Atmos. Chem. Phys. Discuss., 3, 3543–3588, 2003
www.atmos-chem-phys.org/acpd/3/3543/

Atmospheric Chemistry and Physics Discussions

Validation of a 3-D hemispheric nested air pollution model

L. M. Frohn, J. H. Christensen, J. Brandt, C. Geels, and K. M. Hansen

National Environmental Research Institute, Department of Atmospheric Environment, Frederiksborgvej 399, P.O.Box 358, DK-4000 Roskilde, Denmark

Received: 19 December 2002 – Accepted: 27 June 2003 – Published: 8 July 2003

Correspondence to: L. M. Frohn (lmf@dmu.dk)

Abstract

Several air pollution transport models have been developed at the National Environmental Research Institute in Denmark over the last decade (DREAM, DEHM, ACDEP and DEOM). A new 3-D nested Eulerian transport-chemistry model: REGIonal high resolutioN Air pollution model (REGINA) is based on modules and parameterisations from these models as well as new methods.

The model covers the majority of the Northern Hemisphere with currently one nest implemented. The horizontal resolution in the mother domain is 150 km × 150 km, and the nesting factor is three. A chemical scheme (originally 51 species) has been extended with a detailed description of the ammonia chemistry and implemented in the model. The mesoscale numerical weather prediction model MM5v2 is used as meteorological driver for the model. The concentrations of air pollutants, such as sulphur and nitrogen in various forms, have been calculated, applying zero nesting and one nest. The model setup is currently being validated by comparing calculated values of concentrations to measurements from approximately 100 stations included in the European Monitoring and Evalutation Programme (EMEP).

The present paper describes the physical processes and parameterisations of the model together with the modifications of the chemical scheme. Validation of the model calculations by comparison to EMEP measurements for a summer and a winter month is shown and discussed. Furthermore, results from a sensitivity study of the model performance with respect to resolution in emission and meteorology input data is presented. Finally the future prospects of the model are discussed.

The overall validation shows that the model performs well with respect to correlation for both monthly and daily mean values.

1 Introduction

The models developed at the National Environmental Research Institute are used for many different purposes, including air pollution forecasts at regional scales, urban background as well as urban street level, nitrogen load calculations for Inner Danish waters, the Baltic and the North Sea, accidental release modelling as well as extreme air pollution exposures as e.g. ozone episodes. The scales on which the models operate range from 150 km^2 to a few square metres and the domains from the Northern Hemisphere down to individual street canyons. The overall model performance depends in general on the numerical schemes employed, the nesting techniques, initial and boundary conditions, quality of input data as well as chemical and physical parameterisations.

In most cases there is no need for describing all processes from hemispheric scale down to local scale simultaneously However, when addressing problems like coastal eutrophication due to atmospheric deposition of nitrogen, which also includes long-range transport of nitrogen containing species, it is necessary to describe the sources adequately with a sufficiently large domain, yet the resolution in the output need to be high enough for the results to be usable.

In Fig. 1 is shown an example of six-hour mean concentrations of NO_2 calculated with the REGIonal high resolutioN Air pollution (REGINA) model. Furthermore the six-hour mean wind is shown. The concentration pattern is an example of a long-range transport episode, where a low pressure system transports NO_2 from North America towards the Arctic and Europe. Similar episodes occur where ozone or carbondioxide is transported across very large distances e.g. across the Atlantic ocean. Episodes like these will not be included in the results if the domain is limited to e.g. the European area as is the case in regional air pollution models, e.g. DEOM (Brandt et al., 2001) or ACDEP (Hertel et al., 1995). One of the methods for obtaining high resolution output with a large domain setup is to use a nested grid formulation as has been done in the work with this new model.

Over the recent years calculations of concentrations and depositions of many different chemical compounds have been carried out within the Danish Background Monitoring Programme using the ACDEP model (Skjøth et al., 2002). The new model, REGINA, is intended to be applied for studying air pollution phenomena (forecasts, scenarios and assessment) over Denmark with high resolution (Frohn et al., 2002). Furthermore it will complement and on long term substitute the ACDEP model in the monitoring programme. When high resolution modelling is carried out, the input data (emissions, land use, meteorology) must also have high resolution and sufficient quality. In order to test the sensitivity of the results to the resolution in input data, several emission and meteorology data scenarios with the new model (and one nest) have been carried out. The base year for these runs is 1998, for which the most recent emission and measurement data base is available.

2 Model description

The full mathematical model describing the rate of change in the mixing ratio of a chemical compound can be expressed as

$$\begin{aligned}\frac{\partial c_i}{\partial t} &= -\left(u\frac{\partial c_i}{\partial x} + v\frac{\partial c_i}{\partial y} + \dot{\sigma}\frac{\partial c_i}{\partial \sigma}\right) \\ &+ K_x\frac{\partial^2 c_i}{\partial x^2} + K_y\frac{\partial^2 c_i}{\partial y^2} + \frac{\partial}{\partial \sigma}\left(K_\sigma\frac{\partial c_i}{\partial \sigma}\right) \\ &+ E_i(x,y,\sigma,t) - \Lambda_i c_i \\ &+ Q_i\left(c_1, c_2, \ldots, c_q\right) \quad (i = 1,2,\ldots,q)\end{aligned} \tag{1}$$

where $(u, v, \dot{\sigma})$ are wind speed components in the (x, y, σ) directions, c_i are the mixing ratios for the q different species, K_x and K_y are horizontal diffusion coefficients, assumed constant and K_σ is the vertical diffusion coefficient, dependent on space and time coordinates. E_i is the emission of the chemical species i, Λ_i is the scavenging coefficient for wet deposition of species i and Q denotes the chemical reactions.

The mathematical model is split into five sub-models using a simple splitting procedure (McRae et al., 1984). The five sub-models are:

– Sub-model 1: Three-dimensional advection

$$\frac{\partial c_i^1}{\partial t} = -u\frac{\partial c_i^1}{\partial x} - v\frac{\partial c_i^1}{\partial y} - \dot{\sigma}\frac{\partial c_i^1}{\partial \sigma} \tag{2}$$

– Sub-model 2-4: Diffusion

$$\frac{\partial c_i^2}{\partial t} = K_x\frac{\partial^2 c_i^2}{\partial x^2} \tag{3}$$

$$\frac{\partial c_i^3}{\partial t} = K_y\frac{\partial^2 c_i^3}{\partial y^2}$$

$$\frac{\partial c_i^4}{\partial t} = \frac{\partial}{\partial \sigma}\left(K_\sigma\frac{\partial c_i^4}{\partial \sigma}\right)$$

– Sub-model 5: Chemistry, emissions and wet deposition

$$\frac{dc_i^5}{dt} = Q_i\left(c_1^5, c_2^5, \ldots, c_q^5\right) + E_i - \Lambda_i c_i \tag{4}$$

Sub-model 1 is solved using a newly modified accurate space derivatives scheme for the horizontal advection, combined with a finite elements scheme for the vertical advection (Frohn et al., 2002; Pepper et al., 1979). The temporal integration of advection is carried out using a Taylor series expansion to third order. The sub-models 2–4 are solved using a finite elements scheme for the spatial discretization and the θ

method for the temporal integration (Lambert, 1991). Sub-model 5 is solved using a new combination of the Euler Backward Iterative method (Hertel et al., 1993) and the two-step method (Frohn et al., 2002).

Details on the implementation of the numerical methods for solving advection and chemistry can be found in Frohn et al. (2002). Details on the implementation of the numerical methods for solving the diffusion can be found in Christensen (1993) and Christensen (1997).

The model domain covers the majority of the Northern Hemisphere with a resolution for the mother domain of 150 km × 150 km. The grid is an extension of the original 150 km^2 EMEP (European Monitoring and Evaluation Programme) grid. The nest covers the European area with a resolution of 50 km × 50 km (Fig. 2. A second nest covering the Scandinavian area with a resolution of 16.67 km × 16.67 km is planned but not yet implemented.

The model has 18 vertical layers and extends up to 15 km on average. The meteorological input is taken from the MM5v2 model (Grell et al., 1995) run operationally at NERI with one nest. The applied landuse data are derived from a global inventory with eight categories (Wilson and Henderson-Sellers, 1985). The applied chemical scheme is a modified version of the chemical scheme published in Strand and Hov (1994). The modifications consist of the inclusion of ammonia (NH_3) and related species, ammonium nitrate (NH_4NO_3), ammonium bisulphate (NH_4HSO_4), ammonium sulphate ($(NH_4)_2SO_4$) and particulate nitrate (NO_3^-) formed from nitric acid (HNO_3). The modifications have been implemented in order to improve the description of the transformations of nitrogen containing compounds (Hertel et al., 1995).

3 Emission and meteorology scenarios

Three emission data sources were available for the current study.

– GEIA (Global Emissions Inventory Activity); Graedel et al. (1993)

- EDGAR (Emission Database for Global Atmospheric Research); Olivier et al. (1996)
- EMEP (European Monitoring and Evaluation Programme); Vestreng (2001); Vestreng and Støren (2000)

The inventories vary in coverage, resolution and in number of chemical compounds for which data are available (Table 1). No information of seasonal or diurnal variations are included in the emission data used in this study.

Two scenarios have been based on these three emission inventories. The first scenario is constructed using data bases with hemispheric coverage only, i.e. EDGAR and GEIA data only. The EDGAR data in the part of the model domain covered by the EMEP has been replaced with the more recent EMEP data base in the second scenario. The EMEP data have a three times higher resolution and the data covers an area corresponding to the first nest in the model. The properties of the two emission scenarios can be seen in Table 2.

The model has been run with no nests for the entire year 1998 and another run has been carried out with one nest for the months February and August 1998. The data for the nested model runs have been prepared in two different ways in order to test sensitivity to resolution in the meteorology and emission data. Either the data are read directly into the nested domain (and in this case the resolution of the input data matches the resolution of the nest, i.e. 50 km) or data are copied from the mother domain to the nested domain (and in this case the resolution of the input data is three times lower than for the nested domain, i.e. 150 km). This results in eight different model scenarios (Table 3).

4 Results

The results of this study have been validated with data from approximately 100 EMEP measurement stations across Europe (Fig. 3). The model is validated for February

and August for all eight model scenarios considering both monthly and daily mean concentrations.

This validation was performed for the gaseous species nitrogen dioxide (NO_2), ozone (O_3) and sulphur dioxide (SO_2) and the particulates ammonium (NH_4^+), nitrate (NO_3^-) and sulphate (SO_4^{2-}). Furthermore for the sum of ammonia (NH_3) and ammonium (the sum is denoted SNH) and for the sum of nitric acid (HNO_3) and nitrate (the sum is denoted SNO_3).

Only stations with a data coverage of more than 90% have been included in the validation. Furthermore only stations located less than 500 m above sea level are considered, due to the coarse topography in the model.

Three statistical parameters have been calculated for the validation; the correlation (Corr), the fractional bias (FB) and the normalised mean square error (NMSE). The correlation provides a good first impression of the results, describing how well the variations of the measured data are reproduced by the model calculations. The fractional bias gives information on the bias between the calculated and measured data and the normalised mean square error provides the mean square error between the two data series. The advantage of the FB and the NMSE compared to (usual) bias and root mean square error is, that they can be compared for different chemical species, where bias and root mean square error can only be compared for results for the same chemical component.

A ranking procedure has been applied in order to determine the best performing model scenario (Mosca et al., 1997). Taking one statistical parameter and one chemical component at the time a rank between one and eight is assigned to each model scenario for that specific chemical component and statistical parameter. The lowest rank is assigned to the model scenario with the best performance.

For each of the three statistical parameters the ranks are summed on the level of model scenarios and a total rank is determined by summing the ranks of the model scenarios for all three statistical parameters. The scenario with the lowest total rank is the best performing.

4.1 February 1998

4.1.1 Monthly values

The correlation of modelled and measured concentrations of NO_2 (model scenario eight) is good (0.70). However, the model tends to underestimate the concentrations (Fig. 4). For O_3 the correlation is less good (0.47) and the level is underestimated, especially for the high concentrations (Fig. 5).

The correlations show that the performance of the model in general is good for monthly mean values (Table 4). Correlations for the secondary pollutants NH_4^+ and NO_3^- and for the primary pollutants NO_2 and SO_2 all increase when the EDGAR emission data base (model scenario one) is replaced by the EMEP emission data base (model scenario four) in the nested area. There is no noticeable difference for the remaining species SNH, SNO_3, O_3 and SO_4^{2-}. Correlations for O_3 are not impressive, because the model results presented here corresponds to background values which are better represented at the stations located at higher altitudes.

The resolution in the emission and meteorology input data does not have any clear influence on the calculated results for this month. Model scenario number eight performs best, closely followed by scenarios four, six, five and seven, when the ranking of the model scenarios according to correlation is performed and summed over different chemical components. The best performing scenarios with respect to fractional bias and normalised mean square error are number one and five, respectively.

The five best performing model scenarios, when total rank is considered are the ones where the EMEP emission data base is used (Table 5). The model scenarios with the best performance is scenario number six including the EMEP data base (coarse resolution) and nested resolution meteorology.

4.1.2 Daily values

The calculated concentrations of SO_2 are in good agreement with the measured values (Fig. 6), whereas concentrations of SO_4^{2-} are underestimated (Fig. 7). The temporal evolution of the calculated concentrations of NO_2, O_3, SO_2 and SO_4^{2-} for the German station Langenbrügge and the Finnish station Oulanka are in good agreement with the measured concentrations (Figs. 8 and 9).

The calculated statistical parameters for the daily mean concentrations are given in Table 6 for all eight model scenarios. The correlations for the daily concentrations are not as good as for monthly values, except for O_3, where the correlation is better, however the model still performs quite good in general. When the EDGAR data base is replaced with the EMEP data base in the nested domain, the correlation improves for NH_4^+, SNH, NO_2 and SO_2 although not so prominent for the last two, decreases for SO_4^{2-} and remains more or less unchanged for the rest of the chemical compounds. There are no noticeable differences in the correlations when the resolution of the emission and meteorology input data is enhanced.

Model scenario eight performs the best when ranking according to correlation is considered. Scenario three and five perform the best with respect to fractional bias and normalised mean square error, respectively. The overall best performing model scenario is scenario five followed by six, seven and eight, i.e. the scenarios with the EMEP data base (Table 7).

4.2 August

4.2.1 Monthly values

The model overestimates the SO_2 concentrations by a factor of two whereas the calculated SO_4^{2-} concentrations are in good agreement with the measured concentrations for low values and somewhat overestimated for the high values (Figs. 10 and 11).

The performance of the model with respect to correlation is good for all chemical compounds except O_3 (Table 10). Correlations of the primary pollutants NO_2 and SO_2 increase when the EDGAR data base (model scenario one) is replaced by the EMEP data base (model scenario four) in the nested area, however there is no noticeable effect for the other chemical components. Model scenarios with high resolution meteorology (scenario three, six and eight) tends to perform better than scenarios with low resolution meteorology when SO_2 is considered. This tendency is also valid for O_3, however not as pronounced. For NH_4^+ there is also a tendency that higher resolution in the emission input data (scenario seven and eight) results in better correlations.

Applying the ranking procedure on the correlations it is seen that scenario one (rank equal to 26), eight (rank equal to 27) and four (rank equal to 28) performs the best. For the ranking with respect to fractional bias and normalised mean square error the best performing scenarios are number seven and eight, respectively. Similar to the results for February, the five best performing scenarios are the five scenarios with the EMEP emission data base used as input in the nested domain. The best performing model scenario is scenario eight with high resolution in both emissions and meteorological input data (Table 9).

4.2.2 Daily values

The model underestimates the NO_2 values even though the correlation is good (Fig. 12). A number of stations have high measured concentrations, probably corresponding to stations located in or close to urban areas. These high concentations are not captured by the model due to the coarse resolution (50 km). The concentrations of O_3 are overestimated for some stations and underestimated for some stations and the correlation is not as good as for the other chemical components (Fig. 13).

The temporal evolution of the calculated concentrations of NO_2, O_3, SO_2 and SO_4^{2-} at the German station Langenbrügge and the Finnish station Oulanka are in good agreement with the measured concentrations (Figs. 14 and 15).

The correlations of daily concentrations are not as good as for the monthly values, except for O_3 where the correlation calculated from daily concentrations is better (Ta-

ble 10). There is a large increase in the correlation of NO_2 and SO_2 concentrations when the EMEP emission data base (model scenario four) replaces the EDGAR emission data base (model scenario one). The correlations for all the other chemical components remain more or less unchanged.

There does not seem to be any visible difference in the results when the resolution in the input data is changed, except for SO_2 where there is a very small increase in the correlations for model scenario three, six and eight as compared to model scenario two, five and seven, when the resolution of the meteorological input is enhanced.

Considering the ranking of the scenarios with respect to correlation, model scenario four performs the best. For fractional bias and normalised mean square error the best performing scenario is number seven and eight and number four, respectively.

The overall best performing scenario is model scenario number eight and the five scenarios with EMEP emissions data constitutes the top five (Table 11).

5 Conclusions

A 3-D nested Eulerian transport-chemistry model covering the majority of the Northern Hemisphere has been developed. An existing chemical scheme has been extended to include species relevant for nitrogen chemistry.

The model has been run using eight different scenarios combining two emission scenarios and different resolutions in the input data. The results have been validated with measurements from approximately 100 measurement stations across Europe for February and August 1998.

The model performs well with respect to correlation, when calculated concentration levels are compared to measurements, both for monthly and daily values. The model tends to underestimate the NO_2 and O_3 concentrations for the winter month February whereas the SO_4^{2-} concentrations are overestimated for this month. The SO_2 concentrations are overestimated and the NO_2 concentrations are underestimated in the summer month August.

The EDGAR data base is from 1995 whereas the EMEP data base is from 1998. It is therefore expected that the model gives better results for the primary pollutants with the EMEP data base implemented compared to the calculations with the EDGAR data base implemented. This is the case when correlations for the primary pollutants NO_2 and SO_2 are considered. For February and August the correlations for these two chemical components increase when the EDGAR emission data are replaced by the EMEP emission data in the nested area, both for monthly and daily values. The correlations of NH_4^+ and SNH also increase for the monthly mean values for February, indicating that the emissions of the primary pollutant NH_3 which acts as a precursor, are better described when using the EMEP data base compared to when the EDGAR data base is used.

Apparently the resolution of the input is not especially important for this model setup. The resolution of the emission and meteorology input has no influence on the correlation of calculated and measured concentrations for February. It appears that the resolution in the meteorological data input is influencing on the correlation of SO_2 for August (both monthly and daily values improve with higher resolution), independent of the resolution in and source of emission data. The correlation of NH_4^+ increases with increasing resolution of the emission data for monthly values, however there is no influence of emission data resolution for any of the other chemical components.

The explanation for the small response when the resolution of input data is increased could be that the resolution still is very coarse and the difference in resolution is not more than a factor of three. Furthermore a resolution of 50 km is still much too coarse for modelling regional air pollution processes adequately. Another explanation could be the lack of seasonal and diurnal variations in the emission data base. The improvement of the results when increasing the resolution in the input data could be suppressed by the inaccuracies in the results arising from using the same emission data regardless of season and time of day.

A ranking procedure has been applied in order to determine the best performing scenario for all chemical components and statistical parameters. All the best performing

scenarios has the EMEP emissions in the nested area. The best performing scenario for monthly mean values is scenario six with high resolution in meteorology and coarse resolution in emissions for February. Scenario five with coarse resolution in both meteorology and emissions performs the best for the daily mean values. The best performing scenario for both monthly and daily mean values is scenario eight with high resolution in both meteorology and emissions for August. This is in agreement with the conclusion that the resolution in the input only has influence on the results for August.

Future investigations with the REGINA model includes model runs with two nests. Emission data from the GENEMIS data base with a resolution of 16.67 km for the year 1994 are available for these studies.

Acknowledgements. The Danish Research Academy and Risø National Laboratory are acknowledged for financial support.

References

Brandt, J., Christensen, J. H., Frohn, L. M., Palmgren, F., Berkowicz, R., and Zlatev, Z.: Operational air pollution forecasts from European to local scale, Atmospheric Environment, 35, 1, 91–98, 2001. 3545

Christensen , J. H.: Testing Advection Schemes in a Three-Dimensional Air Pollution Model, Mathematical and Computational Modelling, 18, 2, 75–88, 1993. 3548

Christensen , J. H.: The Danish Eulerian Hemispheric Model – A three-dimensional air pollution model used for the Arctic, Atmospheric Environment, 31, 24, 4169–4191, 1997. 3548

Frohn, L. M., Christensen, J. H., and Brandt, J.: Development of a High-Resolution Nested Air Pollution Model. The Numerical Approach. Journal of Computational Physics, 179, 1, 68–94, 2002. 3546, 3547, 3548

Graedel, T. E., Bates, T. S., Bouwman, A. F., Cunnold, D., Dignon, J., Fung, I., Jacob, D. J., Lamb, B. K., Logan, J. A., Marland, G., Middleton, P., Pacyna, J. M., Placet, M., and Veldt, C.: A Compilation of Inventories of Emissions to the Atmosphere, Global Biogeochemical Cycling 7, 1–26, 1993. 3548

Grell, G. A., Dudhia, J., and Stauffer D. R.: A Description of the Fifth-Generation Penn State/NCAR Mesoscale Model (MM5), Mesoscale and Microscale Meteorology Division, National Center for Atmospheric Research, Boulder, Colorado, NCAR Technical Note, NCAR/TN-398+STR, pp. 114, 1995. 3548

Hertel, O., Berkowicz, R., Christensen, J. H., and Hov, Ø.: Test of two numerical schemes for use in atmospheric transport-chemistry models, Atmospheric Environment, 27A, 16, 2591–2611, 1993. 3548

Hertel, O., Christensen, J. H., Runge, E. H., Asman, W. A. H., Berkowicz, R., Hovmand, M. F., and Hov, Ø.: Development and Testing of a new Variable Scale Air Pollution Model – ACDEP, Atmospheric Environment, 29, 20, 1267–1290, 1995. 3545, 3548

Lambert, J. D.: Numerical Methods for Ordinary Differential Systems: The Initial Value Problem, John Wiley & Sons, Chichester, 1991. 3548

McRae, G. J., Goodin, W. R., and Seinfeld, J. H.: Numerical solution of the atmospheric diffusion equations for chemically reacting flows. Journal of Computational Physics, 45, 1–42, 1984. 3547

Mosca, S., Graziani, G., Klug, W., Bellasio, R., and Bianconi, R.: ATMES-II – Evaluation of Long-range Dispersion Models using 1st ETEX release data. Volume 1, Draft report prepared for ETEX Symposium on Long-range Atmospheric Transport, Model Verification and Emergency Response, Vienna, 13–16 May, 1997. 3550

Olivier, J. G. J., Bouwman, A. F., Van der Maas, C. W. M., Berdowski, J. J. M., Veldt,C., Bloos, J. P. J., Visschedijk, A. J. H., Zandveld, P. Y. J., and Haverlag, J. L.: Description of EDGAR Version 2.0. A set of global emission inventories of greenhouse gases and ozone-depleting substances for all anthropogenic and most natural sources on a per country basis and on 1° x 1° grid, RIVM, Bilthoven, RIVM report no. 771060 002, [TNO MEP report nr. R96/119], 1996. 3549

Pepper, D. W., Kern, C. D., and Long, Jr., P. E.: Modeling the dispersion of atmospheric pollution using cubic splines and chapeau functions, Atmospheric Environment, 13, 223–237, 1979. 3547

Skjøth, C. A., Hertel, O. and Ellermann, T.: Use of the ACDEP trajectory model in the Danish nation-wide Background Monitoring Programme, Physics and Chemistry of the Earth, 27, 35, 1469–1477, 2002. 3546

Strand, A. and Hov, Ø.: A 2-Dimensional Global Study of Tropospheric Ozone Production, Journal of Geophysical Research, Atmospheres, 99, D11, 22877–22895, 1994. 3548

Vestreng, V.: Emission data reported to UNECE/EMEP: Evaluation of the spatial distributions of emission, EMEP/MSC-W Note 1/01, July 2001. 3549

Vestreng, V. and Støren, E.: Analysis of UNECE/EMEP Emission Data MSC-W Status Report 2000. EMEP/MSC-W Note 1/00, July 2000. 3549

Wilson, M. F. and Henderson-Sellers, A.: A global archive of land cover and soils data for use in general circulation models. Journal of Climate 5, 119–143, 1985. 3548

Table 1. Emission data available for the REGINA modelling

Database	GEIA	EDGAR	EMEP
Year	1990	1990, 1995	1990, 1997, 1998
Species	NO, NO_2, Isoprene, Terpene other biogenic VOC's	NO, NO_2, SO_X, NH_3, CO, CH_4, NMVOC	NO, NO_2, SO_X, NH_3, NMVOC
Resolution	150 km	150 km	50, 150 km
Domain	N. Hemisphere	N. Hemisphere	EMEP (Nest 1)
NMVOC-split	-	Yes	No
Ships	No	No	Yes
Lightning	Yes	No	No
Soil	Yes	No	No

Table 2. Properties of the two emission scenarios used in the sensitivity study

Source	Scenario 1	Scenario 2
NO_x	EDGAR (1995)	EDGAR (1995), EMEP (1998)
NO_x, Lightning	GEIA	GEIA
NO_x, Soil	GEIA	GEIA
NO_x, Ships	EDGAR (1995)	EDGAR (1995), EMEP (1998)
SO_x	EDGAR (1995)	EDGAR (1995), EMEP (1998)
SO_x, Seasonal	GEIA	GEIA
SO_x, Ships	EDGAR (1995)	EDGAR (1995), EMEP (1998)
NH_3	EDGAR (1990)	EDGAR (1990), EMEP (1998)
CO	EDGAR (1995)	EDGAR (1995)
CH_4	EDGAR (1995)	EDGAR (1995)
NMVOC	EDGAR (1995, 1990)	EDGAR (1995, 1990), EMEP (1998)
Isoprene	GEIA	GEIA
Terpene	GEIA	GEIA
Other VOC	GEIA	GEIA

Table 3. The setup of the eight model scenarios used in the validation of the hemispheric model

Model scenario	Nest	Emission scenario	Emission data resolution	Meteorology data resolution
1	No	1	150 km	150 km
2	Yes	1	150 km	150 km
3	Yes	1	150 km	50 km
4	No	2	150 km	150 km
5	Yes	2	150 km	150 km
6	Yes	2	150 km	50 km
7	Yes	2	50 km	150 km
8	Yes	2	50 km	50 km

Table 4. Correlation coefficients (Corr), fractional bias (FB) and normalised mean square error (NMSE) for calculated and measured monthly mean concentrations for all model scenarios for February 1998. Values which are ascribed a rank equal to one are **boldface**

No. of stations	12	24	9	24	34	57	40	49	
Corr									
	NH_4^+	SNH	NO_3^-	SNO_3	NO_2	O_3	SO_2	SO_4^{2-}	Rank
Model scenario									
1	0.59	0.64	0.67	0.79	0.63	0.41	0.77	0.65	41
2	0.59	0.59	0.57	0.76	0.64	0.45	0.76	0.62	50
3	0.57	0.60	0.55	0.73	0.64	**0.47**	0.76	0.62	52
4	**0.69**	0.66	**0.78**	**0.81**	**0.71**	0.41	0.88	0.66	21
5	0.63	0.64	0.69	0.77	0.70	0.45	0.89	0.65	24
6	0.62	**0.67**	0.67	0.76	0.70	**0.47**	0.89	0.66	21
7	0.63	0.64	0.68	0.78	0.70	0.44	**0.90**	0.65	25
8	0.62	**0.67**	0.67	0.76	0.70	**0.47**	0.89	**0.67**	20
FB									
	NH_4^+	SNH	NO_3	SNO_3^-	NO_2	O_3	SO_2	SO_4^{2-}	Rank
Model scenario									
1	**0.005**	-0.378	-0.484	-0.398	-0.455	-0.341	0.997	-0.042	19
2	0.227	**-0.202**	-0.236	-0.192	-0.332	-0.359	1.106	0.156	37
3	0.166	-0.219	-0.150	**-0.134**	**-0.330**	-0.349	1.135	**0.015**	27
4	-0.166	-0.497	-0.534	-0.503	-0.685	**-0.260**	**-0.012**	-0.097	42
5	-0.051	-0.361	-0.175	-0.241	-0.503	-0.299	0.356	-0.492	31
6	-0.109	-0.379	**-0.128**	-0.207	-0.494	-0.294	0.394	-0.627	32
7	-0.114	-0.404	-0.249	-0.313	-0.553	-0.284	0.254	-0.534	40
8	0.184	-0.432	-0.213	-0.286	-0.540	-0.279	0.287	-0.671	42

Table 4. Continued

No. of stations	12	24	9	24	34	57	40	49	
NMSE									
	NH_4^+	SNH	NO_3	SNO_3^-	NO_2	O_3	SO_2	SO_4^{2-}	Rank
Model scenario									
1	0.28	0.88	0.61	0.77	0.91	0.18	2.50	0.56	44
2	0.44	**0.75**	0.35	0.50	0.70	0.19	3.60	0.83	43
3	0.43	**0.75**	0.30	**0.47**	**0.69**	0.18	4.00	0.59	31
4	**0.20**	1.10	0.61	0.90	1.40	**0.12**	**0.20**	**0.45**	35
5	0.23	0.86	0.25	0.50	0.89	0.15	0.40	0.59	27
6	0.25	0.84	**0.24**	0.48	0.86	0.14	0.50	0.72	28
7	0.24	0.93	0.30	0.58	0.99	0.13	0.26	0.64	36
8	0.27	0,92	0.28	0.57	0.97	0.13	0.33	0.79	37

Table 5. Overall ranking of the eight model scenarios for February 1998, by adding the ranks obtained for the three statistical parameters for the montly mean values. Lowest rank corresponds to best performance

Model scenario	1	2	3	4	5	6	7	8
Corr	41	50	52	21	24	21	25	20
FB	19	37	27	42	31	32	40	42
NMSE	44	43	31	35	27	28	36	37
Overall	104	130	110	98	82	81	101	99

Table 6. Correlation coefficients (Corr), fractional bias (FB) and normalised mean square error (NMSE) for calculated and measured daily mean concentrations for all model scenarios for February 1998. Values which are ascribed a rank equal to one are **boldface**

No. of stations	12	24	9	24	34	57	40	49	
No. of observations	330	668	247	669	935	1567	1114	1366	
Corr									
	NH_4^+	SNH	NO_3^-	SNO_3	NO_2	O_3	SO_2	SO_4^{2-}	Rank
Model scenario									
1	0.51	0.58	0.59	0.68	0.62	0.49	0.62	0.48	40
2	0.50	0.51	0.54	0.62	0.60	0.53	0.60	0.46	52
3	0.49	0.51	0.54	0.60	0.61	**0.55**	0.60	0.47	51
4	**0.61**	**0.64**	0.62	**0.69**	**0.66**	0.45	0.65	0.43	30
5	0.56	0.56	**0.65**	0.66	0.64	0.51	0.70	0.51	28
2	0.54	0.57	0.63	0.65	0.65	0.53	0.70	0.53	26
7	0.55	0.56	0.64	0.67	0.65	0.51	**0.72**	0.51	24
8	0.54	0.58	0.64	0.65	0.65	0.53	0.71	**0.54**	20
FB									
	NH_4^+	SNH	NO_3	SNO_3^-	NO_2	O_3	SO_2	SO_4^{2-}	Rank
Model scenario									
1	**0.007**	-0.380	-0.484	-0.398	-0.457	-0.342	0.998	-0.041	36
2	0.230	**-0.205**	-0.236	-0.192	-0.334	-0.360	1.108	0.158	37
3	0.169	-0.222	-0.148	**-0.134**	**-0.332**	-0.349	1.136	**0.017**	28
4	-0.165	-0.499	-0.533	-0.504	-0.685	**-0.260**	**-0.012**	-0.095	42
5	-0.050	-0.363	-0.173	-0.241	-0.504	-0.300	0.356	-0.491	31
6	-0.108	-0.381	**-0.125**	-0.207	-0.495	-0.294	0.394	-0.627	32
7	-0.113	-0.406	-0.246	-0.314	-0.554	-0.285	0.255	-0.533	40
8	-0.182	-0.434	-0.210	-0.286	-0.541	-0.280	0.288	-0.670	42

Table 6. Continued

No. of stations	12	24	9	24	34	57	40	49	
No. of observations	330	668	247	669	935	1567	1114	1366	
NMSE									
	NH_4^+	SNH	NO_3	SNO_3^-	NO_2	O_3	SO_2	SO_4^{2-}	Rank
Model scenario									
1	0.67	0.18	1.20	1.80	1.40	0.23	3.40	**1.10**	37
2	0.84	**1.70**	0.86	1.40	**1.10**	0.24	5.30	1.50	37
3	0.86	**1.70**	0.77	**1.30**	**1.10**	0.24	5.80	1.30	33
4	**0.58**	2.00	1.30	2.00	2.00	**0.17**	1.20	1.70	45
5	0.61	1.80	0.66	1.40	1.40	0.19	1.10	1.30	22
6	0.67	1.80	**0.64**	**1.30**	1.40	0.19	1.20	1.50	25
7	0.65	1.90	0.73	1.50	1.50	0.18	**0.94**	1.30	29
8	0.72	2.00	0.70	1.50	1.50	0.18	1.00	1.60	38

Table 7. Overall ranking of the eight model scenarios for February 1998, by adding the ranks obtained for the three statistical parameters for the daily mean values. Lowest rank corresponds to best performance

Model scenario	1	2	3	4	5	6	7	8
Corr	40	52	51	30	28	26	24	20
FB	36	37	28	42	31	32	40	42
NMSE	37	37	33	45	22	25	29	38
Overall	113	126	112	117	81	83	93	100

Table 8. Correlation coefficients (Corr), fractional bias (FB) and normalised mean square error (NMSE) for calculated and measured monthly mean concentrations for all model scenarios for August 1998. Values which are ascribed a rank equal to one are **boldface**

No. of stations	9	24	9	24	37	53	36	45	
Corr									
	NH_4^+	SNH	NO_3^-	SNO_3	NO_2	O_3	SO_2	SO_4^{2-}	Rank
Model scenario									
1	**0.86**	**0.70**	0.77	**0.84**	0.63	**0.42**	0.63	0.81	26
2	0.81	0.69	0.82	0.82	0.62	0.39	0.62	0.78	36
3	0.78	0.68	0.84	0.82	0.63	**0.42**	0.66	0.78	35
4	0.84	0.66	0.74	0.82	**0.71**	0.39	0.81	**0.84**	28
5	0.83	0.64	0.80	0.80	0.70	0.33	0.81	0.80	41
2	0.81	0.65	0.84	0.79	**0.71**	0.37	**0.85**	0.80	32
7	**0.86**	0.67	0.84	0.79	0.69	0.35	0.79	0.79	36
8	0.85	0.68	**0.86**	0.79	0.69	0.38	0.84	0.80	27
FB									
	NH_4^+	SNH	NO_3	SNO_3^-	NO_2	O_3	SO_2	SO_4^{2-}	Rank
Model scenario									
1	0.359	-0.076	0.635	0.880	-0.438	0.131	1.450	0.799	46
2	0.429	**-0.069**	0.689	0.840	**-0.341**	0.101	1.489	0.879	45
3	0.381	-0.102	0.679	0.826	-0.346	0.103	1.500	0.785	45
4	-0.126	-0.304	**0.531**	0.716	-0.673	0.121	0.776	**0.172**	36
5	**-0.053**	-0.272	0.596	0.688	-0.572	0.083	0.835	0.281	28
6	-0.115	-0.304	0.586	0.676	-0.577	0.085	0.870	0.182	31
7	**-0.053**	-0.266	0.606	0.684	-0.595	**0.082**	**0.769**	0.273	26
8	-0.117	-0.300	-0.598	**0.673**	-0.589	0.083	0.818	0.176	28

Table 8. Continued

No. of stations	9	24	9	24	37	53	36	45	
NMSE									
	NH_4^+	SNH	NO_3	SNO_3^-	NO_2	O_3	SO_2	SO_4^{2-}	Rank
Model scenario									
1	0.20	**0.38**	0.65	1.40	1.20	0.072	8.10	1.60	43
2	0.33	0.41	0.69	1.40	**0.98**	0.069	11.00	2.10	46
3	0.29	0.43	0.66	1.30	**0.98**	0.065	11.00	1.50	39
4	0.13	0.62	0.54	**1.00**	1.60	0.072	**1.40**	**0.27**	32
5	0.11	0.61	0.57	**1.00**	1.30	0.070	1.90	0.43	33
6	0.14	0.63	**0.53**	**1.00**	1.30	0.066	2.10	**0.27**	28
7	**0.092**	0.58	0.55	**1.00**	1.40	0.068	1.50	0.42	26
8	0.12	0.59	**0.53**	**1.00**	1.40	**0.64**	1.70	**0.27**	21

Table 9. Overall ranking of the eight model scenarios for August 1998, by adding the ranks obtained for the three statistical parameters for the monthly mean values. Lowest rank corresponds to best performance

Model scenario	1	2	3	4	5	6	7	8
Corr	26	36	35	28	41	32	36	27
FB	46	45	45	36	28	31	26	28
NMSE	43	46	39	32	33	28	26	21
Overall	115	127	119	96	102	91	88	76

Table 10. Correlation coefficients (Corr), fractional bias (FB) and normalised mean square error (NMSE) for calculated and measured daily mean concentrations for all model scenarios for August 1998. Values with rank equal to one are **bold**

No. of stations	9	24	9	24	37	53	36	45	
No. of observations	276	735	243	736	1126	1621	1096	1379	
Corr									
	NH_4^+	SNH	NO_3^-	SNO_3	NO_2	O_3	SO_2	SO_4^{2-}	Rank
Model scenario									
1	**0.74**	**0.63**	0.65	**0.67**	0.58	**0.48**	0.48	0.68	29
2	**0.74**	0.61	0.63	0.63	0.58	0.42	0.47	0.68	42
3	0.72	0.62	0.66	0.65	0.58	0.43	0.50	0.67	39
4	0.73	0.61	0.66	**0.67**	**0.66**	0.46	0.60	**0.71**	20
5	0.73	0.60	0.66	0.63	0.65	0.39	0.60	0.69	37
2	0.71	0.61	0.70	0.64	0.65	0.41	**0.63**	0.69	29
7	**0.74**	0.61	0.69	0.62	0.64	0.40	0.60	0.69	33
8	0.72	**0.63**	**0.71**	0.64	0.65	0.42	**0.63**	0.68	24
FB									
	NH_4^+	SNH	NO_3	SNO_3^-	NO_2	O_3	SO_2	SO_4^{2-}	Rank
Model scenario									
1	0.362	-0.077	0.637	0.879	-0.434	0.131	1.450	0.799	46
2	0.431	**-0.071**	0.691	0.839	**-0.336**	0.101	1.489	0.879	45
3	0.383	-0.104	0.681	0.825	-0.341	0.103	1.499	0.785	45
4	-0.124	-0.307	**0.533**	0.713	-0.668	0.121	0.775	**0.173**	37
5	**-0.051**	-0.274	0.599	0.686	-0.567	0.083	0.832	0.281	29
6	-0.114	-0.306	0.588	0.674	-0.571	0.085	0.868	0.183	31
7	-0.052	-0.269	0.608	0.682	-0.590	**0.082**	**0.767**	0.274	27
8	-0.115	-0.303	0.600	**0.671**	-0.583	**0.082**	0.817	0.177	27

Table 10. Continued

NMSE									
	NH_4^+	SNH	NO_3	SNO_3^-	NO_2	O_3	SO_2	SO_4^{2-}	Rank
Model scenario									
1	0.43	**0.80**	1.10	2.40	1.50	**0.11**	9.60	2.50	30
2	0.59	0.87	1.30	2.70	**1.30**	0.12	13.00	3.30	49
3	0.55	0.87	1.20	2.60	**1.30**	**0.11**	13.00	2.5	38
4	0.43	1.10	**0.97**	**1.90**	2.00	**0.11**	**2.00**	**0.68**	22
5	0.40	1.10	1.10	2.30	1.70	0.12	2.80	0.93	36
6	0.46	1.10	1.00	2.10	1.70	**0.11**	2.90	0.72	28
7	**0.38**	1.00	1.10	2.30	1.70	**0.11**	2.40	0.92	24
8	0.44	1.00	1.00	2.10	1.70	**0.11**	2.50	0.72	23

Table 11. Overall ranking of the eight model scenarios for August 1998, by adding the ranks obtained for the three statistical parameters for the daily mean values. Lowest rank corresponds to best performance

Model scenario	1	2	3	4	5	6	7	8
Corr	29	42	39	20	37	29	33	24
FB	46	45	45	37	29	31	27	27
NMSE	30	49	38	22	36	28	24	23
Overall	105	136	122	79	102	88	84	74

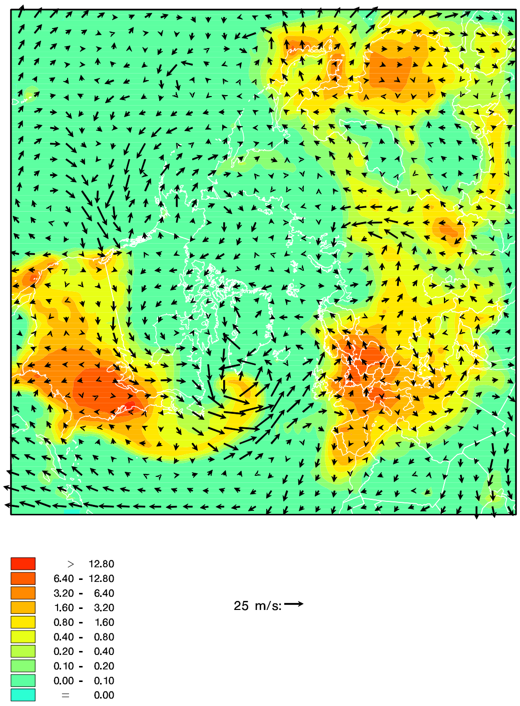


Fig. 1. Six-hour mean concentrations of NO_2 in ppb at December 24, 1998 calculated with the REGINA model. A plume of NO_2 from North America is transported across the Atlantic Ocean towards the Arctic and Europe due to a low pressure system moving east off the east coast of Canada.

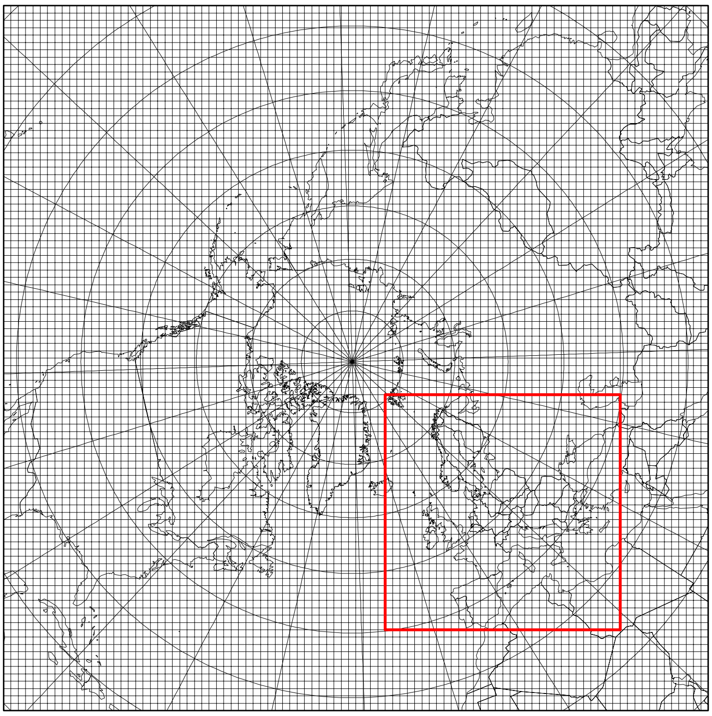


Fig. 2. The model domain is a polar stereographic projection of the Northern Hemisphere. First nest is placed over Europe and second nest is placed over Scandinavia. The increase in resolution is a factor of 3 for the nest. The number of grid points is 96 x 96 in the mother domain as well as in the nest.

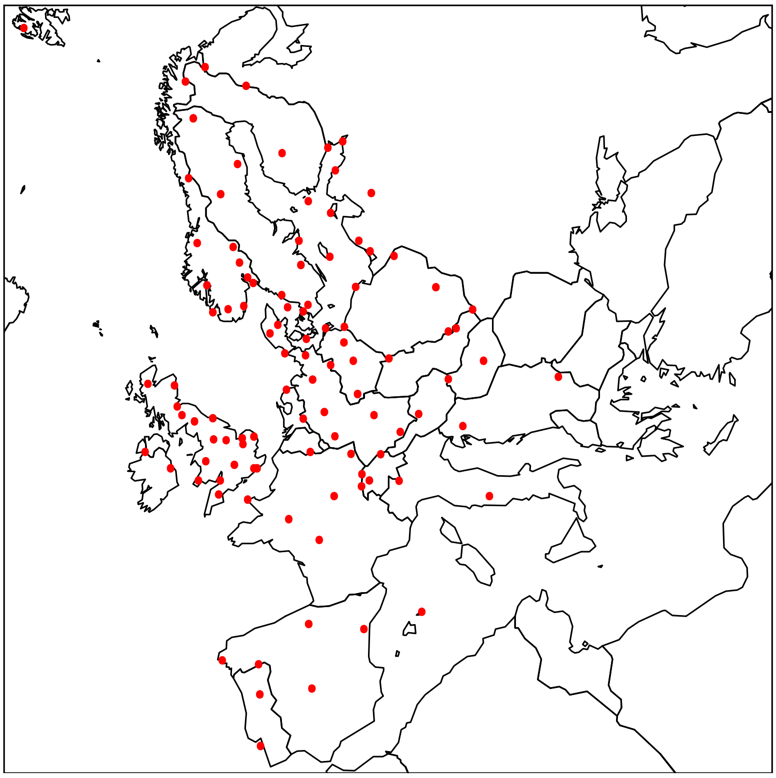
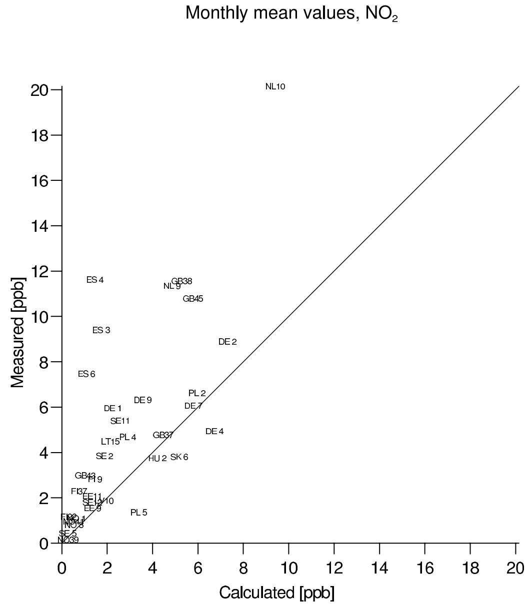
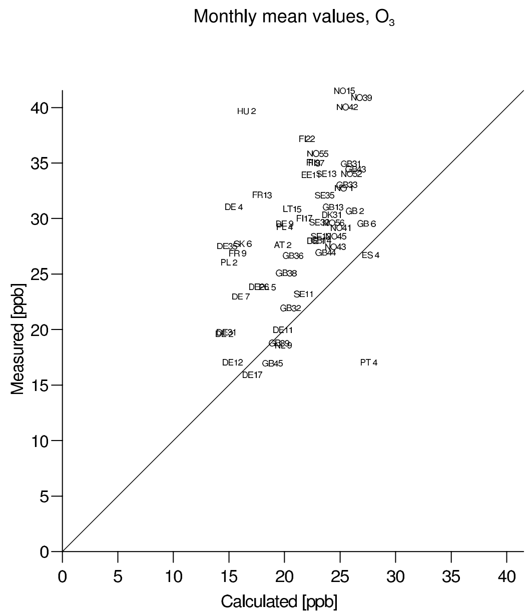


Fig. 3. European monitoring stations from which data are used in the validation of the REGINA model.



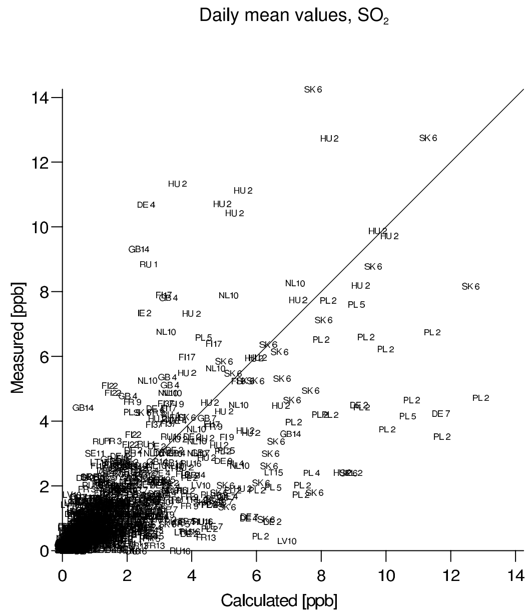
N = 34, means: calculated = 2.93, measured = 5.10
Standard deviations: calculated = 2.34, measured = 4.27
correlation = 0.70, test (Hypotesis: correlation=0): t = 5.55
bias = -2.169, FB = -0.540, FSD = -1.076, NMSE = .97E+00

Fig. 4. Comparison of measured and calculated monthly mean concentrations of NO_2 for February 1998 for model scenario eight. The model tends to underestimate the concentrations. The statistical parameters shown are mean values, standard deviations, correlation, test statistic for the student t-test, bias, fractional bias, fractional standard deviation and normalised mean square error.



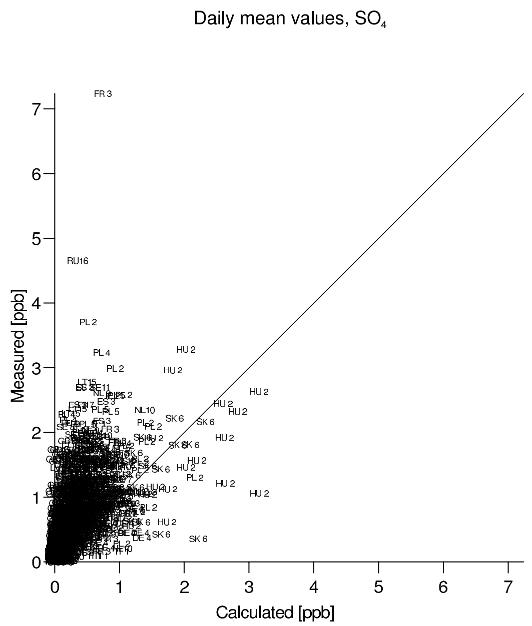
N = 57, means: calculated = 21.57, measured = 28.57
Standard deviations: calculated = 3.82, measured = 6.24
correlation = 0.47, test (Hypotesis: correlation=0): t = 3.92
bias = -7.005, FB = -0.279, FSD = -0.910, NMSE = .13E+00

Fig. 5. Comparison of measured and calculated monthly mean concentrations of O_3 for February 1998 for model scenario eight. The model tends to underestimate the concentrations, especially for the high values. The statistical parameters shown are mean values, standard deviations, correlation, test statistic for the student t-test, bias, fractional bias, fractional standard deviation and normalised mean square error.



N = 1114, means: calculated = 1.71, measured = 1.28
Standard deviations: calculated = 1.98, measured = 1.77
correlation = 0.71, test (Hypotesis: correlation=0): t = 33.84
bias = 0.430, FB = 0.288, FSD = 0.224, NMSE = .10E+01

Fig. 6. Comparison of measured and calculated daily mean concentrations of SO_2 for February 1998 for model scenario eight. The calculated concentrations are in good agreement with the measurements. The statistical parameters shown are mean values, standard deviations, correlation, test statistic for the student t-test, bias, fractional bias, fractional standard deviation and normalised mean square error.



N = 1366, means: calculated = 0.35, measured = 0.71
Standard deviations: calculated = 0.38, measured = 0.60
correlation = 0.54, test (Hypotesis: correlation=0): t = 23.40
bias = -0.355, FB = -0.670, FSD = -0.858, NMSE = .16E+01

Fig. 7. Comparison of measured and calculated daily mean concentrations of SO_4^{2-} for February 1998 for model scenario eight. The concentrations are underestimated by the model. The statistical parameters shown are mean values, standard deviations, correlation, test statistic for the student t-test, bias, fractional bias, fractional standard deviation and normalised mean square error.

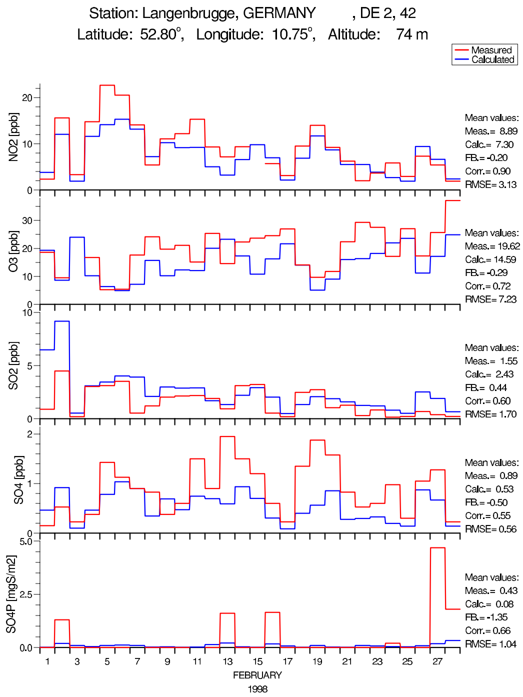


Fig. 8. Timeseries for the German station Langenbrügge for February 1998. Panels show from the top: atmospheric concentrations of NO_2, O_3, SO_2, SO_4^{2-} and SO_4^{2-} in precipitation.

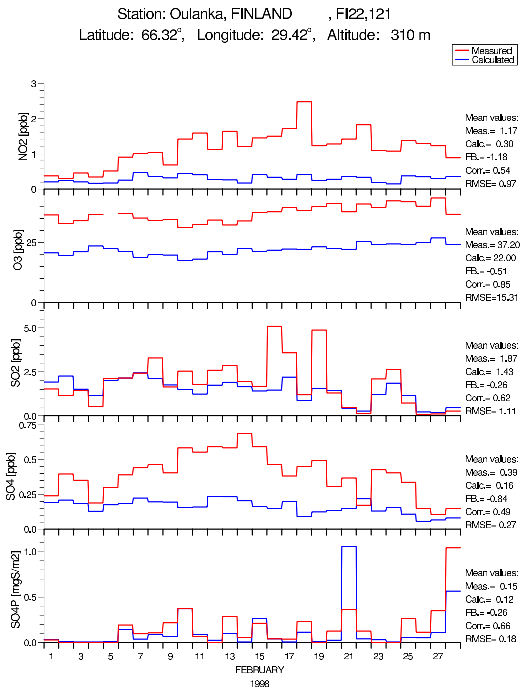
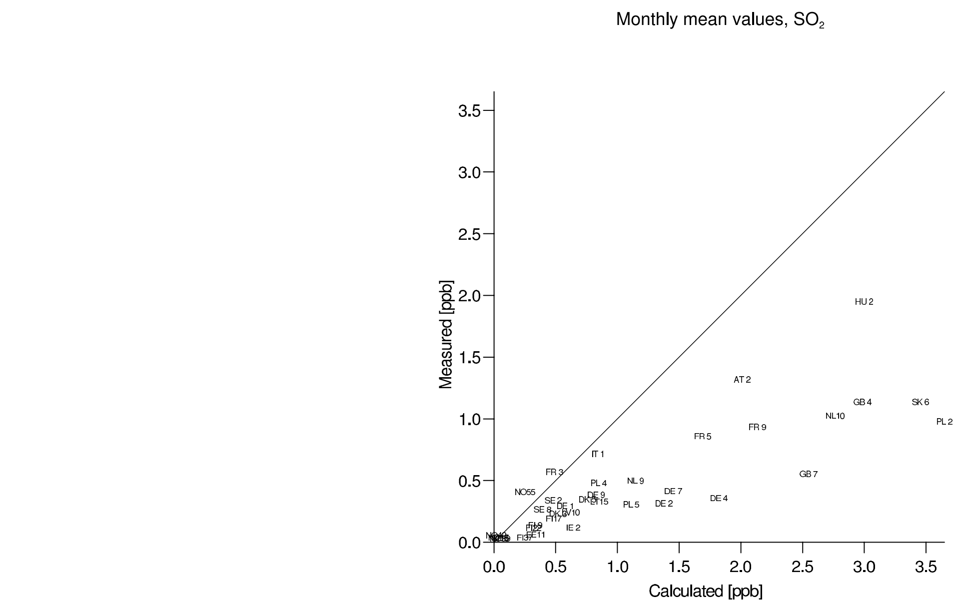
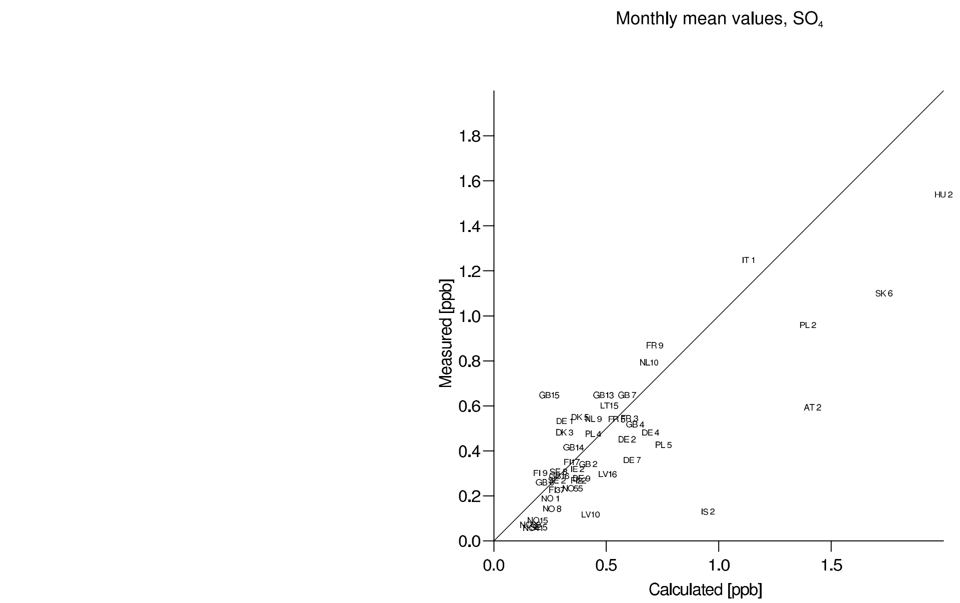


Fig. 9. Timeseries for the Finnish station Oulanka for February 1998. Panels show from the top: atmospheric concentrations of NO_2, O_3, SO_2, SO_4^{2-} and SO_4^{2-} in precipitation.



N = 36, means: calculated = 1.14, measured = 0.48
Standard deviations: calculated = 1.02, measured = 0.43
correlation = 0.84, test (Hypotesis: correlation=0): t = 8.90
bias = 0.663, FB = 0.818, FSD = 1.401, NMSE = .17E+01

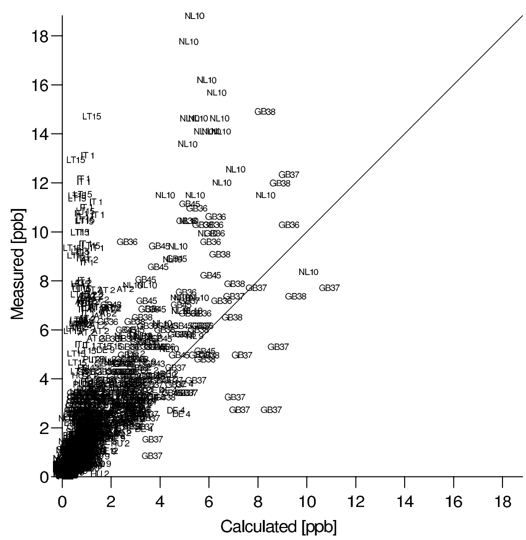
Fig. 10. Comparison of measured and calculated monthly mean concentrations of SO_2 for August 1998 for model scenario eight. The concentrations are overestimated by the model with a factor of two. The statistical parameters shown are mean values, standard deviations, correlation, test statistic for the student t-test, bias, fractional bias, fractional standard deviation and normalised mean square error.



N = 45, means: calculated = 0.55, measured = 0.46
Standard deviations: calculated = 0.40, measured = 0.31
correlation = 0.80, test (Hypotesis: correlation=0): t = 8.62
bias = 0.088, FB = 0.176, FSD = 0.513, NMSE = .27E+00

Fig. 11. Comparison of measured and calculated monthly mean concentrations of SO_4^{2-} for August 1998 for model scenario eight. The calculated concentrations are in good agreement with measured values for low concentrations and overestimated for the high concentrations. The statistical parameters shown are mean values, standard deviations, correlation, test statistic for the student t-test, bias, fractional bias, fractional standard deviation and normalised mean square error.

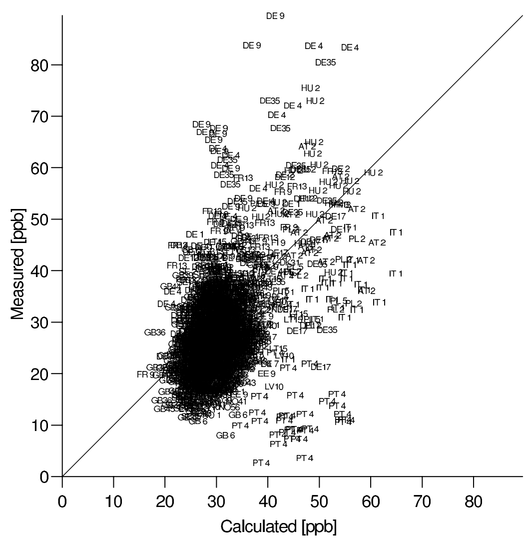
Daily mean values, NO_2



N = 1126, means: calculated = 1.45, measured = 2.65
Standard deviations: calculated = 1.63, measured = 2.96
correlation = 0.65, test (Hypotesis: correlation=0): t = 28.35
bias = -1.197, FB = -0.583, FSD = -1.071, NMSE = .17E+01

Fig. 12. Comparison of measured and calculated daily mean concentrations of NO_2 for August 1998 for model scenario eight. The concentrations are underestimated by the model. The statistical parameters shown are mean values, standard deviations, correlation, test statistic for the student t-test, bias, fractional bias, fractional standard deviation and normalised mean square error.

Daily mean values, O_3



N = 1621, means: calculated = 31.75, measured = 29.24
Standard deviations: calculated = 6.88, measured = 10.23
correlation = 0.42, test (Hypotesis: correlation=0): t = 18.64
bias = 2.512, FB = 0.082, FSD = -0.754, NMSE = .11E+00

Fig. 13. Comparison of measured and calculated daily mean concentrations of O_3 for August 1998 for model scenario eight. The concentrations are overestimated for some stations and underestimated by the model for others. The statistical parameters shown are mean values, standard deviations, correlation, test statistic for the student t-test, bias, fractional bias, fractional standard deviation and normalised mean square error.

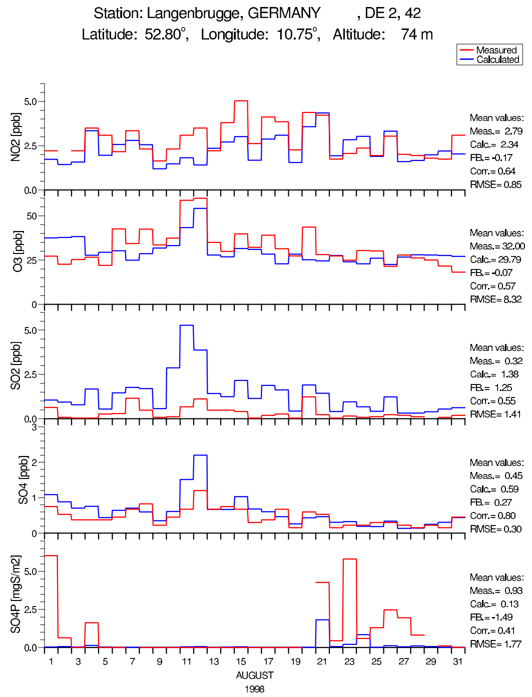


Fig. 14. Timeseries for the German station Langenbrügge for August 1998. Panels show from the top: atmospheric concentrations of NO_2, O_3, SO_2, SO_4^{2-} and SO_4^{2-} in precipitation.

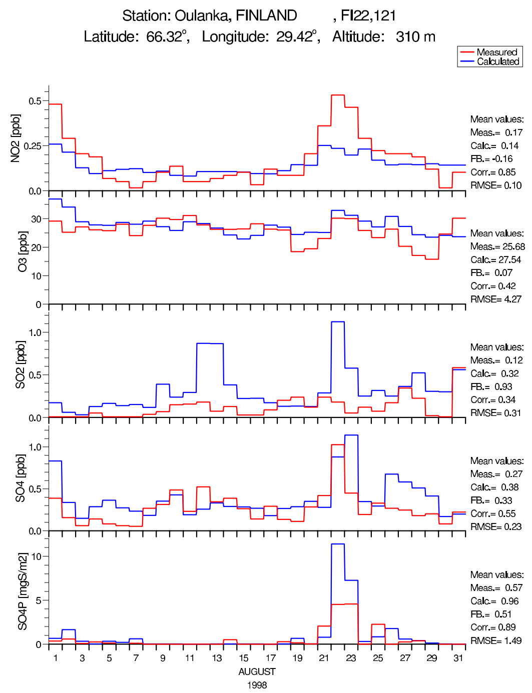


Fig. 15. Timeseries for the Finnish station Oulanka for August 1998. Panels show from the top: atmospheric concentrations of NO_2, O_3, SO_2, SO_4^{2-} and SO_4^{2-} in precipitation.