Basic data assimilation methods:

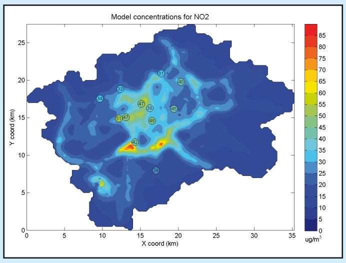
Application to the urban scale

Bruce Denby (bde@nilu.no)
Norwegian Institute for Air Research (NILU)

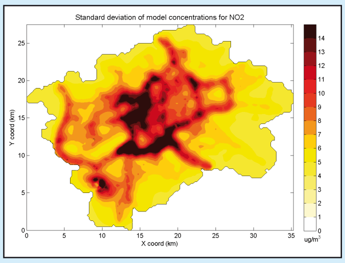
Summary

This case study examines the use of basic data assimilation methods and how they can improve the spatial assessment of air quality on the urban scale. Annual mean model fields of PM_{10}, NO_2 and SO_2, calculated by the ATEM model (*Brechler, 2000*), for the year 2003 have been combined with observations from 12 monitoring sites in the city of Prague. A number of basic data assimilation methods were investigated including linear regression and the Bayesian combination of kriged and model fields. The methodology and results outlined below describe how basic data assimilation can be achieved for the example of NO_2. In addition to the assimilated maps, maps of uncertainty are produced. These verify the improvement obtained with the use of data assimilation.

Modelling map NO_2

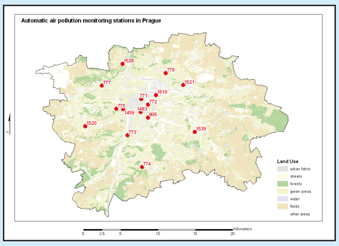


Annual mean field of NO_2 as calculated by the ATEM model. Included are the monitoring station concentrations for comparison



Indicative uncertainty map of the annual mean field for NO2 as calculated by the ATEM model. This map is derived from the normalised root mean square error of the model at the monitoring station sites

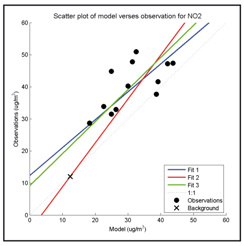
Monitoring NO_2, PM_{10} and SO_2



Monitoring stations in Prague used in this case study

STATION CODE	STATION NUMBER	Annual mean NO_2 (µg/m³)	Annual mean PM_{10} (µg/m³)	Annual mean SO_2 (µg/m³)
AREPA	771	47.24	46.41	8.42
ARIEA	772	37.698	45.61	10.45
ABRAA	773	41.6	45.18	6.68
ALIBA	774	28.66	35.7	6.42
AMLYA	775	50.97	42.68	7.27
ASMIA	1459	47.38	58.58	8.37
ASANA	776	32.92	41.72	8.11
AVELA	777	33.85	37.77	7.27
AKOBA	779	31.46	44.25	7.74
AVYCA	780	40.25	40.6	7.64
APOCA	804	44.85	45.32	8.91
AVRSA	805	47.8	43.58	7.38

Annual mean concentrations as measured by the monitoring stations in Prague. Shown are values for NO_2, PM_{10} and SO_2.



Three different types of regression models

Fit 1: Linear regression

$$M_{reg} = a + b\,M$$

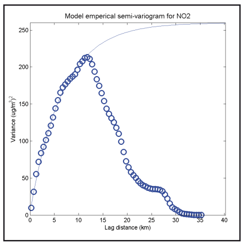
Fit 2: Linear regression with known background (BG)

$$M_{reg} = O_{BG} + b\,(M - M_{BG})$$

Fit 3: Unknown background and no regression

$$M_{reg} = a + M$$

where **M** is the model field, **O** the observations and M_{reg} the resulting regression model field

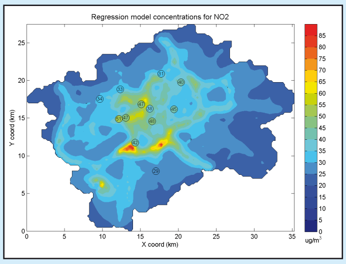


Kriging semi-variogram showing the increase of model variance (uncertainty) as a function of lag distance (distance from a point in space). The semi-variogram is used to construct a variogram model, solid line, used in the kriging interpolation. The interpolation minimises the variance at the interpolation position.

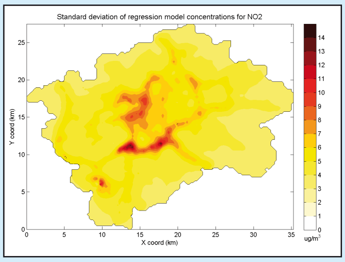
Variogram parameters, spherical model, used for this fit:

Range = 10 km
Sill = 260 (µg/m³)²
Nugget = 5 (µg/m³)²

Model regression map NO_2

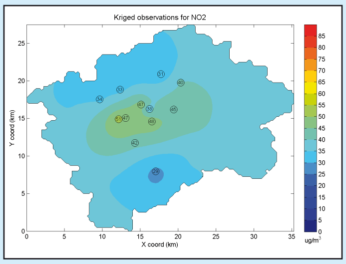


Annual mean model regression field of NO_2 as calculated by the ATEM model using the linear regression **fit 1** where a = 12.3 ug/m³ and b = 0.87. The fit tends to increases the background levels but reduces the maximum levels.

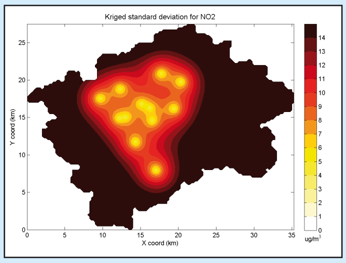


Indicative uncertainty map of the annual mean field for NO_2 as calculated by the ATEM model. Derived from the normalised root mean square error of the model at monitoring station sites

Kriging map NO_2



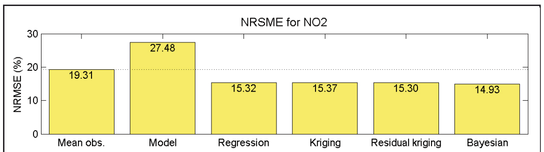
The interpolated monitoring field for the annual mean NO_2. Far from stations the concentration approaches the mean of all observations. No details smaller than the station separation can be discerned.



Uncertainty map of the annual mean interpolated observational field of NO_2 as derived from the semi-variogram model used in the kriging. Close to stations the the result is most certain.

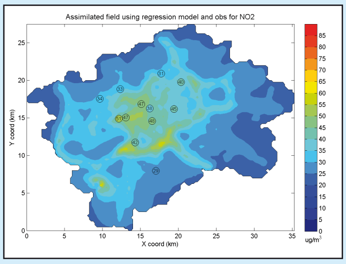
Bayesian assimilation

The assimilated map is created by the weighted combination of the model regression and kriged interpolation fields. The weighting, according to Bayesian theory, minimises the uncertainty in the final map.

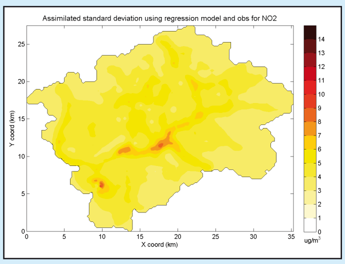


The normalised root mean square error (NRMSE%) for the different methodologies applied in this case study. The final NRMSE for the Bayesian combination is 14.9% compared with the original model NRMSE of 27.5%.

Assimilated map NO_2



The final annual mean map of NO_2 resulting from the Bayesian combination of the model regression map and the kriged interpolation map.



The final uncertainty map of the annual mean field of NO_2 as calculated using the Bayesian combination of model regression and kriged interpolation. Uncertainty is reduced in the region covered by the monitoring stations

Conclusion

- Linear regression is an effective and simple methodology for improving air quality assessment maps providing there are:
  - sufficient numbers of representative monitoring stations available for the analysis covering a wide range of concentration levels
  - a sufficiently good correlation for the regression
- Combining these fields with interpolated monitoring fields will also improve the assessment map when the combination takes into account, through Bayesian statistics, the uncertainty of both the model and monitoring fields.
- The inclusion of uncertainty maps is a useful indicator of the uncertainty in the assessment maps and is highly recommended for any air quality mapping.

References: Brechler , J., 2000. *Model assessment of air pollution in Prague.* Environmental monitoring and assessment 65: 269-276.

NILU PP 27/2006

For more information about Air4EU please visit www.air4eu.nl