Interpolation methods for European scale air quality mapping: Application to European population exposure estimates for PM_{10}

Jan Horálek[1], Bruce Denby[*], Pavel Kurfürst[1], Peter de Smet[2], Frank de Leeuw[2], Marek Brabec[3]

* Corresponding Author: Norwegian Institute for Air Research (NILU), P.O. Box 100, 2027 Kjeller, Norway

[1] Czech Hydrometeorological Institute (CHMI); [2] Netherlands Environmental Assessment Agency (MNP); [3] National Institute of Public Health, Praha

European Topic Centre on Air and Climate Change

The project is fully financed by the European Environmental agency (EEA)

NILU

Poster produced by the Norwegian Institute for Air Research (NILU)

Background

The European Topic Centre for Air and Climate Change (ETC/ACC) is currently carrying out a research task for the European Environmental Agency (EEA), in support of the Structural Indicator work of EEA to DG ENV, that reviews and further develops interpolation methods for use in European wide air quality mapping (*Denby et al., 2005; Horálek et al., 2005*). In the work carried out so far emphasis has been placed on the development of interpolation methodologies for both ozone and PM_{10}, with the subsequent preparation of high-resolution maps covering all of Europe sufficient to resolve urban agglomerations. For PM_{10} both the annual mean and 36'th highest daily average indicator have been interpolated onto a 10 ×10 km² grid.

The ultimate aim of the task is to produce maps that show human and ecosystem exposure. The PM_{10} maps, the derivation of which is presented here, will in the next stage be combined with population density maps to produce population at risk and population weighted indicators on a European wide basis.

Interpolation methodologies

Three basic types of interpolation methods are studied and applied

1. **Standard spatial interpolation methods** - including inverse distance weighting (IDW), ordinary and log-normal kriging and ordinary and log-normal co-kriging.
2. **Regression models** - whereby measurements are correlated with other spatial parameters including chemical transport models, climatic parameters and elevation.
3. **Combined methods** - that involve the use of regression models (2) and the spatial interpolation (1) of the residuals.

Rural and urban measurement stations are interpolated separately to account for the different nature of these stations, i.e. spatial representativeness. The two interpolation fields are combined using a population density weighting.

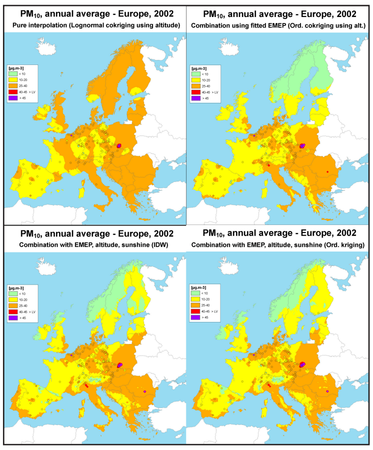


Figure 1. Maps showing the annual average PM_{10} concentrations ($\mu g.m^{-3}$) on European scale for rural areas in a 10 km x 10 km grid resolution as a result of four different interpolation Methods 1e (top left), 3d (top right), 3e (bottom left) and 3f (bottom right), using 2002 rural background monitoring data.

Data used

The basic observational data used for the interpolation are taken from the AirBase (*AirBase, 2006*) and EMEP (*EMEP, 2006*) databases which provide data for the study years (2000-2003). 205 rural background and 724 urban/suburban PM_{10} stations are used. Supplementary data, used for co-kriging and regression models, include concentration fields calculated with the EMEP unified model, climatic parameters, altitude and population density.

	Interpolation method for rural maps	Average RMSE	Ranking
	Standard spatial interpolation methods		
	Inverse distance weighting (IDW)	6.59	13
	Ordinary kriging	6.29	12
	Ordinairy co-kriging (altitude)	5.43	5
	Log-normal kriging	6.15	11
	Log-normal co-kriging (altitude)	5.19	3
	Residual interpolation with respect to the EMEP model		
2a	No residual interpolation (pure EMEP comparison)	12.78	15
2b	Inverse distance weighting (IDW)	5.91	10
2c	Ordinary kriging	5.86	9
2d	Ordinairy co-kriging (altitude)	5.48	6
3	**Combination of regression and interpolation of the residual**		
3a	Regression (EMEP model) no interpolation	6.96	14
3b	Regression (EMEP model) and IDW of residual	5.72	8
3c	Regression (EMEP model) and ordinary kriging of residual	5.59	7
3d	Regression (EMEP model) and ordinary co-kriging (altitude) of residual	5.37	4
3e	Regression (EMEP model + altitude + sunshine) and IDW of residual	4.96	2
3f	Regression with (EMEP model + altitude + sunshine) and ordinary kriging of residual	4.84	1

Table 1. Comparison of the different interpolation methods showing the average RMSE (in $\mu g.m^{-3}$) over the 4 year period 2000-2003 for annual average PM_{10} concentrations. The method ranked 1 (3f) gives the best result, the method ranked 15 (2a) the worst.

	Interpolation method for urban maps	Average RMSE	Ranking
1	**Regression methods**		
1a	Regression with population density	10.20	7
1b	Regression with EMEP model + altitude + sunshine + wind speed + longitude	8.93	6
1c	Regression with rural background field	8.33	5
2	**Spatial interpolation across city borders**		
2a	Interpolation of urban concentrations using IDW	7.31	4
2b	Interpolation of urban concentrations using ordinary kriging	7.12	2
3	**Spatial interpolation of deltas across city borders + rural background fields**		
3a	Interpolation of deltas using IDW + rural background fields	7.15	3
3b	Interpolation of deltas using ordinary kriging + rural background fields	7.09	1

Table 2. Comparison of different interpolation methods for cross city border interpolation of rural PM_{10} showing the RMSE of cross-validation for the year 2003. The method ranked 1 (3b) gives the best result, the method ranked 7 (1a) the worst.

Testing of rural interpolation methodologies

A number of different interpolation methods are tested and assessed using cross-validation. This involves the successive removal of each measurement from the interpolation and comparison of that measurement with the resultant interpolation field. The summed root mean square error (RMSE) is then calculated for each interpolation method. Results from these tests are shown in table 1 and the resulting maps from the 4 best ranked methods are shown in figure 1. The lowest RMSE (method 3f) is found using a regression analysis with the EMEP model, elevation and sunshine duration and then interpolating the difference using ordinary kriging. Log-normal co-kriging with altitude (method 1e) also provides a low RMSE.

Testing urban interpolation methodologies

Interpolation of urban stations is more difficult due to their limited spatial representativeness. This problem is compounded because measurements are not available in all cities so some form of interpolation must be carried out over city borders. A number of interpolation methods are tested. These include direct spatial interpolation, regression models based on supplementary data (including population density) and spatial interpolation of deltas (the difference between urban measurements and rural fields). The results of these tests are shown in table 2. The best method is found to be the interpolation of the urban delta using ordinary kriging with the addition of the rural background field.

Merging of rural and urban concentration fields

The interpolated rural and urban concentration fields are merged together based on a population density weighting. In areas where population densities are less than 100 inhabitants per km² the rural concentration field is used and in areas where population densities are greater than 500 inhabitants per km² the urban concentration field is used. Between these values a linear combination is used. The method is based on an analysis of the convergence of rural and urban mean concentrations as a function of population density as shown in figure 2.

Final interpolation methodology

Based on the above tests an interpolation methodology is chosen for use. This involves:

1. The splitting of the rural and urban measurement stations and the interpolation of these separately

2. A regression model is established using rural station measurements, the EMEP model calculations (50×50 km²), elevation (30"×30") and climatological sunshine duration (10'×10'). This model is used to map the PM_{10} fields at 10×10 km² resolution

3. The difference between this regression field and the measurements is then interpolated using ordinary kriging and added to the regression field to produce an interpolated rural concentration field

4. The interpolated rural concentration field is subtracted from the urban station measurements (delta) and this delta is then interpolated over city borders using ordinary kriging with the addition of the rural background field

5. The interpolated rural and urban pollutant fields are then combined using population weighting into a complete European map. Interpolation accuracy is assessed by examination of the root mean square error (RMSE) using cross-validation.

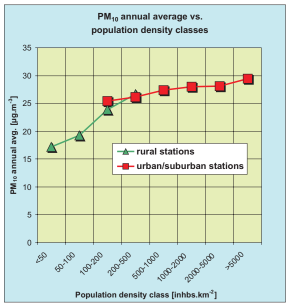
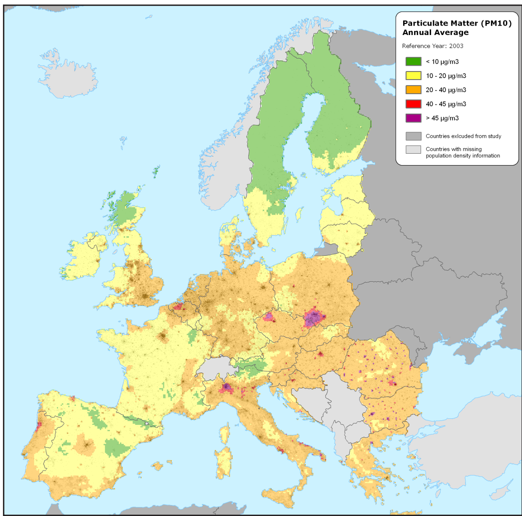


Figure 2. PM_{10} annual average vs. population density classes, for rural and urban/suburban stations. For every class the average value of all the stations in that class is considered for the years 2000-2003.

Final maps of PM_{10}



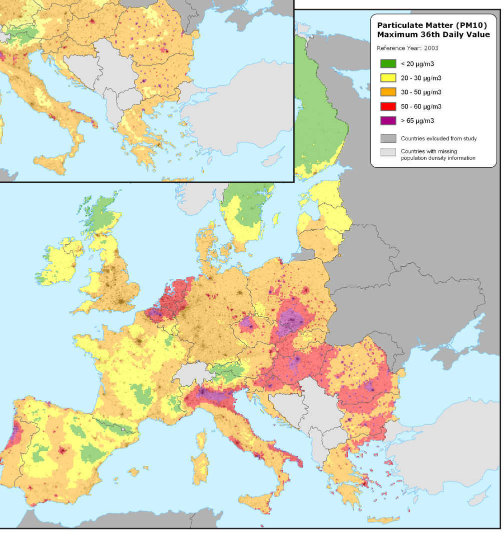


Figure 3. Final interpolated rural and urban concentration maps of PM_{10} (2003) overlaid by population density grid. Left: annual average, right: 36th maximum daily average. Units: $\mu g.m^{-3}$. These concentration maps are created by merging the rural map (Method 3f in table 1: Regression with EMEP model + altitude + sunshine and ordinary kriging of residual) and the urban map (Method 3b table 2: Interpolation of deltas using ordinary kriging + rural background fields). Countries with interpolation based on additional data only: BG, GR, HR, HU, RO. Countries with missing population density information and therefore excluded: AD, AL, BA, CH, CS, CY, IS, LI, MK, NO, TR.

Conclusions

The current work explores a number of interpolation methods for producing high resolution European maps for PM_{10}, made by merging spatially interpolated rural and urban concentration fields. Regression techniques, based on meaningful supplementary spatial data, have been shown to improve the spatial interpolation significantly. The combination of regression models and spatial kriging of the residuals has also been shown to be the most effective method for spatial interpolation yet tested.

The methodology developed will be further improved in 2006, including uncertainty assessment and the establishment of routine map production, and applied to human and ecosystem exposure calculations.

References

Denby, B., J. Horálek, S.E. Walker, K. Eben and J. Fiala, 2005. *Interpolation and assimilation methods for European scale air quality assessment and mapping. Part I: review and recommendations.* ETC/ACC Technical paper 2005/7. *http://air-climate.eionet.eu.int/docs/ETCACC_TechnPaper_2005_7_SpatAQ_Interpol_Part_I.pdf*

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

AirBase, 2006. *The European Air Quality Database.* URL: *http://air-climate.eionet.eu.int/databases/airbase*

EMEP, 2006. *EMEP measurement data online.* URL: *http://www.nilu.no/projects/ccc/emepdata.html*