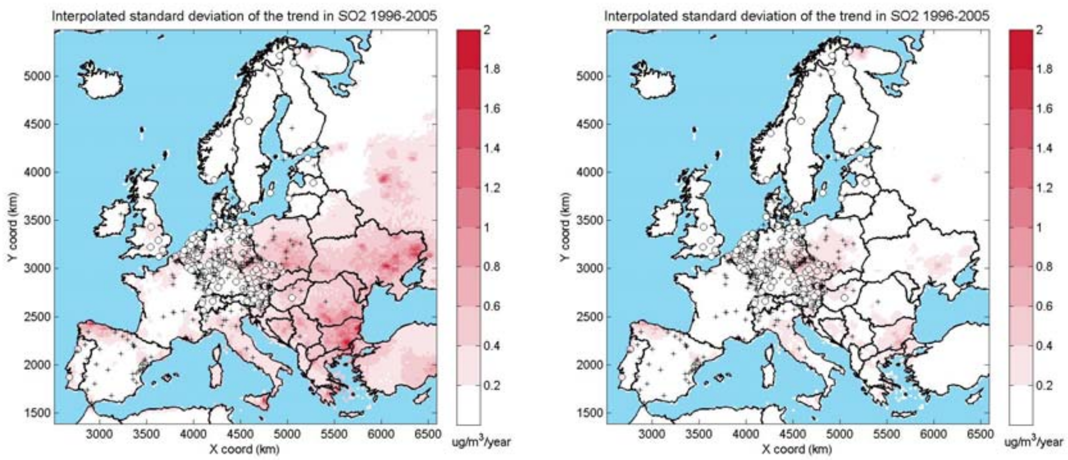


Figure 6.9. Trend uncertainty maps showing the uncertainty (standard deviation) in the linear regression trend for annual mean SO_2 (1996-2005) using the Monte Carlo methods (left) and the residual standard deviation method (right). The maps are made using all available station data. Residual kriging after regression with the EMEP model is the interpolation method. Also shown for comparison are the station trend uncertainties. Circles with colour indicate stations with at least 8 years of data and crosses indicate other stations used for the interpolations but with less than 8 years of data.

Discussion: The results presented here are similar to those found for AOT40, presented in section 5.6. The major difference here is that the uncertainty in the trend is significantly lower in comparison to the magnitude of the trend. This is shown, as in section 5.6, by comparing the ration of the trend to the standard deviations of the trends for the two methodologies (figure 6.10). For the Monte Carlo method a large area of Europe shows trends in annual mean SO_2 that are significantly larger than the calculated standard deviation.

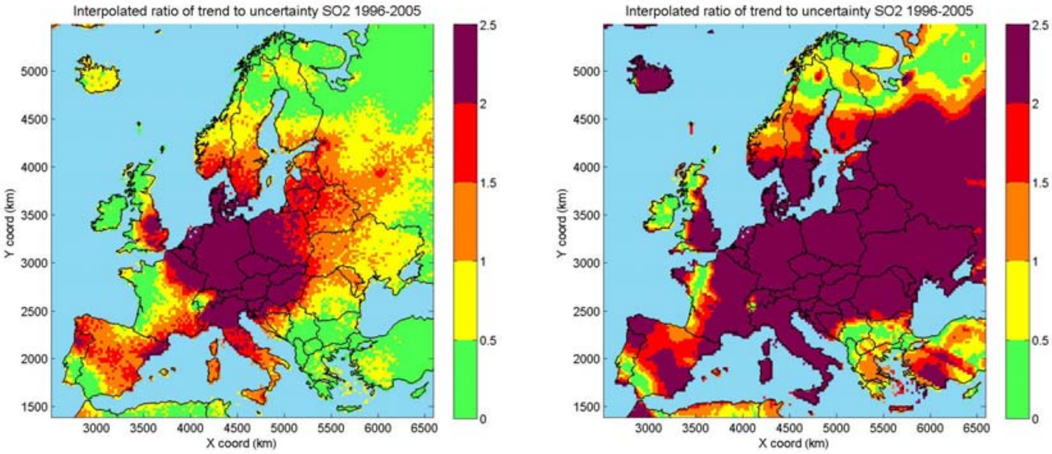


Figure 6.10. Maps showing the ratio of the absolute trend to the standard deviation of the trend for annual mean SO_2 (1996-2005) using the Monte Carlo methods (left) and the residual error method (right). The maps are made using all station data. Residual kriging after regression with the EMEP model is the interpolation method.

It is possible to compare the results presented in figure 6.10 with the Mann Kendall significance test. In this case the alpha parameter is set at 0.05, meaning that there is at a maximum a 5% chance that the trend is the result of a random process. Figure 6.11 shows this result for the same case as in figure 6.10. As in the case with AOT40, the significance map is very similar to the residual error ratio map above at the contour value of 2, but covering a slightly reduced area. This is not surprising since they both represent a 90% percentile based on the same available data distribution.

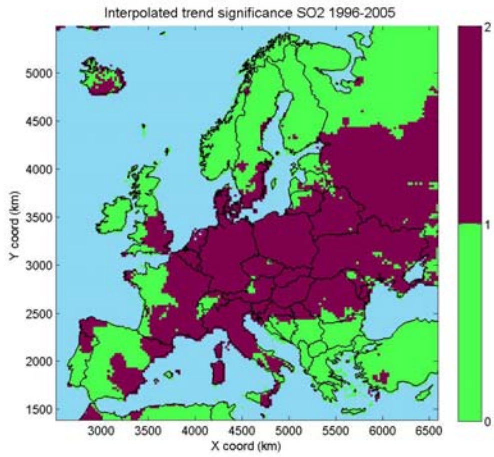


Figure 6.11. Map showing the Mann Kendall significance test for annual mean SO_2 (1996-2005). Significance is based on the assumption of a maximum 5% chance that the trend is the result of random processes. Areas in dark purple are considered significant, i.e. the trend is considered not to be the result of random sampling processes. The map is made using all station data. Residual kriging after regression with the EMEP model is the interpolation method.

Conclusion: The Monte Carlo method for uncertainty assessment indicates higher levels of uncertainty than does the residual error or Mann Kendall methods. This is the direct result of the inclusion of the interpolation uncertainty in the data used for the trend, rather than just the use of the data variability to indicate uncertainty. However, in all cases the trend in SO_2 is shown to be significant for most of Europe. Further development and comparison of the uncertainty and significance methodologies would be necessary to determine the best methodology for representing and communicating trend uncertainties.

7 Conclusions and recommendations

In this section we present specific conclusions from the study, some general conclusions concerning the methodology and trend assessment in general. We then provide a list of recommendations on further application and development necessary for improving the spatial assessment of trends.

7.1 Conclusions concerning AOT40 and the spatial assessment of trends

The following conclusions concerning the methodological application of trend assessment to AOT40 are reiterated from the various sections in this report.

- A comparison of modelled and observed average trends provides the following conclusions (Section 5.2):
  - The EMEP model underestimates the observed AOT40 by more than a factor of two.
  - The model trend in AOT40 is decreasing whilst the observational trend in AOT40 is increasing, for both station selections.
  - The observed trend is a factor of three higher when all the available stations are used, compared to the 8 year station selection.
  - The modelled trend is less negative when all the available stations are used, compared to the 8 year station selection.
  - The inter-annual variability of the observations is larger than the inter-annual variability of the EMEP model. This is attributed to variability in meteorological conditions.
- Linear regression and Sen's method give very similar estimates for the trend. However, Sen's method appears to be less sensitive to outliers and may be the preferred method of trend assessment in future studies (Section 5.3)
- The trend maps can be very sensitive to the introduction of new stations, particularly when the new data differs strongly from the previously interpolated data and when the maps are based on observational data only (Section 5.4).
- In this study two station selections were used; all available monitoring data and all stations with monitoring data available for at least 8 years. The interpolated trend maps are sensitive to the station selection, particularly when only kriging of observations is used. For this reason it is not recommended to use pure kriging of observations for trend analysis. When residual kriging is applied the spatial distribution of the trends is less sensitive to station selection. In areas where the station coverage is good for at least 8 years there is little difference between the two station selection methods tested here (Section 5.5).
- Monte Carlo is not the only method for determining uncertainty in the trends but it provides a more realistic expression of the uncertainty than just using the regression residual standard deviation as it includes more information concerning the uncertainty of the data used for the trend. The usefulness and application of significance tests needs to be further considered, specifically in regard to the information conveyed in presentation maps (Section 5.6)

7.2 Conclusions concerning SO_2 and the spatial assessment of trends

The following conclusions concerning the methodological application of trend assessment to SO_2 are reiterated from the various sections in this report.

- A comparison of modelled and observed average trends provides the following conclusions (Section 6.2):
    - The EMEP model underestimates the observed annual mean concentration of SO_2 by 1/3.
    - There is a significant negative trend in SO_2, for both model and observations. However, the model trends underestimate the magnitude of the observed trends in SO_2 by a factor of two.
    - Both station selection methods applied (all available stations and only stations with 8 years or more of data) give very similar average trends.
    - The trend over the 10 year period is non-linear. Before 2000 the trend was large but after this period the trend reduces significantly in magnitude
- In the case of SO_2, Sen's methodology provides lower values for the magnitude of the negative trend. This is likely due to the non-linear nature of the trend. More attention should be given to these differences in any further assessment (Section 6.3).
- The kriging of observations trend maps do not provide good spatial information on the trend in areas where there are few stations for a good interpolation (Section 6.4)
- The Monte Carlo method for uncertainty assessment indicates higher levels of uncertainty than does the residual error or Mann Kendall methods. This is the direct result of the inclusion of uncertainty in the data used for the trend, rather than just the use of the data variability to indicate uncertainty. Further development and comparison of the uncertainty and significance methodologies would be necessary to determine the best methodology for representing and communicating trend uncertainties (Section 6.5).

7.3 General conclusions concerning the spatial assessment of trends

The following conclusions are more general in nature and present an interpretation of both the results of this scoping study and other comments and input.

- The methodology applied here for spatially mapping trends of AOT40 and annual mean SO_2 concentrations makes use of both models, with good spatial distribution, and observations, with realistic measurements. There is a significant difference between the modelled and observed trends when viewed separately. Any spatial assessment of trend cannot be based on modelling alone and the interpolation methodology applied here for producing combined concentration maps is considered to be suitable for providing both spatial coverage and observed trends.
- AOT40 is very sensitive to model bias and may also not be a good indicator for ozone trends. Use of other ozone indicators for trends may be more appropriate, as discussed in Solberg et al. (2008). The 26th highest daily 8 hour running mean, included in the air quality directive, may be a more appropriate indicator for trend assessment.
- There is a need for better quality control of the monitoring data applied in the trend assessment. Solberg et al. (2008) carry out visual inspection of the data as a quality control. The results of that study, and other methods, can be used to help improve the quality of the maps presented here.
- Two methodologies for calculating trends were presented here. These were linear regression and Sen's method. Both methods gave similar results but Sen's method was seen to be less sensitive to outliers. Though linear regression was applied more extensively in this study, Sen's method may be more appropriate for further assessment.
- The trend in SO_2 was found to be non-linear in nature. Division of the period before and after 2000 would give more information concerning the development of the trend in more recent years, compared to pre 2000 levels.

- The spatial mapping of the trends was found to be sensitive to the inclusion or selection of observational data. This is particularly the case when a monitoring site is introduced in an area where little monitoring is available and this new data differs from the interpolated value at that site. This was found to be most sensitive when the interpolation was based on kriging of the observations alone but less sensitive when interpolation also made use of the model results.

7.4 Improvements for the spatial assessment of trends

The current study was considered from the outset to be a feasibility study. Within the scope of this report it could not deal with the complete range of methodological or assessment possibilities. The following points discuss further improvements or questions that need to be addressed if the methodology is to be further developed and implemented.

- How to deal with non-linear trends? Can these be assessed using the currently employed trend analysis methods? Are other analysis methods required?
- How to best calculate and to represent the uncertainty, or significance, of the trends? In the current study two methods for determining uncertainty have been applied. From this the Monte Carlo method, which includes the interpolation uncertainty of each of the maps, gives larger uncertainty levels than methods based on the trend analysis alone. Are there other methods for determining the uncertainty and what is the best method for presenting the results? Is the calculated uncertainty appropriate for regions far from available observations?
- How best to select the stations? What is the minimum level or requirements concerning the spatial distribution of stations in such a selection? Station selection has been shown in this study to influence the results. Though in an ideal world the use of stations with continuous monitoring is the most appropriate for trend analysis it would be a mistake to reduce the analysis to just these stations when added information is available from stations with shorter monitoring histories. How best to optimise this combination is not clear.
- How many years are required for trend analysis? No tests were carried out concerning this aspect in the current study. PM_{10} was not assessed due to the shorter period of available data. PM_{10} could be assessed for a 6 year period but is this sufficient?
- In regard to the spatial assessment of trends, what is the minimum required quality of each of the spatial maps to be included in the trend analysis?
- What are the differences in trends resulting from subtle changes in the interpolation method, e.g. log-normal versus normal residual kriging or the inclusion of other supplementary data such as meteorology in the interpolations?
- To enhance the assessment of uncertainty in the trends use can be made of cross-validation methods to assess the predicted trends at sites with available observations, thus allowing a comparison of 'predicted' and 'observed' trends.
- Improved station selection. In this report all stations were used without question. There are large discrepancies between the cross-validation RMSE values obtained here and those obtained in Horálek et al. (2005; 2007; 2008) where more extensive station selection was carried out. In addition the ozone report from Solberg et al. (2008) has provided a list of ozone stations that should be excluded from the analysis.
- The current study has been applied to rural stations only. The same methodology can also include urban and suburban stations in the analysis at a 10 km resolution. What will be the impact of this on the maps generated? Do trends differ spatially for the (sub)urban and rural stations?

References

AirBase (2008), European air quality database. http://airbase.ionet.europa.eu/

Baldsano, J. M., M. Goncalves and P. Jimenz (2005). A review of background air quality trends in Europe using EMEP data. Air pollution XIII: Thirteenth international conference on modelling, monitoring and management of air pollution. vol. 82, pp. 129-138.

Box, G.E.P. and G. M. Jenkins (1976) Time series analysis: Forecasting and Control, 2nd Edition, Holden-Day, San Francisco.

Bronnimann S., B. Buchmann and H. Wanner (2002). Trends in near-surface ozone concentrations in Switzerland: the 1990s, Atmospheric Environment. 36, 2841-2852.

Chandler R. (2002). Trend analysis for the environmental sciences – a review. ESSG meeting, March 2002. http://www.homepages.ucl.ac.uk/~ucakarc/work/trends/0203talk.pdf

Cressie, N. (1993). Statistics for spatial data. Wiley series, New York.

De Leeuw, F.A.A.M (2000), Trends in ground level ozone concentrations in European Union, Environmental Science & Policy 3, 189-199

Denby B., J. Horálek, P. de Smet, F. de Leeuw, P. Kurfürst (2008). European scale exceedance mapping for PM10 and ozone based on daily interpolation fields. ETC/ACC Technical paper 2007/8. URL: http://air-climate.eionet.europa.eu/reports/ETCACC_TP_2007_8_spatAQmaps_dly_interpol

EC (2008) Directive 2008/50/EC of the European Parliament and of the Council of 21 May 2008 on ambient air quality and cleaner air for Europe. Official Journal L 152, 11.6.2008, p. 1–44 http://eur-lex.europa.eu/LexUriServ/LexUriServ.do?uri=OJ:L:2008:152:0001:0044:EN:PDF

Gilbert, R.O. (1987). Statistical Methods in Environmental Pollution Monitoring, Van Nostrand Reinhold, New York (1987).

Gröger J. and H. Rumohr (2006). Modelling and forecasting long-term dynamics of Western Baltic macrobenthic fauna in relation to climate signals and environmental change. Journal Sea Research 55: pp 266-277.

Hess A., Hari Iyera and William Malm (2001). Linear trend analysis: a comparison of methods. Atmospheric Environment 35 (2001) 5211–5222.

Hirsch, R. M. and J. R. Slack (1984). Non-parametric trend test for seasonal data with serial dependence, Water Resource Res. 20(6), 727–732.

Horálek, J., P. Kurfürst, B. Denby, P. de Smet, F. de Leeuw, M. Brabec, J. Fiala (2005). Interpolation and assimilation methods for European scale air quality assessment and mapping. Part II: Development and testing new methodologies. ETC/ACC Technical paper 2005/8. http://air-climate.eionet.europa.eu/docs/ETCACC_TechnPaper_2005_8_SpatAQ_Part_II.pdf

Horálek, J., B. Denby, P. de Smet, F. de Leeuw, P. Kurfurst, R. Swart and T. van Noije, (2007). Spatial mapping of air quality for European scale assessment. ETC/ACC Technical paper 2006/6. URL: http://air-climate.eionet.europa.eu/reports/ETCACC_TechnPaper_2006_6_Spat_AQ

Horálek, J., P. Kurfurst, P. de Smet, F. de Leeuw, R. Swart, B. Denby and J. Fiala, (2008). European air quality maps for 2005 including uncertainty analysis. ETC/ACC Technical paper 2007/7. http://air-climate.eionet.europa.eu/reports/docs/ETCACC_TP_2007_7_SpatialAQmapping2005_annual_interpolations.pdf

Hu, W., S. Tong, K. Mengersen and B. Oldenburg (2006). Rainfall, mosquito density and the transmission of Ross River virus: a time-series forecasting model. Ecological Modelling vol. 196 505-514.

Ibarra-Berastegi et al (2001), Long-term changes in Ozone and traffic in Bilbao, Atmospheric Environment 35, 5581-5592.

Kendall, M. G. (1975). Rank Correlation Methods, Griffin, London.

Mann, H. B., (1945). 'Nonparametric tests against trend', Econometrica 13, 245–259.

Mol, W.J.A., P.R. van Hooydonk and de F.A.A.M. de Leeuw (2008). European exchange of monitoring information and state of the air quality in 2006. ETC/ACC Technical paper 2008/1 http://air-climate.eionet.europa.eu/reports/ETCACC_TP_2008_1_EoI_AQ_meta_info2006

Mosmann et al. (2004). Detection of statistically significant trends in the summer precipitation of mainland Spain. Atmospheric Research , 70 , 43-53.

OSPAR Commission (2005), Assessment of trends in atmospheric concentration and deposition of hazardous pollutants to the OSPAR maritime area – Evaluation of the CAMP network. London, 2005, OSPAR publication 234/2005.
http://www.ospar.org/documents/DBASE/Publications/p00234_camp%20assessment%20report.pdf

Partal T. and E. Kahya (2006). Trend analysis in Turkish precipitation data. Hydrological Processes 20: 2011–2026.

Porter, P. S., S. T. Rao and C. Hogrefe (2002). Linear trend analysis: a comparison of methods. Discussion paper: Atmospheric Environment 36 (2002) 3055–3056.

Rao, T.S. and I.G. Zurbenko (1994). Detecting and tracking changes in ozone air quality. Journal of the Air & Waste Management Association 44, 1089-1092.

Salmi T. et al. (2002). Detecting trends of annual values of atmospheric pollutants by the Mann-Kendall test and Sen's slope estimates- the Excel template application MAKESENS, Finnish Meteorological Institute, Publication on Air Quality No. 31, (FMI-AQ-31).

Schär and Frei (2001). Detection Probability of trends in Rare Events: Theory and Application to Heavy Precipitations in the Alpine Region. Journal of Climate, 14, 1568-1584.

Sen, P. K. (1968). 'Estimates of the regression coefficient based on Kendall's tau', J. American Statist. Assoc. 63, 1379–1389.

Simpson, D., H. Fagerli, J.E. Jonson, S. Tsyro and P. Wind (2005). Transboundary Acidification, Eutrophication and Ground Level Ozone in Europe: PART I. EMEP Status Report 1/2005.
http://www.emep.int/publ/reports/2003/emep_report_1_part1_2003.pdf

Solberg S., J. Horalek, J. E. Jonson, S. Larssen, F. de Leeuw (2008). Assessment of ground-level ozone within the EEA Member Countries with focus on long-term trends. First draft of an ETC/ACC technical report, in preparation.

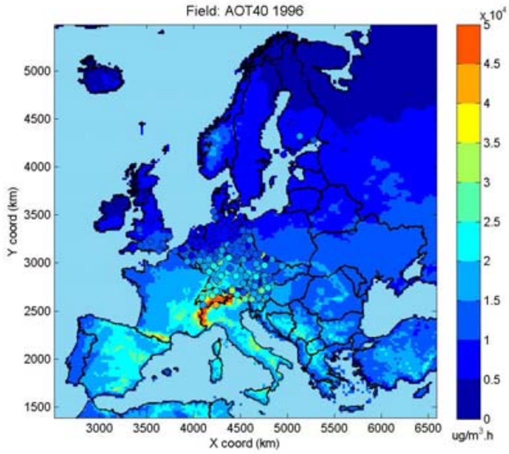
Theil, H., (1950). 'A rank-invariant method of linear and polynomial regression analysis, I, II, III', Nederl. Akad. Wetensch. Proc. 53, 386–392; 512–525; 1397–1412.

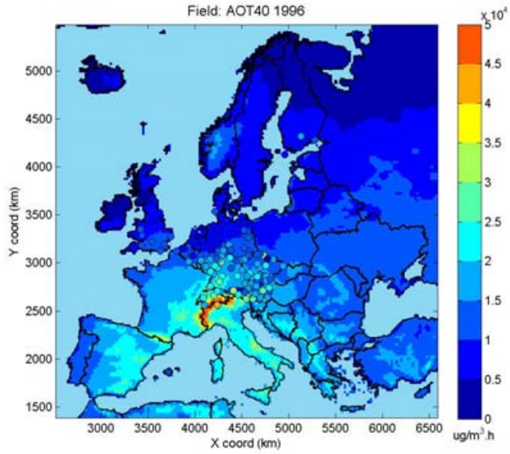
Zurbenko, I. G., (1986). The spectral Analysis of Time Series. North-Holland Series in Statistics and Probability, 248pp.

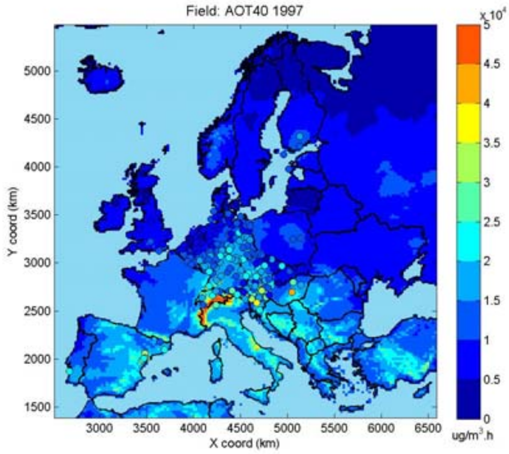
Annex I: Yearly maps of AOT40

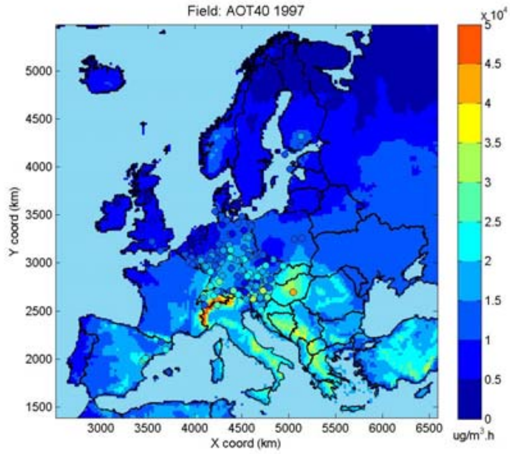
To aid visualisation each individual map produced when using the residual kriging method for AOT40 are presented. These maps show the effect of including all (left) or a selection of the stations (right).

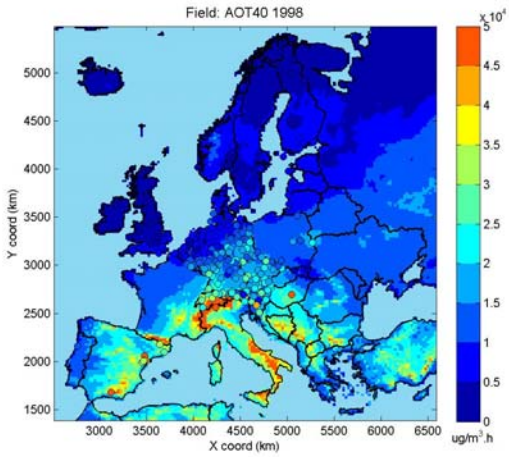
All stations	Selected stations >= 8 years

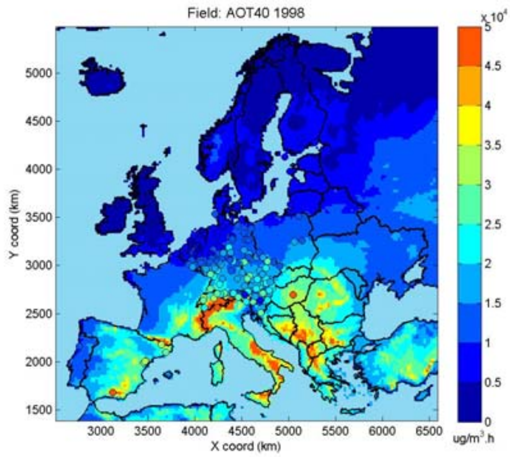




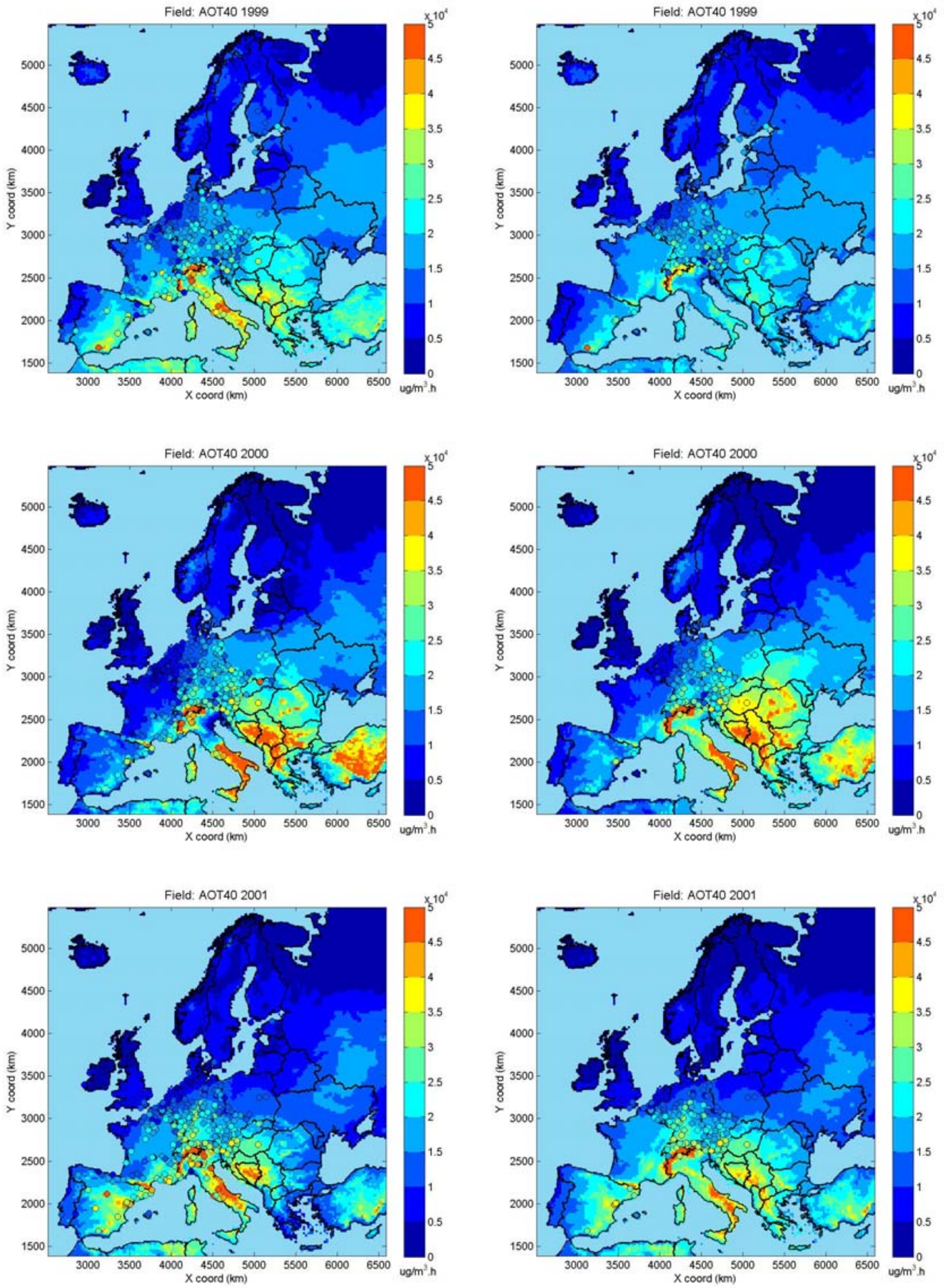




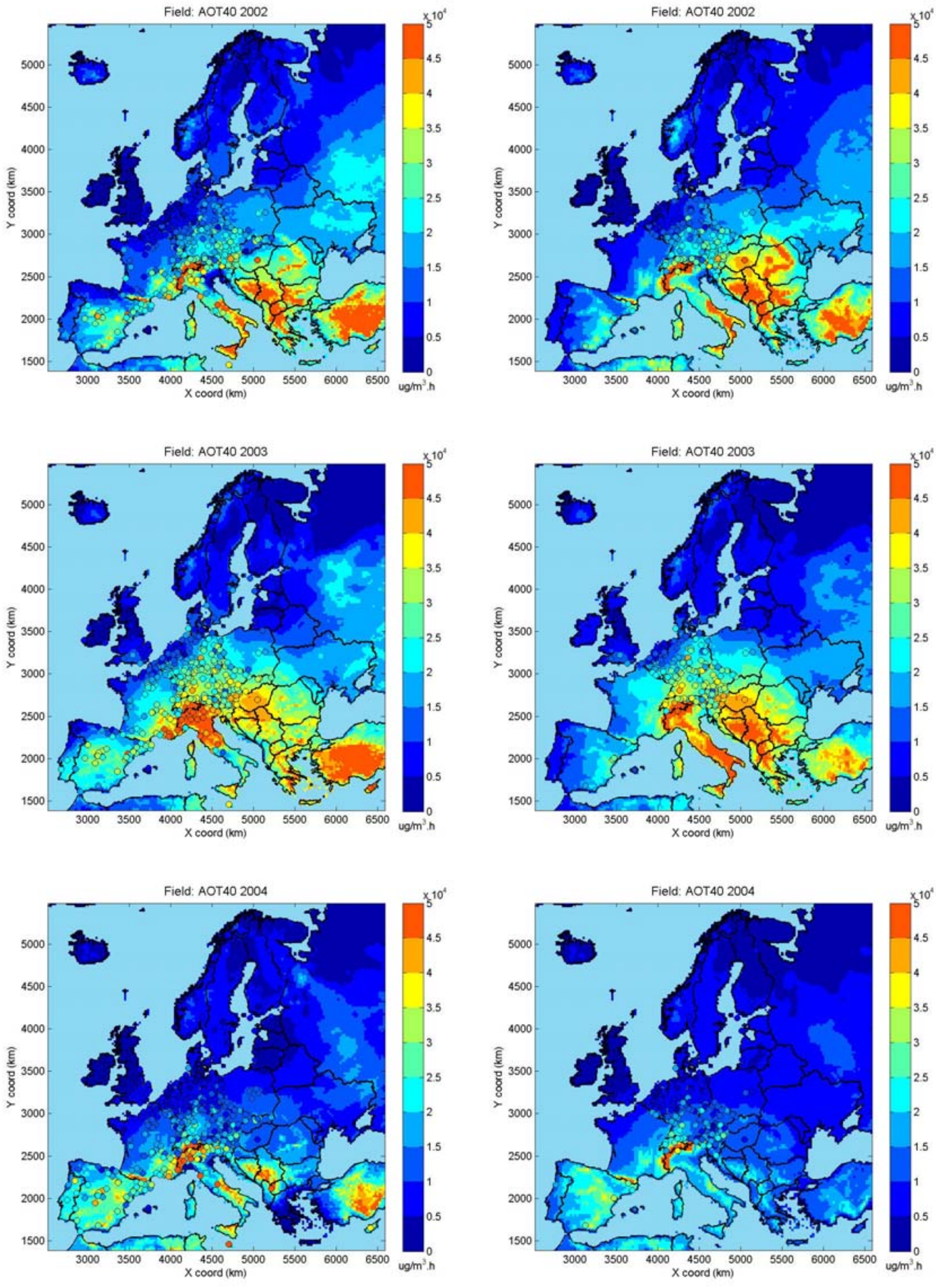




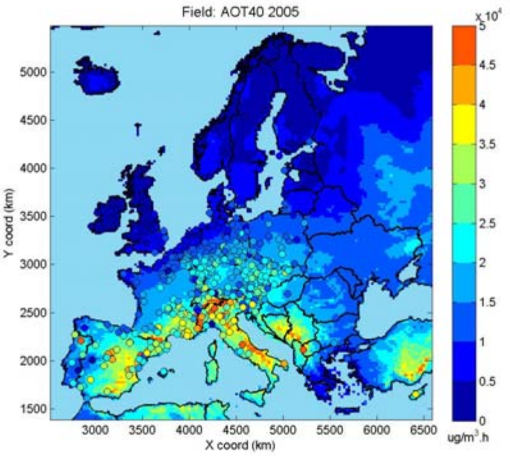
All stations | Selected stations >= 8 years



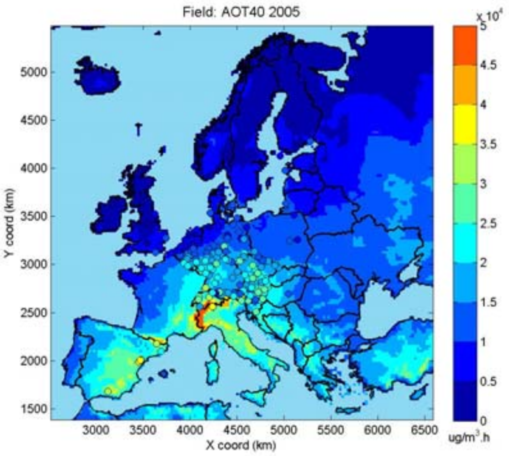
All stations | Selected stations >= 8 years



All stations



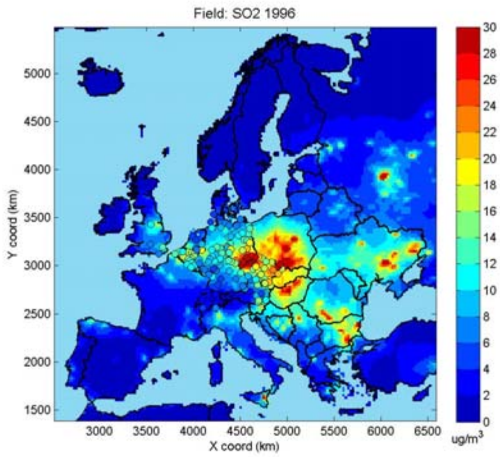
Selected stations >= 8 years

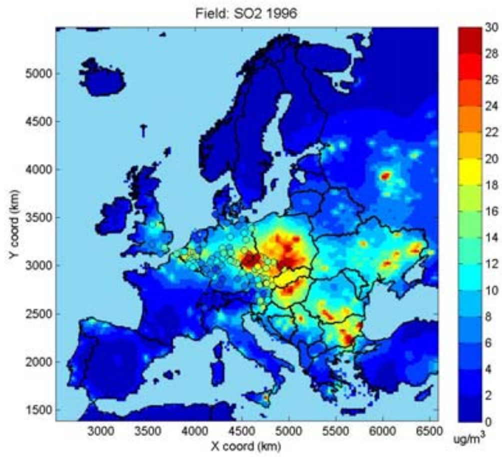


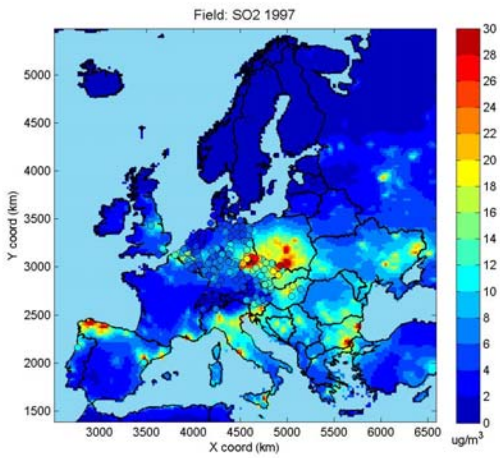
Annex II: Yearly maps of annual mean SO_2

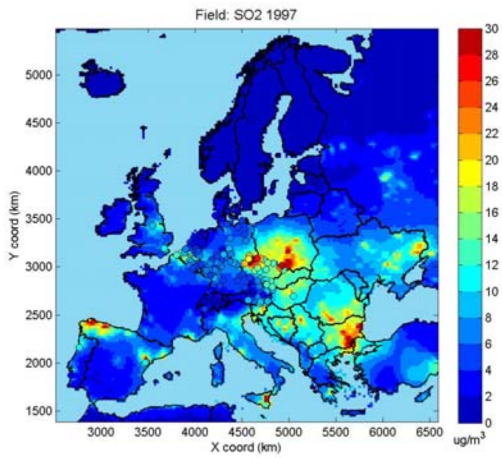
To aid visualisation each individual map produced when using the residual kriging method for SO_2 are presented. These maps show the effect of including all (left) or a selection of the stations (right).

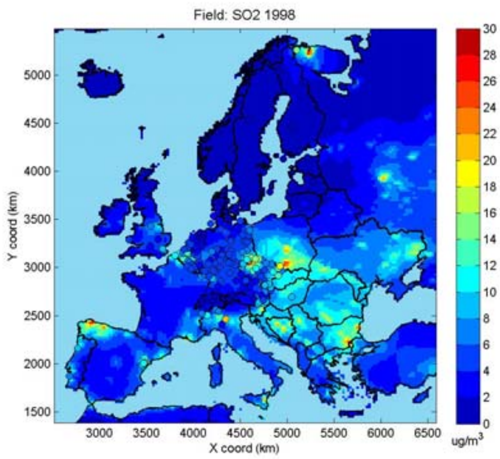
All stations | Selected stations >= 8 years

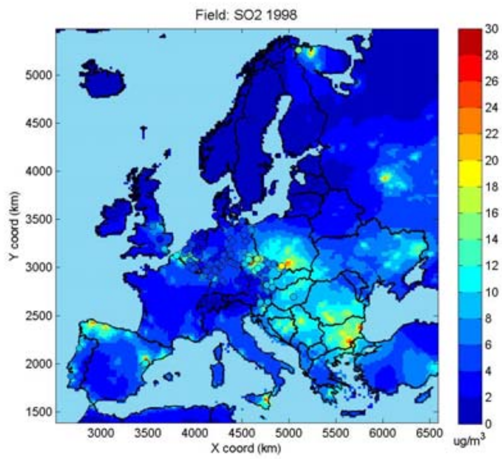




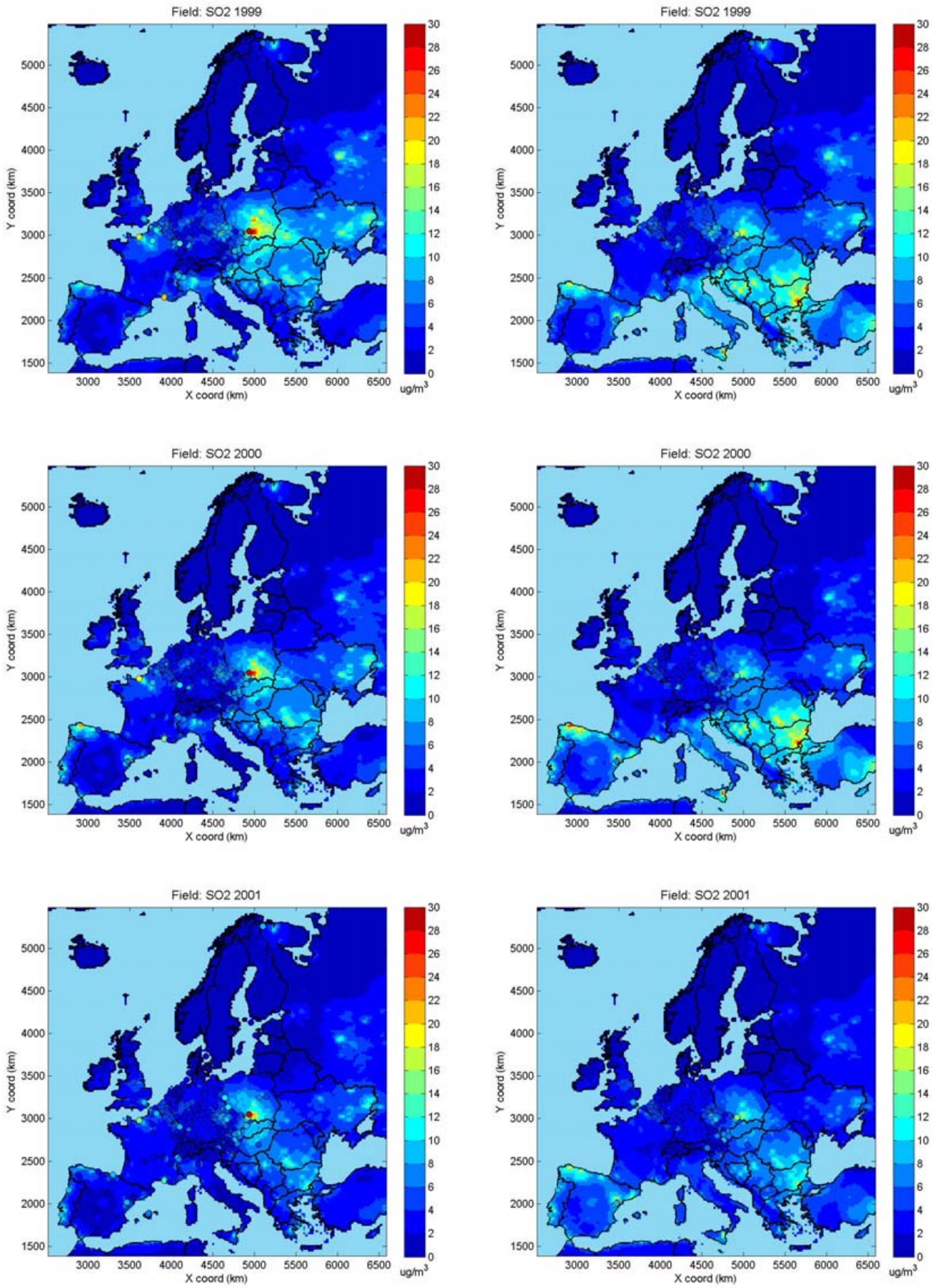




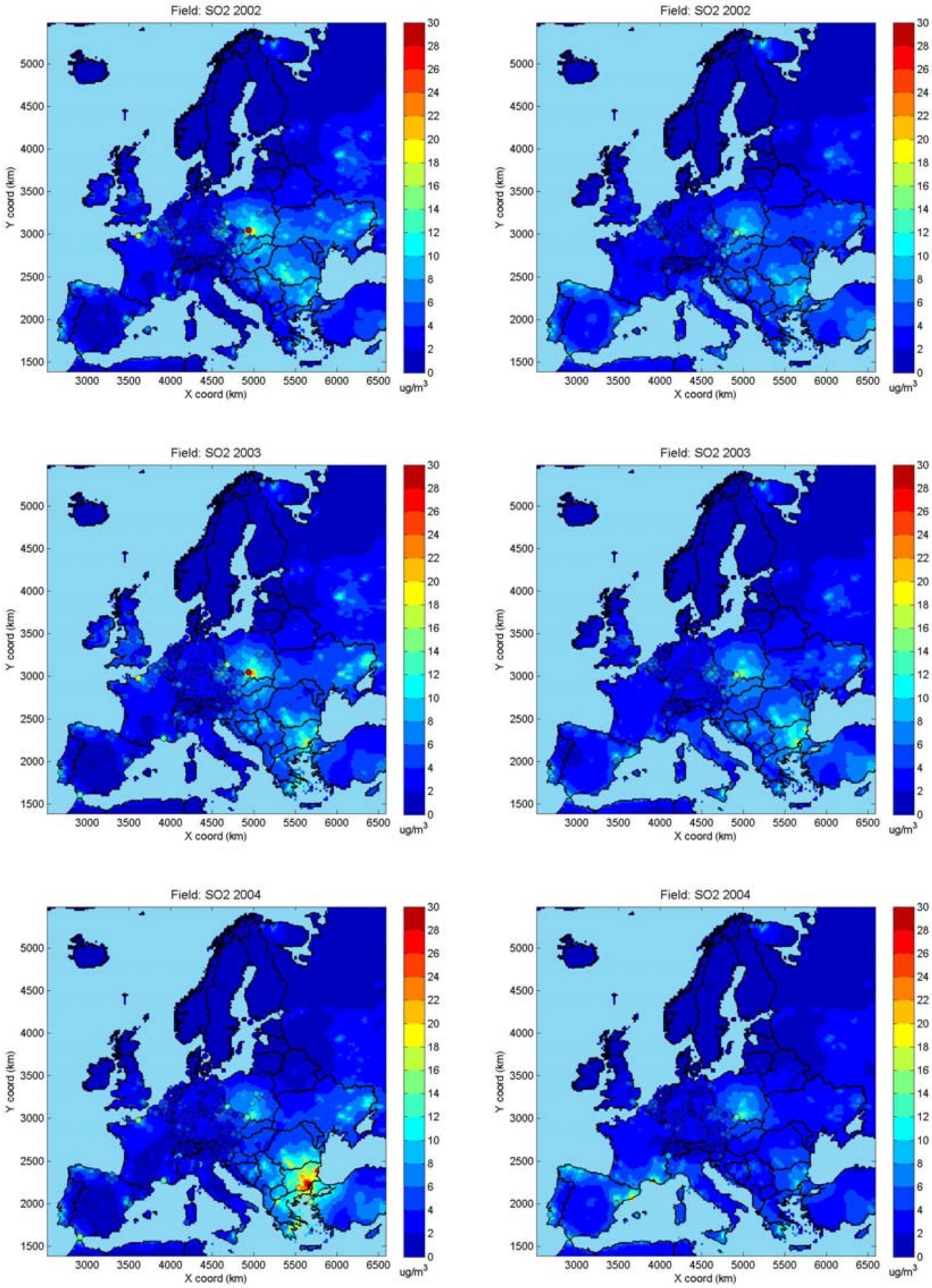




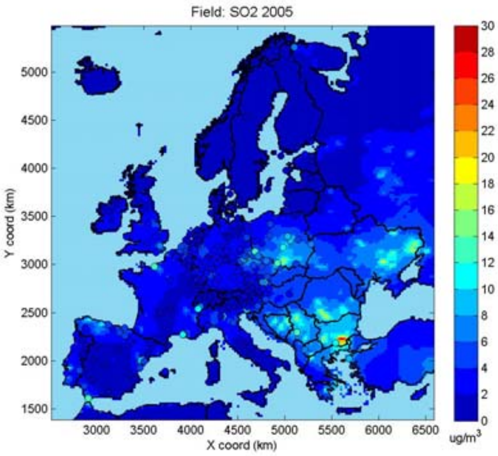
All stations | Selected stations >= 8 years



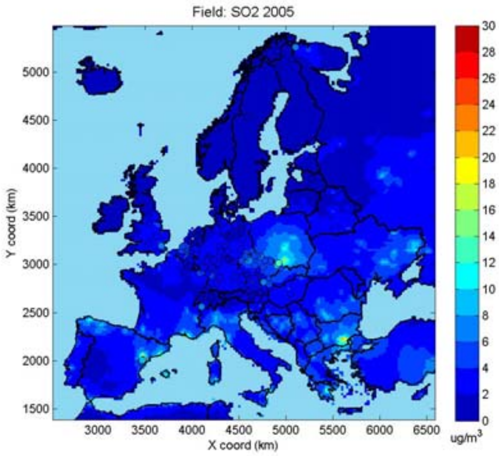
All stations | Selected stations >= 8 years



All stations



Selected stations >= 8 years



Annex III: Tables of monitoring data availability

Country	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
FR	0	0	0	242	284	313	332	362	365	368
DE	273	274	267	248	281	280	284	261	248	237
ES	0	16	17	30	35	53	71	81	91	114
IT	0	0	0	50	39	57	59	72	91	105
AT	80	84	78	87	91	93	93	91	83	90
GB	37	53	63	64	62	63	68	69	74	75
CZ	21	26	28	29	29	30	29	35	39	53
PL	0	13	17	18	21	20	20	20	27	45
PT	0	2	1	6	7	11	14	20	28	31
NL	13	28	29	29	28	29	26	26	27	29
BE	14	17	18	18	22	22	23	24	25	27
SK	0	12	11	0	16	11	16	19	18	21
CH	19	19	18	18	18	17	17	16	17	17
FI	9	8	10	11	11	11	10	14	14	14
HU	0	5	1	1	1	1	1	3	6	14
SE	0	0	2	6	7	7	7	7	12	12
GR	0	2	0	1	3	12	11	11	9	9
NO	0	0	10	10	10	10	10	9	9	9
SI	0	3	4	4	4	3	7	7	8	8
BG	0	0	2	3	3	2	0	3	6	8
DK	0	2	7	6	0	6	6	6	6	7
RO	0	0	0	0	0	0	0	0	7	7
IE	0	0	0	5	5	5	6	6	5	6
LV	0	1	1	2	7	3	2	4	4	4
EE	0	1	1	2	3	4	4	4	4	4
LT	0	4	1	1	3	2	3	3	4	3
IS	0	0	0	0	0	0	0	1	1	1
RS	0	0	0	0	0	0	0	0	1	1
CY	0	0	0	0	0	0	0	1	1	1
MK	0	0	0	0	0	0	0	0	1	1
MT	0	0	0	0	0	0	1	1	1	0
LI	0	0	0	0	0	0	0	0	1	0

Ozone: Total of all rural, suburban and urban stations available for the calculation of AOT40 crops with an hourly average coverage > 75%.

Country	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
FR	0	0	0	216	236	246	242	235	209	207
DE	326	301	289	274	292	279	221	197	184	178
ES	0	14	16	27	29	56	55	68	67	111
PL	0	10	11	23	30	44	41	32	40	94
IT	0	0	32	26	27	35	31	48	58	90
AT	67	87	83	90	95	89	83	77	72	76
CZ	46	46	47	47	46	45	45	47	49	70
GB	30	37	48	49	49	59	58	57	57	59
NL	12	34	34	32	32	32	27	28	26	44
PT	0	5	5	8	8	13	14	20	27	35
BE	33	36	35	33	36	37	37	38	33	34
SK	3	18	6	5	5	5	5	18	18	18
HU	0	7	1	1	1	1	1	5	6	16
RO	0	0	0	13	14	13	16	15	18	16
SE	0	6	8	7	8	7	8	8	8	14
SI	0	1	2	0	2	1	6	6	6	13
RS	0	0	0	0	0	0	0	5	12	13
FI	1	2	8	7	7	7	7	5	6	12
NO	0	0	8	8	8	8	8	7	7	12
CH	15	15	14	14	11	11	10	8	10	10
BG	0	0	3	4	0	4	0	4	6	10
MK	0	11	10	4	7	6	8	9	7	9
LV	0	2	2	4	6	4	3	5	6	8
EE	0	1	2	2	3	4	4	4	4	4
LT	0	0	1	1	1	1	1	3	3	4
IE	0	0	0	0	0	0	1	2	3	3
DK	0	1	3	3	1	3	2	5	2	3
GR	0	2	0	2	1	8	6	7	3	2
BA	0	0	0	0	0	0	1	1	1	0
LV	0	0	0	0	0	0	0	1	1	0
LI	0	0	0	0	0	0	0	0	1	0

SO_2: Total of all rural, suburban and urban stations available for annual mean concentrations with a daily mean coverage > 75%.

Country	1996	1997	1998	1999	2000	2001	2002	2003	2004	2005
FR	0	0	0	0	0	108	182	207	213	222
DE	0	0	44	16	76	166	166	223	228	217
PL	0	4	3	13	18	24	28	27	54	98
CZ	46	46	47	47	46	45	41	43	44	82
ES	0	2	1	9	12	21	38	50	53	72
IT	0	0	0	0	2	12	12	33	45	54
AT	0	0	0	0	3	16	33	40	49	49
GB	18	31	38	40	41	50	46	45	49	46
PT	0	0	0	3	3	7	9	15	19	28
BE	1	4	4	6	10	13	13	17	23	23
NL	14	14	14	14	13	14	13	14	21	23
SK	0	0	0	2	2	3	5	18	18	18
CH	0	6	12	11	13	12	13	14	14	14
BG	0	0	1	0	1	1	0	5	4	13
HU	0	0	0	0	0	0	0	2	4	10
NO	1	0	0	0	0	1	1	2	4	9
SE	0	0	3	2	4	3	3	5	7	8
SI	0	0	0	0	0	1	5	5	5	7
FI	0	0	0	1	2	6	7	8	7	6
DK	0	0	0	0	0	1	4	3	2	5
RO	0	0	0	0	0	0	0	3	4	4
LT	0	0	0	1	0	0	0	1	2	3
IE	0	0	0	1	0	2	1	2	6	3
GR	0	0	0	0	0	4	4	5	5	3
EE	0	0	0	0	0	1	1	1	1	1
IS	0	0	0	0	0	0	0	1	1	1
RS	0	0	0	0	0	0	0	0	1	1
CY	0	0	0	0	0	0	0	1	1	1
MK	0	0	0	0	0	0	0	0	0	1
LV	0	0	0	0	0	0	0	1	1	0
LI	0	0	0	0	0	0	0	0	1	0

PM_{10}: Total of all rural, suburban and urban stations available for annual mean and percentile concentrations with a daily mean coverage > 75%.