INVERSE MODELLING OF EMISSIONS AND THEIR TIME PROFILES

Jaroslav Resler*, Krystof Eben, Pavel Jurus, Jitka Liczki
Institute of Computer Science
Academy of Sciences of the Czech Republic, Prague, Czech Republic

*1. INTRODUCTION

In our earlier work [Resler 2008] we set up a 4DVar data assimilation scheme based on CMAQ adjoint model [Hakami 2007]. This scheme enables assimilation of in-situ and satellite observations of NO_2. Initial conditions and emissions of NO and NO_2 are optimized; for the case of emissions a simple parametrization with one multiplicative coefficient per gridpoint has been used. The experiments resulted in a better agreement between the model and the observations of NO_2 for both analysis and forecast. However, this assimilation scheme achieved no improvement in modelling and forecasting of ozone concentrations.

This contribution shows further modification and improvements of the method (section 2). The main goal of these changes is to improve analysis and forecast of other species, in particular O_3. The results of a pilot experiment are discussed in sections 3 and 4, while the first results from a long-term experiment and the outlook are discussed in section 5.

We follow the terminology and notation used in [Resler 2008].

2. ENHANCEMENTS OF THE METHOD

As a results of the experiments in [Resler 2008], key issues to be addressed have been identified:

- insufficiency of the chemical mechanism CB4
- no direct information about O_3 has entered the assimilation routines
- optimization of other species is necessary
- diurnal profiles of emissions are burdened with errors

2.1 Chemical mechanism

The chemical mechanism CB4 is obsolete and it is not included in recent version of CMAQ model. Since the SAPRC99 mechanism is more suitable for modelling ozone chemistry, we developed the SAPRC99 version of the CMAQ adjoint model. The implementation was done with the help of the KPP tool (see [Sandu 2003]).

2.2 In-situ and satellite observations of ozone

The assimilation scheme was extended for inclusion of observations of different species. In particular, observations of O_3 from ground level stations from the database of EEA (European Environment Agency) and the first layer of the O_3 profile retrieved from IASI satellite instrument are assimilated together with NO_2 observations in the experiments described below.

2.3 Optimisation of other species

It is well known that formation of ozone depends strongly not only on concentrations of NO and NO_2, but also on other species, especially on species belonging to the VOC group. This fact is well seen from adjoint sensitivity experiments, too. An improvement in modeling ozone concentrations thus requires optimization of more species. The assimilation scheme was enhanced and all species with significant backward sensitivities (except radicals and other short living species) have been included into optimization.

The inclusion of species with different value ranges and statistical characteristics in minimization scheme led to necessity of regularization of the minimization scheme. This was reached by applying transformation of the variables following [Courtier 1997].

2.4 Diurnal profile optimization

From the previous experiments it follows that emission diurnal time profiles generated by the emission model are burdened with inaccuracies in many places and multiplication by one coefficient per day gives no possibility to correct it. Therefore we decided to generalize optimization of daily emission totals to optimization of the profile itself. To avoid creation of an ill-posed problem, we model the diurnal profile correction as a linear combination of five fixed base functions. Each member of this base characterizes one important part of the emission time profile: night, morning peak, midday trough, evening peak and late evening (see Fig. 1). The gradient with respect to these

*Corresponding author: Jaroslav Resler, Institute of Computer Science Academy of Sciences of the Czech Republic, Pod Vodarenskou vezi 2, 182 07 Prague 8, Czech Republic; e-mail: resler@cs.cas.cz

parameters was calculated and minimisation of the cost function was performed. Another experiment in which emission for every hour of the day is optimised independently was also performed and the result is briefly discussed in section 5.

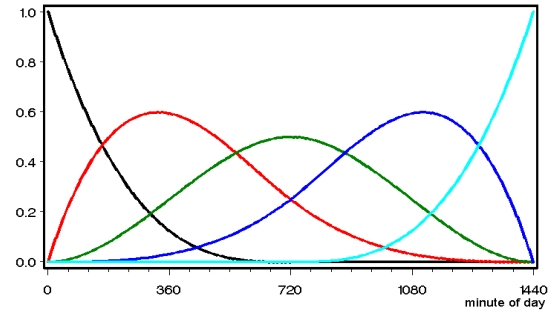

Fig 1: Base for parametrization of diurnal profile correction.

In order to see the effect of this generalisation, a comparison experiment was performed. We have kept the assimilation setup used in [Resler 2008] and enhanced it with emission profile optimization.

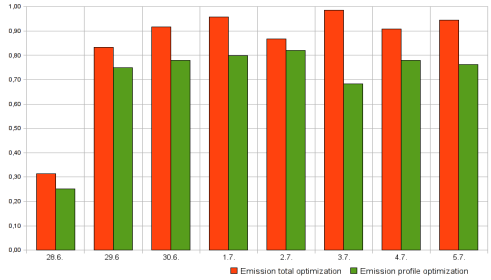

Fig 2: Relative decrease of the cost function for particular days. Red: optimization of emission daily total, green: optimization of daily profile.

Fig. 2 shows a comparison of relative decrease of the cost function in the course of the assimilation process in particular days for the original scheme (red columns) and enhanced scheme (green columns). The result shows significantly improved decrease of the cost function in the first day of experiment. This represents higher ability of the assimilation to bring the model closer to the observations. However, the higher decrease of the cost function in the next days shows slightly smaller stability of the emission correction parameters between particular days.

3. EXPERIMENT SETUP

To test further the proposed enhancements, we performed an experiment with domains and episode selected as in [Resler 2008]. The coarse assimilation domain has horizontal resolution 27 km and covers most of the important emission sources in Central and Western Europe (see Fig. 3). Data assimilation is performed on a period of eight days from June 28 to July 5, 1998 covering a short ozone episode.

The initial and boundary conditions are taken from a run of the CMAQ model on a large domain covering most of Europe and the north of Africa.

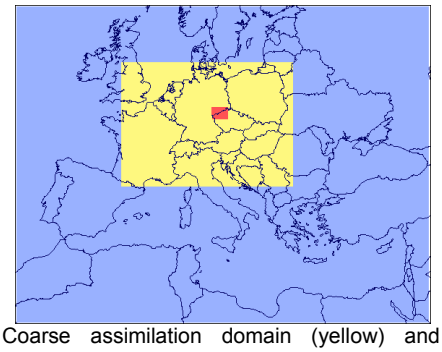

Fig 3: Coarse assimilation domain (yellow) and large simulation domain (blue).

The ground level observations are taken from cca 280 background monitoring stations of NO_2 and 360 stations of O_3 [1].

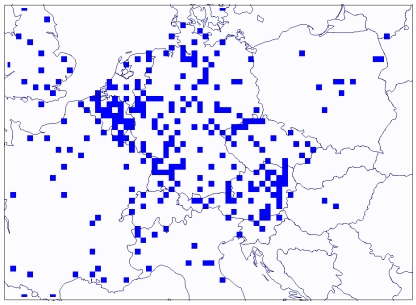

Fig 4: Ground level O_3 observation stations included in the experiment.

Satellite observations of NO_2 have the form of tropospheric columns of NO_2 retrieved from satellite instruments OMI and GOME2 from the TEMIS service ([Boersma 2007]). For O_3 we use the lowest layer of the O_3 profile retrieved from the IASI instrument.

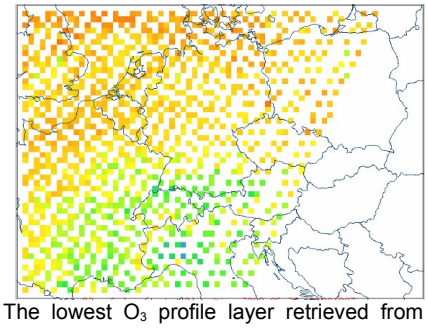

Fig 5: The lowest O_3 profile layer retrieved from the IASI instrument on July 1 2008 at 12:00.

[1] *The selection of data has been dictated by their availability. We have used NO_2 data from Germany (provided by the Umweltbundesamt www.uba.de) and Czech republic (provided by the Czech Hydrometeorological Institute). The NO_2 observations from UK and Belgium and all of O_3 observations are supplied by courtesy of the European Environment Agency.*

The error covariance matrices are assumed to be diagonal in this pilot experiment. We have performed some tests with covariance structure generated with the help of the diffusion operator, which has been used as covariance function (see [Weaver 2001], [Elbern 2007]). However, due to small contribution to the improvement in 27 km resolution and due to high demands on computing resources we did not use it in pilot experiments. The standard deviation of each in-situ observation is calculated as proportional to the observed value. It accounts for the representativity error of different types of stations and different spatial and time variability of the individual species. The standard deviations of the columns retrieved from satellite observations were taken from the data supplied. These values originate in the retrieval process.

First experiments with the enhanced assimilation scheme suggested that the optimization algorithm adjusts mainly the initial conditions of ozone. This leads to a dramatical improvement of agreement of the analysis with observations in the first part of the day. The adjustment, however, has a very small influence to the evening peaks and to the one-day-ahead forecast (cf. [Eben 2005]). Since the main goal of this study is an improvement of the model forecast by means of optimization of emissions, we decided to decrease the variance of the initial conditions artificially. As a result the decrease of the cost function was not as big as before, but on the other hand, changes occurred in emissions which gives an expectation of an improvement of the model forecasting capability.

4. RESULTS OF THE EXPERIMENT

An example of the resulting emission coefficients with respect to individual members of the emission profile parametrization base is shown in Fig. 6. From the maps it can be seen that in some regions, a part of the emission is translated from evening peak to morning or noon.

Fig. 7 shows time series of NO_2 concentrations from the original CMAQ model, optimized model and the forecast from optimized model compared with observed values for Brno, CZ (the second largest town in Czech Republic). Here we can see the "learning process" of the assimilation.

The forecast is constructed by adopting emission correction factors from the current day and running the model for the next day using these correction factors (persistent forward model). Although this works reasonably for weekdays, in places with significant weekly periodicity the persistent model for corrections is not satisfactory.

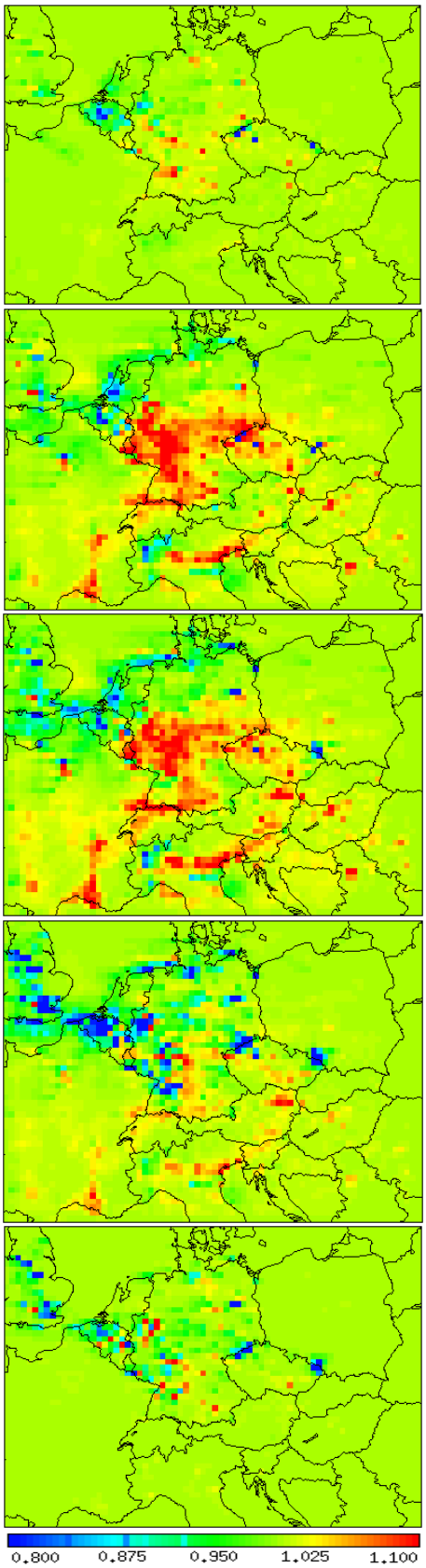

Fig 6: Emission coefficients with respect to the emission profile base for day July 2, 2008. From above downwards: night (black in Fig. 1), morning (red), midday (green), evening (blue) and late evening (cyan). The images are scaled to the same scale.

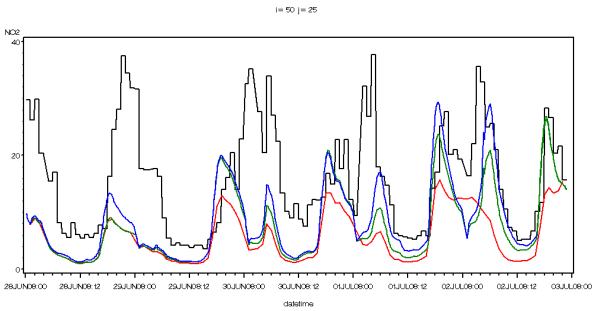

Fig 7: Graphs of the modelled and observed concentrations of NO_2 in $\mu g/m^3$ in Brno,CZ. Red: modelled concentrations, blue: optimized concentrations, green: forecast from optimized emissions, black: observed concentrations (weighted average of values from several background stations at the gridpoint).

Fig. 8 shows the concentrations in Prague, CZ. Here we can notice significant differences between observed concentrations during the weekend (Saturday, Sunday and Monday morning) and following weekdays. This was not well reflected in the original emission model and the assimilation routine produced corrections. The persistent model, however, used Sunday correction for Monday, giving rise to errors. Similar situation arises in many other places. We plan to develop a better forward model for the corrections in the future.

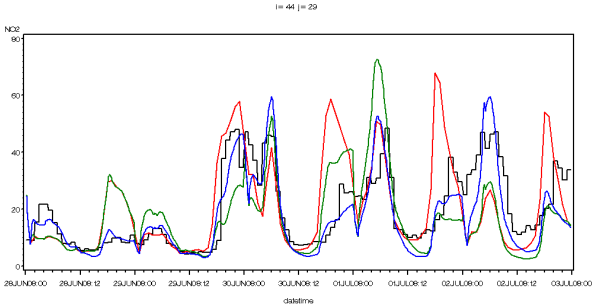

Fig 8: Graphs of the modelled and observed concentrations of NO_2 in $\mu g/m^3$ in Prague,CZ. The meaning of the colours is the same as in Fig 7.

The remaining graphs show some examples of the diurnal profile of the emission correction factor and of the original and optimized emission. For July 2 the assimilation process decreases the emissions of NO_2 in Prague,CZ over all day except late evening (Fig. 9). A different situation arises in Brno,CZ (Fig. 10) where the emissions have been increased for the whole day. An interesting situation occurs in places where the emissions are moved from one part of the day to another one (cf. Fig. 6 and the discussion there). Some examples are shown in Fig. 11 (Usti nad Labem,CZ) and Fig. 12 (Koln,DE) where a part of the emissions is shifted from the evening peak to morning and noon times.

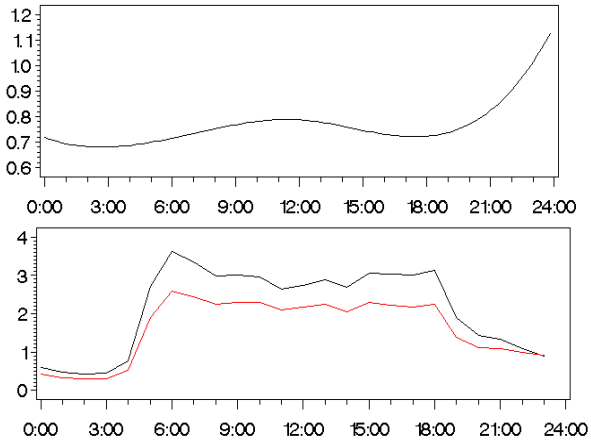

Fig 9: Graphs of the diurnal profiles of the emission correction factor of NO_2 (upper) and of the original (black) and optimized (red) emissions of NO_2 (lower) in Prague,CZ on July 2, 2008.

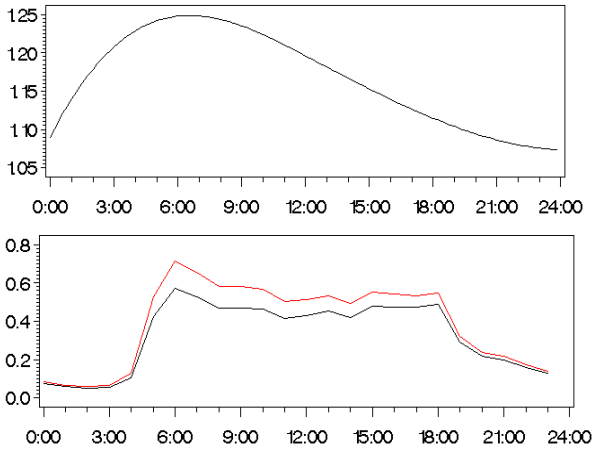

Fig 10: Graphs of the diurnal profiles of the emission correction factor of NO_2 (upper) and of the original (black) and optimized (red) emissions of NO_2 (lower) in Prague,CZ on July 2, 2008.

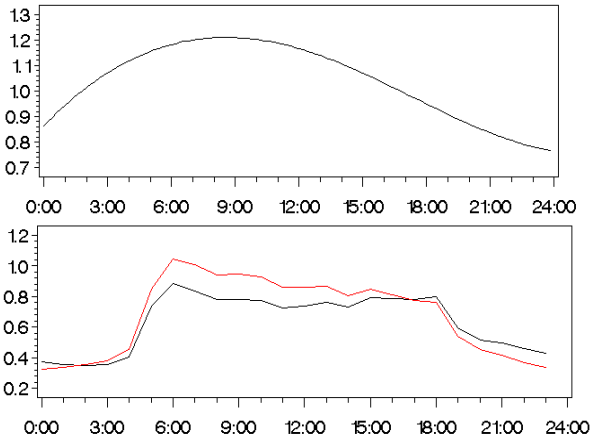

Fig 11: Graphs of the diurnal profiles of the emission correction factor of NO_2 (upper) and of the original (black) and optimized (red) emissions of NO_2 (lower) in Usti nad Labem,CZ on July 2, 2008.

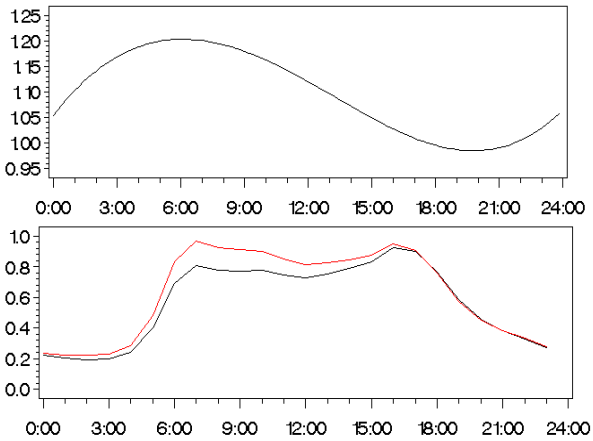

Fig 12: Graphs of the diurnal profiles of the emission correction factor of NO_2 (upper) and of the original (black) and optimized (red) emissions of NO_2 (lower) in Koln,DE on July 2, 2008.

An evaluation of performance of the method for the tested episode is given in Table 1 and Table 2 which show some summary statistics of the improvement of agreement between observations and modelled/analysed/forecasted concentrations of NO_2 resp. O_3.

NO_2	28.6.	29.6.	30.6.	1.7.	2.7.	3.7.	4.7.	5.7.
Model	5,4	5,9	7,3	9,4	10,0	9,4	6,8	6,6
Analysis	3,9	3,6	5,3	7,3	7,3	6,9	5,1	---
Forecast	---	4,3	6,2	8,3	8,3	7,6	6,0	5,2
Analysis %	**27,9**	**39,4**	**26,5**	**22,0**	**26,6**	**26,0**	**25,6**	**---**
Forecast %	**---**	**27,6**	**14,7**	**11,7**	**17,3**	**18,9**	**11,0**	**20,8**

Table 1: Mean absolute error of NO_2 compared with observations of the model, analysis and forecast in $\mu g/m^3$ and percentage decrease of the error of the analysis and forecast against error of the model.

O_3	28.6.	29.6.	30.6.	1.7.	2.7.	3.7.	4.7.	5.7.
Model	18,6	20,8	24,4	29,9	29,4	26,4	20,7	24,4
Analysis	16,4	17,4	18,4	23,5	24,9	21,6	16,0	---
Forecast	---	20,7	24,0	27,6	27,6	24,8	18,2	22,0
Analysis %	**11,9**	**16,4**	**24,9**	**21,4**	**15,2**	**18,2**	**22,8**	**---**
Forecast %	**---**	**0,3**	**1,6**	**7,8**	**6,2**	**6,1**	**12,1**	**9,7**

Table 2: Mean absolute error of O_3 compared with observations of the model, analysis and forecast in $\mu g/m^3$ and percentage decrease of the error of the analysis and forecast against error of the model.

The tables show a good efficiency and a stable performance of the method for both NO_2 and O_3, which thus confirms the ability of data assimilation to bring the model solution closer to the observations by means of changes of the emissions. An improvement of the forecast of NO_2 can be seen too. The contribution of the method to improvement of the forecast of O_3 is still only moderate but the increasing trend may be attributed to slower "learning" process of the method for O_3 then for NO_2.

5. BUILDING AN OPTIMIZED EMISSION MODEL

5.1 A long term simulation

The optimized emission fields can ultimately be used for building a data dependent top-down emission model. In Europe it is difficult to collect a detailed emission source inventory on a continental level. Instead, emission models based on coarse gridded yearly totals are being used. The proposed method can be used for correcting the totals and assigning each gridpoint its specific model of daily profiles. This task naturally requires a long term simulation. At present we run such a study for four months of 2007, with slightly different observation datasets. We run several different configurations of the method. First of all, different versions of parametrization of daily profiles are tested. As an example of early outputs of this study a comparison of different parametrizations of correction factors for three particular days is given in Fig. 13. This graph shows also the differences between emission corrections of weekend days and weekday (compare with discussion of Fig. 8).

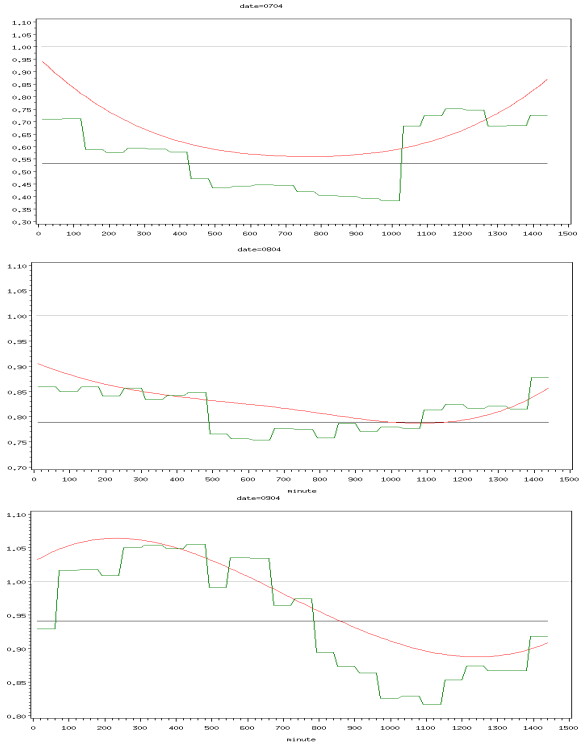

Fig 13: Graphs of the diurnal profiles of the emission correction factor of NO in Usti nad Labem,CZ on April 7.-9. 2007 (Saturday,Sunday,Monday). Optimization of emission daily total (black), five spline parametrization (red) and independent optimization of every hour (green).

A long term study should also clarify the relation of contributions of different data sources to the changes in emissions. In particular, we run a number of separate experiments where we assimilate just one source or group of sources (ground level/satellite, NO_2/O_3) and we compare the induced changes for synergy and influence. An example is in Fig. 14.

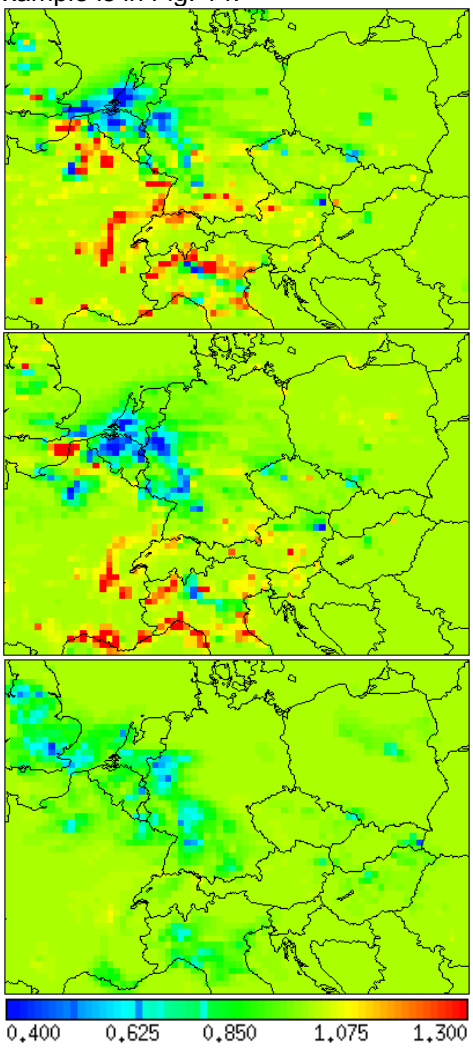

Fig 14: Emission coefficients with respect to the "midday" member of emission profile base (green in Fig. 1) for day April 8, 2008. Corrections from assimilation of NO_2 in-situ observations (upper), O_3 observations (middle) and satellite columns from OMI and GOME-2 (lower). The images are scaled to the same scale.

There is a generally good correspondence between changes induced by assimilation of ground level observations of NO_2 and O_3 in most places, whereas the changes induced by satellite observations show more substantial differences. These differences may be caused by insufficient satellite data for a given region and given time, by inadequate modelling of vertical transport or even by the retrieval process of satellite columns. These issues require further investigations.

5.2 Further considerations

Besides emission profile estimation, error statistics of the correction factor has to be derived from the experiment in order to construct a reasonable forward model for one-day ahead prediction of the correction factors. Using the optimized profile and improved forward model should bring some improvements in forecasting performance of the model, too.

6. CONCLUSIONS

Proper emission modelling is essential for a good performance of real-life air quality models. Although many issues still remain to be solved and long-term evaluation needs to be completed, the method gives a powerful tool for inverse modelling of time profiles of emissions. It can serve a base for building more sophisticated data-driven emission models.

7. LITERATURE

[Boersma 2007] *Boersma, K.F., H.J. Eskes, J.P. Veefkind, E.J. Brinksma, R.J. van der A, M. Sneep, G.H.J. van den Oord, P.F. Levelt, P. Stammes, J.F. Gleason and E.J. Bucsela, Near-real time retrieval of tropospheric NO_2 from OMI, Atm. Chem. Phys., 2013-2128, sref:1680-7324/acp/2007-7-2103, 2007*

[Courtier 1997] *Courtier, P.: Dual formulation of four-dimensional variational assimilation, Q. J. R. Meteorol. Soc., 123, 2449–2461, 1997*

[Eben 2005] *K.Eben, P.Jurus, J.Resler, M.Belda, E.Pelikan, B.C.Kruger and J.Keder: An ensemble Kalman filter for short term forecasting of tropospheric ozone concentrations, Q. J. R. Meteorol. Soc., 131, No. 613, 3313-3322.*

[Elbern 2007] *H. Elbern, A. Strunk, H. Schmid and O. Talagrand: Emission rate and chemical state estimation by 4-dimensional variational inversion, Atmos. Chem. Phys., 7, 3749–3769, 2007, 2007*

[Hakami 2007] *Amir Hakami, Daven K. Henze, John H. Seinfeld, Kumaresh Singh, Adrian Sandu, Soontae Kim, Daewon Byun, and Qinbin Li: The Adjoint of CMAQ, Environ. Sci. Technol., 2007, 41 (22), pp 7807–7817, 2007*

[Resler 2008] *J. Resler, K. Eben, P. Jurus and P. Krc: Inverse modeling of emissions using the CMAQ adjoint model, Extended abstracts, 7th Annual CMAS Conference, Chapel Hill, NC, USA, October 6-8, 2008*

[Sandu 2003] *Adrian Sandu, Dacian N. Daescu, Gregory R. Carmichael: Direct and adjoint sensitivity analysis of chemical kinetic systems with KPP: Part I—theory and software tools, Atmos. Environ. 37 (2003) 5083-5096., 2003*

[Weaver 2001] *Weaver, A. and Courtier, P.: Correlation Modelling on the Sphere Using a Generalized Diffusion Equation, Q. J. R. Meteorol. Soc., 127, 1815–1846, 2001*

Acknowledgement This work was supported by the Grant Agency of the Academy of Sciences of the Czech Rep. (grant No. 1ET400300414, framework „Information Society"), by the grant No. SP/1a4/107/07 of the Ministry of Environment of the Czech Rep. and by the Institutional research plan AV02 10300504 Computer Science for the Information Society "Models, Algorithms, Applications".