Inverse modeling of emissions using the CMAQ adjoint model

Jaroslav Resler, Kryštof Eben, Pavel Juruš and Pavel Krč

E-mail: resler@cs.cas.cz

Summary
The CMAQ adjoint model has been used to set up a data assimilating system based on the 4DVar method. Ground-level observations of NO2 and tropospheric NO2 columns retrieved from the satellite instruments GOME2 and OMI have been assimilated into the model. The 4DVar method has been used to optimize both initial conditions and emission factors. Simulations and forecasting experiments have been performed on mesoscale domains with different horizontal resolution.

Motivation
Chemistry transport models often show substantial bias and errors due to insufficient specification of inputs, lack of information in higher levels of troposphere, simplified chemistry etc.

Possible directions towards an improvement of performance:
- Optimization of the emission inventory and emission model
- Using the information contained in satellite data. They have better spatial coverage but about two overpasses per day only. The essential feature of satellite data is that they bring an important information about vertical distribution of species.

Means to achieve these tasks:
- Data assimilation and inverse modeling of emission

Adjoint of the CMAQ CTM
- Implemented in Cahets and Virginia Tech
- Includes adjoint code for gas phase (chemical mechanism CB4)
- Implementation by similar means as in STEM

Our contribution:
- Finalized parallelization of the code
- Fixed some bugs
- Implementation of the observation operator for satellite data
- What is still missing:
  - Some scientific processes
  - More chemical mechanisms (SAPRC99, CB5)
  - Aerosol processes
- Parallelization of the L-BFGS-B code
- Standard build process

Effectivity of the parallelization
The parallelized version of the adjoint and 4DVar code was tested for both correctness and performance in different configurations. Parallel efficiency of the code is about 0.85 for 16 parallel MPI processes. This test was performed for a 4DVar run on coarse assimilation domain on a Linux computer with 4xAMD Opteron 8350, 16GB RAM.

4DVar design
- We follow partially the approach of H. Eiblern
- This work is performed by D. Bombar et al. for the ESA’s Living Marine Resources project (LIFE project) Within the framework of the European In-ITN (HyAdjoint)
- Emisions optimized by one multiplicative factor for each gridpoint and day
- Imaging cost function: 
  \[ J = \sum_{i=1}^{N_t} \sum_{j=1}^{N_c} \left( y_{ij} - H(x_{ij}) \right)^2 + \lambda \sum_{i=1}^{N_t} \sum_{j=1}^{N_c} \left( x_{ij} - x_{ij}^* \right)^2 \]
- \( y_{ij} \) are the observed values, \( x_{ij} \) are observed values, \( x_{ij}^* \) are optimized and first guess concentrations in time \( t \)
- \( x_{ij}^* \) are optimized and first guess emission multiplicative factors
- \( H(x_{ij}) \) is the observation operator
- \( \lambda \) is the observation operator
- \( \lambda \) is the observation operator
- \( \lambda \) is the observation operator
- \( \lambda \) is the observation operator

Gradient with respect to \( x_{ij} \):
\[ \nabla J(x_{ij}) = \sum_{i=1}^{N_t} \sum_{j=1}^{N_c} \left( y_{ij} - H(x_{ij}) \right) \nabla x_{ij} + \sum_{i=1}^{N_t} \sum_{j=1}^{N_c} \lambda \left( x_{ij} - x_{ij}^* \right) \nabla x_{ij} \]

Results
Decay of the cost function
After the first two days, the decrease of the cost function is rather small.

Performance of the forecast
A forecast of next day’s concentrations was run each day, starting from optimized initial conditions and emission factors obtained by data assimilation for the previous day.

Optimized emission factors
After the first two days of assimilation, the emission factors converged to a fairly stable solution. The changes in some regions (e.g., France, Italy, Austria) can be attributed to satellite observations only.

Conclusions and outlook
The assimilation experiment shows the stability and good performance of this particular setting of the 4DVar method. The satellite data seem to be useful for constraining emissions and initial conditions even in the short term and fine resolution. The improvement in forecasting NO2 concentrations due to assimilation of in-situ and satellite observations is moderate so far. Longer experiments, more flexible parameterization of the emission corrections, adaptations of the emission model itself and a deeper investigation of complementarity between in-situ and satellite observations are required before using the technique in operational forecasting.

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. SP2016/1707/0 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”.