Sources of uncertainty

and their

assessment in spatial mapping

ETC/ACC Technical Paper 2008/20

April 2009

Bruce Denby, Frank De Leeuw,

Peter de Smet, Jan Horálek

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
PBL UBA-D UBA-V NILU AEAT AUTh CHMI MET.NO ÖKO TNO REC

Front page picture:

Google Earth screenshot of the village Uncertain, Texas, USA

Author affiliation:

Bruce Denby: Norwegian Institute for Air Research (NILU), Norway

Frank De Leeuw, Peter de Smet: The Netherlands Environmental Assessment Agency (PBL), The Netherlands

Jan Horálek: Czech Hydrometeorological Institute (CHMI), Praha

DISCLAIMER

ETC/ACC Technical paper 2008/20
European Topic Centre on Air and Climate Change
PO Box 303
3720 AH Bilthoven
The Netherlands
Phone +31 30 2743550
Fax +31 30 2744433
Email etcacc@mnp.nl
Website http://air-climate.eionet.europa.eu/

Contents

1 Introduction

The European Environmental Agency (EEA) and the European Topic Centre for Air Quality and Climate Change (ATC/ACC) have in the past few years developed methodologies for the spatial assessment of air quality on a European wide basis (Horálek et al., 2005; 2007; 2008, De Smet et al., 2009 and Denby et al. 2005; 2008a). 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.

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 has been assessed for the years 2003 – 2005. These mapping methods include the use of monitoring data (provided by the AirBase and EMEP databases), additionally supported by 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 data are combined, firstly by using a multiple linear regression of all the data with the observations to provide a baseline spatially resolved map, and then secondly ordinary kriging is applied to the resulting residuals, differences between the baseline map and the observations, to obtain the final maps.

Provided with these maps are also maps of uncertainty. To date these maps are based on the residual kriging variance, which is itself dependent on the fitted variogram parameters of nugget, sill and range. It has been recognised that the uncertainty mapping is limited in its extent and does not explicitly include a number of other sources of the uncertainty. This report outlines and to some extent deals with other possible sources of uncertainty in the air quality maps.

1.1 Aim

To define and improve the uncertainty assessment carried out in the mapping activities of the ETC/ACC. This will also hopefully lead to improvement in the assessment itself.

1.2 Scope

Initially this document will list all possible relevant sources of uncertainty (Chapter 2) and the limitations of the methodologies used to assess these (Chapter 2 and 3). It will then propose a number of possible improvements in the methodology needed to assess these sources of uncertainty and methods for including these in the uncertainty mapping methodology (Chapter3). Conclusions will then be made and recommendations for implementation and further work will be given (Chapter4).

2 Overview of the sources of uncertainty

The following sections list and comment on the various sources of uncertainty related to the data available for the interpolation as well as the mapping methodology itself. This is intended to be as wide and comprehensive as possible.

2.1 Monitoring uncertainty

Examples of the monitoring uncertainties include:

- Errors in instrumental calibration, leading to bias
- Differences in monitoring methods (e.g. automatic or gravimetric or other differences)
- Innate variation in analysis and reporting methods (e.g. PM_{10} correction factor, level of quality control)
- Differences in instrumental placing and set up at stations
- Human error in reporting and analysis (e.g. lack of quality control)
- Stochastic errors related to hourly and daily means and local fluctuations (this is also connected to the spatial representativeness)
- Errors in the aggregated data (e.g. annual means and percentiles, respectively the x^{th} highest values), caused by the missing primary data (e.g. under- or over-estimation of PM_{10} 36^{th} highest daily mean in the case of missing daily values)
- Uncertainty in metadata for stations, e.g. station position and station classification, resulting in misleading information concerning the spatial representativeness of the station. See section 2.3
- Differences in the time coverage (e.g. measurements from automatic stations in AirBase are aggregated from 0 to 24 UTC, while the measurements from manual stations in some countries are measured from 6 to 6 UTC).

2.2 Modelling uncertainty

Model uncertainty deals with the process description and the input data as follows:

- Model formulation and parameterisation. Missing processes and approximations within the model that do not take into account all the real processes and effects. These can be
    - Chemical schemes, including rate constants and unaccounted reactions and process descriptions in both gas and aerosol phases
    - Transport and dispersion (e.g. boundary layer description and vertical exchange)
    - Surface/air interaction and deposition rates
    - Sub-grid effects, higher order chemical processes associated with non-homogenous concentration distributions
- Finite numerical scheme. Approximations associated with grid sizes (especially mean grid concentrations) and time steps
- Emissions including:
    - Missing sources
    - Emission rates
    - Emission timing

    - Horizontal and vertical position of emissions, including stack heights and plume rise models
    - Direct fraction NO_2 in NO_x
    - VOC speciation and reactivity
    - Size distribution of primary PM
- Meteorological input data, including uncertainties in grid mean values and spatial representativeness
- Land use data
- Boundary conditions

2.3 Spatial representativeness uncertainty

This is a general problem concerning modelling, monitoring and other supplementary data and relates to the spatial region that the information is intended to represent.

- Representativeness of a measurement value for a spatial region (i.e. area for which the monitoring data is representative)
- Representativeness of a model concentration for any measurement within a grid (i.e. associated sub-grid variability of a model, both temporal and spatial)
- Representativeness of supplementary spatial data

2.4 Statistical interpolation uncertainty

The interpolation methodology currently employed, i.e. residual kriging after multiple linear regression, has a number of sources of uncertainty in the individual steps and data used. The following lists these.

2.4.1 Regression

The regression provides two types of information for the interpolation. This includes:

1. A spatial baseline for the residual kriging. Use of regression for this purpose is intended to make the residual kriging more stationary (see next section) or spatially homogenous.
2. A higher resolution structure, not obtainable through the kriging, is included in the maps as a result of the regression. Kriging uses spatial ranges of the order of 50 km and up whilst the regression uses information in the range of 10 - 50 km resolution.

In regard to the first of these, the uncertainty associated there is most relevant to the assumption of spatial homogeneity. If it is not the case then the assumptions in kriging are not valid and the effect of this on the interpolation and its derived uncertainty needs to be assessed.

In regard to the second, this aspect is most clearly seen in the regression with altitude, which is included in the regression at high resolution, 10 km. There is an associated error in any of these regression parameters and this also needs to be dealt with. Currently this baseline error is not being taken into account in the residual kriging methodology.

2.4.2 Residual kriging methodology

Ordinary kriging, when used within the framework in which it was developed, provides good estimates, based on spatial statistics, of the uncertainty. These estimates, of the kriging variance, are currently used to provide uncertainty estimates of the air quality maps.

There are, however, a certain number of assumptions involved in the ordinary kriging used for the interpolation. When these are not met then the statistical basis fails, or is compromised, in the interpolation. This will introduce methodological uncertainty into the interpolation that cannot be ignored. These include:

- Stationarity (i.e. that the region is spatially homogenous from a statistical view point)
- Normally distributed data (not an absolute necessity but the methodology is optimal for this)
- Isotropy (if anisotropic variogram is not taken into account)
- Validity of the variance model (e.g. spherical model) and the adequacy of the fit made to this

When these conditions are not met then the uncertainty estimates may be poor. All of the above may not be valid for air quality assessment on a European scale.

In addition the ordinary kriging carried out is point kriging. Since we are generally interested in the mean of an area, block kriging should perhaps have been used. This is generally not important when applying variograms that have significant nugget values. However, if nugget values are low and the grid point of interpolation is close to a monitoring site then this may lead, in some instances, to large differences between the area averaged and point interpolation.

Another point, usually not dealt with in the statistical analysis, is the physical nature of the problem. Examples of this occur when natural barriers, e.g. mountain ranges, physically separate two monitoring stations from each other so that they are in actual fact completely independent of each other. In the current methodology this will be, to a small degree, included in the spatially variance but not explicitly resolved on the map itself.

2.4.3 Supplementary data

The supplementary data used for the multiple linear regression also has a level of uncertainty. When meteorological data is used then this has a similar source of uncertainty to the meteorological data provided for the model calculations. The spatial representativeness of the data is also important, e.g. the meaning of a mean altitude for a 10 km grid.

2.5 Uncertainty in the urban maps and the combination with rural maps

The methodology by which urban maps are created and how they are combined with the rural maps is a separate study. Urban maps are also made using multiple linear regression of the monitoring data with supplementary data. These supplementary data also include the EMEP model, altitude and meteorological data, to produce the baseline map before residual kriging is applied. The urban interpolation is carried out independent of any rural station and is interpolated to cities, sometimes far away, without stations. Uncertainty is currently provided by the ordinary kriging variance. There are a number of questions concerning this methodology and its uncertainty that need to be addressed:

- The urban and rural maps are combined using a weighting according to population density. No consideration is given to their respective uncertainties when doing this weighting. Including their uncertainty when combining the urban and rural maps may be a suitable alternative.
- The weighting according to population density used as a merging methodology itself contains uncertainties that need to be explored further. Other possible methodologies, other than merging, should be examined. E.g. Janssen et al. (2008a; 2008b)
- Possible supplementary data not currently being used for the urban interpolation needs also to be considered.
- The classification of the urban/rural stations should be verified, e.g. in comparison with the relevant population density.

2.6 Additive versus multiplicative errors (normal or log-normal distributions)

In almost all of the above methodologies there is an innate acceptance that the errors being dealt with are in some way normally distributed. This is very often the case when dealing with additive errors. However when the errors are multiplicative then the distribution is more likely to be log-normal in character. If this is the case, as it usually seems to be for concentration data, then the interpolation should be carried out after logarithmic transformation and then back-transformed appropriately.

In the work carried out in APMoSPHERE (Briggs et al., 2005) logarithmic distributions were assumed. In Denby et al. (2008a) logarithmic transformations were also applied to daily mean PM_{10} data. However, it is not necessarily the case that integrated indicators such as AOT40, SOMO35 or indeed percentiles are log-normally distributed. The advantages of using the logarithmic transformation are:

- If the data is log-normally distributed then the transformation leads to a normal distribution for the interpolation
- Avoids unwanted negative values
- Leads to relative uncertainties, rather than absolute, in the interpolated fields. This is intuitively a more correct description of the uncertainty when dealing with a wide range of concentrations.

2.7 Uncertainty in exposure and effects calculations

The concentration maps are combined with population density to provide exposure estimates. The uncertainties will be dependent on the exposure and concentration-response method applied, e.g. uncertainty in total population weighted exposure will be different to spatially distributed exposure calculations. The following aspects should be taken into account:

- Uncertainty in the population density maps
- Uncertainties associated with the application of exposure to static populations, i.e. home residents used rather than mobile commuting populations
- Uncertainties resulting from differences in activities and micro-environments including indoor and outdoor concentrations and activities.
- Uncertainty in the population exposure due to sub-grid variability (both in concentrations and population) and/or uncertainty in the concentration values
- Uncertainties when spatially aggregating exposure calculations, e.g. spatial correlation needs to be taken into account when aggregating mean exposures or concentrations spatially
- Uncertainties in demographic and health data (e.g. age distribution, incidence baseline) including its spatial variability
- Uncertainties in the concentration-response models
- In cases where a no-effect level is assumed (e.g. the non-anthropogenic background) the results of the health impact assessment will strongly depend on value and spatial distribution of the assumed baseline.

Currently some of these aspects are being dealt with in an alternative health related ETC/ACC task, looking at the sensitivity of the health outcomes based on a number of these aspects.

3 Methods for dealing with the uncertainties

The following section indicates possible methods for analysing and mapping the uncertainty sources provided in the list in section 2. An assessment is given on the importance of the sources, whether they need to be given extra attention and how these contributions to the uncertainty can be best analysed, as well as reduced.

3.1 Monitoring uncertainty

3.1.1 How to reduce the monitoring uncertainty?

The monitoring uncertainty, as described in section 2.1, can be seen to be the sum of a range of uncertainties. There is little that can be done for the mapping activities in regard to the implicit instrumental error, or the errors associated with station positioning or subtleties in monitoring methodologies. However, there remain gross errors, such as quality control of the data, missing data and misrepresentation of the station type that can still be dealt with before, or during, the implementation of the monitoring data in the interpolation. This may be achieved through the following types of methods:

- Assessment of AirBase data at higher temporal resolution than just annual statistics. Removal of outliers, visual inspections, e.g. Mol and van Hooydonk (2005) and Solberg et al. (2008)
- The use of imputation methods, e.g. Plaia and Bond'i (2006) and Hien et al. (2007). To indicate when measured concentrations deviate from there expected values.
- Adopt methodologies for correcting (scaling) to compensate for the PM correction factors (e.g. Van de Kassteele et al., 2006).
- Assessment and possible removal of stations that deviate significantly from the predicted concentrations, using cross-validation or other techniques.

3.1.2 How to define the monitoring uncertainty?

The air quality directive (EC, 2008) provides guidelines for quality control and assurance of monitoring data. According to the requirements there, uncertainties in monitoring data should be no more than 15% - 25%, dependent on the compound. These can be considered to be the upper limits of the direct monitoring uncertainty, excluding the range of factors connected to placement and user error.

Note: Uncertainties defined in the directive are considered to describe the 95% confidence level. For a normally distributed uncertainty this is equivalent to 2σ, where σ is the standard deviation.

Standard instrumental error requirements are probably, or at least should be, lower than this upper limit. Individual studies, that take place in most countries, would need to be assessed. E.g. the instrumental error for the gravimetric instruments in the Czech Republic is stated to be ± 8% and routine comparison between radiometric and reference gravimetric measurements indicates uncertainties of 9 – 15%. Equivalent tests in the UK (Harrison, 2006) also provide details of the uncertainty of the monitoring methods.

3.1.3 How to include monitoring uncertainty into the uncertainty maps?

The interpolation methods used, kriging, do not directly consider the monitoring uncertainties used. However they do implicitly include these uncertainties in the interpolation. As an example, consider the nugget value of the variogram, which describes the variance of the spatial data in close proximity to the measurements. If a number of stations are placed close together (e.g. metres) and give identical measurement values then the variogram nugget will be 0. If the measurement values deviate from each other in some way then the nugget value would be representative of the variance of this deviation. This deviation may be the result of spatial variance but it may also be due to instrumental variance. Both will

be explicitly included in the nugget, and thus variogram assessment, and it is the variogram that is used for the final interpolation uncertainty.

The answer is thus that monitoring uncertainty, or correctly the variability between monitors, is implicitly included in the current interpolation uncertainty. It will not address, however, any general bias to all the monitoring data, such as incorrect PM correction factors for all stations.

3.2 Model uncertainty

3.2.1 How to reduce the model uncertainty?

It is not possible, for the mapping activities, to reduce model uncertainties from the EMEP model. These are provided as is. The only possible method available is to use the output from an ensemble of models. Van Loon et al. (2007) showed that the use of model ensembles, in the EuroDelta project, can lead to improved model performance. Also Pagowski et al. (2006) have shown that the use of regression to combine the models can be used to improve forecasts of ozone. Currently such ensembles should be available through activities related to GEMS where ensemble forecasts using 7 different models are made every day (http://gems.ecmwf.int/d/products/raq/). Though this is a possible path for improving the model assessment fields, tests carried out in Denby et al. (2008a) using the multiple linear regression of LOTOS-EUROS and EMEP did not yield improved results.

3.2.2 How to determine the model uncertainty?

Model errors are usually assessed through comparison with observations. Whenever such comparisons are made it is necessary to match the spatial representativeness (both horizontal and vertical) of the modelling and monitoring data. In addition any uncertainties in the monitoring data itself will manifest in a model error. Such assessments can be carried out of the model, as described in Denby et al. (2008b), to provide an overall estimate of the model uncertainty. In that case the model uncertainty was described using just two parameters, as a function of the absolute uncertainty σ_A and the relative uncertainty σ_R in the following form

$$\sigma_M(x,y) = \sqrt{\sigma_A^2 + \sigma_R^2 M(x,y)^2} \qquad (1)$$

where *M(x,y)* is the model concentration. This provides a general structure for calculating uncertainty as a function of model concentrations. The estimated values of σ_R and σ_A may be calculated by fitting equation 1 to scatter plots representing the standard deviation (SD) as a function of model concentration (figure 1). In this case the residual kriging result is used for the estimate but the principle is the same.

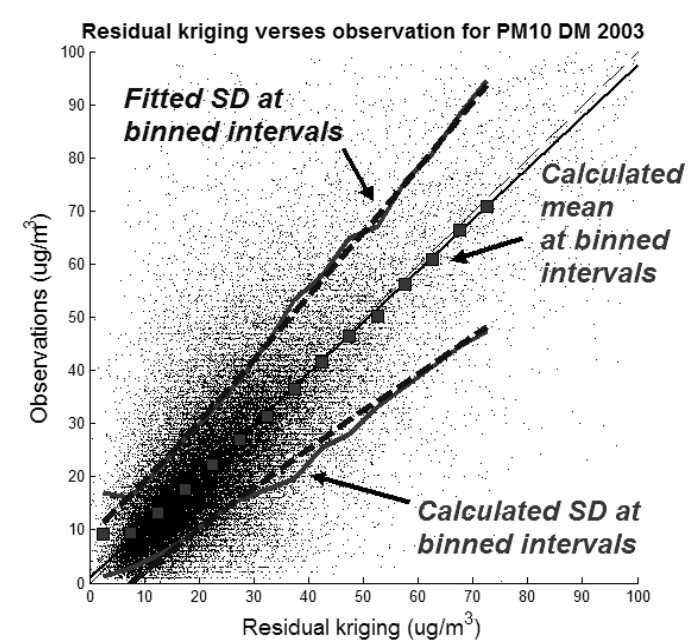


Figure 1. Scatter plot of assimilated model results (Residual kriging using the Unified EMEP model) and observations for daily mean PM_{10} in Europe. The model results are binned in 10 $\mu g.m^{-3}$ bins and at these binned intervals the SD is calculated (solid thick grey line) as well as the mean binned values (squares). Equation 1 is fitted to the binned SD (dashed thick black line) to obtain σ_R and σ_A. The resulting values found are σ_A=8.0 $\mu g.m^{-3}$ and σ_R= 0.30. Data based on the statistical interpolation study from Denby et al. (2008a).

Such a method is a simplification, but on a general statistical level it does provide one estimate of the model uncertainty. Even simpler methods, such calculating the root mean square error (RMSE) may also be used, but this does not provide a spatial element to the uncertainty.

Another method for estimating uncertainty is to carry out ensemble runs by perturbing model parameters and input data (within their uncertainty) to obtain estimates of the intrinsic model uncertainty (e.g. Rodriguez, 2007). This can be carried out using various Monte Carlo methods where the Probability Density Functions (PDFs) of the model parameters are used as prior distributions for the ensembles. This will give a spread of model results that can be mapped. It will not provide information on bias unless direct comparisons are made with observations. Unfortunately such ensembles must be carried out by the modelling groups.

Similar to Monte Carlo simulations within a model, simulations using an ensemble of models, as previously mentioned (e.g. Van Loon et al., 2007), can provide estimates of uncertainty, though this has yet to be carried out in such studies.

3.2.3 How to include model uncertainty in the uncertainty maps?

Model fields, or the regression of the models fields, are currently used in the residual kriging interpolation as the baseline for determining the residual fields. No use is made of the model uncertainty in this methodology. In Denby et al. (2008b) modelling uncertainty was included, based on the uncertainty assessment described in equation 1, by using a weighted combination of modelled and kriged fields. In that study such fields never showed improved cross-validation statistics when compared to the residual kriging fields and so the methodology was not further elaborated. However, the use of the kriged fields based on observations only, was probably the limiting factor in that methodology. As an alternative we suggest the following.

Given an observational field *O(x,y)*, determined for example by kriging, and a model field *M(x,y)* then their weighted combination, using the spatially dependent weighting function λ will produce a combined, or assimilated, field *C(x,y)*.

$$C(x,y) = \lambda O(x,y) + (1-\lambda)M(x,y) \quad (2)$$

Various methods may exist for producing the optimal weight but it is normal to apply Bayesian methods to this, such that the weighting is dependent on the uncertainties of both the model and the observations, where less weight is given to the most uncertain element. In Denby et al. (2008c) this is described and was achieved by the use of the kriging variance, to describe the observational uncertainty, and equation 1, to describe the model uncertainty. Since there were few data available on a daily basis to provide information on the model uncertainty, the relative and absolute uncertainties in equation 1 were defined pragmatically by finding the values that gave the minimum cross-validation RMSE.

As a further development of this method, and one that can be readily implemented, equation 2 can be rewritten using the definition of the residual R given by

$$R(x,y) = O(x,y) - M(x,y) \quad (3)$$

so that the assimilated field becomes the sum of the model field and the weighted residual field by combining equations 2 and 3.

$$C(x,y) = \lambda R(x,y) + M(x,y) \quad (4)$$

Since the weighting parameter λ is always less than or equal to 1 the assimilated field will always lie between the sum of the residual kriging and model fields and the model field. It is expected that λ will generally be close to 1 except in areas far from observations where the large observational uncertainty from the kriging will lead to a lower weighting value and there the model results can be the most dominant. However, the model uncertainties in areas remote from observations may also be significant and unvalidated. Indeed when regression is used to create the model fields applied in equation 4, there is an unknown and unsampled element to this.

The weighting parameter λ is given by:

$$\lambda = \frac{\sigma_M^2}{\sigma_R^2 + \sigma_M^2} \quad (5)$$

The final uncertainty of the assimilated field is given as:

$$\sigma_C^2 = \frac{\sigma_R^2 \sigma_M^2}{\sigma_R^2 + \sigma_M^2} \quad (6)$$

In this way the model uncertainty will be included in the assimilation.

3.3 Spatial representativeness uncertainty

3.3.1 How to determine the spatial representativeness of monitoring stations?

The question of representativeness has been discussed in several reports. In Larssen et al. (1999) this was discussed in relation to EUROAIRNET. In that report it was noted that a station's area of representativeness (AoR) varies significantly, from traffic stations (a few metres) to remote rural stations (100's of kilometres). The AoR is defined there to be the area in which the actual pollutant concentration varies by less than a predefined range. A value of 10 - 20% is considered to be indicative but of course a relative variability is not sufficient for low pollutant levels, and this will also require an absolute concentration level, e.g. 5 ppb. It is recommended that such ranges be linked to the quality objectives of the measurements. Larssen et al. (1999) note the following:

'"*The area of representativeness is not easily determined. It requires either extensive monitoring at several adjacent sites covering an area around the station, or rather detailed dispersion model calculations based upon detailed emission inventories, both for the area in question and the larger surrounding area.*"

In the GEOMON project (http://geomon.empa.ch) a report was recently written (Henne et al., 2008) to help assess the representativeness of 34 EMEP super sites. Two methodologies were applied. The first involved the use of high resolution population, land use and topographical data. These data were used as a proxy to indicate the variability of what was considered to be relevant information for the representativeness of the stations. The values of these data over a 10 km and a 50 km radius, surrounding each station, were compared. Using this, an assessment of relative representativeness was given. The second method applied in that report involves the use of Lagrangian particle models to map the 'footprint' of contributing areas to stations.

In neither of these reports does a methodology materialise for assessing the AoR of individual stations in an efficient or effective way. This aspect of the uncertainty remains problematic and will continue to reply, without intensive studies, on estimates of representativeness based on expert knowledge of the individual sites. However, it should be noted that the question of spatial representativeness is addressed in the kriging methodology. The nugget value, as previously mentioned, contains variance related to spatial representativeness as well as monitoring uncertainty.

A number of other suggestions for obtaining information concerning representativeness include the following. These would all require research efforts of sizable proportions.

1. Carry out monitoring campaigns that can actually measure the spatial variability. Passive sampling techniques are probably the best methods available for estimating the spatial representativeness by taking many samples over a large area. However, the monitoring uncertainty of these instruments is in the range of 20% so even this type of campaign will only provide information concerning AoR at a level greater than the monitoring uncertainty.
2. Satellite data may be used, in the future, to provide estimates of the spatial variability of various pollutants. At the moment the horizontal resolution and uncertainty in the satellite measurements themselves limits their use in such assessments.
3. The assessment of the temporal correlation between stations at varying distances also provides information concerning their representativeness in relation to each other. This may be most helpful in verifying the station classifications.
4. Visual, or even automated, inspection of satellite maps (e.g. through Google Earth) may be applied generically on all stations provided in AirBASE to provide a 'second' opinion on station classification.
5. Other land use data, similar to that used in GEOmon, of the surrounding area may also be applied to assess the station classification. Such data may include population data, high resolution emission databases (e.g. 5 km emission database developed within GEMS http://gems.ecmwf.int/do/get/Themes/RAQ), proximity to major roads, proximity to major industries.

3.3.2 How to determine the sub-grid variability of the model and other supplementary data?

The sub-grid variability of the chemical transport model is a current area of research. One of the most straight forward methods for assessing this is to continuously increase the resolution of the model in steps seeing how the variability increases with increasing resolution. Such a study has been carried out by Ching et al. (2006) in an urban area. Application of this method implies that both meteorological and emission data is concurrently available at all resolutions. Such studies reveal not only the variability but also biases introduced in the models that are dependent on the resolution applied.

In regard to the sub-grid variability of the supplementary data used in the interpolations, this is easily assessed using ArcGIS to spatially aggregate the data. This can be applied to altitude and population data, which are available on a finer resolution than are currently used in the mapping, and this would also apply to land use or emission data if these were to be used at any further period in time. Unfortunately meteorological data is not available at higher resolutions than the mapping is currently carried out. These data may include significant sub-grid variability, especially in regions of complex terrain.

3.3.3 How to include the spatial representativeness into the uncertainty maps?

As previously mentioned the spatial representativeness, along with other monitoring uncertainty aspects, lies inherently in the nugget value of the variogram. In the case of residual kriging this will also include sub-grid variability of the chemical transport model and the supplementary data used. However, since there is an assumption of stationarity in the kriging interpolation this spatial representativeness uncertainty is described by a single value. I.e. it does not vary in space. To some extend this can be addressed through log-normal kriging which considers the uncertainty to be relative to the concentrations, rather than an absolute concentration value.

3.4 Statistical interpolation uncertainty

3.4.1 Regression

As outlined in section 2.4 the multiple linear regression carried out using the EMEP model and the supplementary data has an uncertainty attached to it. This uncertainty is expressed in terms of the uncertainty of the regression coefficients and is the result of a limited dataset as well as the non-applicability of a linear regression model. Statistically there are different methods available for calculating the uncertainty in the regression coefficients. These include both analytical statistical assessments (e.g. Brown et al. 1998) and Monte Carlo methods providing uncertainty in the regression coefficients.

Even if these have been determined the question remains on how to implement these uncertainty estimates further in the regression model results. Of particular concern are the uncertainties related to extrapolated data, i.e. values that are determined from the regression that lay outside of the available data set. A good example of this is altitude where there are few, if no, stations with altitude over 1500m, whilst maps use altitudes over 3000 m. These values are of high uncertainty because it is not known if the linear regression model for altitude outside of the data range is applicable.

3.4.2 Residual kriging

The sources of uncertainty in the residual kriging were described in section 2.4.2. If all the assumptions on which kriging is based are fulfilled, as listed there, then the uncertainty of the kriging will be well represented by the kriging variance. The question of how to deal with uncertainties related to non-fulfilment of the assumptions is not easily answered. One way of at least assessing if the uncertainty is being well described, and this includes all uncertainties, is to assess the performance of the uncertainty estimates using the cross-validation prediction errors. For each prediction at a measurement site there is an associated uncertainty prediction. By aggregating these uncertainty predictions and the actual cross-validation errors a statistical distribution of the errors can be derived. If this is similar to the predicted uncertainty distribution then we can consider the uncertainty estimates to be valid.

On a more complex level this would require the construction of a cross-validation error distribution, and compare this to the predicted distribution function. On a simpler level the cross-validation RMSE can be used to represent the standard deviation of the cross-validation error distribution (this can be done because the kriging should be unbiased and so the mean of cross validation errors should be close to 0). For example, the maps shown in figure 3 had a cross-validation RMSE of 6.0 μgm^{-3}. This corresponds well with the predicted uncertainty close to the majority of monitoring sites in both the uncertainty maps.

3.4.2 Supplementary data

The level of uncertainty in altitude data, especially aggregated data, is assumed to be negligible. There is however a high level of uncertainty in the meteorological data used. Assessing this is a subject in itself and beyond the realms of this task. A range of projects and programmes, such as COST728, address this sort of question.

3.5 Uncertainty in the urban maps and the combination with rural maps

The uncertainty in the urban scale maps is probably larger than the rural maps, especially for PM_{10}, due to the inherent spatial variability in urban regions and due to the interpolation of urban measurements to cities without observations. In essence the urban scale maps are subject to the same types of uncertainties described earlier.

In order to produce 10 km resolution maps of Europe rural and urban maps are produced separately and merged together, to reduce the extended influence of urban stations in the interpolation. The current methodology used for merging the rural and urban maps under ETC/ACC (e.g. Horálek et al., 2007) uses the population density classification criterion of rural areas (< 100 inhbs.km^{-1}), urban areas (> 500 inhbs.km^{-1}), and a weighting function for the 'mixed' areas in between. Potential weak elements of the current merging methodology that could be improved have been examined in a parallel report (de Horálek et al., 2009). The elements considered there are:

- The level of correlation between the station type and the population density area type in the corresponding grid cell
- The merging function itself
- The grid resolution

The non-correspondence of the station type with the underlying grid cell has its cause in the map quality. E.g. a weak agreement of the measured and estimated values was found for the stations classified as urban (under EoI classification), which are in the rural or mixed areas according to the population density grid (in 10x10 km resolution).

An alternative methodology has been tested, that being a population density weighted linear regression using the same population density classes, but accounting for both rural and urban/suburban stations simultaneously in the regression, followed with the spatial interpolation of their common regression residuals. This way the population density based weighting function is applied more directly than in the presently used methodology. According to the first results the suggested method shows a reduced uncertainty in the urban areas, however, with a co-existing uncertainty increase at the rural areas. (A week point of this new proposed methodology seems to be the application of the same variogram for both rural and urban areas in the interpolation of the regression residuals.)

3.6 Additive versus multiplicative errors (normal or log-normal distributions)

In Section 2.6 a small description is provided concerning the use of normal or log-normal transformations of the data. The fact that the residual data has a log-normal distribution, as illustrated by figure 2, infers that that log-normal distributions should be used in the interpretation.

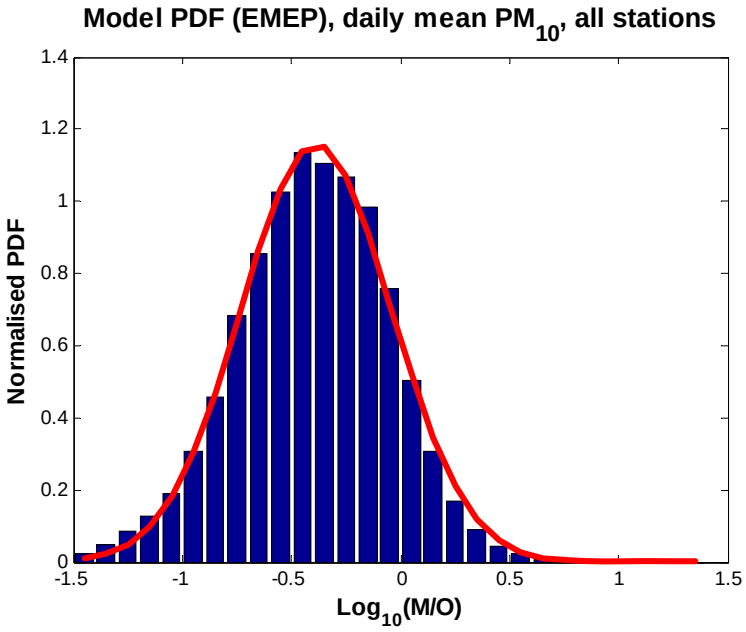
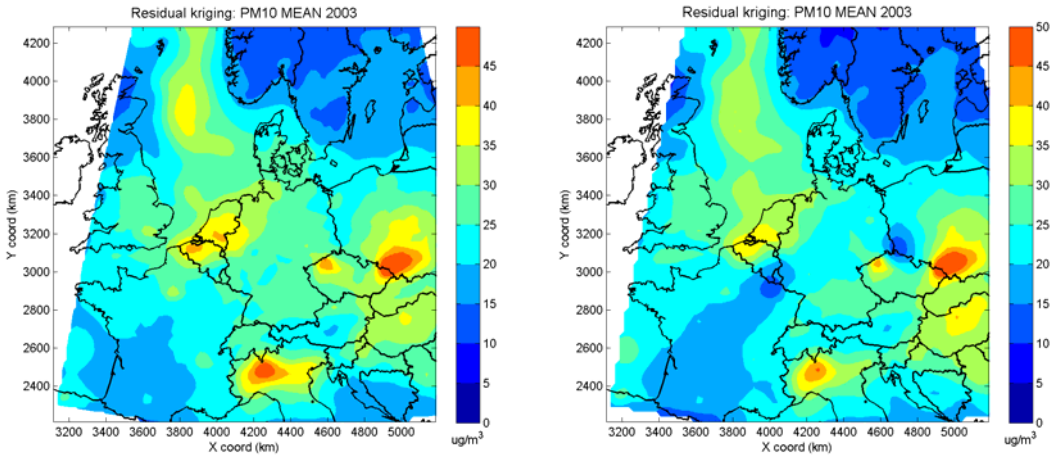
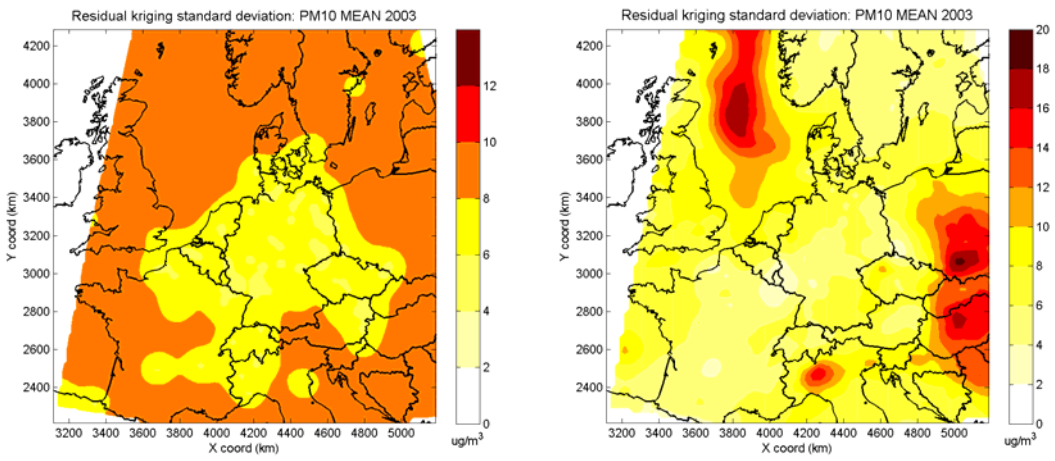


Figure 2. Frequency (probability) distribution of the EMEP model daily mean PM_{10} calculations for 2003 (Rural stations available from AirBase only). Shown is the log of the ratio of the modelled to observed concentrations $log_{10}(M/O)$, which is the same as log_{10} (M) – log_{10} (O). In other words the negative of the residual after log-normal transformation. The red line is a log-normal fit to the data. The data clearly has a log-normal distribution.

The use of the log-normal transformation will provide relative instead of absolute uncertainty maps, which can be quite different. An example of how different these are is provided in figure 3. These maps show the annual mean PM_{10} concentration fields and associated uncertainty, when using the residual kriging combined with regression of the EMEP model only, taken from Denby et al. (2008c). The maps on the left do not use the log transformation (not published in that paper) whilst the maps on the right do use the log-normal transformation (published). Though the concentration fields are very similar the uncertainty fields are not.





Normal (no transformation) Log-normal transformation

Figure 3. Top are the residual kriging fields, after interpolated after regression with the EMEP model., for annual mean PM_{10}. Bottom are the corresponding uncertainty fields based on the kriging variance. The left fields are not log-normally transformed whilst the right fields are. The log transformation occurs with the initial concentrations, observations and models, and the data are back transformed.

Cross-validation error assessments made for the mapping (e.g. Denby et al., 2008a) show a variable result in regard to the use of log-normal transformations for the pure kriging, i.e. kriging with monitoring data only. Sometimes the log-normal distribution had a higher cross-validation RMSE and sometimes a lower for the daily interpolation of PM_{10}. This is shown in figure 4 where a histogram of the relative RMSE difference between normal and log-normal pure kriging is shown. When normal kriging performs better for the cross-validation then the difference is positive, when log-normal performs better then the difference is negative.

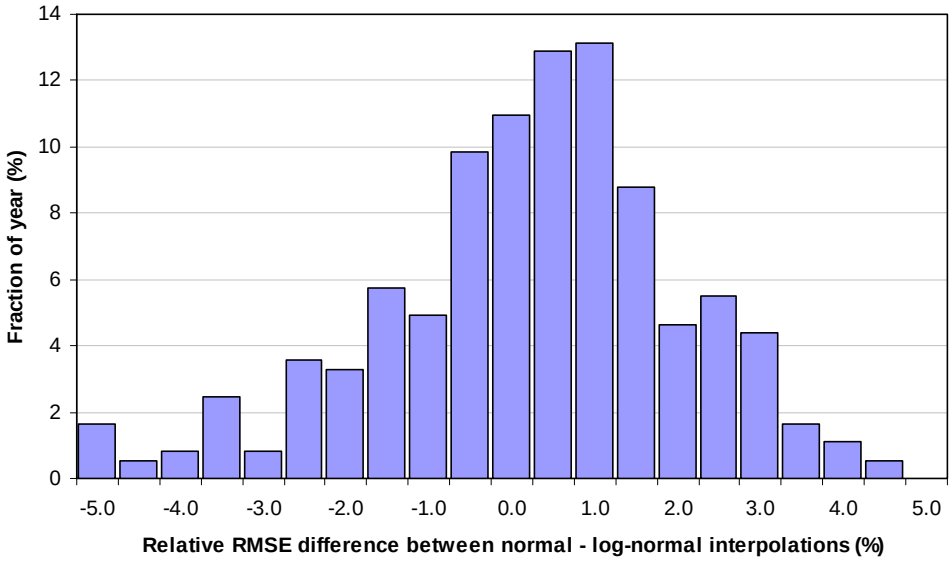


Figure 4. Histogram showing the fraction of the year (%) for which the normal interpolation performed better than the log-normal interpolation. Normal interpolation performed better when the relative RMSE difference is positive, log-normal interpolation performed better when the relative RMSE difference is negative. Data is for PM_{10} daily mean, based on daily mean interpolations for the year 2003.

3.7 Uncertainty in exposure and effects calculations

In section 2.7 a number of sources of uncertainty in regard to the health impacts are listed. These are further addressed here.

3.7.1 Population density and demographic data

The population density map used is given at a resolution of 100 x 100 m (Source EEA, pop01c00v3int, official version Aug. 2006; Owner: JRC). These data are based on the degree of urbanisation from Eurostat and the population census of the European communes 2001, mapped on the basis of CLC2000 land cover. As a result there is fairly high certainty in the census data, covering a larger area, but higher uncertainty in regard the disaggregation to 100 m resolution using land use data. The expected uncertainty in the 10 km aggregated population maps are then expected to be very low and these will not be significant for the health calculations. Uncertainty will increase with time if the population data is not regularly updated.

The uncertainty in demographic data, however, may be significantly larger. Indeed, this is provided on a country wide basis and disaggregation of these data will lead to high uncertainty. This is discussed further in section 3.7.3.

3.7.2 Accounting for sub-grid variability and covariance when calculating population exposure and population weighted average concentrations

The population weighted average concentrations are currently calculated for individual countries and Europe as a whole. The following formula is applied to calculate, for any given region, the population weighted average concentration:

$$C_{pop} = \frac{\sum_{i}^{n} C_i P_i}{\sum_{i}^{n} P_i} \qquad (7)$$

where i is a grid point, n the number of grids, C_i is the concentration in a grid and P_i the population in that grid. In addition to the population weighted average concentration an exposure distribution is also made, i.e. the number of people exposed between two defined levels. The population weighted average concentration is simply the mean of this distribution.

It is necessary, when determining the uncertainty of these estimates, to differentiate between two different aspects. These are

a) Sub-grid variability, and

b) Grid concentration uncertainty

The first is intrinsic and cannot be reduced by aggregation, the second is dependent on the mapping methodology and may be reduced by aggregation. Most of the uncertainty will lie, as described in section 3.7.1, with the concentration data, rather than with the 10 km population data. If the uncertainty of the concentration maps is provided (e.g. as residual kriging variance) then this value will indicate both the sub-grid variability and/or the uncertainty in the mean grid values.

Unfortunately it is not easy to distinguish between the two sources of variance, as discussed in previous sections. In this, and the following section we treat the two aspects separately. I.e. the variance is assumed to be due entirely to sub-grid variability (this section) or due to uncertainty in the grid concentration (section 3.7.3). For the case of sub-grid variability addressed in this section there will be a probability distribution that describes the probability that any person living within a grid will be exposed to a given concentration range. This probability distribution is defined by the mean and the standard deviation (variance), and will be dependent on the type of distribution assumed for the concentrations, i.e. normal or log-normal distributions.

By way of example, if at any grid point the annual mean concentration of PM_{10} is calculated to be 20 μg/m^3 and the standard deviation is calculated to be 10 μg/m^3 then for a normal distribution there is a 5% probability that a person living in that grid is exposed to an annual mean concentration above the limit value of 40 μg/m^3. (Using the same logic this also tells us also that there is a 5% chance that a person is exposed to negative concentrations, but this reflects the advantage of using log-normal distributions rather than normal).

In any case, use of the uncertainty maps interpreted as sub-grid variability, will provide a distribution of population exposure at each grid point, rather than just a single value. These distributions can then be summed to provide a more realistic distribution of population exposure than is currently done. Practically this is achieved by:

a) Defining discrete concentration levels of interest, e.g. by establishing concentration bins in the case of PM_{10} of 5 μg/m^3.

b) For each grid point the mean and standard deviation, either based on normal or log-normal distributions, is established. The distribution is then used to determine how many people, of the population within that grid, are exposed to each of the binned concentrations. This will involve integration of the distribution between the binned minimum and maximum values, e.g. 5 – 10 μg/m^3.

c) The bins are then summed over the area of interest, e.g. country or Europe, and the weighted concentration is then determined for each of the bins by dividing by the total population. The resulting distribution will show the number of people, fraction of the total population, exposed to that binned concentration level.

d) The total population weighted average concentration will then be calculated by finding the mean of the distribution.

The resulting distribution will be much broader than is currently calculated, since it includes a level of variability in the results. The average, however, will be the same if normal distributions are applied but will be to some degree different for log-normal distributions.

The above discussion assumes that there is no covariance between concentrations and population density. However, since many pollutant sources are related to population density this may be the case. To illustrate this take as an example a 10x10 km grid, with average concentration *C*, that is subdivided in *n* sub-grids (e.g. 100 1x1 km sub-grid cells). In each sub-grid cell *j* the concentration is given by $C + c_j$ where c_j is the difference with the averaged and $\Sigma\, c_j = 0$. In the large cell we have a population density *P* and in each sub-grid cell $P + p_j$ people are living, where $\Sigma\, p_j = 0$.

The population weighted concentration for the large cell, calculated using the sub-grid cells is given by:

$$\begin{aligned} C_{pop} &= \frac{1}{P}\sum\left(C + c_j\right)\left(P + p_j\right) \\ &= \frac{1}{P}\left(\sum CP + \sum Cp_j + \sum c_j P + \sum c_j p_j\right) \\ &= C + \frac{1}{P}\left(\sum c_j p_j\right) \\ &= C\left(1 + \frac{\sigma_P}{P}\frac{\sigma_C}{C}\rho_{PC}\right) \end{aligned} \qquad (8)$$

In the final equation above we have inserted σ to indicate the standard deviation of the concentration and population in the large cell and ρ_{PC} for the correlation between them in the sub-grid cells. For PM_{10} *C* and *P* are most likely positively correlated and as a result we will systematically underestimate the population weighted concentration when using a 10 x 10 km grid cell.

We can further estimate the effect of correlation on the population weighted concentration if we assume values for these parameters (or better still determine them). Given realistic relative standard deviations

of 50% for both concentration and population density, and a correlation of 0.2 between these two we arrive at an increased estimate of population weighted concentrations by a factor of 5%. This factor is not large and a correlation of 0.2 may be a high value.

As a final point, for population exposure there is an uncertainty introduced by population movements when calculating exposure at higher resolutions since people will move between grid cells when travelling form home to work. As cities (with relative high PM-concentrations) attract a lot of people (workers, tourist, shopping) from the more rural surroundings one may expect that for those commuters the actual exposure is higher than estimated for their home address. This may be as high as 100% (Barrett et al, 2008). For other population groups (children, elderly) the home address concentration might be more representative as they tend to travel less than commuters.

3.7.3 Accounting for uncertainty and correlation when aggregating spatial data

If the mapped concentration uncertainty is assumed to be only due to uncertainty in the methodology, rather than variability, then the uncertainty in the aggregation of data, such as the population weighted averages, needs to be addressed. This is not straight forward since there is a degree of spatial correlation involved that affect the calculation of the aggregated uncertainty.

For simplicity of demonstration let us assume we have, for each grid point, just one value for the concentration with an accompanying uncertainty, rather than a distribution. To determine the total, for all of Europe or for a given country, average concentration then we need to aggregate (sum) these values. Aggregation of grid point concentrations is equivalent to the spatial integration of these concentration values. This can be easily done, but what is the uncertainty in the aggregated value? If the grid point concentrations were totally uncorrelated then the aggregated uncertainty would simply be given by:

$$\sigma^2_{uncor} = \frac{1}{n^2}\sum_i^n \sigma_i^2 \qquad (8)$$

where *i* represents a grid point and *n* is the total number of grid points. For the case where the grid point concentrations are totally correlated then the uncertainty would be written as

$$\sigma^2_{cor} = \frac{1}{n}\sum_i^n \sigma_i^2 \qquad (9)$$

which is the same as the average variance. There is a large difference between these two uncertainty estimates, a factor of *n* where *n* is a large number (number of grid points).

In order to assess the actual uncertainty correctly then the spatial correlation must be included in the assessment. If kriging or residual kriging has been used to determine the concentration uncertainty at each grid point then the correlation is known and can be calculated from the parameters used to describe the variogram (nugget, sill and range).

A methodology for determining the spatial uncertainty after aggregation is described in Rolstad et al., 2009). The application in that case was for Glacier mass balance calculations, but the theory is exactly the same. It is based on information contained in the spherical variogram and on the area over which the aggregation takes place (*A*). The aggregated uncertainty can, in that case, be written as:

$$\begin{aligned}
\frac{\sigma_A^2}{\sigma_{cor}^2} &= 1 && \mathrm{L} \le \Delta h \\
&= \frac{c_0}{c}\frac{\Delta h^2}{L^2} + \frac{c_1}{c}\left(1 - \frac{L}{a_1} + \frac{1}{5}\left(\frac{L}{a_1}\right)^3\right) && \Delta h < L < a_1 \qquad (10) \\
&= \frac{c_0}{c}\frac{\Delta h^2}{L^2} + \frac{1}{5}\frac{c_1}{c}\frac{{a_1}^2}{L^2} && L > a_1
\end{aligned}$$

where L represents the radius over which the aggregation takes place, a_1 the variogram range, c_0 the nugget value and c_1 the sill value and $c = c_0 + c_1$. Δh represents the size of a single grid cell by the form $\Delta h = \Delta x / \sqrt{\pi}$ where Δh is the grid spacing. The above equation is based on a circularly symmetric integration but the integration distance L can be simply estimated for non-circular integrations from the integration area, A, using the formula

$$L = \sqrt{\frac{A}{\pi}} \qquad (11)$$

A simple example, describing the spatially averaged annual mean PM_{10} concentrations in Europe, would be similar to an integration carried out over a radius $L = 2000$ *km*, with a variogram range $a_1 = 500$ *km*, and a ratio of nugget to sill variance (c_1/c) being 0.5. Applying these values to equation 9 provides the total uncertainty, compared to the correlated uncertainty σ_{cor} (average grid uncertainty, equation 9), would be:

$$\frac{\sigma_A}{\sigma_{cor}} \approx 10\% ,$$

which is significantly smaller (10%) than the typical uncertainties found at the individual grid points. Aggregation over a country, however, can give significantly higher uncertainties. For example a country of area 100 000 km^2 (effective $L = 180$ *km*), given the same variogram parameters, would have a total uncertainty of

$$\frac{\sigma_A}{\sigma_{cor}} \approx 57\% ,$$

which is roughly half of the correlated uncertainty (typical grid point uncertainties).

The above analysis can be easily coded for providing estimates of uncertainties in aggregated values, e.g. average population weighted concentrations, for all of Europe or for any particular country.

3.7.4 Distinguishing between variability and uncertainty

In order to proceed in calculating the uncertainty in health risks it is necessary to deal with the separation of variability and uncertainty. In many texts on this subject, reference will often be made to this, but there is little practical advice on distinguishing the two. It is at least recommended in most texts to treat them separately where possible.

If a practical and useable distinction is to be made between variability and uncertainty in the residual kriging framework then this may be made by defining the nugget value, which is representative of the local variability, as the sub-grid variability. The remaining variance in the variogram can be defined as the uncertainty. This implies that uncertainty is 0, or at least small, near measurement stations and increases with distance from them. This is in contrast to the sub-grid variability, which is considered to be constant everywhere. This is of course a grave simplification, since the sub-grid variability depends on the variability of emissions sources, land use as well as topography within in a grid and this will not be constant everywhere. However the delineation will allow for a separation of the two so that they may both be determined for health assessments using the methods set out in the previous sections. The division of urban and rural maps appears to be favourable for this approach since the nugget values will most likely be different in urban regions, reflecting different sub-grid variability's.

The practical application for the determination of uncertainty in the aggregation of exposure distributions can then be achieved by assigning uncertainties, based on the aggregated uncertainties with the removal of the nugget (section 3.7.3), to the exposure distributions described in section 3.7.2. Exactly how this is done will require more thought and effort but the basis will lie in this distinction. In essence two-dimensional probability distributions will be created based on the variability and the uncertainty.

3.7.5 Concentration-response functions and calculation of premature deaths

In order to calculate the number of premature deaths attributable to a pollutant more specific information is required concerning age distribution and baseline mortality. These data are provided as country-specific data and have been taken from the UN Population Division (UN, 2005) and the WHO Burden of Disease project (WHO, 2004). The uncertainty in these data is not known but they are based on large samples of statistics and at the national level the uncertainty is likely to be negligible. In addition to these demographic data, concentration-response functions are required (Künzli et al., 2000), which have their own significant uncertainty, and assumptions concerning the non-anthropogenic levels of the pollutants are also required if 'premature' deaths are to be calculated.

Given that the age and baseline mortality data is provided per country then it is recommended to apply the calculations for health risks on the national level, i.e. to use concentration data that is aggregated to the country level. This should take the form of exposure distributions, described in sections 3.7.2 – 3.7.4, based on sub-grid variability as well as uncertainty. In any case, assuming we have an exposure distribution and associated uncertainty at the national level, then this can be combined with the age distribution, concentration-response functions and baseline mortality data to determine the number of premature deaths.

One method for achieving this, which will include the dimension of uncertainty, is to use Monte Carlo sampling techniques. These are straight forward but a variety of improved versions, such as Markov Chain Monte Carlo methods, also exist. In fact such methods are regularly used in probabilistic models for exposure assessments and they may also be successfully applied here. Their application, however, will require significant testing and assessing but the results of such an approach will provide a much better estimate of the health impacts of air quality in Europe.

4 Conclusions and recommendations

This short report has tried to cover, and provide some methods for dealing with, the various sources of uncertainty in the mapping methodology currently implemented by ETC/ACC and EEA for producing air quality maps and related health and ecosystem effects and impact assessments for Europe. This is a very rich area, with a large range of uncertainty sources that cannot always be easily determined. Indeed the area becomes more difficult to define as some uncertainties will overlap or are included in other uncertainty estimates. For example the residual kriging variogram effectively includes a wide range of uncertainties, from monitoring to modelling, as well as spatial variability. Some uncertainties are not easily determined, e.g. the uncertainty resulting from the non-fulfilment of the kriging interpolation assumptions, and some may be unknown without further extensive research, e.g. meteorological fields.

Table 1 provides and overview of the uncertainty sources and methodologies for analysing and mapping these.

Given the broad range of sources and the level of knowledge on these a pragmatic approach must be taken. In this case we make the following recommendations for further work to improve the representation of uncertainty in the mapping methodology.

1. The residual kriging variance, based on the available spatial statistical residual measurement data is a good first estimate of the uncertainty of the interpolated maps.
2. Some effort should be made to better quantify the uncertainty related to extrapolation, particularly for the regression, and how this can be used in defining the total regression uncertainty.
3. The quality of the uncertainty estimate based on the variogram variance will generally decrease in areas where no observational data are available. In this regard the uncertainty estimates should be considered valid for interpolation, rather than extrapolation, only. This is also true for the cases where regression models are extrapolated into areas where no data exists.
4. A brief description is given on how to include the model uncertainty in the residual kriging uncertainty. This methodology would need to be properly assessed before it could be implemented further in the mapping methodology.
5. It is generally recommended to use the log-normal transformed concentrations for the interpolation, particularly when the range of data is large. The major argument for this is that the frequency distributions of measured and modelling concentrations are seen to be closer to log-normal distributions than to normal distributions. The major consequence of using the log-normal transformation is that uncertainty estimates become relative, rather than absolute, values.
6. An assessment of the predicted uncertainty can be made, using cross-validation prediction errors, to determine if the predicted uncertainty is in fact realised.
7. When aggregating data for health risk assessment then it is beneficial to separate variability from uncertainty. A basic method is suggested to do this when using the residual kriging method. In addition, the spatial correlation of data makes the estimate of spatial aggregation uncertainty more complicated. A methodology for determining this is presented.
8. When calculating population exposure use should be made of the uncertainty and variability within grid cells. Though there may be some correlation between population density and concentrations, particularly for pollutants such as NO_2, the inclusion of this correlation in the population exposure calculations needs to be assessed.
9. It would be beneficial for the task team if a new member could be found with a strong statistical and spatial statistical background.

Table 1. Summary list of uncertainty types and sources including their individual estimated uncertainty and an assessment of their potential for improvement, where relevant.

Source of uncertainty	Comment / method for assessment	Estimated uncertainty (σ)	Potential for reduction of uncertainty
Monitoring			
Instrumental and calibration error	Assume to be equivalent to the quality objectives as laid out in the EU directive	5%	Very low
Monitoring analysis, reporting and quality control	Comparative methods, e.g. Van de Kassteele et al. (2006) can be used to compare the reporting from different countries. Cross border comparisons. Inter-national calibration programmes. Improved monitoring methods.	20%	High
Spatial representativeness of station data	Difficult to quantify. Contributes to the sub-grid variability of the maps. Sub-grid variability is not a source of uncertainty that can be reduced, only assessed. Directly related to the station classification.	unknown	Low
Temporal coverage of station data	Exclusion of sites with missing data to minimise this. Currently carried out by requiring 75% temporal coverage.		
Station classification	Potential for large errors in the mapping, especially where limited monitoring sites are available. A range of suggested methods for 'checking' the classification are suggested in section 3.3.1	unknown	High
Modelling			
Total model uncertainty	The complexity of the modelling system and its reliance on uncertain input and output data makes it difficult to assess individual aspects of uncertainty. Total model uncertainty is still best assessed through comprehensive model validation exercises by comparison with observations.	20 – 100%	Medium
Process formulation and model assumptions	In order to indicate the degree to which model formulation and process descriptions effect model uncertainty then model ensembles can be applied, either by using models or using separate modules within models.	50%	Medium
Emissions	Independent and comparative checks on emission inventories are required to indicate the uncertainties in emissions	10 – 200 % depending on emission	High
Meteorology	There has been a slow and steady improvement in meteorological modelling, at least for forecasting, over the past decades. This is a large research area.	10 – 200 % depending on model	Medium
Representativeness			
Spatial representativeness	Raising awareness concerning station classifications and their spatial representativeness is generally a good idea. Extra checks concerning discrepancies in observed concentrations in regard to station classification can be applied to AirBase data.	unknown	
Sub-grid variability	It would be useful to try to estimate the sub-grid variability and separate this from the uncertainty. Can use spatial statistics from observations and/or high resolution modelling studies to assess this part.	20 – 50%	Medium

Mapping methodology			
Non-fulfilment of the assumptions used for the interpolation method	It is possible to assess the fulfilment of the criteria statistically. This may be done to some degree but the effect of any non-fulfilment is more difficult to define	unknown	
Extrapolation of regression relationships	Extrapolation outside of the regression data limits. Can be assessed through Monte Carlo or other statistical methods	unknown	
Inclusion of modelling uncertainty in the interpolation method	By assuming that the baseline model fields, including regression, have a given uncertainty, this may be combined with the residual kriging using weighting functions	50%	
Uncertainty in the monitoring	General monitoring uncertainty, or variability between monitoring sites and methods, will become part of the general variance identified in the variogram.	20%	
Normal or log-normal transformation	Important difference when calculating the uncertainty, but less important for the actual values. Should be assessed by looking at the distributions of the monitoring and modelling data		High
Urban/rural merger	Non-correspondence of the station type with the underlying grid cell leads to uncertainty increase. Can be solved by the new proposed methodology; however, problematic is the use of the same variogram for both rural and urban areas in the interpolation of the residuals.	20%	Medium
Exposure and health			
Population and demographic data	Very low uncertainty at aggregated levels but with increasing uncertainty at higher resolutions	2%	Low
Concentration-response functions	Uncertainty in the health risk, based on exposure	30%	Low
Reference pollutant level	Levels that need to be defined if 'premature' deaths are to be determined. Can be improved through improved non-anthropogenic modelling studies	20%	High
Sub-grid variability of the concentration fields	Can be estimated from the nugget. This is natural variability and cannot be reduced.	20-50%	None
Uncertainty of the concentration fields	Can be estimated from the variogram and the nugget. Uncertainty at this level may be improved by any of the methods previously described.	20 – 50 %	Medium
Population exposure	Can be determined by separating the variability and uncertainty. Monte Carlo methods may be applied	20 – 50 %	Medium
Health risks (premature deaths)	Result of the other uncertainties, being the end point of the assessment. Can be determined by separating the variability and uncertainty and combining with the uncertainty in concentration-response functions. Monte Carlo methods may be applied.	20 – 100 %	

References

Barrett, K., Steinar Larssen, Ingrid Sundvor, Lise Fjellsbø, Maria Dusinka, Frank de Leeuw, Jaroslav Fiala, Jana Ostatnicka, Jan Horálek, Libor Cernikovsky, Fotios Barmpas, Nicolas Moussiopoulos, Cristos Vlahacostas (2008). Health impacts and air pollution: An exploration of factors influencing estimates of air pollution impact upon the health of European citizens. ETC/ACC Technical Paper 2008/13. Available at: http://air-climate.eionet.europa.eu/reports/ETCACC_TP_2008_13_HealthImpact_AirPoll

Briggs, D. et al. (2005). APMOSPHERE – Air Pollution Modelling for Support to Policy on Health and Environmental Risks in Europe. http://www.apmosphere.org/

Brown, Coleman, and Steele (1998). A Methodology for Determining the Experimental Uncertainty in Regressions. J. of Fluids Engineering, Vol. 120, No. 3, 1998, pp. 445-456.

Ching, J., Herwehe, J., and J. Swall (2006) On joint deterministic grid modeling and sub-grid variability conceptual framework for model evaluation, Atmos. Envir., 40, 4935-4945.

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

Denby B., A. M. Costa, A. Monteiro, A. Dudek, S.E. Walker, S. van den Elshout, B. Fisher. (2008b) Uncertainty Mapping For Air Quality Modelling And Assessment. Accepted for publication in the International Journal of Environmental Pollution.

Denby B., M. Schaap, A. Segers, P. Builtjes and J. Horálek (2008c). Comparison of two data assimilation methods for assessing PM10 exceedances on the European scale. Atmos. Environ. 42, 7122-7134.

Denby, B., Horálek, J., Walker, S.E., Eben, K. and Fiala, J. (2005) Interpolation and assimilation methods for European scale air quality assessment and mapping. Part 1: Review and recommendations. ETC/ACC Technical Paper 2005/7. December 2005. http://air-climate.eionet.europa.eu/reports/reports/ETCACC_TechnPaper_2005_7_spatial_AQ_interpol_Part_I

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

Harrison, D. (2006) UK Equivalence Programme for Monitoring of Particulate Matter. Final Report for: Department for the Environment, Food and Rural Affairs; Welsh Assembly Government; Scottish Executive; Department of Environment for Northern Ireland. Ref: BV/AQ/AD202209/DH/2396 www.airquality.co.uk/archive/reports/cat05/0606130952_UKPMEquivalence.pdf

Henne, Stephan, Doris Folini, Dominik Brunner, Brigitte Buchmann and Sverre Solberg (2008). Report on supersite representativeness and representativeness assessment method. D1.1.2 of GEOmon. http://geomon.empa.ch/D2.1.2_StationRepresentativeness_StephanHenne_EMPA_M18.pdf

Hien Q. Le, Stuart A. Batterman and Robert L. Wahl (2007) Reproducibility and imputation of air toxics data. J. Environ. Monit., 9, 1358 - 1372, DOI: 10.1039/b709816b

Horálek et al. (2009). Improved urban and rural mergers evaluation for European scale assessment. ETC/ACC Technical Paper 2008/19. April 2009. http://air-climate.eionet.europa.eu/reports/ETCACC_TP_2008_19_urban_rural_aq_merger

Horálek, J., de Smet, P., de Leeuw, F., Denby, B., Kurfürst, P. and Swart, R. (2008). European air quality maps for 2005 including uncertainty analysis. ETC/ACC Technical Paper 2007/7. http://air-climate.eionet.europa.eu/reports/ETCACC_TP_2007_7_spatAQmaps_ann_interpol

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

Horálek, J., Kurfürst, P., Denby, B., de Smet, P., de Leeuw, F., Brabec, M. and Fiala, J. (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. December 2005. http://air-climate.eionet.europa.eu/reports/reports/ETCACC_TechnPaper_2005_7_spatial_AQ_interpol_Part_II

Janssen, S., Dumont G., Fierens F. and Mensink C., (2008a) Spatial interpolation of air pollution measurements using CORINE land cover data. Atmospheric Environment, 42, 4884–4903.

Janssen, S., Frans Fierens, Gerwin Dumont and Clemens Mensink (2008b) RIO: A NOVEL APPROACH FOR AIR POLLUTION MAPPING. Proceedings of the 12th International Conference on Harmonisation within Atmospheric Dispersion Modelling for Regulatory Purposes, Cavtat, Croatia, October 2008. http://www.harmo.org/conferences/Proceedings/_Cavtat/publishedSections/O_S4-04.pdf

Kassteele J. van de, Koelemeijer R.B.A., Dekkers A.L.M., Schaap M., Homan C.D., Stein A. (2006) Statistical mapping of PM10 concentrations over Western Europe using secondary information from dispersion modeling and MODIS satellite observations, Stochastic Environmental Research and Risk Assessment 2006; 21(2):183-94.

Künzli N., Kaiser R., Medina S., Studnicka M., Chanel O., Filliger P., Herry M., Horak F., Jr., Puybonnieux-Texier V., Quenel P., Schneider J., Seethaler R., Vergnaud J.C., Sommer H. (2000). Public-health impact of outdoor and traffic-related air pollution: a European assessment. Lancet 2000;356(9232):795-801.

Larssen, S., R. Sluyter, and C. Helmis (1999), Criteria for EUROAIRNET - The EEA Air Quality Monitoring and Information Network, 56 pp, European Environment Agency, Copenhagen. http://reports.eea.europa.eu/TEC12/en/tech12.pdf

Loon, M. van, Vautard, R., Schaap, M., Bergstrom, R., Bessagnet, B., Brandt, J., Builtjes, P.J.H., Christensen, J.H., Cuvelier, K., Foltescu, V., Graff, A., Jonson, J.E., Kerschbaumer, A., Krol, M., Langner, J., Roberts, P., Rouil, L., Stern, R., Tarrason, L., Thunis, P., Vignati, E., White, L., Wind, P (2007), Evaluation of long-term ozone simulations from seven regional air quality models and their ensemble average, Atmospheric Environment, 41 (10), 2083-2097

Mol, W.J.A., Hooydonk P.R. van (2005b): QA/QC checks on air quality data in AirBase and on the EoI2004 data – Procedures and results. ETC/ACC Technical Paper 2005/3. Available at http://air-climate.eionet.europa.eu/reports/ETCACC_TechnPaper_2005_3_QAQC_feedback_EoI2004

Pagowski, M; Grell, GA; Devenyi, D; Peckham, SE; McKeen, SA; Gong, W; Delle Monache, L; McHenry, JN; McQueen, J; Lee, P, (2006): Application of dynamic linear regression to improve the skill of ensemble-based deterministic ozone forecasts. Atmos. Environ.: 40, 3240-3250

Plaia A. and Bond`ı A.L. (2006) Single imputation method of missing values in environmental pollution data sets, Atmospheric Environment, 40, 7316–7330.

Rodriguez, M.A. , J. Brouwer, G.S. Samuelsen and D. Dabdub (2007). Air quality impacts of distributed power generation in the South Coast Air Basin of California 2: Model uncertainty and sensitivity analysis. Atmos. Environ., 41: 5618-5635.

Rolstad, C., T. Haug and B. Denby (2009) Geodetic glacier mass balance and its uncertainty based on a spatial statistical analysis: application to the Western Svartisen ice cap, Norway. Submitted to the Journal of Glaciology.

Smet, P. de, Jan Horálek, Markéta Coňková, Pavel Kurfürst, Frank de Leeuw, Bruce Denby (2009) European air quality maps of ozone and PM10 for 2006 and their uncertainty analysis. ETC/ACC Technical Paper 2008/8. http://air-climate.eionet.europa.eu/reports/ETCACC_TP_2008_8_spatAQmaps_2006

Solberg S., Jan Horálek, Jan Eiof Jonson, Steinar Larssen, Frank 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.

UN (2005). United Nations, Department of Economic and Social Affairs, Population Division (2005). World Population Prospects: The 2004 Revision. CD-ROM Edition - Extended Dataset (United Nations publications, Sales No. E.05.XIII.12).

WHO (2004). Global Burden of Disease project. http://www.who.int/topics/global_burden_of_disease/en/

Additional links for related reading not referenced in the report

http://www.mnp.nl/en/publications/2006/Guidance_for_uncertainty_assessment_spatial_information.html

http://www.mnp.nl/en/publications/2003/Uncertainty_analysis_for_NOx_emissions_from_Dutch_passenger_cars_in_1998.html

http://www.mnp.nl/nl/publicaties/2004/TNOreport__Uncertainty_assessment_of_NOx__SO2_and_NH3_emissions_in_the_Netherlands.html