SPECIAL PROJECT FINAL REPORT

Project Title:	GEMS: Global and Regional Earth-System Monitoring using Satellite and in situ data Work Package WP RAQ 2 MACC: Monitoring Atmospheric Composition and Climate
Computer Project Account:	SP DEFRIU
Start Year - End Year :	2009 - 2011
Principal Investigator(s)	Hendrik Elbern
Affiliation/Address:	Rhenish Institute for Environmental Research at the University of Cologne, RIUUK Aachener Str. 209 50931 Cologne Germany
Other Researchers (Name/Affiliation):	Elmar Friese Rhenish Institute for Environmental Research at the University of Cologne, RIUUK

Summary of project objectives

Project work is exclusively restricted to the RIUUK contribution to the work packages WP_RAQ2, WP_RAQ4, and WP_RAQ5 of GEMS and to the work packages R-EDA, R-ENS, and R-EVA of MACC. The continental-to-local scale chemical data assimilation EURAD-IM system with both its full 4d-var and 3d-var assimilation options is implemented. The system operates to perform the tasks invoked by GEMS and MACC for RIUUK. The tasks include development and delivery of improved data assimilation methods to assure optimal exploitation of observations, assimilation of available data from both in situ and satellite devices, provision of pre-operational NRT daily air quality forecasts, and performance of regional air quality analyses for the years 2007-2009.

Summary of problems encountered

Additional computing resources were needed to enable delivery of the air quality services until the end of 2009. A total amount of about 4,000 SBUs per day is needed for the complete model chain. The underestimation of compute resources was caused mainly by two issues: first, the computing time needed for the air quality simulations has been extrapolated too optimistically from our experiences with a coarser horizontal model resolution of 45 km. Second, about 85,000 SBU's were wasted by incompatible interaction with the file system monitoring software, when eventually switched on at ECMWF compute facilities. The problem was solved by the user support. Thirdly, following the encouragement to simulate the Eyjafjallajökull volcanic ash dispersion over Europe, expressed by EC officials, a further consumption of 15,000 SBUs has occurred

Experience with the Special Project framework

We are satisfied with the application procedure and with the requirements on project reporting.

Summary of results

1. European environmental data assimilation

GEMS sub-project WP_RAQ2 and MACC sub-project R-EDA developed and delivered improved data assimilation methods to assure optimal exploitation of observations. The ability to provide analysed chemical fields along with locally and temporally resolved error estimates is a key objective. The tasks scheduled include provision of observation forward operators and its adjoint operators for various types of routine observations and retrievals, and related error covariance.

1.1 Observation forward operator for the aerosol optical depth

An observation forward operator for the aerosol optical depth has been developed, which calculates aerosol optical properties during model execution. Aerosol particles are treated as internally mixed. Refractive indices of the individual aerosol species are averaged according to their weight fraction. Inorganic aqueous species are merged to one component whose refractive index is computed with the Biermann (2000) mixing rule. A fast approximation according to Evans and Fournier (1990) is used for calculation of extinction efficiency and a 50 point Gauss-Legendre quadrature is used to integrate the extinction efficiency with respect to particle size. Figure 1 shows the aerosol optical depth at 550 nm for 6th August 2007 at 18:00 UTC calculated with the developed observation operator.

1.2 Improvement of air quality forecasts by data assimilation techniques.

Principally two data assimilation experiments were conducted to investigate to what extent air quality forecasts can be improved due to application of advanced data assimilation techniques. The data assimilation techniques applied, are intermittent 3d-var and 4d-var optimization of initial values and emission factors. The data assimilation system, including the formulation of the background error covariance matrix, is described in Elbern et al. (2007).

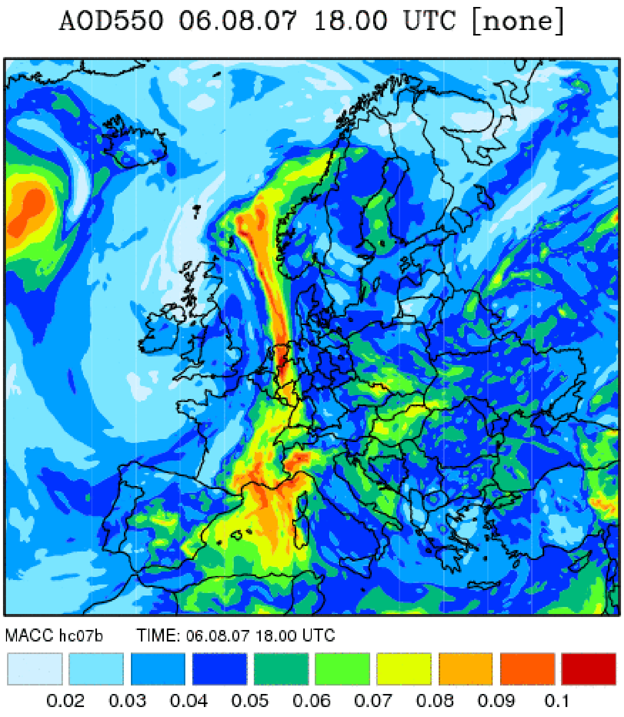

Figure 1: Aerosol optical depth at 550 nm for August 6, 2007 at 18:00 UTC.

1.2a 3d-var assimilation of in situ data and OMI NO_2 columns

The impact of chemical initial values originating from a 3d-var analysis of the previous day on today's forecast quality has been studied for the period September 16 to September 22, 2008. Horizontal model resolution for this experiment has been set to 45 km. In situ measurements of O_3, NO, NO_2, SO_2, and CO in conjunction with OMI NO_2 column retrievals were assimilated each hour using the intermittent 3d-var method. Concentrations of O_3, NO_2, SO_2, and CO derived first from a free forecast initialized with the model state D+1 of the free forecast from the previous day, and second derived from a free forecast initialized with the final model state of the 3d-var analysis of the previous day are compared with hourly measurements at all available stations from the federal environment agency (UBA). Figure 2 show time-series of spatially and temporal averaged concentrations of O_3, NO_2, SO_2, and CO. After six hours the impact of data assimilation on the forecast quality of O_3 and NO_2 is insignificant due to advection of air masses which were not subject of the assimilation procedure as well as due to fast chemical reactions. The influence of chemical reactions cannot be covered by initial values derived from a 3d-var analysis. In case of the low reactive species SO_2 and the persistent species CO the impact of data assimilation on the forecast quality is significant during the first 12 hours at least.

1.2b 4d-var assimilation of NO_2 column retrievals

In contrast to stratospheric constituent data assimilation and general meteorological data assimilation, the evolution of the tropospheric state of air pollutants is not primarily controlled by the initial state. Instead, emissions are a strong controlling factor, and exert a direct influence over short timescales (ranging from seconds to days). Furthermore emission rates are not sufficiently well known. Thus, emission rates must be considered as another parameter to be optimized in the data assimilation process, as shown by the work of Elbern and Schmidt (2001) and Elbern et al. (2007) with the EURAD-IM, and as it is applied in this study.

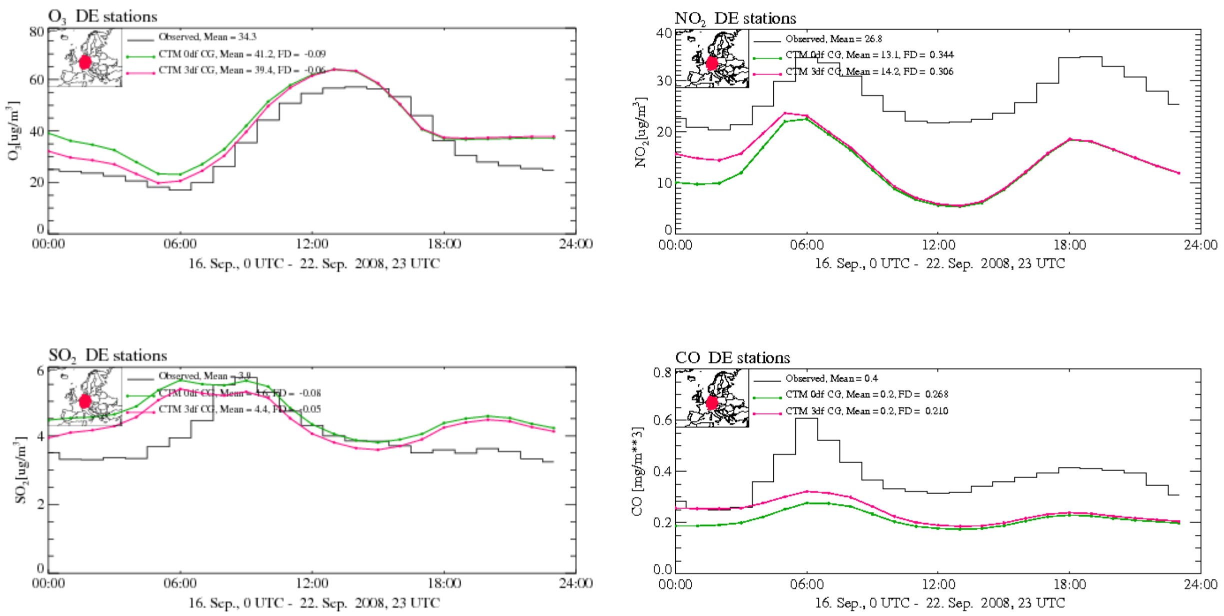

Figure 2: Hourly O_3 (upper left), NO_2 (upper right), SO_2 (lower left), and CO (lower right) concentrations averaged over the period from September 16 to 22, 2008 averaged for all available stations from the German UBA. Green: Model results from a free forecast initialized with the model state D+1 of the forecast from the previous day, red: model results from a free forecast initialized with the final model state of a 3d-var analysis of the previous day, black: measurements.

Description of the experiment

Both SCIAMACHY and OMI satellite retrievals from KNMI were assimilated for the period July 2006 by averaging kernels, using error information from the data provider. In an attempt to provide a horizontal model resolution comparable to the minimal OMI 24x13 km^2 footprint, the horizontal model resolution was set to 15 km. The data assimilation configuration has a time window of three hours (09:00–12:00 UTC) to include Envisat data with a late morning overpass over Europe and OMI overpass in early afternoon over Eastern Europe. After assimilation, a 24-h forecast is made, starting at 08:00 UTC; the analysis produced by the assimilation is the initial field, and an emission rate correction factor is applied. After assimilation, an a posteriori analysis is performed.

While the SCIAMACHY footprint pattern covers the western model domain due to its late morning orbit, covers the OMI footprint pattern eastern parts due to its early afternoon overpass. Figure 3 exhibits these conditions, along with retrievals (**y**), forecasted retrievals of NO_2 columns ($\mathbf{Hx}_b$), and analysed tropospheric NO_2 columns ($\mathbf{Hx}_a$). The analysis result can be clearly identified as a weighted combination of both information sources.

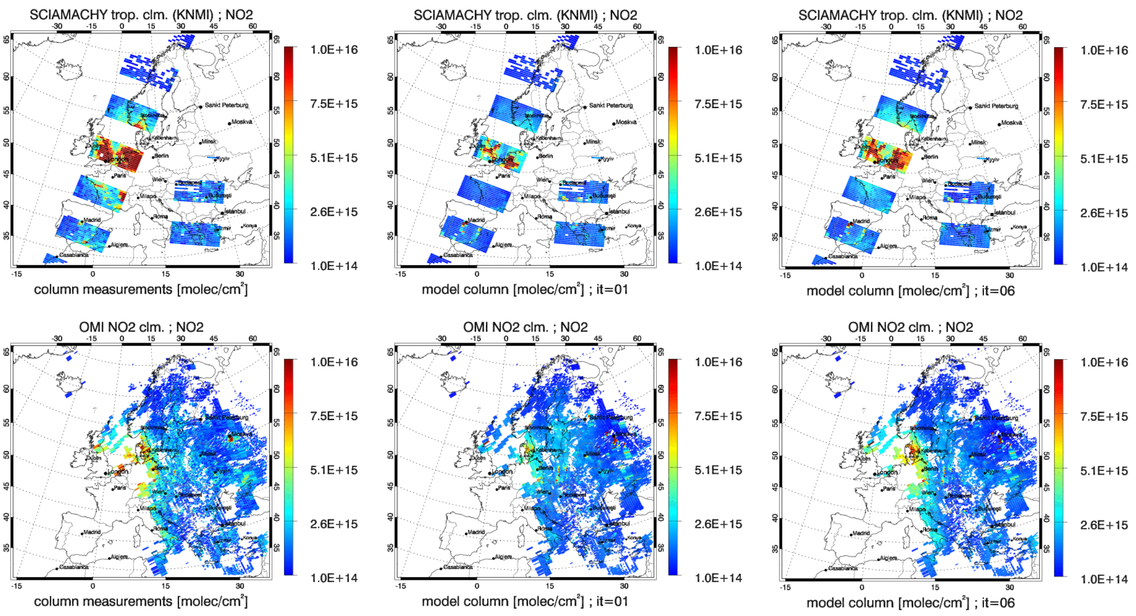

Figure 3: Comparison of NO_2 tropospheric columns in molecules/cm2 for July 6, 2006. Left panel column: KNMI retrieved and assimilated values ($\mathbf{y}$); middle panel column: EURAD-IM forecasted values ($\mathbf{Hx}_b$); right panel column analyses ($\mathbf{Hx}_a$).

Figure 4 show the effect of data assimilation for July 6, 2006. Clearly major increments can be observed in western England and in the area of north western Russia. Both these signals are visible for surface concentrations. Significant tropospheric colum amounts over central Europe are probably a retrieval phenomenon, as assimilation increments of opposite signs follow two neighbouring OMI tracks.

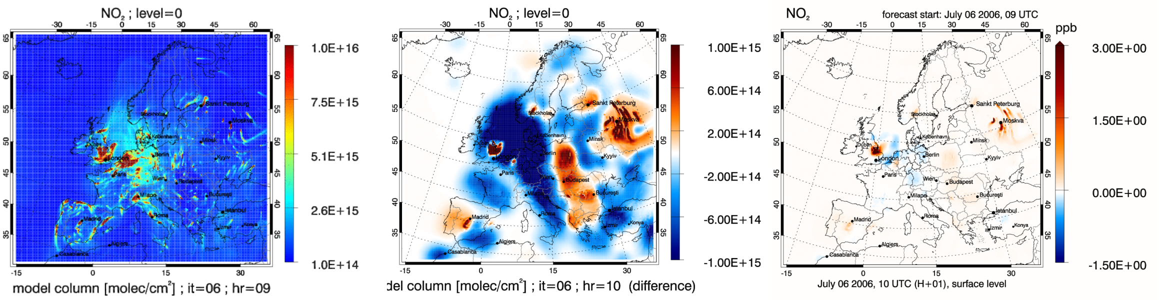

Figure 4: Data assimilation result in terms of tropospheric columns for July 6, 2006. NO_2 model columns based on OMI and SCIAMACHY assimilation within the assimilation interval 09-12 UTC. Units in molecules/cm^2 (left panel). Difference field giving implied changes for tropospheric columns by assimilation (middle panel), and induced surface concentration changes in ppb.

Depending on the evaluation space, the impact of the earth observation data on the model performance is different:

1. *Minimisation of costs with averaging kernel based cost function*: There is a fully satisfying result for ingesting averaging kernel based NO_2 information. Figure 5 presents the cost improvement for forecasts and analyses. After two days of spin up run, July 3 is assimilated first. Analysis reduction to 65% can be attained, as both NO_2 initial values and emission rates could as yet not benefit from any observational information. For the following days, it can be seen that there is only a rather small forecast improvement from day to day, (with

even an outlier on July 7). However, the analyses improve, approaching a cost reduction to 30% with respect to the control run.

2. *Minimisation of costs with averaging kernel based cost function*: There is a fully satisfying result for ingesting averaging kernel based NO_2 information. Figure 5 presents the cost improvement for forecasts and analyses. After two days of spin up run, July 3 is assimilated first. Analysis reduction to 65% can be attained, as both NO_2 initial values and emission rates could as yet not benefit from any observational information. For the following days, it can be seen that there is only a rather small forecast improvement from day to day, (with even an outlier on July 7). However, the analyses improve, approaching a cost reduction to 30% with respect to the control run.

3. *OmF and OmA pdfs*: The overall result presented in Figure 5 can be corroborated quantitatively by the OmA and OmF pdfs displayed in Figure 6. As two typical days, July 6 and July 8, 2006 are selected. The marginal forecast improvement is reflected by the close coincidence of the observation-minus-control run pdf (OmC). However, it is visible that a slight bias reduction occurs. Clearly, the analysis shows a significant improvement in terms of both, bias and variance reduction.

4. *Forecast improvement validated against surface data*: A clear validation result indicating forecast skills improvements by surface in situ stations is hard to obtain. On the one hand the averaging kernel with only 10% average signal strengths at the ground implies only minor impact in near surface layers, while on the other hand NO_2 is reacting too fast for displaying a direct benefit for the next days forecast. As concerned the indirect effect via ozone improvement, two cases must be distinguished: In the rural case, ozone formation is often controlled by low NOx levels, where relative variance of retrievals is high and assimilation is difficult. In the case of urban areas, the spatial resolution problem of the model and its emissions prevail.

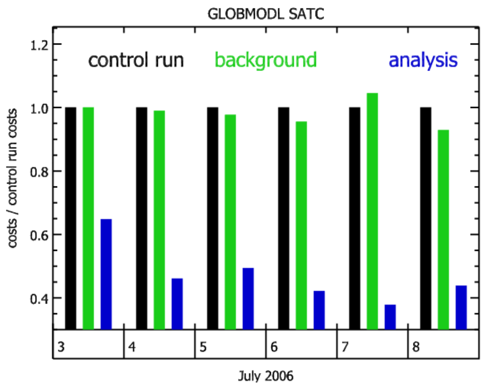

Figure 5: Objective function based normalized costs of combined OMI and SCIAMACHY assimilation runs between July 3 to 8, 2006. Black bars: control run without any data assimilation, for reference and normalisation to value 1 only. Green bars: one day forecast costs. Blue bars: analyses costs.

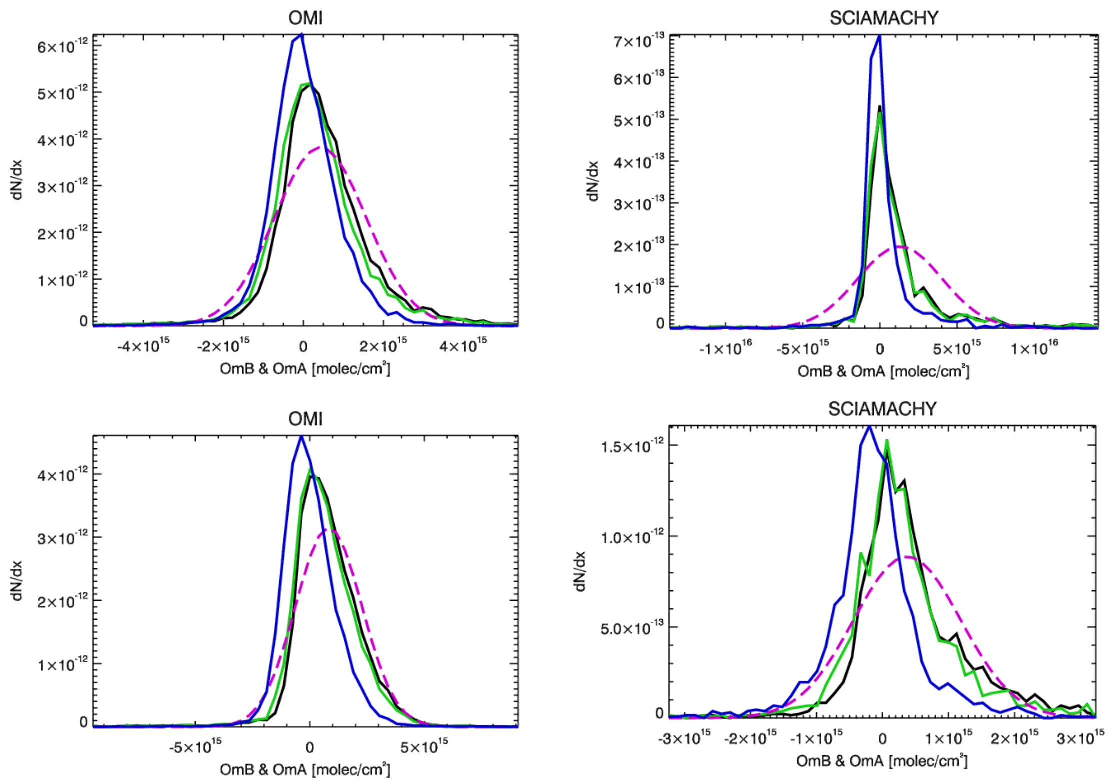

Figure 6: Probability density functions (pdf) of differences between observations and model runs for July 6 (top line), and July 8, (bottom line). Left panels show results based on OMI retrievals, right panels show results for SCIAMACHY retrievals. Each panel displays observation differences with the control run (OmC) (no data assimilation at all,) black bold line, differences between observations and forecasted values (OmF) green bold line, differences between observations and analyses (OmA) blue bold line. For comparison Gaussian fit to OmF pdf by mean and standard deviation given by broken purple line.

2. Pre-operational air quality forecast and analysis

GEMS sub-project WP_RAQ4 and MACC sub-project R-ENS primarily focus on delivering a pre-operational European-scale regional air quality service. The tasks performed include

- Provision of daily 3d-var air quality analyses with 15 km resolution over Europe for the previous day, using the system developed and tested within sub-projects WP_RAQ2 and R-EDA. For this purpose near real-time surface in situ measurements of O_3, NO_2, NO, SO_2, CO, and PM_{10} from the France and German environmental protection agencies as well as satellite derived NO_2 columns from OMI, GOME-2, and SCIAMACHY are hourly assimilated using the intermittent 3d-var method. The EURAD-IM daily air quality analysis is running pre-operational and results are accessible on the MACC web page.

- Provision of pre-operational daily air quality forecasts for Europe. This continental scale near real-time air quality forecast, based upon the IFS operational meteorological forecast, has been set up using the EURAD-IM model system with 15 km resolution. A 72 h forecast of the concentrations of O_3, NO, NO_2, CO, SO_2, PM_{10} and optionally concentrations of the particulate matter constituents $PM_{2.5}$, SO_4^{2-}, NO_3^-, NH_4^+, secondary organics, sea salt, mineral dust and primary anthropogenic particles are regularly delivered. Predicted concentrations of O_3, NO, NO_2, SO_2, and PM_{10} contribute to a multi-model ensemble forecast. Chemical initial values are provided by the pre-operational 3d-var analysis of the previous day. Chemical boundary values are taken from the pre-operational global

MOZART-IFS air quality forecast. The EURAD-IM daily air quality forecast is running and results are accessible on the MACC web page.

MACC sub-project R-ENS furthermore includes continued development of the pre-operational air quality forecast and assimilation systems. Within this scope the previously used GEMS TNO emission inventory for the base year 2000 has been replaced by the newly available TNO MACC emission inventory for 2005. As the most striking result, the new emission inventory prevents over estimation of SO_2 near surface concentrations especially over Eastern Europe. Figure 7 exemplary shows time series of hourly SO_2 concentrations averaged over all available Polish measurement sites for the time period between August 1 and 15, 2007. The forecast of SO_2 has been significantly improved due to application of the MACC TNO emission inventory with base year 2005.

A further step in development of the EURAD-IM air quality forecast and analysis systems is the ability to use aerosol boundary values taken from the pre-operational IFS aerosol forecast. Concentrations of mineral dust (represented by using three size bins), hydrophobic and hydrophilic organic matter and black carbon, and sulphate from the IFS aerosol model are used to prepare boundary values for the corresponding EURAD-IM aerosol species with a time-step of 6 hours. The influence of aerosol boundary values taken from the IFS aerosol model on PM_{10} concentrations has been studied in a EURAD-IM forecast of August 2007.

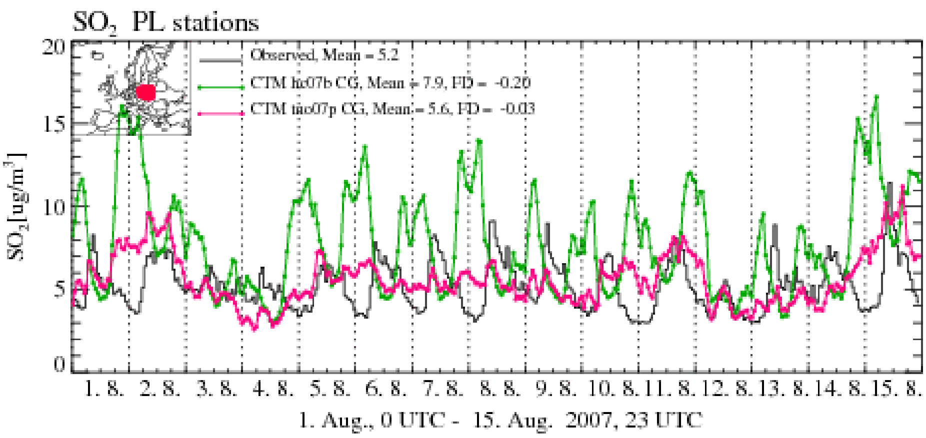

Figure 7: time-series of SO_2 concentrations averaged over all available Polish measurement sites for the time period between August 1 and 15, 2007. Black: observations, green: EURAD-IM forecast based on the TNO GEMS emission inventory with base year 2000, red: EURAD-IM forecast based on the TNO MACC emission inventory for 2005.

Figure 8 shows daily averages of PM_{10} concentrations averaged over all available Spain measurement sites for August 2007. During the first as well as last week of August 2007, the PM_{10} forecast is improved due to the use of aerosol boundary values taken from the IFS aerosol re-analysis. Both time periods exhibit significant transport of mineral dust across the southern model boundary.

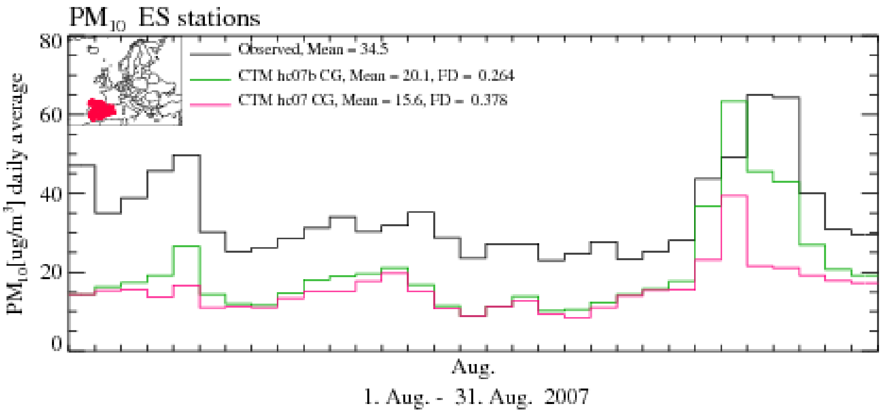

Figure 8: Daily averages of PM_{10} concentrations averaged over all available Spanish measurement sites for August 2007. Black: observations, green: EURAD-IM forecast with aerosol boundary values taken from the IFS aerosol re-analysis, red: EURAD-IM forecast with climatological aerosol boundary values.

3. Validated assessment of air quality in Europe

MACC sub-project R-EVA is focused on the evaluation of air quality in Europe, established on validated observational data, and also on operational air quality model results. The overall objective of this sub-project is an operational chain using models and data assimilation methodologies for simulating realistic air quality patterns in Europe. The final products of the sub-project are annual assessment reports, describing the state and the evolution of background concentrations of air pollutants in European countries. Special attention is given to pollutants characterized by the influence of long range transport, correctly caught by European scale modelling: O_3, NO_2, $PM_{2.5}$, and PM_{10}. An ensemble approach is used to provide the best possible representation of these pollutants. The EURAD-IM 3d-var assimilation system contributes to the ensemble air quality re-analyses for the years 2007-2009. The re-analyses for the three years have been accomplished. In situ data (Airbase surface measurements and air-borne MOSAIC data) as well as satellite retrievals (NO_2 column retrievals from OMI, GOME-2, SCIAMACHY, and MOPITT CO profiles) are assimilated every hour using the intermittent 3d-var technique. About 30% of surface in situ background stations were held back from assimilation. Figure 9 shows bias and root mean square error of daily averaged O_3, NO_2, and PM_{10} concentrations averaged over all Airbase background measurement sites, which were held back from assimilation for the year 2009.

4. Technical description of code

The model for the computational work done in the work packages of GEMS and MACC is the EURAD-IM model system. The EURAD-IM model system consists of 5 major parts: the meteorological driver MM5 (version 3), the pre-processors EPC and PREP for preparation of meteorological fields and observation data, the EURAD-IM Emission Model EEM and the chemistry transport model EURAD-IM. The data flow of the EURAD-IM system is depicted in Figure 10. The EURAD-IM is a mesoscale-α chemistry transport model involving transport, diffusion, chemical transformation, wet and dry deposition and sedimentation of tropospheric trace gases and aerosols.

EURAD-IM allows for three-dimensional variational data assimilation (3d-var) and for optimization of chemical initial values and/or emission factors using the four-dimensional variational data assimilation (4d-var) method.

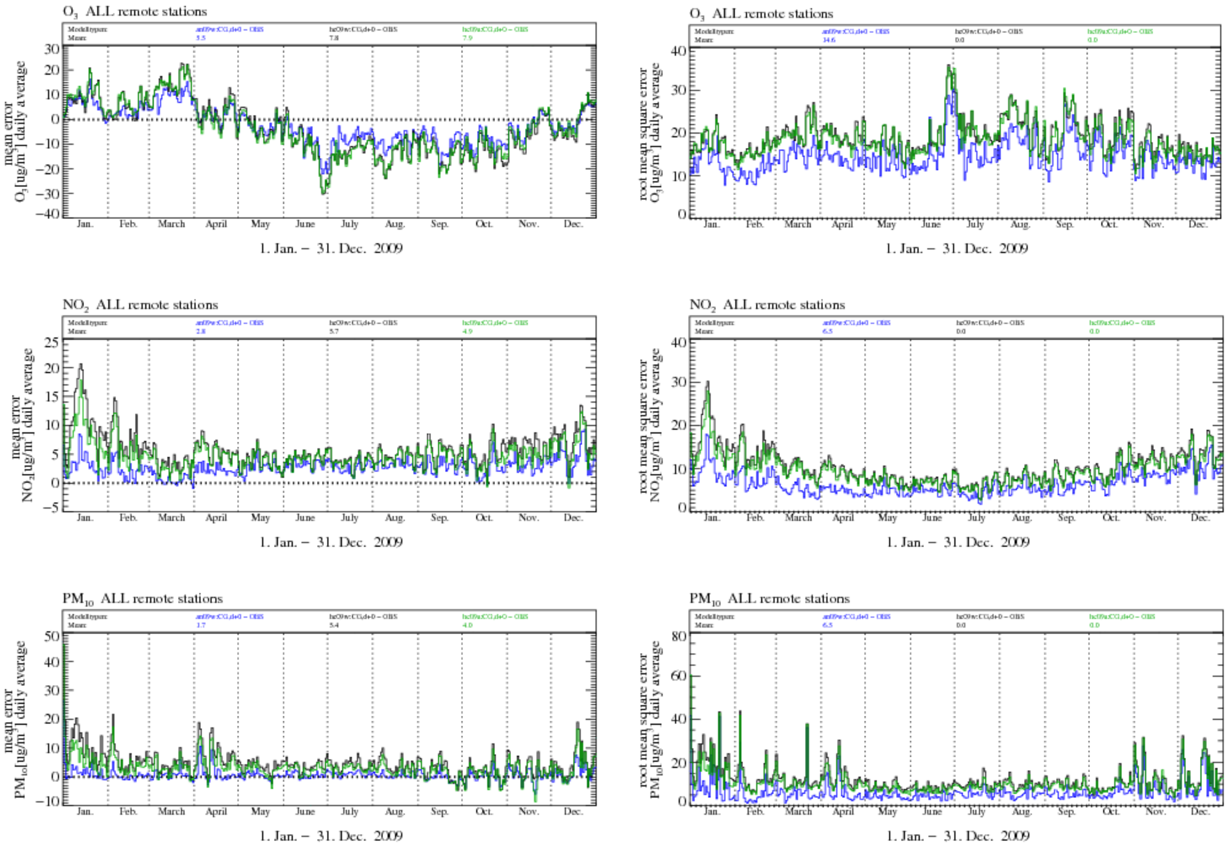

Figure 9: Bias (left) and root mean square error (right) of daily averaged O_3, NO_2, and PM_{10} concentrations averaged over all Airbase measurement sites, which were held back from assimilation for the year 2009. Blue: EURAD-IM 3d-var re-analysis, 30% of background stations held back from assimilation, black: control run (no data assimilation at all), green: control run initialized with the re-analysis for the previous day.

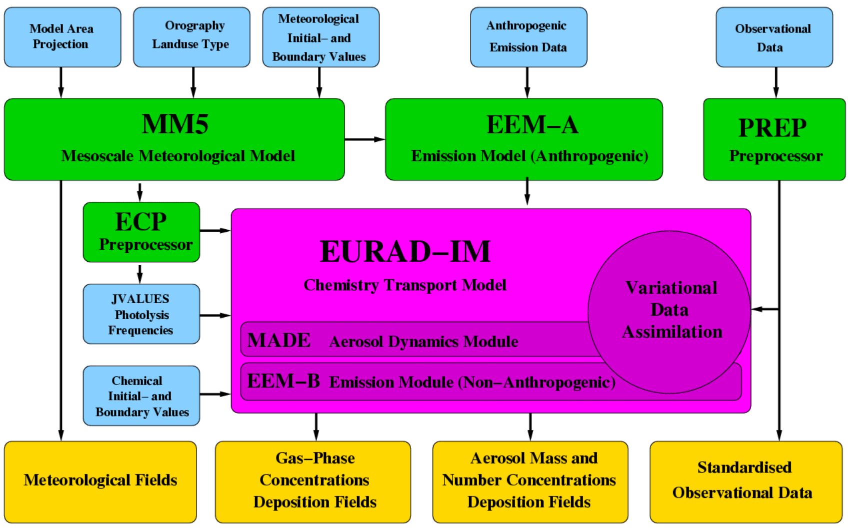

Figure 10: Flowchart of the EURAD model system containing the meteorological driver MM5, the pre-processors ECP and PREP, the emission model EEM and the chemistry transport model EURAD-IM (input parameters are shaded in blue, output parameters are shaded in yellow and procedural parts are shaded in green or magenta).

On ECMWF's compute server, the EURAD-IM uses 121 CPU's, which are parallelised with MPI. MM5 is currently parallelised with the OMP interface using 64 CPU's.

The Supervisor Monitor Scheduler (SMS) is used to control EURAD-IM model runs.

5. Justification of requested computer resources

The computer resources requested are mainly due to the following tasks**:**

5.1 Air quality forecasts
The pre-operational air quality forecast for Europe, based upon the 12:00 IFS operational meteorological forecast, has been set up using the EURAD-IM system with 15 km resolution. A 72h forecast of O_3, NO, NO_2, CO, SO_2, $PM_{2.5}$, and PM_{10} is regularly delivered. Computing resources needed are 860 SBU's per day for the meteorological driver MM5 and about 2300 SBU's per day for the chemistry transport model EURAD-IM. This results in an amount of about 1,200 kSBU's per year.

5.2 Air quality analyses
Daily 3d-var air quality analyses for Europe are provided with 15 km horizontal resolution. Near real time surface in situ measurements of O_3, NO_2, NO, SO_2, CO, and PM_{10} as well as satellite derived NO_2 column retrievals from OMI, GOME-2, and SCIAMACHY are hourly assimilated using the intermittent 3d-var method. This task needs about 210 SBU’s per day for the meteorological driver MM5 and about 1,300 SBU's per day for the chemistry transport model EURAD-IM, which results in an amount of about 550 kSBU's per model year.

References

Biermann, U.M., B.P. Luo, and T. Peter, Absorption spectra and optical constants of binary and ternary solutions of H_2SO_4, HNO_3, and H_2O in the mid infrared at atmospheric temperatures, *J. Phys. Chem. A*, **104**, 783-793, 2000.

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

Elbern, H. and H. Schmidt, Ozone episode analysis by four–dimensional variational chemistry data assimilation, *J. Geophys. Res*., **106**, 3569–3590, 2001.

Elbern, H., A. Broad, R. Engelen, P. Husson, S. Scollo, P. Seibert, A. Stohl, S. Tait, T. Thordarsen, and S. Varghese, How can the R&D community best contribute to improve VAAC analysis and prediction of volcanic ash plume? In: C. Zehner, Ed. (2010). Monitoring Volcanic Ash from Space. Proceedings of the ESAEUMETSAT workshop on the 14 April to 23 May 2010 eruption at the Eyjafjoll volcano, South Iceland. Frascati, Italy, 26-27 May 2010. ESA-Publication STM-280. doi:10.5270/atmch-10-01<92>.

Elbern, H., A. Strunk, L. Nieradzik, Data assimilation, in: The remote sensing of tropospheric composition from space, J. Burrows, U. Platt, P. Borrell (Eds.), Springer, 2010.

Elbern, H. A. Strunk, L. Nieradzik, Inverse modelling and combined state-source estimation for chemical weather, In: Data Assimilation: Making sense of observations, Springer, eds. Lahoz, Khattatov, Menard, pp 491<96>513, doi 10.1007/978-3-540-74703-1 19, 2010.

Evans, T.N. and G.R. Fournier, Simple approximation to extinction efficiency valid over all size parameters, *Applied Optics*, **29**, 4666-4670, 1990.

Martynenko, D., T. Holzer-Popp, H. Elbern, and M. Schroedter-Homscheidt, Understanding the aerosol information content in multi-spectral reflectance measurements using a synergetic retrieval algorithm, *Atmos. Meas. Tech. Discuss*., **3**, 2579-2602, 2010 doi:10.5194/amtd-3-2579-2010 .

Schroedter-Homscheidt, M., H. Elbern, and T. Holzer-Popp, Observation operator for the assimilation of aerosol type resolving satellite measurements into a chemical transport model, *Atmos. Chem. Phys*., **10**, 10435-10452, doi:10.5194/acp-10-10435- 2010, 2010.

List of publications/reports from the project with complete references

Huijnen, V., H.J. Eskes, B. Arnstrup, R. Bergstrom, K.F. Boersma, H. Elbern, J. Flemming, G. Foret, E. Friese, A. Gross, M. D'Isidoro, I. Kioutsioukis, A. Maurizi, D. Melas, V.-H. Peuch, A. Poupkou, L. Robertson, M. Sofiev, O. Stein, A. Strunk, A. Valdebenito, C. Zerefos, and D. Zyryanov, Comparison of OMI NO_2 trpospheric columns with an ensemble of global and European regional air quality models, *Atmos. Chem. Phys. Discuss.*, **9**, 22271-22330, 2009.

Future plans

MACC-II (Monitoring Atmospheric Composition and Climate – Interim Implementation) will provide continuity and refinement of the atmospheric services provided by MACC.

In the data assimilation part of MACC-II it is planned to establish an extended suite of data to be assimilated from space-borne sensors, including IASI data for O_3 and AOD data from MODIS, MSG/SEVIRI. The emission data inversion procedure based on 4d-var will be applied to identify emission correction factors. The continued inversion of the model with respect to emissions should provide emission factors, which are able to improve forecast quality even in a free model run.

The transition from prototype to fully operational service is projected for the European scale air quality forecast and analysis.

In MACC-II, the objectives of the successor of the R-EVA sub-project are the development and implementation of an operational process dedicated to the yearly production of European air quality assessment reports. EURAD-IM will contribute to the assessment reports for the years 2010-2012.

We have placed a request for a new special project, which may allow for fulfillment of the work envisaged in MACC-II.