REGIONAL AIR QUALITY FORECASTING FOR THE CZECH REPUBLIC

Linton Corbet[1], Ondřej Vlček[1], Kryštof Eben[2], Jitka Liczki[2], Nina Benešová[1], Miloslav Modlík[1]

[1]Czech Hydrometeorological Institute, Prague, Czech Republic
[2]Institute of Computer Science, Academy of Sciences of the Czech Republic, Prague, Czech Republic

Abstract: This paper describes a 48-hour air quality forecast system that is targeted towards the Czech Republic. It is currently being tested for operational usage at the Czech Hydrometeorological Institute (CHMI). The air quality forecast (AQF) is designed to provide air quality predictions for the Czech Republic, and thereby compliment forecasts from MEDARD which also focuses on a similar domain. The forecast will principally be operated to supplement existing air quality information that is provided by the CHMI to the public.

This CHMI AQF is based on the chemical transport model CAMx. Meteorological inputs are obtained from the ALADIN NWP model, run operationally at the CHMI. Anthropogenic emissions are prepared from a number of sources including the Czech Republic's national emission inventory. This, in particular, allows detailed emission information to be processed for the Czech Republic. A modified version of SMOKE is used to process emissions. The emission processing work has allowed some emission sources to be released into above-ground layers with the aim of reflecting the real situation as best as practicable. Biogenic emissions are prepared using BEIS. Boundary conditions for the model are extracted from MACC products.

Performance evaluation of the forecast has not been completed and therefore results are not yet available. Nonetheless, the chosen configurations are documented.

Key words: *Air quality forecast, CTM, emissions, verification*

INTRODUCTION

The regulatory motivation behind operating an air quality forecast lies in the current European directive on ambient air quality and cleaner air for Europe (2008/50/EC) which directs member states to communicate air quality information (actual or predicted exceedances of alert or information thresholds) to a wide range of people and organizations (article 26, 1(a), which directs readers to Annex XVI). A forecast allows information on future air quality to supplement actual information coming from monitoring data.

Currently, and specifically for the Czech Republic, two operational AQFs exist. MEDARD (www.medard-online.cz) provides weather and air quality forecasts and operates with one 9km-resolution nested grid focusing on the Czech Republic. In the near future an enhanced version with two nested grids (the innermost down to 3km-resolution) should become publicly available.

The second AQF that we discuss here is focussed on Prague and has been developed within the framework of the PASODOBLE project (http://www.myair.eu). This forecast also employs a nesting approach. The innermost nest is at 1km-resolution and covers the Capital City of Prague. It uses different emissions, meteorology and a different chemical transport model to MEDARD.

The CHMI AQF is designed to compliment the existing provision of air quality information to the public. The CHMI, as the national institute for meteorology, hydrology and air quality, operates a network of monitoring stations and collaborates closely with local municipalities, the Health Institute and neighboring countries to richly observe air quality in the country and around selected border regions. To enhance the information that is provided to the public, and make further use of meteorological forecasts produced by the institute, an air quality forecast system has been developed.

Here, the configuration and development process of the CHMI's AQF, which uses CAMx as the chemistry transport model (CTM), is discussed.

THE CHEMICAL TRANSPORT MODEL

Domain and horizontal grid configuration

A coarse grid of resolution 14.13 km and an inner, nested, grid of resolution 4.71 km is used. The coarse grid domain is based on a Lambert projection having a single standard parallel at 46.245 N and the central meridian at 17.0 E, it has 171 columns and 135 rows and is centered over Rappottenstein in northern Austria. The coarse

CAMx domain is set to match the meteorological data domain. The inner grid has 62 rows and 131 columns and is centered over the state. The domains are pictured in Figure 1. No nesting is done in the vertical grid.

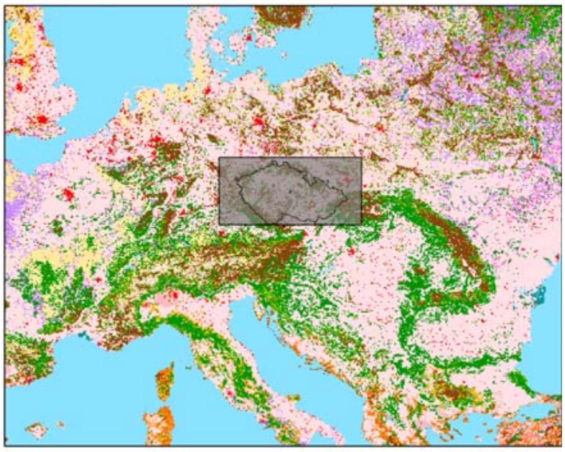

Figure 1. Coarse model domain and inner domain (shaded). Colored pixels denote different landcover categories from the GLC dataset. These are displayed for orientation only.

Model settings
CAMx v5.40 is used (Environ, 2011). Gas-phase chemistry is simulated using the SAPRC99 model. Aerosol chemistry is modelled using the coarse-fine approach (CF mode). Dry deposition is simulated using the Zhang03 resistance model for gasses (Zhang, 2001, 2003) however default leaf area index values are used.

EMISSIONS

Emission preprocessor
A system to prepare anthropogenic emissions was developed using US EPA emission model SMOKE (UNC, 2010). Besides SMOKE, this system contains several preprocessors that convert the emission inventories produced according to the European methodology into the format required by SMOKE. Programming these preprocessors was inevitable since SMOKE was intended as a powerful and flexible tool to generate gridded, temporally disaggregated and chemically speciated emissions in the USA and so it was strictly oriented to US standards.

Anthropogenic emissions
Czech national emission inventory REZZO for the year 2010 (CHMI, 2010) for locations inside the Czech Republic and TNO/MEGAPOLI emission database (Dernier van der Gon, 2009) for the year 2005 for the rest of the model domain were used.

The emission inventories gathered in REZZO database provide detailed information on emission sources that allows precise emission modelling. From the point of view of emission modelling there are three big groups of sources in REZZO database. Each of these groups was treated in a different way. Large and mid-size sources of emissions from combustion processes and technologies were regarded as point sources. Besides exact geographical location, other attributes (as stack height and diameter, gas velocity and temperature) are at our disposal. These attributes are crucial for vertical allocation of emissions. The emissions from small sources (local heating, solvent use) are estimated for cadastral units and they were treated as area emissions. Their spatial location onto the model domain was made using surrogates based on data from census return (housing). Mobile source emissions include emissions from both on-road and non-road mobile sources. They were prepared applying vehicle census, fuel sale and consumption. Then they were spatially disaggregated onto the grid with 0.01 degree horizontal resolution.

TNO/MEGAPOLI database contains surface gridded emissions with resolution cca 7 x 7 km and emissions from major European point sources. Since TNO/MEGAPOLI database does not contain stack parameters, these had to be completed according to the data from the REZZO database. Large point sources from the REZZO database were divided into 18 groups. Each group was characterized by both SNAP code and the amount of emissions of dominant pollutants. Emission-weighted average of stack height and diameter, gas velocity and temperature were then calculated and assigned to the point sources from the corresponding groups in the TNO/MEGAPOLI database.

Figure 2 provides an indication of the spatial differences between the two emission databases. VOC emissions into first model layer at 7 UTC July 1st are displayed. TNO/MEGAPOLI emissions include small point sources, area sources and transport whereas for REZZO just emissions from local heating, solvent use and transport are displayed. Even for this coarse 9-km resolution major roads can be clearly seen when REZZO database is used.

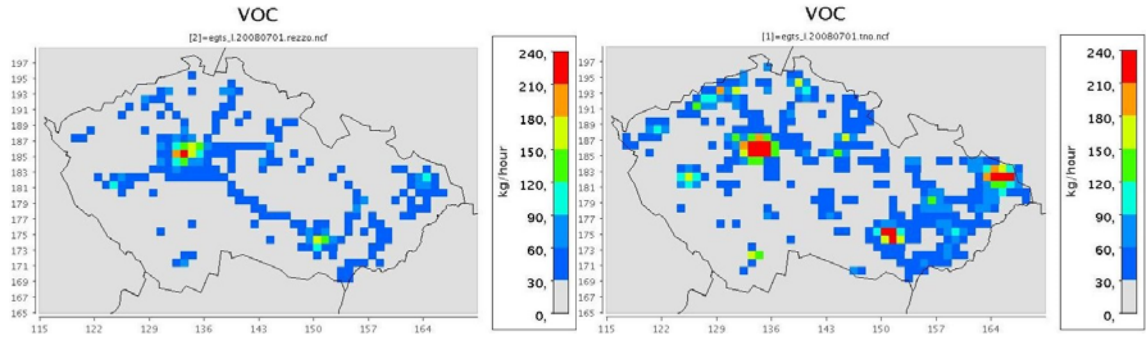

Figure 2. VOC emissions into first model layer at 7 UTC July 1st. TNO/MEGAPOLI emissions (right) include small point sources, area sources and transport. Czech national emissions REZZO (left) include local heating, solvent use and transport.

Biogenic emissions

Biogenic emissions were prepared using BEIS model v3.09, which is a part of SMOKE. BEIS requires a special format of input data that consists of: gridded land cover data, emission factors for vegetation types and actual meteorological values. Land cover data for the domain of interest were prepared from the grid with resolution of 0.01 degree, with the main type of vegetation on each cell of the grid. It was based on the data gathered during AFOLU project for EU (Köble and Seufert, 2001). Using a wide range of literature, the emission factors for all of 230 types of vegetation were computed (Zemankova , 2010).

Temporal processing

Temporal processing of the emission information takes advantage of source characteristics for large point sources from REZZO, otherwise it follows temporal profiles given in Dernier van der Gon (2011, Appendix 1). Hourly biogenic emissions are computed from actual meteorological values (temperature, radiation, humidity).

Chemical speciation

The chemical speciation for SAPRC99 mechanism was done as follows: gas pollutants were speciated according to the emission model of B. Krüger from Universität für Bodenkultur, Wien (personal communication). PM speciation was based on the first order bulk composition profile for the gridded PM_{10} and $PM_{2.5}$ data delivered by TNO under the GEMS project. For biogenic emissions SMOKE's standard speciation profile was applied.

METEOROLOGY

At the CHMI, the limited area numerical weather prediction model ALADIN is run operationally and produces 54h forecasts 4 times a day (00, 06, 12, 18 UTC). The current configuration of the AQF makes use of the first 48 hours of the 00UTC weather forecast to produce the forecast results.

The AQF coarse domain has the same extent as the ALADIN data, and the ALADIN data resolution matches that of the inner grid of the AQF. Therefore, in the horizontal, ALADIN data is aggregated for the coarse domain and nothing is needed to be done for the inner domain. The first 68 vertical ALADIN levels are aggregated into 20 vertical levels for CAMx which means the highest CAMx level is approximately 10km above ground level. The height of the lowest CAMx level is approximately 10m.

Vertical diffusivity fields are calculated from the ALADIN meteorological data using the integration methodology employed in CMAQ model (Byun and Schere 2006). A future task will be to test other formulation methods for this important input field however more work is needed to resolve details in the validation process that will be used.

To generate initial conditions for the next air quality forecast cycle, the other meteorological forecasts (06, 12, 18 UTC) are used to bring the CAMx model state to 00UTC using the most recent, and therefore most accurate, meteorology that is available. In other words, the 48-hour forecast is run from 00UTC. Separate to that, the 06 UTC forecast meteorology is used to move the predicted CTM state at 06 UTC (initial conditions taken from the main forecast run) to 12 UTC, the 12 UTC forecast is used to move the CTM from 12 to 18 UTC, and the 18

UTC forecast is used to move the CTM to the beginning of the next day. This leaves a more accurate model state (in comparison with using the main forecast's 23h prediction) to begin the next forecast with.

OTHER INPUT DATA

Boundary conditions

Boundary conditions are gathered from the Monitoring Atmospheric Composition and Climate (MACC) project, made available by the Forschungszentrum Jülich Web Coverage Service (WCS). The MACC project is a quasi-operational service to provide analyses and forecasts of the global atmospheric chemical composition (greenhouse gases, aerosols and reactive gases) based on a sophisticated modelling and data assimilation system running at the European Centre for Medium Range Weather Forecast (ECMWF).

The *MACC fnyp 3-hourly forecast* data set is used in our configuration. Experiment *fnyp* is a data assimilation/forecasting system providing a 5-day forecast. In *fnyp* MOZART v3.5 is coupled to the ECMWF's IFS to provide chemical tendencies. The IFS operates at a resolution of roughly 80 x 80 km and the MOZART is operated at 1.25° resolution (roughly 95 x 110 km over Europe). The dataset made available by the Jülich WCS does not provide all CAMx species however it does provide many of the most important ones: black carbon, organic carbon, dust in three sizes, seasalt in three sizes, sulfate, ethane, formaldehyde, methane, carbon monoxide, nitric acid, isoprene, nitrogen monoxide, nitrogen dioxide, ozone, hydroxyl radical, peroxyacytyl nitrate and sulfur dioxide. Seasalt and dust concentrations from the MACC dataset are apportioned into the relevant CAMx size bins (coarse/fine) using a log-normal distribution equating cumulative mass with particle diameter. It is noted that CAMx will use the minimum values for species not described in the boundary condition file. Should this cause problems a constant climatological file will be prepared for boundary conditions.

The WCS service has been found to be convenient as a source for boundary condition data because MACC forecast results are transferred from ECMWF to Juelich where they are reformatted and published, and the service allows users to extract only a portion of the global MACC domain to expedite download times and other processing tasks. Furthermore, 3-hourly data is available which allows the temporal profile to be better represented at the boundary, when compared to other global model datasets that provide 6-hourly fields.

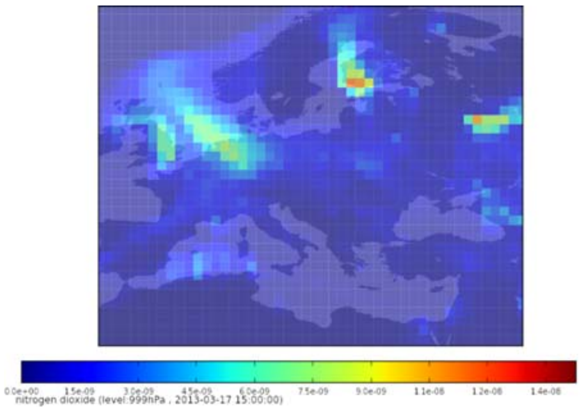

Figure 3. Example of forecast mixing ratio of NO_2 from the MACC project (18/03/2013 15:00).

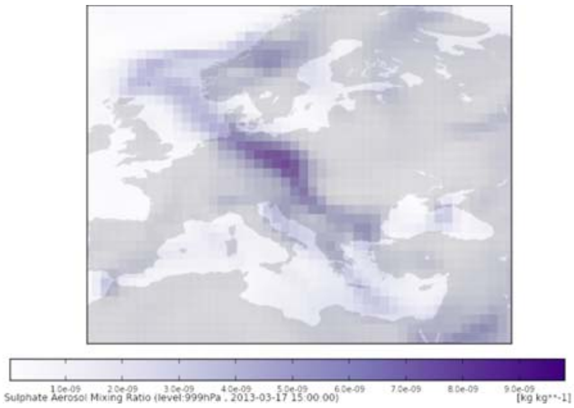

Figure 4. Example of forecast mixing ratio of sulfate aerosol from the MACC project (18/03/2013 15:00).

Landuse information

Three landuse files have been prepared. A 26-category and 11-category file were derived from USGS landuse data via a WRF to CAMx processing tool. The 26-category file will allow the dry deposition model based on the work of Zhang (2001, 2003) to be employed. The 11-category version will enable the older model based on Wesely (1989) and Slinn and Slinn (1980) to be used (see CAMx user's guide). The third landuse file is also of the 26-category type and was derived using, primarily, CORINE 2006 landcover information. Global Land Cover (GLC) 2000 landcover information was used to extend the spatial extent of the CORINE database for the entire CAMx domain. This third file includes relative percentages of each landuse category (based on the finer resolution CORINE data (250 m), and GLC data (1000 m)) in the grid cell therefore it better represents the sub-grid cell landuse variability. As yet, no custom leaf area index information has been included in the dataset.

Use of the Zhang model for dry deposition allows users to provide specific leaf area index (LAI) values to CAMx. A future task is to process MODIS LAI tiles to produce a climatological dataset for the domain. The key hurdle here is the filling of gaps in the satellite data which can be caused by many things, including clouds. At this stage it is proposed to use both high and good quality pixels and prepare daily LAI values for each pixel using the average of values in moving 25-day window. Multiple years of data would be processed to ensure the highest likelihood of having valid data for each pixel in the time window.

The landcover categories of CORINE and GLC required translation into the 26 CAMx categories. Generally, equivalent categories were self-evident. When this was not the case, only the key features driving differences in deposition were considered (in lieu of conducting a detailed assessment of equivalence) therefore differentiation between deciduous and evergreen vegetation and between bare land, grasslands, shrubs and forests were concentrated on.

AVAILABILITY OF RESULTS

The results of the AQF are expected to be ready for processing and dissemination around 8 UTC. This accommodates delivery of meteorological fields at approximately 04 UTC, processing of CTM inputs (meteorology is used in emission calculations, for example) taking 30 minutes, runtime of the CTM (approximately 3.5 hours). Processing and quality control tasks might also require some time leading to the results eventually becoming publicly available by 9 UTC, 10 CET.

REFERENCES

Byun D., K. L. Schere, 2006: Review of the governing equations, computational algorithms, and other components of the models-3 Community Multiscale Air Quality (CMAQ) modeling system. *Applied Mechanics Reviews*, **59**, 51–77

CHMI, 2010: REZZO Czech National Register of Emissions, Czech Hydrometeorological Institute

Denier van der Gon, H. A. C., A. J. H.Visschedijk , H. van der Brugh and R. Dröge, 2009: A base year (2005) MEGAPOLI European gridded emission inventory (1st version), MEGAPOLI Scientific Report 09-02, on-line at http://megapoli.dmi.dk/publ/MEGAPOLI_sr09-02.pdf

Dernier van der Gon, van der, H. D., C. Hendriks, J. Kuenen, A. Segers and A. Visschedijk, 2011: Description of current temporal emission patterns and sensitivity of predicted AQ for temporal emission patterns. EU FP7 MACC deliverable report D_D-EMIS_1.3

ENVIRON, 2011. User's guide to the Comprehensive Air Quality Model with Extensions (CAMx). Version 5.4. Available at: http://www.camx.com.

Köble, R. and G. Seufert, 2001: Novel maps for forest tree species in Europe. Proceedings of the 8th European Symposium on the Physico-Chemical Behaviour of Air Pollutants: "A Changing Atmosphere!", Torino (It) 17-20 September 2001. Data set available online at: http://afoludata.jrc.ec.europa.eu/index.php/dataset/detail/66

Slinn, S. A. and W. G. N. Slinn, 1980: Predictions for particle deposition on natural waters. *Atmos. Environ.*, **24**, 10131016.

UNC, 2010: SMOKE v2.7 User's Manual, University of North Carolina at Chapel Hill

Wesely, M. L., 1989: Parameterization of Surface Resistances to Ga seous Dry Deposition in RegionalScale Numerical Models. *Atmos. Environ.*, **23**, 1293–1304.

Zemankova, K., 2010: Study of links between biogenic VOC emissions and concentration of tropospheric ozone, Dissertation thesis, Charles University in Prague, http://is.cuni.cz/studium/eng/dipl_st/index.php?doo=detail&did=43503

Zhang, L., J. R. Brook, and R. Vet, 2003: A revised parameterization for gaseous dry deposition in airquality models. *Atmos. Chem. Phys.*, **3**, 2067–2082.

Zhang, L., S. Gong, J. Padro and L. Barrie, 2001: A sizesegregated particle dry deposition scheme for an atmospheric aerosol module. *Atmos. Environ.*, **35**, 549–560.