SmogProg

Towards operational smog forecasts
using near-real-time satellite measurements

D.P.J. Swart,
S. Jongen, H. Eskes,
M. Schaap, R.M.A. Timmermans, A. Seegers,
A.M.M. Manders, F.J. Sauter, J.P.J. Berkhout, D.E. Lolkema

NIVR project no. 53615RI
August 2008

P.O. Box 35
2600 AA Delft
The Netherlands
Kluyverweg 1
2629 HS Delft
The Netherlands
Tel. +31 (0)15 278 80 25
Fax +31 (0)15 262 30 96
Email info@nivr.nl
www.nivr.nl

NATIONAL USER SUPPORT PROGRAMME (NUSP) 2001-2005

http://www.ao-go.nivr.nl

The National User Support Programme 2001-2005 (NUSP) is executed by the Netherlands Agency for Aerospace Programmes (NIVR) and the SRON Netherlands Institute for Space Research. The NUSP is financed from the national space budget. The NUSP subsidy arrangement contributes to the development of new applications and policy-supporting research, institutional use and use by private companies.

The objectives of the NUSP are:

- To support those in the Netherlands, who are users of information from existing and future European and non-European earth observation systems in the development of new applications for scientific research, industrial and policy research and operational use;
- To stimulate the (inter)national service market based on space-based derived operational geo-information products by means of strengthening the position of the Dutch private service sector;
- To assist in the development of a national Geo-spatial data and information infrastructure, in association with European and non-European infrastructures, based on Dutch user needs;
- To supply information to the general public on national and international space-based geo-information applications, new developments and scientific research results.

The National Institute for Public Health and the Environment (RIVM) is a recognized leading centre of expertise in the fields of health, nutrition and environmental protection. We work mainly for the Dutch government. We also share our knowledge with governments and supranational bodies around the world. The results of our research, monitoring, modeling and risk assessment are used to underpin policy on public health, food, safety and the environment. We employ over 1500 employees, many of whom work in multidisciplinary fields.

Our mission is to benefit people, society and the environment, matching our expertise, knowledge and research with that of colleagues from around the world

Contact information relating this report:

RIVM,
Laboratory for Environmental Monitoring,
M: P.O. Box 1, 3720 BA Bilthoven, the Netherlands.
T: +31 30 274 26 77
F: +31 30 228 75 31
E: daan.swart@rivm.nl
W: www.rivm.nl

KNMI

KNMI is the national institute for weather, climate research and seismology. It disseminates weather information to the public at large, the government, aviation and the shipping industry in the interest of safety, the economy and a sustainable environment. To gain insight into long-term developments, KNMI conducts research on climate change. Making the knowledge, data and information on hand at KNMI accessible is one core activity. KNMI is an agency of the Ministry of Transport, Public Works and Water Management (Ministerie van Verkeer en Waterstaat). KNMI's duties are set forth in KNMI Act (Wet op het KNMI).

Contact information relating this report:

KNMI,
M: P.O. Box 201, 3730 AE De Bilt, the Netherlands.
T: +31 30 220 63 52
F: +31 30 221 04 07
E: henk.eskes@knmi.nl
W: www.knmi.nl

TNO's mission is to apply scientific knowledge with the aim of strengthening the innovative power of industry and government.

The combination of developing, integrating and applying knowledge differentiates us from other knowledge institutions. By encouraging the effective interplay of knowledge areas, we generate creative and practicable innovations: new products, services and processes, fully customized for business and government.

TNO Built Environment and Geosciences is an innovative partner for government and industry for sustainable organization, use and management of the built environment, infrastructure, subsurface and the environment in general.

Contact information relating this report:

TNO Built Environment and Geosciences,
Department of Air Quality and Climate Change,
M : P.O. Box 80015, 3508 TA Utrecht, the Netherlands
T: +31 30 256 42 56
F: +31 30 256 44 75
E: peter.builtjes@tno.nl
W: www.tno.nl

Contents

Abstract

Smog forecasting in the Netherlands is a legal task of RIVM, performed in commission of and paid by the Dutch Ministry of Environment. The current smog forecast is based on hourly measurements in the Dutch Air Quality Monitoring Network, temperature forecasts by the Dutch Met Office KNMI, and a statistical model PROZON maintained by RIVM.

This project aims to develop and demonstrate a new instrument to improve the existing smog forecasts in the Netherlands. It extends the chemical transport model LOTOS-EUROS with Ensemble Kalman Filter data assimilation to improve the forecast performance of the free running model. Assimilated data include surface ozone data from ground based networks in the Netherlands and neighbouring countries, and tropospheric nitrogen dioxide column data over Western Europe measured by OMI on board of NASA's AURA satellite.

LOTOS-EUROS is an operational Chemical Transport Model that is used, maintained and developed further by a consortium of Dutch research institutes: TNO, RIVM, PBL and KNMI. Model results and validation studies have been reported in literature and the model has participated in model comparisons for ozone and aerosols.

The groundbased data used in this project are acquired by the member states under the EU "Directive on ambient air quality and cleaner air for Europe". Member states are obliged under EU legislation to perform these measurements and report the results to the European Union. The directive will be extended to include near-real-time data in the near future. RIVM operates the Dutch air quality monitoring network and contributes its data to the EU. It has access to all similar European data at no cost.

The model has been adapted, and the necessary real time data streams have been implemented. Ground based data are assimilated and used in the forecast. Satellite data can technically be assimilated and used in the forecast but suffer from discrepancies between observations and model. Some flaws remain in the stream of European scale ground based data, but these are outside the scope of this project and will be resolved soon.

Performance
The performance of the new system has been validated using a historical dataset, and results were compared to the statistical model PROZON. Results show that the new tool shows higher spatial and temporal resolution (13 km, 1 hour), and provides the new capability to generate maps and movies of the evolution of air quality over a large part of Western Europe, including a good quality *nowcast*. The quality of its *one-day forecast* is better than PROZON with respect to timing issues: in a changing situation PROZON is often seen to lag behind, where SmogProg is in better synchronisation with the observations. On the other hand, SmogProg at this stage of development still has considerable difficulty to predict next day's absolute smog levels correctly, and is generally underestimating levels above 150 $\mu g/m^3$. As the relevance of smog forecasting is closely linked to predicting correctly the crossing of the European information and alarm thresholds of 180 and 240 $\mu g/m^3$, this is still a major flaw. It is however anticipated that the new model's performance on this point can be improved considerably in the coming months.

Cost-benefit

Due to its simplicity, the operational costs of the old tool PROZON are very low. As a consequence, the new tool SmogProg will inevitably be more expensive to run on an operational basis. A cost-benefit analysis should therefore focus on the question if the additional benefits obtained outweigh the additional costs.

The additional benefits lie in a better scientific understanding of the smog situation and in forecasts with better timing and higher spatial resolution, which enable sensitive people to avoid adverse situations more effectively. While these are clear benefits for scientists, policymakers and society, it is hard to express them directly in economic value. From economic point of view the increased communication power to non-experts is considered the most direct asset of the new system: Smog-induced healthcare costs amount to at least 100 million euro per year in the Netherlands. Part of these costs can be avoided by "smog-aware" behaviour of the population. If the new tool is not considered cost-effective on the less material grounds mentioned above, just a minor swing of the order of 0.1 % in health related costs will make the new tool cost-effective on a pure economic basis.

Executive summary

Monitoring and forecasting air quality in the Netherlands is a legal task of RIVM, required by Dutch and European law and international treaties. RIVM's operational forecast is currently based on a statistical model, PROZON. RIVM would like to improve on the current statistical approach for four reasons:

1. Quality: especially the beginning and end of the episode are difficult to predict accurately with the current system.
2. Spatial resolution: the actual smog distribution over the Netherlands is often not adequately represented by the three regions north, middle, and south used in the current forecast. It would be attractive to give a geographically more detailed prognosis.
3. Communication: like in the weather forecast, maps showing the development of air quality are easily understood, and are a good way to communicate an adverse situation to non-experts and the general public, inducing "smog-aware" behaviour and reducing health effects and costs.
4. Scientific understanding: having a model that describes a smog situation physically explicit allows better understanding of the situation. Studies can be undertaken on major causes, and possible abatement strategies.

In this project, we develop a new tool, named SmogProg, which aims to make progress along these four lines. Principle user of the new SmogProg machine is the RIVM smog team, responsible for the Dutch smog forecast. The RIVM smog team defined in a separate work package demands and desires on the output.

SmogProg is based on the Dutch chemical transport model LOTOS-EUROS that is used, maintained and developed further by a consortium of research institutes: TNO, RIVM, PBL and KNMI. We combine model results with surface ozone measurements from ground based networks in the Netherlands and neighbouring countries, and tropospheric nitrogen dioxide column data over Western Europe measured by OMI on board of NASA's AURA satellite. Combining data from these different sources is done using a data-assimilation technique called Ensemble Kalman filtering.

A real time data stream has been set up transferring Dutch air quality network data from RIVM to KNMI, where the model runs. A similar stream brings surface ozone data from Belgium and Germany from the European Environmental Agency (EEA). This second stream is not yet flawless, but work is under way at EEA to fix this. Satellite data from OMI is available directly at KNMI, as KNMI hosts the Principle Investigator of OMI and is one of the lead institutes in OMI data analysis.

Existing ensemble Kalman filter software for the assimilation of surface observations was further extended and numerous sensitivity experiments have been performed to optimise the settings of the Kalman filter for the assimilation of surface ozone observations from the Dutch air quality network. In our approach, model parameters like NO_x and VOC's source strengths, ozone levels at the top of the model domain, and dry deposition are dynamically adjusted to achieve a better agreement between the model state and the measurements in a balanced way.

A new module has been developed to simulate/predict the OMI NO_2 observations from the LOTOS-EUROS three-dimensional concentrations. This includes the convolution of the model NO_2 profile with the averaging kernel provided by the OMI retrieval, and the error provided in the OMI product is explicitly used in the assimilation. Assimilation experiments have been conducted for the month July 2006.

Measurement data is used in two ways to improve the forecast:
Actual measurements from the near past are used to tune the model away from its natural free-running state, and make the state of the model as realistic as possible.
This in itself provides a better forecast, as the starting point is more accurate.
For the forecast itself, data-assimilation cannot be used as such data does not yet exist. For the forecast, we therefore have to rely again on the free-running model fed by prognostic meteorological fields. Yet, in SmogProg we have extended the use of the assimilated data in the forecast by developing an inheritance scheme that passes down the adjusted model parameters to the forecast run, rather than relaxing fully towards the free-running model. Here we assume that the adjustments made in the past will be beneficial in the future too. We have tested this assumption and studied extensively what works best in this respect. We found that, when assimilating surface ozone data, passing down the parameter adjustments of 15:00 hours of the previous day increases the quality of the forecast most.

Over the last year, the project has demonstrated the capability to run the prognostic system 24/7 in real time, and make the results available on the web. At this stage however this applies to the free-running model with actual and prognostic meteo fields, and in parallel the ground based data from the Netherlands and the satellite data from OMI. In addition, we also run forecasts with the French CHIMERE model on a daily basis. In this way we have created a "mini-ensemble" forecast and we can study model differences. The data assimilation part is not yet included in real time. These developments are under way and will be complete before the end of 2008.
RIVM expects to continue this 24/7 web service after the completion of the project as part of the new smog forecast system, as it is an essential element in communication, and in making the project cost-effective.

The performance of the new system has been validated using a historical dataset of July 2006 that includes several smog episodes. Results were compared to the statistical model PROZON. The comparison shows that the new tool shows higher spatial and temporal resolution (13 km, 1 hour), and provides the new capability to generate maps and movies of the evolution of air quality over a large part of Western Europe, including a good quality *nowcast*. PROZON only generates a prognosis of the daily ozone maximum at 22 monitoring locations. The quality of its *one-day forecast* is somewhat better than PROZON with respect to timing issues, but PROZON is still the clear winner when predicting next day's smog levels correctly. Major shortcoming is that high ozone peaks (above 150 $\mu g/m^3$) are currently underestimated by the SmogProg model, so threshold levels are not reached even if in reality moderate smog levels do occur. Further improvements in the quantitative results of SmogProg are anticipated with the implementation of an improved chemistry scheme, the determination and elimination of possible model biases (especially those that occur specifically during smog episodes), further development of the OMI assimilation part, and the inclusion of a larger number of surface ozone stations if these become available in time. Another possibility to improve the quantitative agreement of the forecast with the actual levels is the inclusion of a statistical post-processing. Budget to implement at least some of these improvements is already allocated by RIVM, TNO and KNMI, and will be part of the work of the coming months.

A cost benefit analysis shows that ozone smog-induced healthcare costs amount to at least 100 million euro per year in the Netherlands. Part of these costs can be avoided by “smog-aware” behaviour of the population. The increased communication power to non-experts is considered a large asset of the new system. It is much better suited to communicate adverse air quality situations to non-experts and the general public. Costs of development, migration and operation are likely to be small compared to the reduced healthcare costs if only a small fraction of the public reduces its exposure to smog during episodes. Just a minor swing of the order of 0.1 % in health related costs will make the new tool cost-effective.

The SmogProg proposal was originally written for a two-year period. This report shows the progress made in only 19 months (Jan 2007-July 2008). We conclude that the project is in good shape but not yet complete, and that in the remaining 5 months further effort is needed to bring forecast results numerically closer to reality during smog episodes. The additional benefit of the assimilation of OMI data must be assessed in more detail. Further work is also needed on the interface side of the model, to make it better usable for the RIVM smog team. RIVM, KNMI and TNO will continue the work on this project, and update this report in approximately 6 months time.

1 Introduction

1.1 What is smog, and why is it relevant?

An introduction adapted from Wikipedia, the free encyclopaedia[1]:

Smog is a kind of air pollution; the word "smog" is a portmanteau of smoke and fog. Classic smog results from large amounts of coal burning in an area and is caused by a mixture of smoke and sulphur dioxide. Modern smog does not usually come from coal but from vehicular and industrial emissions that are acted on in the atmosphere by sunlight to form secondary pollutants that also combine with the primary emissions to form photochemical smog. In the 1950s *photochemical smog* was first described. It forms when sunlight hits various pollutants in the air and forms a mix of inimical chemicals that can be very dangerous. A photochemical smog is the chemical reaction of sunlight, nitrogen oxides (NO_x) and volatile organic compounds (VOCs) in the atmosphere. Nitrogen oxides are released by nitrogen and oxygen in the air reacting together under high temperature such as in the exhaust of fossil fuel-burning engines in cars, trucks, coal power plants, and industrial manufacturing factories. VOCs are released from man-made sources such as gasoline (petrol), paints, solvents, pesticides, and biogenic sources, such as pine and citrus tree emissions.

This noxious mixture of air pollutants can include the following: nitrogen oxides, tropospheric ozone, volatile organic compounds (VOCs), peroxyacyl nitrates (PAN), and aldehydes. All of these chemicals are usually highly reactive and oxidizing. Photochemical smog is therefore considered to be a problem of modern industrialization. It is present in all modern cities, but it is more common in cities with sunny, warm, dry climates and a large number of motor vehicles. Because it travels with the wind, it can affect sparsely populated areas as well.

Smog harms human health (WHO, 2003). Ground level ozone, sulphur dioxide, nitrogen dioxide and carbon monoxide are especially harmful for senior citizens, children, and people with heart and lung conditions such as emphysema, bronchitis, and asthma. It can inflame breathing passages, decreasing the lungs' working capacity, and causing shortness of breath, pain when inhaling deeply, wheezing, and coughing. It can cause eye and nose irritation and it dries out the protective membranes of the nose and throat and interferes with the body's ability to fight infection, increasing susceptibility to illness. Hospital admissions and respiratory deaths often increase during periods when ozone levels are high (Fischer et al, 2004).

1.2 Smog forecasting in the Netherlands: current practice

To avoid exposure of the population to high ozone concentrations, ozone levels are monitored continuously by air quality monitoring networks, and warnings are given when concentrations exceed limit values. In the Netherlands, the Dutch Air Quality Monitoring Network (LML, Landelijk Meetnet Luchtkwaliteit) is operated by RIVM. Apart from these ad hoc warnings, smog forecasts are given. Smog forecasting in the Netherlands is a legal task of RIVM,

[1] See http://en.wikipedia.org/wiki/Smog

performed in commission of and paid by the Dutch Ministry of Environment. The current smog forecast is based on hourly measurements in the Dutch Air Quality Monitoring Network, temperature forecasts by the Dutch Met Office KNMI, and a statistical model PROZON maintained by RIVM.

The current smog forecasting model, PROZON, has been developed around 1990 and has been used since 1992 (Noordijk 2003) as operational model. It is a statistical model which predicts daily maximum ozone concentrations at the measurement locations of the Dutch Air Quality Monitoring Network (see appendix A). Tomorrow's maximum is determined by today's maximum ozone concentration, today's maximum temperature and tomorrow's (expected) maximum temperature:

O_3(tomorrow) = O_3(today) * factor(temperatures, concentration).

The factor is determined by evaluating measurements from the past and constructing statistics from it. Apart from the temperature, also the month of the year, the region (north, middle or south) and the location type (urban, traffic, rural) are taken into account. Measurements of the ozone concentrations are provided by LML, the temperature forecast is given by KNMI.

PROZON is run twice a day: once in the morning (8.30 local time) when yesterday's day maximum concentration and temperature are known, and today's maximum can be forecasted relatively accurately, and once in the afternoon (15.30 local time) to get an indication for tomorrow. This forecast is less accurate, since the forecasted maximum temperature for today and tomorrow have to be used and today's maximum concentration is in general not yet reached. Forecasting further ahead is possible but becomes inaccurate. Running the model is fast: it takes only a few seconds. Producing the statistical factors is the time-consuming part of the model; this is done only once every few years. The statistical parameters of the current version were updated in 2000.

In general, the results of PROZON are reasonable, as shown by Noordijk (2003) and an evaluation of PROZON over 2005 and 2006 (Manders, Nguyen and Hoogerbrugge 2008). But there are some limitations:

- forecasts are only given for monitoring locations
- only daily maximum concentrations are forecasted
- only one temperature is used for the whole country

These limitations were accepted since the monitoring network is distributed quite evenly and relatively dense and high concentrations always occur late in the afternoon when the temperature is high and the process of ozone formation has been acting all day on a warm sunny day. However, the forecast is less accurate at the beginning and end of a smog episode when there are large regional differences due to for example the passage of a front. In some parts of the country the concentrations may still rise and in other parts cool air and clouds may already inhibit the formation of ozone. Apart from the uncertainty in the model itself, the uncertainty in the temperature forecast plays a role.

On the RIVM smog website (www.lml.rivm.nl), actual concentrations are displayed and forecasts are given per station. The forecast is also presented on videotext, there only the maximum value per zone (north, middle and south, as defined in the Smogregeling 2001) is given. The forecast produced by PROZON is always evaluated by an expert of the RIVM smog team before it is published on the web. In few cases, the forecast is overruled by a smog team member and only one value per zone is given on the web site.

1.3 Goal of the project/benefits

Although the present forecasting practice is fairly accurate, there is evident room for improvement: the spatial coverage can be increased and extended and the time evolution can be modelled by using a chemical transport model (CTM). The goal of the SmogProg project is to use a CTM to give a better forecast regarding these aspects and to better communicate the forecasts to stakeholders and the general public.

CTMs have the advantage that the full temporal and spatial evolution of concentrations can be modelled by using emissions of ozone precursor gases and weather to explicitly calculate ozone production using chemical reactions, also the transport of precursor gases and ozone itself are accounted for. Much of the changes in air pollution levels is caused by changes in the weather (temperature, rain, clouds, wind speed and direction, vertical stability) rather than by changes in the emissions. One clear example of the strong link between atmospheric chemistry and the weather is the sharp gradient in ozone concentrations observed during the passage of fronts. With the models we can describe and understand these effects and the impact on the chemistry and air pollution levels.

CTMs take more computation effort which make them more expensive to use, and more important, free-running models are often less accurate in predicting daily maximum concentrations than PROZON one day ahead. For forecasts on a longer term the difference in accuracy becomes smaller.

The reason for this is that PROZON is updated daily with new observations and the underlying statistics include many different circumstances. A CTM has inherent model uncertainties, like the exact emissions of ozone precursors and uncertainties in the chemical reaction rates. In free running mode it will not be corrected towards reality. This can be compensated for, by using data-assimilation: using observations to let the model select optimal settings within the limits of the model and observation uncertainties for example with respect to the amount of emissions. In this way, least the initial conditions for a forecast are improved. When improved emission values are used for the forecast, the forecast should be improved as well. At the start of the project, it was not evident how to use input parameter settings (which were obtained by data assimilation) in making a forecast. Finding this out was the fundamental research part of this project.

Even if the forecast would only be improved slightly, benefits are also an increased time resolution, with the timing of the peak concentrations, better resolution of regional differences and better coverage. These improvements will enable a better communication of the ozone forecast to the general public since the results will be easier to understand. Chapter 2 lays down the requirements for SmogProg of RIVM's smog team. These requirements result from Dutch and European legislation and experiences in the daily forecasting practice.

1.4 European perspective

The SmogProg project is strongly related and tightly linked to the atmosphere services of the European Global Monitoring for Environment and Security (GMES) programme. GMES is a joint initiative of the European Commission (EC) and the European Space Agency (ESA).

The aim of GMES is to optimally combine all available surface, in-situ, remotely sensed and space based observations with state-of-the art models, and to set up information services for specific users, policy makers, scientists and the general public as a whole. Within the GMES programme a series of operational satellites will be launched – the so called sentinels – to provide the space component of the programme. Sentinels 4 and 5 are atmospheric composition missions, and observation of air quality is one of the prime goals of these missions.

The scope of a “GMES Atmosphere Service” (GAS) has recently been defined. One of the main activities will be the provision of European-scale air quality forecast and analysis information. This air-quality service will be based on an ensemble of air quality models which will be run in parallel. These models have been developed in he past in the individual EU/ESA member state countries. The result is an ensemble forecast of aerosols, ozone and other trace gas species. The assimilation of both surface observations and satellite measurements is foreseen.

The “PROtocol MOniToring for the GMES Service Element: Atmosphere” (PROMOTE) project (stage 2, July 2006 - August 2009, http://www.gse-promote.org/) is part of the GMES Service Element programme of ESA. The GMES Service Element for Atmosphere PROMOTE delivers policy-relevant services on multiple atmospheric issues to end-users. With a large consortium of 20 institutes, including service providers and researchers, PROMOTE focuses on stratospheric ozone, surface ultraviolet radiation, air quality, greenhouse gases, and special services. PROMOTE services assist user organizations in the public sector in things such as their monitoring obligations and in their tasks to warn the public for air pollution episodes or enhanced levels of ultraviolet radiation. Currently about 50 such user organisations in 16 European countries have signed a Service Level Agreement with PROMOTE. With these services PROMOTE reaches out to 30 % (air quality) or even 60% (UV) of the population in EU25. International organizations like EEA and WMO are provided with information such as long-term satellite data on stratospheric ozone and up-to-date information, including forecasts, on the Antarctic ozone hole. PROMOTE does not directly fund product development, but there are (modest) funds provided for setting up internet interfaces to existing data and for optimising the services in response to feedback from the users.

The air quality service of PROMOTE has two elements which are tightly related to SMOGPROG. As part of regional air-quality services there is the “Air Quality Forecast for The Netherlands” service. This service provides an extra internet interface to the air quality forecasts developed within SMOGPROG. In this way the Europe-wide visibility of the SMOGPROG development is enhanced. Secondly, PROMOTE has developed an “Integrated Air Quality platform for Europe”, based on forecasts of 5 different air quality models. One of these models is LOTOS-EUROS, and TNO is the service provider for this. LOTOS-EUROS forecasts will be delivered starting from the summer of 2008. These forecasts will come from the operational system developed within SMOGPROG.

The EU-funded GEMS project (2005-2009, http://gems.ecmwf.int/) is developing comprehensive data analysis and modelling systems for monitoring the global distributions of atmospheric constituents, important for climate, air quality and UV radiation, with a focus on Europe. Similar to PROMOTE, the European-scale air quality forecast follows the ensemble of models approach (but LOTOS-EUROS is not part of GEMS). The SMOGPROG and GEMS development of air quality forecasting has many things in common, and the interaction is of mutual benefit. In particular, the near-real time data flow of RIVM LML data to KNMI, set up within SMOGPROG, has also been used to provide GEMS with the needed near-real time data from the individual member state countries.

The “Monitoring Atmospheric Composition and Climate” (MACC) proposal is a next step towards the development of the operational GMES Atmosphere Service. MACC is part of the EC 7th framework programme, and is currently in the final phase of the negotiations. MACC (Monitoring Atmospheric Composition and Climate) is designed to meet the requirements that have been expressed for the pilot Core GAS. The project has been prepared by the consortia of the FP6 project GEMS and the GSE project PROMOTE, whose core service lines will provide the starting point for MACC. From mid-2009 MACC will continue, improve, extend, integrate and validate these service lines, so that the overall MACC system is ready near the end of 2011 for qualification as the operational Core GMES Atmospheric Service. MACC will prepare the core service in terms of implementation, sustained operation and availability. Part of MACC is the Integrated Air Quality platform for Europe, which will again be based on an ensemble of air quality models developed nationally. MACC will merge the air-quality forecasting activities of GEMS and PROMOTE, and LOTOS-EUROS will be part of this ensemble. The SMOGPROG work provides important preparations for the operational forecasting which is planned to start in June 2009.

To summarise, the SmogProg developments are of direct benefit to the (precursor) activities related to the GMES Atmosphere Service:

- The PROMOTE project provides a second interface to the SmogProg forecasts for the Netherlands and surrounding countries. This extends the visibility of the SmogProg work within Europe.
- The SmogProg forecasts of LOTOS-EUROS on the European domain will be used by PROMOTE for their European ensemble forecast.
- SmogProg has helped to set up an operational data supply of the near-real time Dutch LML air quality surface measurements to the European GEMS project. As a result the Netherlands was one of the first countries to deliver this data.
- The SmogProg experiences with the assimilation of satellite OMI measurements and surface observations with the ensemble Kalman filter technique are of direct relevance for projects like MACC and for the exploitation of satellite data within GMES.
- The operational forecast with the LOTOS-EUROS model, set up within SmogProg, is the precursor of the Dutch contribution to the European ensemble air quality forecast of the MACC (GMES Atmosphere Service) project.

1.5 Test episode

We have chosen July 2006 as test period for SmogProg, since it is characterized by a number of episodes with high ozone concentrations. They are related to the high temperatures in this period. On most days it was warmer than 25 ˚C and on some days warmer than 30 ˚C with weak southerly or easterly winds (not shown). Figure 1 clearly shows the day-night cycle with high values late in the afternoon and low values early in the morning. Ozone concentrations are well correlated with the temperature. Three episodes can be distinguished: roughly from 29 June to 6 July, from 17 to 20 July and from 25 to 28 July.

The average picture gives a good indication of the episodes and the day night cycle. But there are clear differences between the individual stations, which can be attributed to their location. In Figure 2 the day maxima of 5 individual stations are shown. These 5 stations show the extremes: station Kollumerwaard is near the Wadden Sea in the North, far away from sources of precursors, and has the lowest peak values (and relatively high background values), station Philippine is in the South-West and has its maximum concentration in the second episode one day earlier than most stations. Station Hellendoorn is in the Eastern part and there the end of the episode is somewhat later than for the other stations.

This one-month period is certainly not representative for an average summer. But from the operational perspective it is especially relevant to forecast high ozone values correctly, as well as the onset and end of an episode with high concentrations. Both underestimations and false alarms must be avoided, since the former has direct consequences for public health and the latter will induce the loss of public interest for the next alarm.

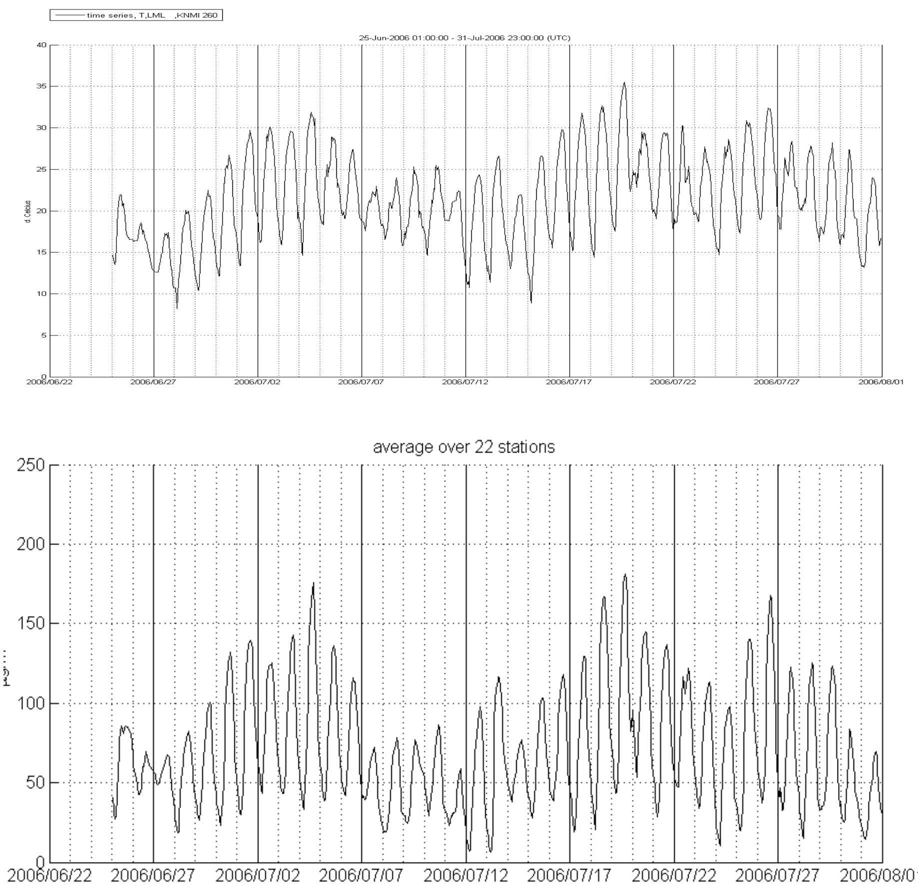


Figure 1. Upper panel: observed temperature in De Bilt. Lower panel: observed ozone concentrations, average over 22 rural stations.

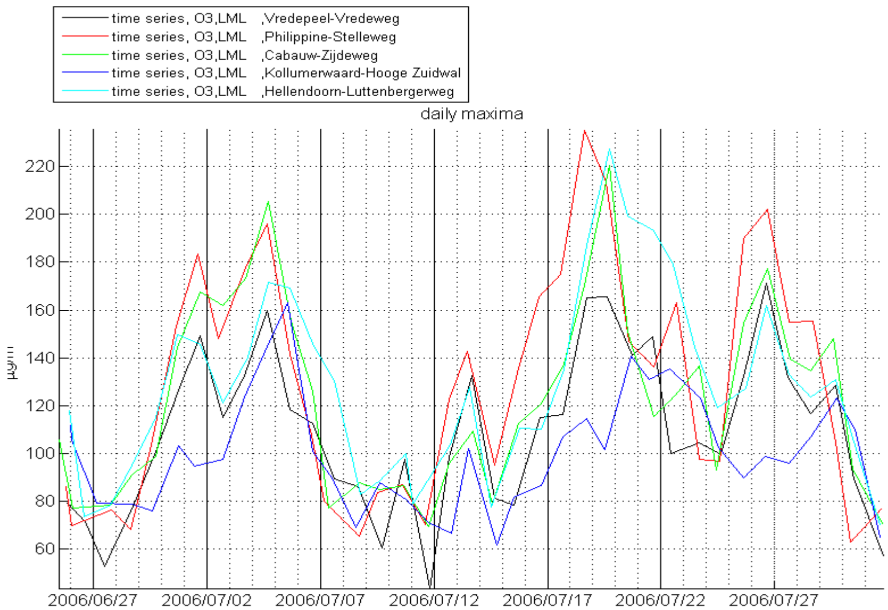


Figure 2. Observed daily maxima for 5 rural stations. These different stations are representative of the relatively large regional differences which exist over the Netherlands.

1.6 Ground and satellite observations

In this project, the LOTOS-EUROS model is used for forecasting ozone. This model has already been in use as a research model at RIVM and TNO since 2004 and KNMI has recently joined the consortium. The model and the model development resulting from the SmogProg project are described in Chapter 3. In this project, data assimilation of ozone ground observations and NO_2 observations from the OMI satellite are used to constrain LOTOS-EUROS' inherent uncertainties.

For validation and assimilation, we will use ozone ground observations from LML. These data become available every hour in real time. An overview of the locations is given in appendix A. Only rural stations will be used for assimilation and validation, since local titration processes on urban and street locations cannot be resolved by LOTOS-EUROS. This is no serious drawback since maximum concentrations do not occur at these locations in the Netherlands. Since ozone is not a local phenomenon but is influenced by transport, also data from rural stations in Germany and Belgium will be used (appendix A). Hourly data from the past can be obtained from AIRBASE and since 2007 on a daily basis from the European Environmental Agency (EEA). These data are not yet available in real time but are at the moment of writing updated once daily. A further improvement of the data availability from EEA is foreseen in the near future. The groundbased data used in this project are acquired by the member states under the EU "Directive on ambient air quality and cleaner air for Europe". Member states are obliged under EU legislation to perform these measurements and report the results to the European Union. The directive will be extended to include near-real-time data in the near future. RIVM operates the Dutch air quality monitoring network, contributes its data to the EU and it has access to all similar European data at no cost.

The Ozone Monitoring Instrument (OMI) has been launched on NASA's EOS-AURA satellite on 15 July 2004 and is working according to expectations. OMI has been built by Dutch Space/TNO-TPD in The Netherlands, in co-operation with Finnish subcontractors VTT and Patria Finavitec. The KNMI (Royal Netherlands Meteorological Institute) is the scientific lead institute. OMI is a UV/VIS nadir solar backscatter spectrograph, which provides near-global coverage in one day with high spatial resolution. OMI measures various key parameters for stratospheric and tropospheric chemistry and for climate research, including ozone (O_3) (column and profile), nitrogen dioxide (NO_2), sulphur dioxide (SO2), OClO, BrO, formaldehyde (HCHO), UV-B radiation at the surface, aerosol and cloud properties. OMI's high spatial resolution leads to a relatively high percentage of cloud-free pixels, thus giving better penetration into the troposphere than any other UV/VIS backscatter instrument flown to date, e.g. SCIAMACHY and GOME, which are, apart from resolution, in many respects similar to OMI.

The Ozone Monitoring Instrument (OMI) is an excellent tool to monitor worldwide NO_2 concentrations, as it achieves coverage of the entire Earths atmosphere in a single day with ground pixel sizes as small as ~24x13km2 at nadir. This small pixel size approaches the size of large cities, enabling the identification of individual major sources. As a result OMI observations are particularly promising for regional air quality applications. The daily global coverage allows for a day-to-day observation of emission sources, as well as the detection and tracking of transport phenomena. Compared to its predecessors like GOME and SCIAMACHY, the improved spatial resolution of OMI together with its daily global coverage is a huge step forward.

Retrievals of tropospheric NO_2 are available in near-real time mode (within 3 h of the actual OMI measurement at 13:40). This near-real time product was developed within the NIVR DOMINO and DOMINO-2 projects. The retrieval is based on the combined retrieval-assimilation-modelling approach originally developed at KNMI for off-line tropospheric NO_2 from the GOME and SCIAMACHY satellite instruments. We refer to the DOMINO final report for more details.

The DOMINO tropospheric NO_2 columns have been validated versus independent measurement during various campaigns. During the INTEX-B campaign in March 2006 over the southern United States, Mexico, and the Gulf of Mexico, DOMINO tropospheric NO_2 column measurements compared favourably to NO_2 columns derived from coinciding in situ aircraft measurements on vertical spirals. Good correlation with no significant bias (r2=0.67, slope = 0.99 ± 0.17, n = 12) was found for the ensemble of comparisons when the aircraft could spiral sufficiently low to sample most of the NO_2 column. During the DANDELIONS campaign in May-June 2005, and September 2006 at Cabauw, the Netherlands, OMI data was validated versus various ground based measurements. NO_2 from OMI showed good agreement with two independent ground based techniques that provide tropospheric and total-column NO_2 measurements.

The effect of the different data sets on the success of data assimilation and forecast is studied separately in Chapter 4. In this chapter also the improvements of SmogProg with respect to PROZON will be discussed. In Chapter 5 the present means of communication of the results are described and an outlook is given for the operational phase of SmogProg. Chapter **Error! Reference source not found.** relates the costs of the development and maintenance of the SmogProg system to expected benefits from the expected effect of better communication on public health. Finally an outreach and conclusions are presented in Chapters 7 and 8.

2 User Requirements for SmogProg

The goal of the SmogProg project is to provide an operational ozone forecast model that can replace the currently used statistical forecast model. Especially, improvements with respect to the statistical model are expected in the following areas:

i. The forecast of the beginning and ending of a smog episode
ii. The geographical details
iii. Information to non experts and the general public

This chapter describes the requirements laid down by the RIVM "smogteam": a group that monitors the air pollution every day and provides information about the air quality conditions to the local authorities and the general public.
The requirements result from legislation, practical use and future changes. The legislation is laid down in the *Smogregeling 2001*[2] and in the implementation of the European Directives[3,4,5] on communication and action plans during smog episodes into Dutch legislation. Additional requirements are based on experience with the current forecast model[6].

2.1 Requirements resulting from legislation

Air quality is classified using four indicators. These indicators are ozone, particulate matter, sulphur dioxide and nitrogen dioxide. The severity of smog in the Netherlands is divided into three classes (see Table 1). The boundaries of the classes are set by the limit values, target values, information threshold and, the alert threshold stated in the European Directives with exception of the alert threshold for PM_{10}. There is no alert threshold defined in the European Directives for PM_{10}. However, to create a uniform warning policy for air pollution, an alert threshold of 200 µg/m^3 for PM_{10} is set. For ozone, the information threshold is set to 180 µg/m^3, the transition from little to moderate smog, the alert threshold to 240 µg/m^3 as laid down in the third Daughter Directive[7] .

Table 1: Classes of smog in the Netherlands, concentrations in µg/m^3

	no or little smog	**moderate smog**	**severe smog**
Ozone (hourly average)	< 180	180 - 240	> 240
PM_{10} (daily average)	< 50	50 - 200	> 200
Sulphur dioxide (hourly average)	< 350	350 - 500	> 500 *
Nitrogen dioxide (hourly average)	< 200	200 - 400	> 400 *

* Severe smog occurs only when levels are exceeded for three consecutive hours

[2] Government Gazette (Staatsblad) 11, June 2001 no. 109 / page 16
[3] 1996/62/EC Official Journal of the European Union, 27 September 1996, no. L 296/ page 55
[4] 1999/30/EC Official Journal of the European Union, 22 April 1999, no. L 163/ page 41
[5] 2002/03/EC Official Journal of the European Union, 9 March 2002, no. L 67/ page 14
[6] Noordijk H (2003), Prozon en Propart; stat. modellen voor smogprognose, RIVM Rapport 725301012
[7] 2002/3/EC Official Journal of the European Union, 9 March 2002, no. L 67/ page 14].

The Daughter Directive also describes the details that have to be supplied as soon as possible to the public when the information or alert threshold is exceeded or is predicted to exceed. Annex II of the Directive states to forecast the next details for the following afternoon and/or day(s):

- the geographical area where the information and/or alert threshold are expected to be exceeded;
- the expected change in pollution.

The *Smogregeling 2001* is the implementation of the European Directives regarding public information into Dutch law. The *Smogregeling* provides additional requirements about the number of updates per day and the communication channels that should be used. This leads together with the *Smogregeling 2001* to the following essential requirements for an ozone forecasting system:

- The model is updated twice a day. Preferably at 8.30 and 15.30 local time or at another time once in the morning and once in the afternoon. The model has to supply at least information about the following afternoon and the next day.
- The model will provide information about the daily maximum one hour mean ozone concentration.
- The model will provide information about the development of the concentration levels, the area in which the information threshold or alert threshold are exceeded and, the duration of the exceedances.
- The model output has a resolution of 10 by 10 kilometres and is representative for heights of 3.5 m.
- The model will provide information on a European (2000x2000 km) and a national scale
- The model output can also be used for videotext, messages on mobile phones and e-mail
- The model uncertainty has to stay within 50% at ozone concentrations around the information and alert threshold.

The above mentioned requirements are *essential* to perform the legal obligation of the RIVM. Other options for SmogProg were also formulated, and indicated as *desirable*:

- The model can provide information about the maximum daily 8-hour mean ozone concentration
- The model can make a reliable forecast on a regional scale (provinces)
- The members of the RIVM smog team can provide the forecast with comments.
- The boundaries between the classes of smog can be easily changed.
- The model can detect outliers within the dataset during data-assimilation
- The reliability of the model output is indicated
- The results of model tests will be described in a validation report
- The model is managed by the SmogProg consortium during the project. Agreement on the management of the operational model will be made in a future stage
- The model will be provided with a user guide
- In the future SmogProg can be extended to forecast other smog indicators

More detailed information about the SmogProg requirements can be found in Dutch in Appendix B.

3 The SMOGPROG machine

3.1 The LOTOS-EUROS model

3.1.1 Model outline

The LOTOS-EUROS model is an operational 3D chemistry transport model aimed to simulate air pollution in the lower troposphere over Europe.
The master domain of LOTOS-EUROS is bound at 35°N and 70°N and 10°W and 60°E. The projection is normal longitude-latitude and the standard grid resolution is 0.50° x 0.25° (approximately 35×25 km in Europe). In this study we use the zooming possibility of the model and zoom in on the domain from 2W to 14E and 46N to 56N at a resolution of 0.25° x 0.125° (approximately 18 x 14 km^2). Lateral boundary conditions are extracted from a run on the larger domain. Boundary conditions at the top are derived from the Logan climatology (Logan, 1998) for ozone and from functions matching observations for a number of other components including the ozone precursors.
In the vertical the model has four layers up to 3.5 km above sea level: a fixed surface layer of 25 meter and three dynamic layers. The lowest dynamic layer is the mixing layer, followed by two equally thick reservoir layers up to the model top. The height of the mixing layer is part of the meteorological input data. The height of the reservoir layers has a minimum of 50m. In some cases when the mixing layer extends near or above 3.5 km the top of the model exceeds the 3.5 km according to the abovementioned description. For output purposes the concentrations at measuring height (reference height is usually 3.6 m) are diagnosed by the constant flux approach which relates the dry deposition speed and the concentration of a pollutant.
The chemical mechanism used in the model is a slightly modified version of the CBM-IV scheme (Whitten et al. 1980). For more details on the different processes and approaches in the model we refer to Schaap et al., 2005.
Model results and validation studies of an earlier model version have been reported in literature (Schaap et al., 2008). Furthermore, the model has participated in model comparisons for ozone and aerosols (van Loon et al., 2007; Vautard et al., 2007). These studies indicate that the performance of LOTOS-EUROS is comparable to that of the other models.
The decision to use the LOTOS-EUROS model for the SMOGPROG machine is furthermore based on the possibility to perform active data assimilation and the expertise on the model, built up at the participating institutes.

3.1.2 Emissions

The major driver of the LOTOS-EUROS system is the anthropogenic emission data of VOC, SO_x, NO_x, NH_3, CO, CH_4 and PM. In this study we used the emissions from the TNO emission database (Visschedijk and Denier van der Gon et al., 2005) updated for a higher resolution of 0.125° x 0.0625° for the year 2003. The annual emission totals are converted to hourly emission estimates using time factors for the emissions strength variation over the months, days of the week and the hours of the day.

3.1.3 Meteorological data

The model is driven by 3-hourly meteorological data provided by European Centre for Medium-range Weather Forecast (ECMWF). These include 3D fields for wind direction, wind speed, temperature, humidity and density, substantiated by 2-d gridded fields of mixing layer height, precipitation rates, cloud cover and several boundary layer and surface variables. The vertical velocity field is calculated using the horizontal wind fields and the mass conservation law of incompressible fluids. Further, the water vapour concentration is calculated using the Claussius-Clapeyron relation. Rain is neglected when the 3-hour accumulated amount is less then 0.3 mm. Linear interpolation is used to derive the meteorological fields at the interval times between the update times.

3.2 Data-assimilation

3.2.1 Principles of data assimilation

Observations of ozone (or any other component) consist of data that are irregularly distributed in space and time. Data assimilation allows the calculation of continuous fields in space and time from observations that are irregularly distributed. It consists of making a best estimate of the state of the atmosphere on the basis of observations and a model prediction of the atmospheric state both of which have associated errors. Data assimilation basically defines a new atmospheric state by making a weighted average of the observed and modelled state in an intelligent and statistically sound way. Hence, if a model value is more uncertain than an observed value, more weight will be put on the observation, and the assimilated value will tend to get closer to the observed value and vice versa (see Figure 3).

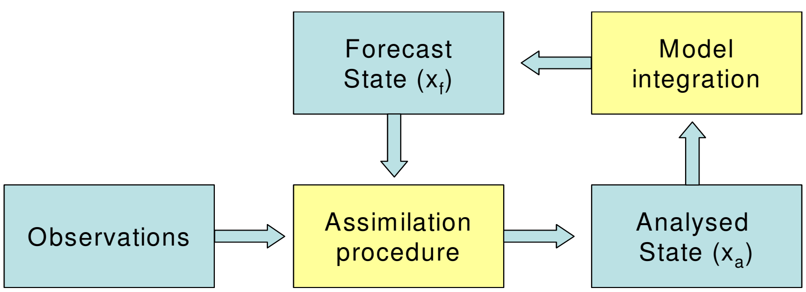


Figure 3: Schematic representation of the data assimilation procedure.

3.2.2 Ground based and satellite observations

In our LOTOS-EUROS runs, the following data were assimilated:

a) *Ozone ground measurements of 22 rural stations of the Dutch LML measurement network.* Rural stations were selected because they are not, like (sub-) urban stations influenced by local emissions and titration processes. The 22 stations were selected in such a way that they cover the Netherlands and are more or less equally spread. The LML data are generated by RIVM and are available to the project on an hourly basis, in real time and at no cost.

b) *Ozone ground measurements of 39 stations of the German and Belgian network.*
German and Belgium stations are assimilated because changes in ozone level in Germany and Belgium can influence air quality in The Netherlands if the wind direction is (south)-east. This wind direction is prevalent in most ozone episodes. These measurements are retrieved from AIRBASE, the database of the European Environmental Agency (EEA). At this moment, they are available at EEA on a daily basis. In the near future, these data will become available in real time.

c) *OMI NO_2 columns*
The Ozone Monitoring Instrument (OMI) monitors worldwide NO_2 concentrations. It achieves coverage of the entire Earths atmosphere in a single day with ground pixel sizes of ~24 x 13 km^2 at nadir. This small pixel size approaches the size of large cities, enabling the identification of individual major sources. The daily global coverage allows for a day-to-day observation of emission sources, as well as the detection and tracking of transport phenomena.
Retrievals of tropospheric NO_2 are available in near-real time mode: within 3h of the actual OMI measurement at 13:40 UTC.
Tropospheric NO_2 columns as retrieved in the DOMINO project have been validated versus independent measurements during various campaigns. During the INTEX-B campaign in March 2006 over the southern United States, Mexico, and the Gulf of Mexico, DOMINO tropospheric NO_2 column measurements compared favourably to NO_2 columns derived from coinciding in situ aircraft measurements on vertical spirals. Good correlation with no significant bias (r2=0.67, slope = 0.99 ± 0.17, n = 12) was found for the ensemble of comparisons when the aircraft could spiral sufficiently low to sample most of the NO_2 column. During the DANDELIONS campaign in May-June 2005, and September 2006 at Cabauw, the Netherlands, OMI data was validated versus various ground based measurements. NO_2 from OMI showed good agreement with two independent ground based techniques that provide tropospheric and total-column NO_2 measurements.

3.2.3 Ensemble Kalman filter

In this study we used an ensemble Kalman filter to assimilate the ozone observations within LOTOS-EUROS. The Ensemble Kalman Filter (EnKF; Evensen 1997) is a data-assimilation technique often applied in geophysical applications. The central part of an EnKF is an ensemble of model states. The ensemble members are created by adding random noise to certain variables in the model, e.g. the emissions. The spread between the ensemble members is then supposed to represent the uncertainty in the model. In particular, each ensemble member is supposed to be one realization out of the stochastic distribution of the unknown true state. Statistical properties are drawn from the ensemble: for example the ensemble mean is often used to represent the most likely value of the state.
Within data assimilation techniques the uncertainties involved with the modelled and observed values determine the weights that are put on the measured and calculated values. With a Kalman filter there is no need to specify the model uncertainties as they are determined by the range of modelled states of the ensemble members. Hence, the specification of the noise on input parameters influences the weights and therewith the results of the procedure. In this work we used an ensemble of 15 members, which was created by adding random noise to the emissions of NO_x and VOC, dry deposition and top boundary conditions of ozone.

3.2.4 Algorithm for the Ensemble Kalman Filter

The first step in order to build the Ensemble Kalman Filter around LOTOS-EUROS is to embed the model and the available measurements in a stochastic environment:

$$x^{k+1} = f^k(x^k, w^k)$$
$$y^k = H^k x^k + v^k,$$

where the superscripts (k) denote the time-steps. The model state vector is denoted by x and the measurements by y. The function f denotes the non-linear model operator which, acts on the state vector and on a white noise vector w with Gaussian distribution and diagonal covariance matrix Q. The measurement vector y is assumed be a linear combination of elements of the state vector and a random, uncorrelated Gaussian error v with (diagonal) covariance matrix R. The matrix H maps the model results on the measurement locations. The basic idea behind the ensemble filter is to express the probability function of the state in an ensemble of possible states $\{\xi_1,..., \xi_N\}$, and to approximate statistical moments with sample statistics:

$$\hat{x} \approx \frac{1}{N}\sum_{j=1}^{N}\xi_j$$
$$P \approx \frac{1}{N-1}\sum_{j=1}^{N}\left(\xi_j - \hat{x}\right)\left(\xi_j - \hat{x}\right)^T$$

where the pair ($\hat{x}$,P) (expectation and covariance matrix) describe the probability of the state vector x completely if x has a Gaussian distribution. Since we are dealing with strongly non-linear models, it cannot be expected that x really has a Gaussian distribution. We assume however that the distribution is at least close to Gaussian so that the bulk of the statistical properties is captured by the pair ($\hat{x}$,P). The filter algorithm consists of three stages:

<u>initialisation</u>:
each ensemble member is set to the initial state:

$$\xi_j = x^0$$

<u>forecast</u>:
each ensemble member is propagated in time by the model, where the noise input w^k is drawn from a random generator with covariance Q;

$$\xi_j^f = f(\xi_j, w^k)$$

<u>analysis</u>:
given an (arbitrary) gain matrix K, each ensemble member is updated according to:

$$\xi_j^a = \xi_j^f + K(y + v - H^T\xi_j^f)$$

where v_j represents a measurement error, drawn from a random generator with zero mean and covariance R. The gain matrix K is by default the optimal gain matrix from the original Kalman Filter:

$$K = PH^T (HPH^T + R)^{-1}$$

The forecast step is the most expensive part of the algorithm, since for each ensemble member the model has to be evaluated one time. Typical ensemble sizes range from 10-100. If the number of measurements is limited (in order of hundreds), the total computation time involved with the ensemble filter is proportional with the ensemble size.

3.2.5 Noise

In the model implementation used in this study, the noise parameters are part of the model state. Hence they are estimated by the filter as well. The noise parameters w_i that are for instance used as emission correction factors for the actual emission field E_j are estimated by the filter as

$$E_i \leftarrow E_j (1 + w_i).$$

To create an ensemble of simulations for the Kalman filter we have selected four parameters to apply noise to (Table 2). We have selected the most important/uncertain processes influencing ozone, i.e. the emissions of NO_x and VOC, dry deposition velocity and top boundary conditions of ozone. All noise factors were applied with a mean of 1 and a standard deviation of 0.5. Hence, the vast majority for the factors are within the range of 0.5 to 1.5. All noise factors were set after an assimilation step and held constant until the next assimilation step.

Table 2. Noise parameters used to create an ensemble of model simulations

Emissions of NO_x
Emissions of VOC
Top boundary conditions of O_3
Dry deposition velocity of O_3

3.2.6 Initialisation of the ensemble

The filter is initialised with a zero spread in the ensemble, that is, all ensemble members are set to the same model state. The variance in the ensemble is built up during a 5-days spin up using random forcing for emissions and/or other uncertain parameters.

For chemical data assimilation this initialisation procedure is sufficient to reach the maximum spread in the ensemble. Variations in concentrations vanish to zero in a few days anyway because of the damping effect of chemistry and deposition. Chemical data assimilation is much different from meteorological assimilation in this way, since for the latter small changes in initial state tend to grow continuously in time.

3.2.7 Spatially limiting the influence of measurements

Spurious correlations may arise between elements of the state vector, because the sample size is finite. Furthermore, undesired correlations arise due to the choice of the noise processes. Noise processes that act on emission fields of various emitted compounds can cause "instantaneous" correlations throughout the domain. For example the ozone concentration at hour t somewhere in The Netherlands becomes correlated with the ozone concentration in, say, the south of France, because noise was added to the NO_x emission field at hour t-1. Although this is exactly what should happen when defining noise in this way, such correlations are not realistic and should be somehow ignored by the filter. The noise processes is chosen this way because it is infeasible to subdivide the emission fields into a number of sub-domains on each of which a different noise parameter is acting. That would increase the dimension of the noise vector dramatically and hence the necessary ensemble size to capture the statistical properties.

One way to ignore unrealistic correlations over large distances is the use of localisation [Houtekamer and Mitchel, 2001]. The gain matrix is only unequal to zero around the locations of observations. Such a gain matrix K may be formed using a covariance matrix which is an elementwise product of the original sample covariance and a correlation function with local support. For a single scalar measurement, the resulting gain matrix is given by (omitting the subscripts):

$$K = I(\rho)PH/(H^T PH + R)$$

where $I(\rho)$ is a diagonal matrix; the diagonal elements are filled with a prescribed correlation between the corresponding grid cell and the grid cell of the measurement. Different choices for the values of ρ_i are possible. In this study we take

$$\rho_i = \exp[-½\,(r_i/L)^2] \quad \text{for } r_i \le 3.5\,L$$

and zero otherwise. r_i denotes the distance from the grid cell considered to the location of the analysed measurement and L is a length scale parameter. See the next section for a choice of L. For satellite retrievals, horizontal correlations are removed completely, such that a single retrieval leads to an update over a single cell in the domain only.

3.3 Kalman filter settings

The following parameters need to be specified in the Ensemble Kalman filter:

1) number of ensemble modes necessary for convergence
2) correlation length and number/choice of stations to be assimilated
3) correlation time τ: the effect of ozone at time t on ozone at time t+τ
4) spread of the noise σ, which is a measure for the model uncertainty
5) measurement uncertainty. For the Dutch LML stations and the German and Belgian stations used here, this uncertainty is estimated at 15 %.

Test runs were performed with LOTOS-EUROS for July 16-21, 2006, in order to evaluate the best choices for these parameters. In this period, high ozone values occurred over The Netherlands. Assimilated stations were either the ground based Dutch LML stations (GBNL) or the Dutch and Belgian/German stations near The Netherlands (GBEU).

All runs were evaluated with respect to the settings of the reference run:

- σ = 0.5
- number of modes = 15
- τ = 0.25
- $\rho = L$ = length scale parameter = 25 km
- assimilation of GBEU stations.

To evaluate the runs, the following plots are shown:

1) Time series of hourly ozone concentrations for the LML stations Vredepeel, Cabauw and Balk, together with the average over all rural LML stations.

2) Scatter plots of ozone maxima for the different sensitivity runs. Modelled ozone peaks (daily maxima) are plotted against measured ozone peaks.

3) Noise factors on NO_x and VOC emissions, top O_3 boundary conditions and dry deposition for the LML station Cabauw

3.3.1 Number of modes

First the number of ensemble modes necessary for convergence was determined. For this purpose, 3 assimilation runs were done for the period July 16-21, 2006, with 12, 15 and 25 modes.

The other settings were kept the same as the reference run.

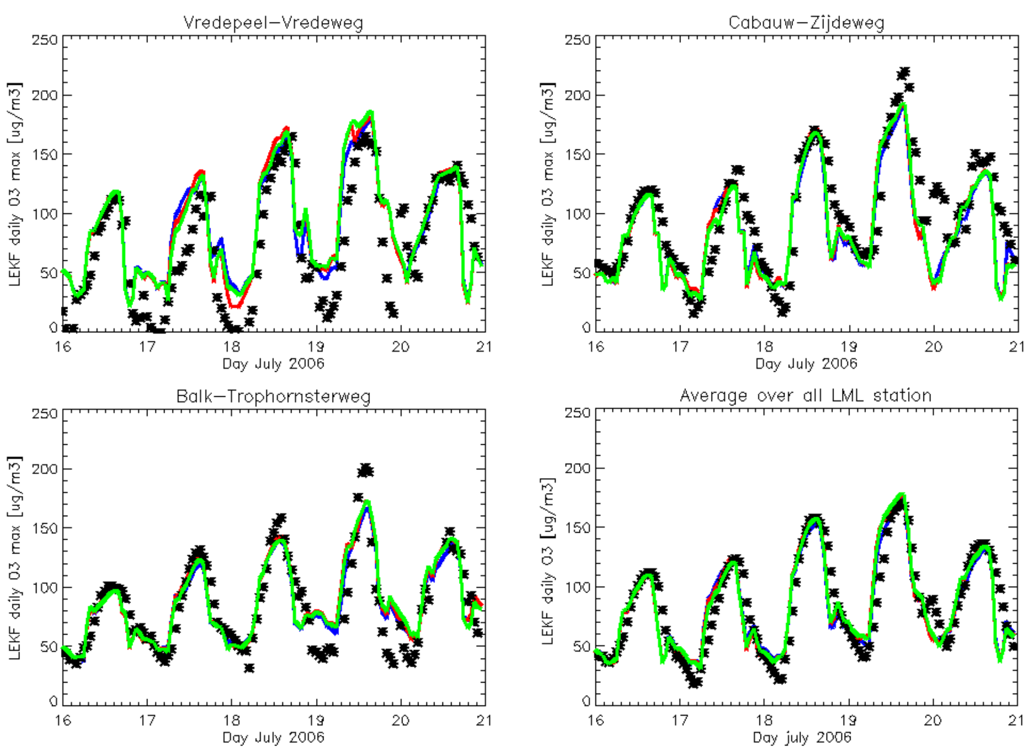
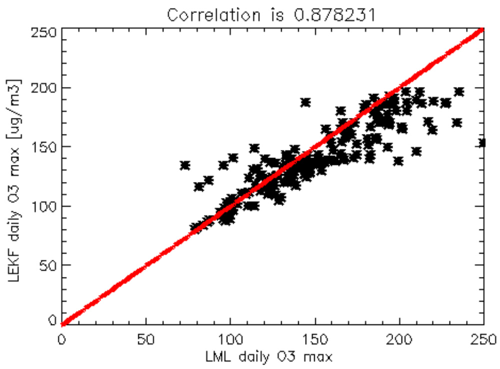
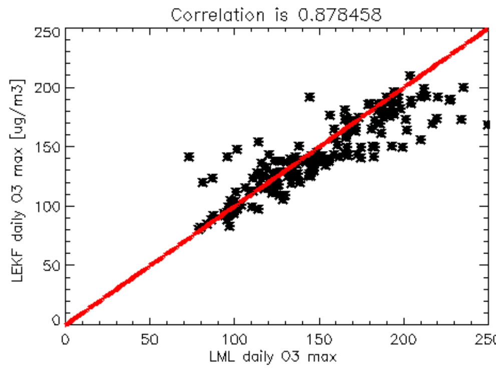


Figure 4. Time series of hourly ozone concentrations for 3 stations and averaged over all rural LML stations.

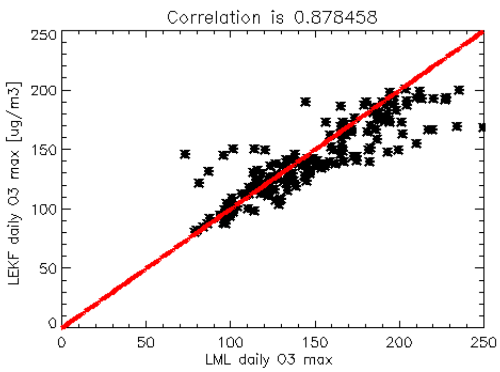
*: LML; lines: LOTOS-EUROS (blue: 12 modes, red: 15 modes, green: 25 modes).



12 modes

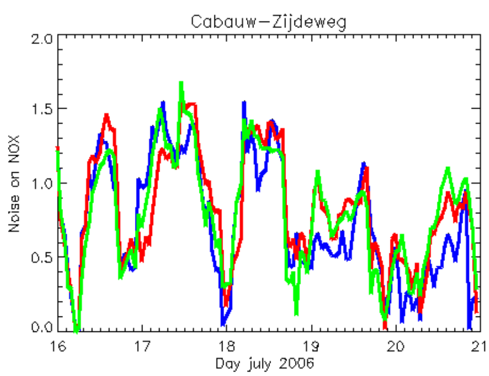


15 modes

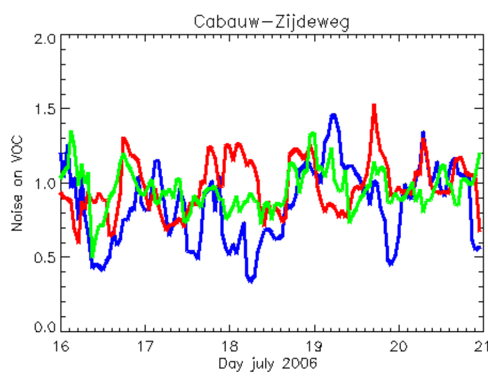


25 modes

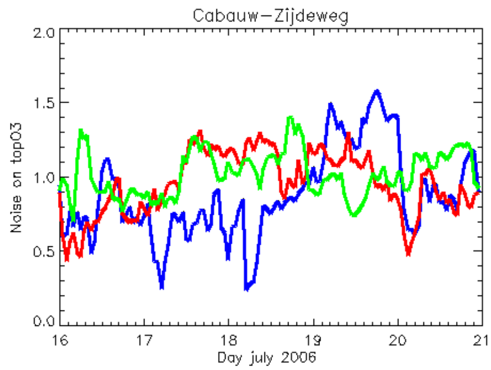
Figure 5. Scatter plots of modelled ozone maxima against LML measurements for different number of modes



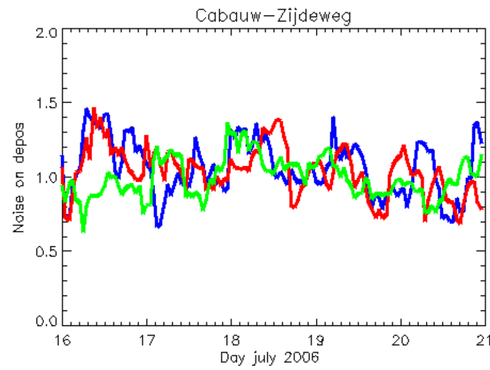
NO_x emission noise factors



VOC emission noise factors



top O_3 b.c. noise factors



deposition noise factors

Figure 6. Noise factors on NO_x and VOC emissions, top O_3 boundary conditions and deposition at the station Cabauw-Zijdeweg. *: LML; lines: LOTOS-EUROS (blue line : 12 modes, red line: 15 modes, green line : 25 modes).

Both time series and scatter plots show that there is not a big difference between the runs with 12, 15 and 25 ensemble modes. Of course the run with 25 ensemble modes has the best correlation coefficients: the higher the number of ensemble modes, the better the ozone modelling. On the other hand, when using 25 ensemble modes, 1 assimilation run takes too much time. In the noise plots, the red line (15 modes) is most in agreement with the green line (25 modes), indicating convergence of the model. The blue line (12 modes) gives highest noise factors and differs significantly from the green line (25 modes) indicating that the model is not sufficiently converged. As 15 modes still gives good results within a reasonable calculation time, it was decided to use 15 ensemble modes for our evaluation runs.

3.3.2 Correlation length / number and choice of stations

The correlation length ρ (or *L*) and number and choice of stations were estimated from 6 runs for July 2006:

1) RUN1: a run with a correlation length of 25 km, assimilating
   a) 22 assimilated GBNL stations, no monitoring stations
   b) 61 assimilated GBEU stations, no monitoring stations
2) RUN2: a run with a correlation length of 50 km, assimilating
   a) 11 assimilated GBNL stations and monitoring 11 stations
   b) 50 assimilated GBEU stations and monitoring 11 stations
3) RUN3: a run with a correlation length of 100 km, assimilating
   a) 5 assimilated GBNL stations and monitoring 16 stations
   b) 44 assimilated GBEU stations and monitoring 16 stations.

See Appendix A for a listing of the different stations and the division between assimilation and monitoring stations.

First, the correlation coefficient between ozone values at different stations as a function of the distance (“radius”) between the stations was calculated according to the A_CORRELATE function: this function computes the autocorrelation $P_x(L)$ of a sample population *X* as a function of the lag *L*:

$$P_x(L) = P_x(-L) = \frac{\sum_{k=0}^{N-L-1} (x_k - \bar{x})(x_{k+L} - \bar{x})}{\sum_{k=0}^{N-1} (x_k - \bar{x})^2}$$

where $\bar{x}$ is the mean of the sample population $x = (x_0, x_1, x_2, \ldots, x_{N-1})$.

.

Results are shown below:

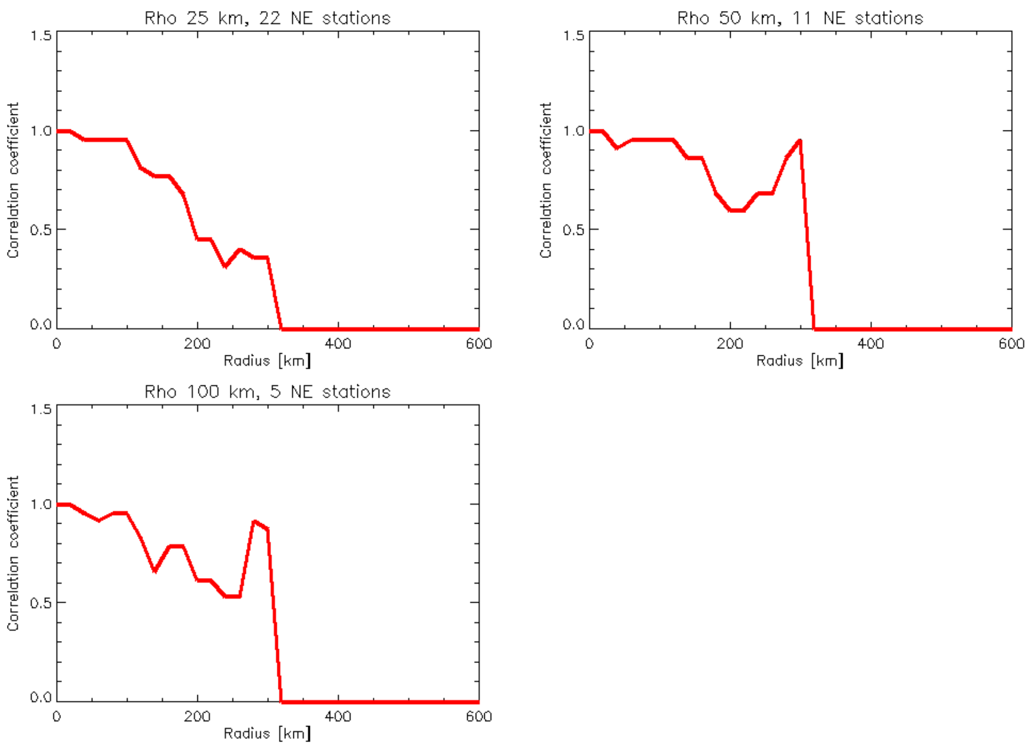


Figure 7.Correlation coefficient between ozone at different LML stations as a function of the distance between two stations ("radius") for different values of ρ and different number of stations. Assimilation of GBNL (Dutch) stations.

Figure 7 shows the correlation coefficient between 2 stations for runs assimilating only data of the Dutch LML stations: above they are referred to as RUN1a, RUN2a and RUN3a.The pictures show that the correlation coefficient decreases significantly around 100 km, so the estimated correlation radius is 100km.

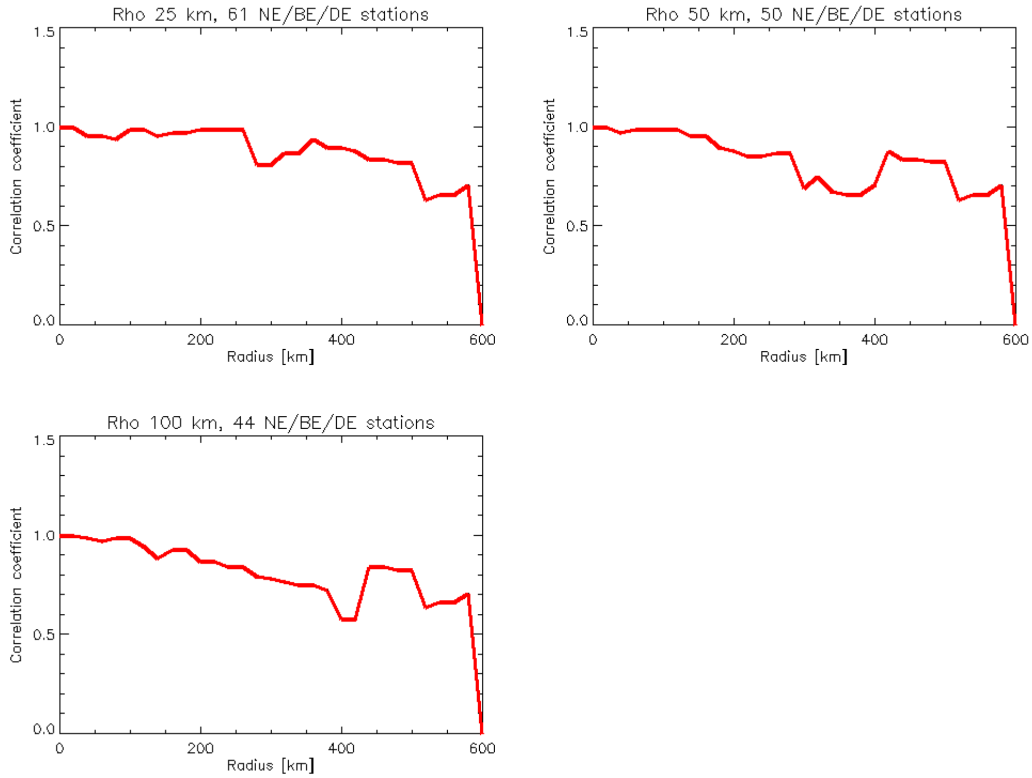
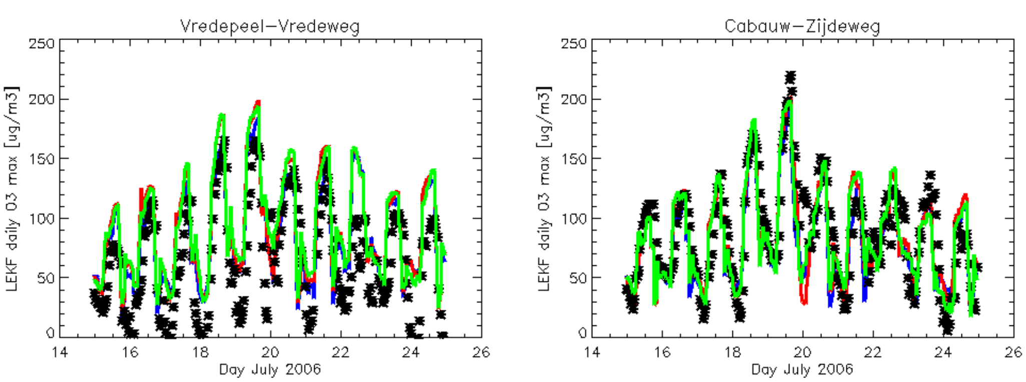


Figure 8. Correlation coefficient between ozone at different LML stations as a function of the distance between two stations ("radius") for different values of ρ and different number of stations. .Assimilation of GBEU (Dutch, Belgian and German) stations.

Figure 8 shows the correlation coefficient between 2 stations for runs assimilating data from German, Belgian and Dutch stations: above they are referred to as RUN1b, RUN2b and RUN3b. The pictures show that the correlation coefficient decreases significantly around 200-300 km, so the estimated correlation radius is 200-300 km.

Next, time series, scatter plots and noise factors on stations are evaluated for RUN1-3. First for the Ground based Dutch, Belgian and German stations (GBEU):



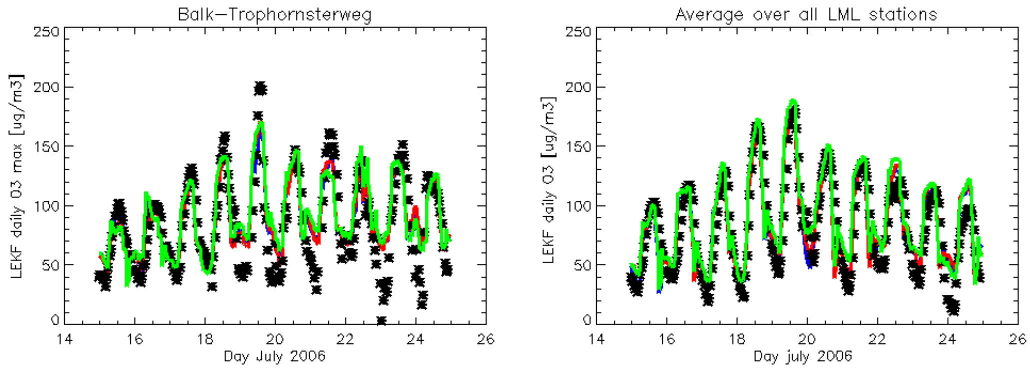
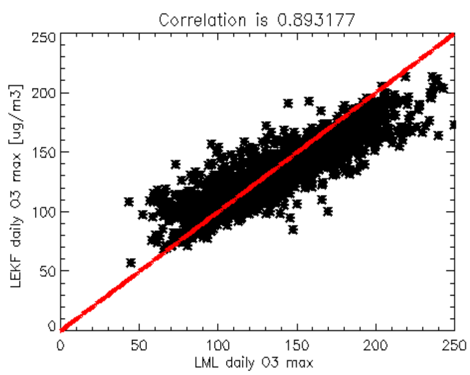
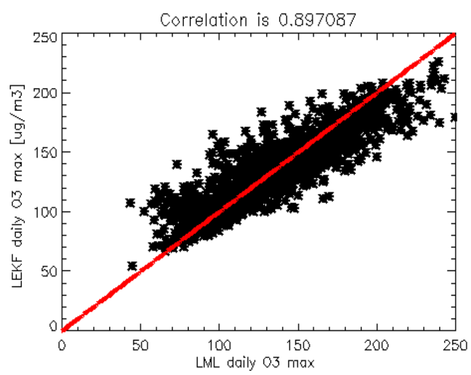


Figure 9. Time series of hourly ozone concentrations at 3 rural LML stations and average over all LML rural stations. Assimilation of GBEU (Dutch, Belgian and German) stations.
*: LML, lines: LOTOS-EUROS (▬: ρ = 25 km, ▬: ρ = 50 km, ▬: ρ = 100 km).

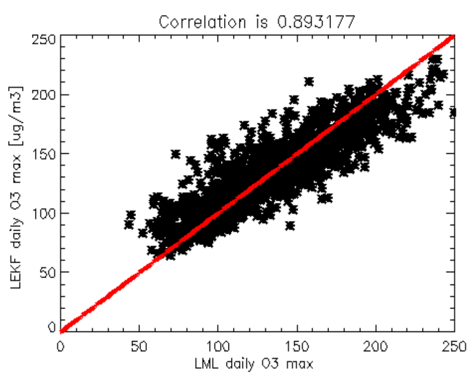
Time-series show that RUN1-RUN3 do not differ a lot in their ozone peak estimate.



RUN1: ρ=25 km; 61 GBEU stations assimilated

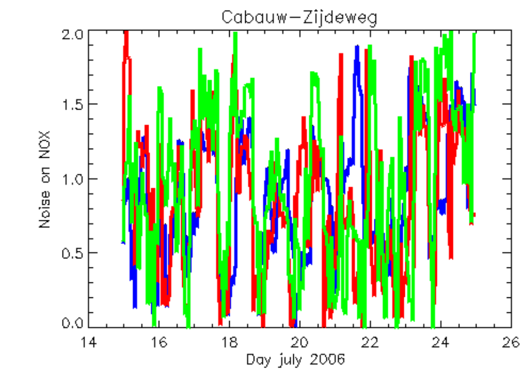


RUN2: ρ=50 km; 50 GBEU stations assimilated

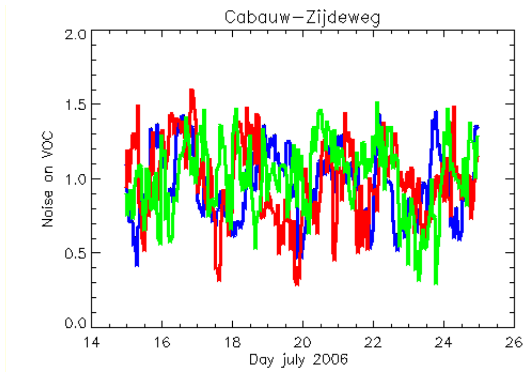


RUN3: ρ=100 km; 44 GBEU stations assimilated

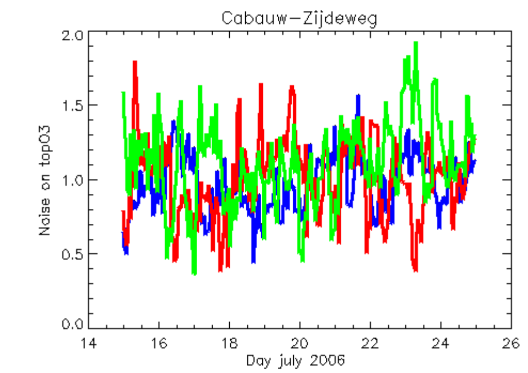
Figure 10. Scatter plots of modelled ozone maxima against LML measurements for different values of the correlation length ρ. Assimilation of GBEU (Dutch, Belgian and German) stations.



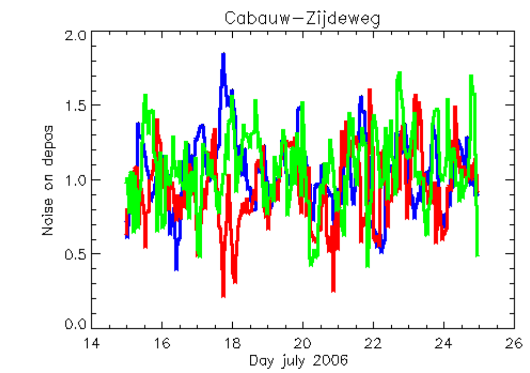
NO_x-emission noise factors



VOC-emission noise factors



top O_3 b.c. noise factors

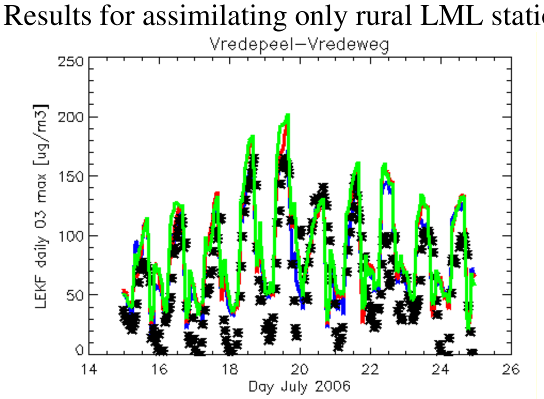


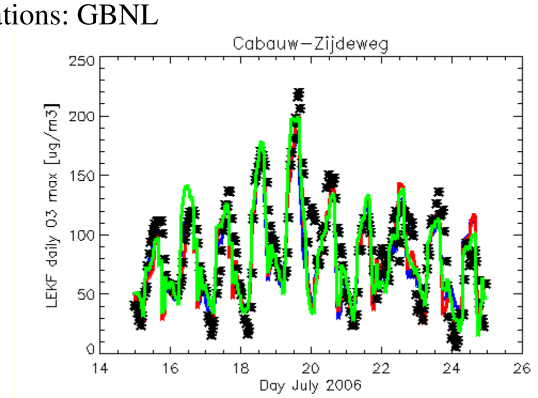
deposition noise factors

Figure 11. Noise factors at Cabauw. Assimilation of GBEU (Dutch, Belgian and German) stations. LOTOS-EUROS (blue: ρ = 25 km, red: ρ = 50 km, green: ρ = 100 km).

All runs show highest noise factors for NO_x emissions.
In general, RUN3 (correlation length of 100 km) gives rather high noise factors (reasonable noise factors lie between 0.5 and 1.5, as noise factors were applied with a mean of 1 and a standard deviation of 0.5). RUN1 (correlation length 25 km) gives lowest and most reasonable values. Even though the estimated correlation radius based on the GBEU run was approximately 200-300 km, evaluation of RUN1b-RUN3b does not show a large difference in (average) time series and scatter plots/correlation coefficients between the runs. Noise series on the other hand show lowest values for the smallest correlation radius of 25 km. Since we don't want the model needing very high noise factors, it was decided to use RUN1's value of ρ = 25 km.

Results for assimilating only rural LML stations: GBNL





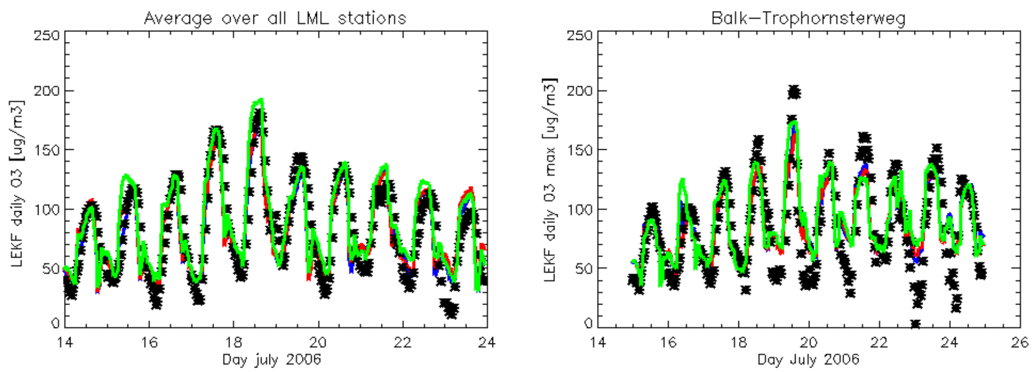
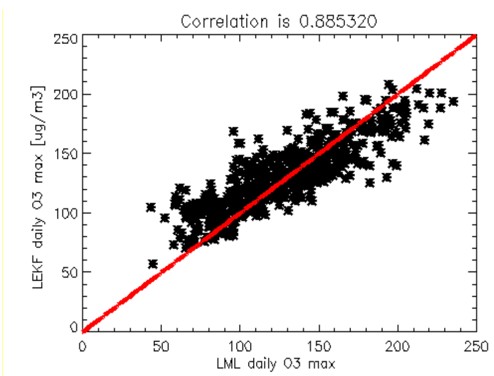
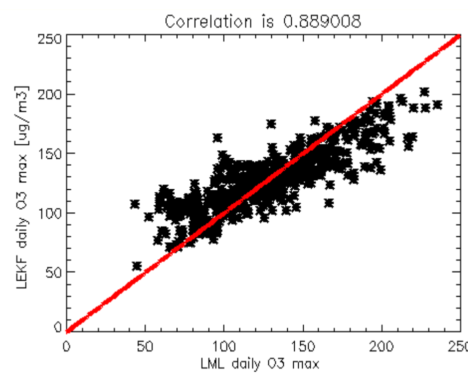


Figure 12. Time series of hourly ozone concentrations at 3 rural LML stations and average over all LML rural stations. Assimilation of GBNL (Dutch) stations.
*: LML; lines: LOTOS-EUROS (blue: ρ = 25 km, red: ρ = 50 km, green: ρ = 100 km).

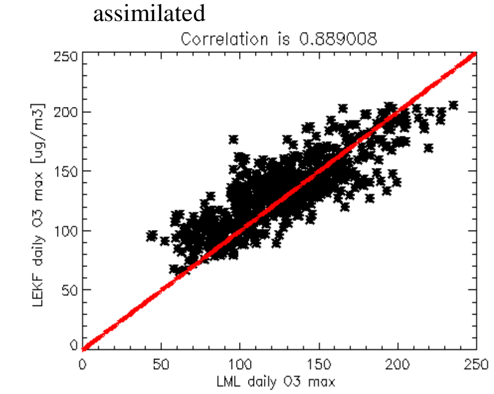
Time-series show that RUN1-RUN3 do not differ a lot in ozone peak estimate.



RUN1: ρ=50 km; 11 GBNL stations assimilated



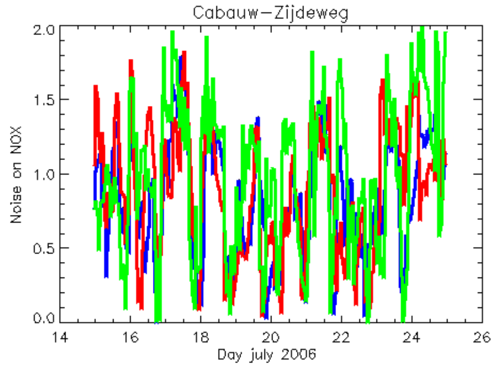
RUN2: ρ=25 km; 22 GBNL stations assimilated



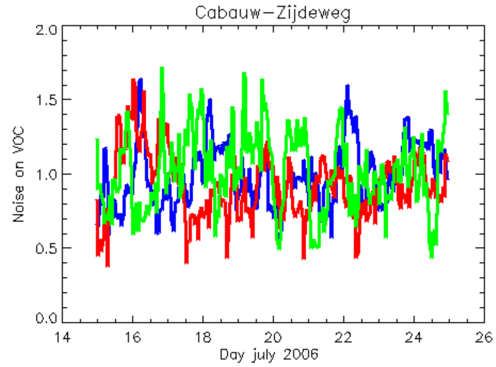
RUN3: ρ=100 km; 5 GBNL stations assimilated

Figure 13. Scatter plots of modelled ozone maxima against LML measurements for different values of the correlation length ρ. Assimilation of GBNL (Dutch) stations.

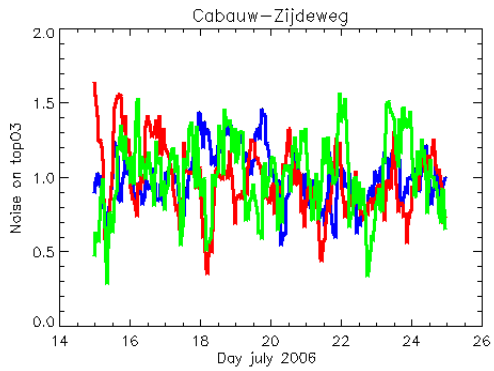
The runs assimilating GBEU stations give little better results than the GBNL runs.



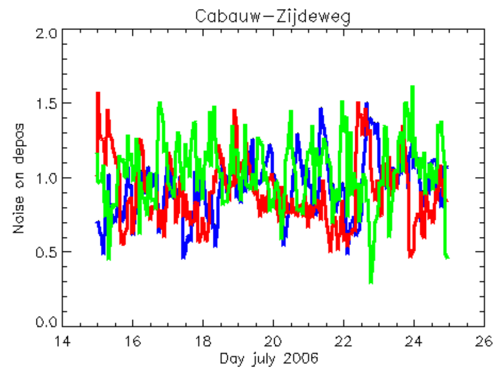
NO_x noise factors



VOC noise factors



top O_3 b.c. noise factors



deposition noise factors

Figure 14. Noise factors at Cabauw. Assimilation of GBNL (Dutch) stations.
LOTOS-EUROS (blue: ρ = 25 km, red: ρ = 50 km, green: ρ = 100 km).

The same conclusions as for the GBEU-runs, hold for the runs assimilation GBNL stations.

3.3.3 Correlation time τ

Next to the correlation in space, the correlation in time between ozone measurements is determined for RUN2b. This coefficient was calculated according to the A_CORRELATE function that computes the autocorrelation or autocovariance as above.

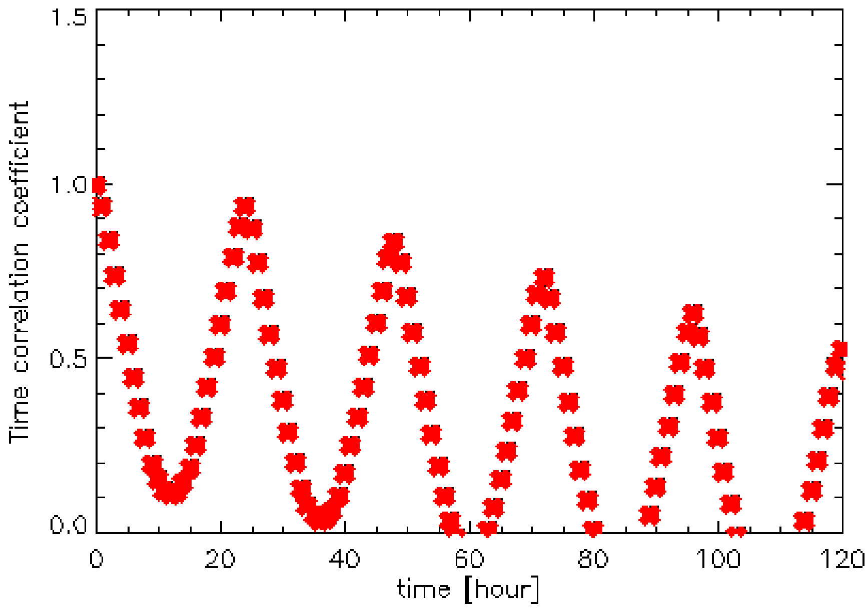


Figure 15. Time correlation for Cabauw as a function of time lag in hour

Figure 15 shows that the correlation decreases periodically: after 6 hour it is approximately equal to 0.5. This result is in agreement with the daily ozone cycle. As a function of days the amplitude of the sinus dampens as expected. It was decided to set a time correlation of 6 hour (0.25 day) in order to take into account the daily ozone cycle.

3.3.4 Standard deviation of noise

In order to determine the best choice of standard deviation on the noise (σ), applied on the NO_x, VOC emissions, top O_3 boundary conditions and dry deposition velocity, 5 assimilation runs with different values of σ (0.2, 0.4, 0.5, 0.8 and 0.9) were done for the period July 16-21, 2006. Other settings were the same as in the reference run.

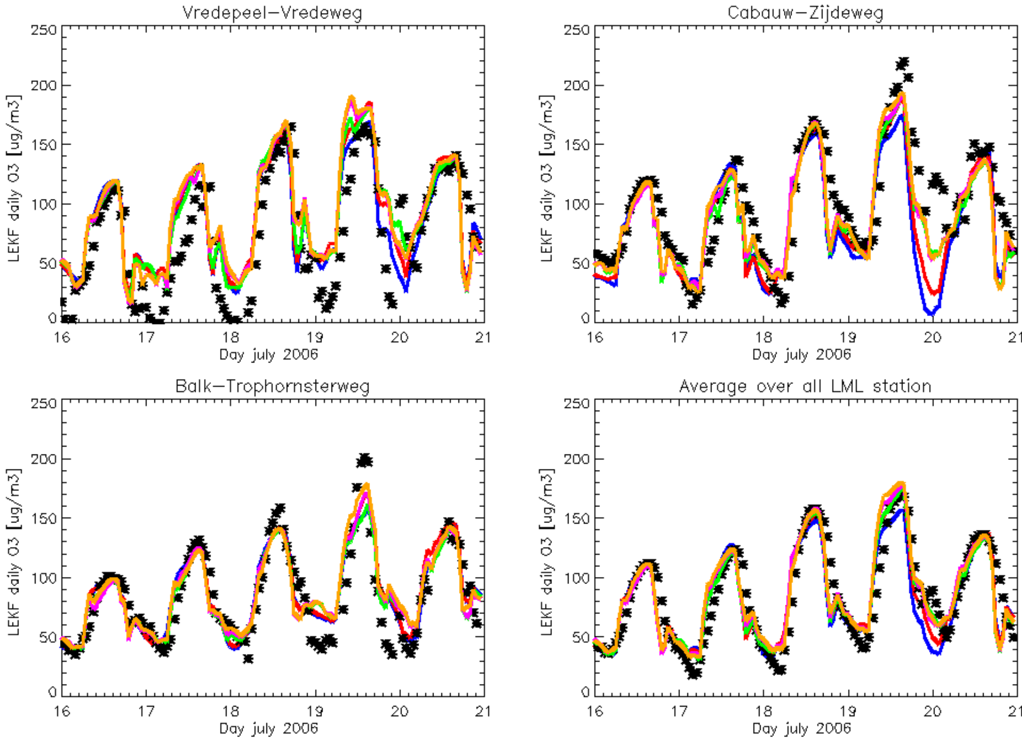
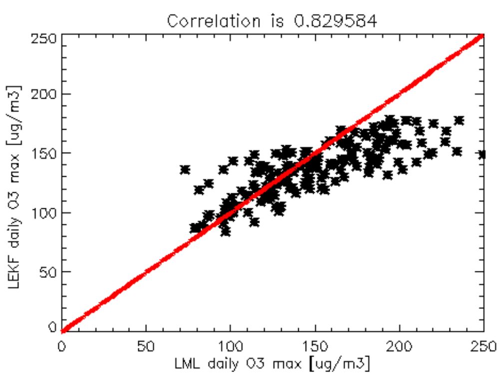
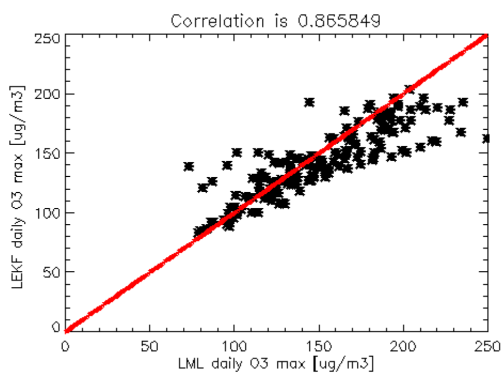


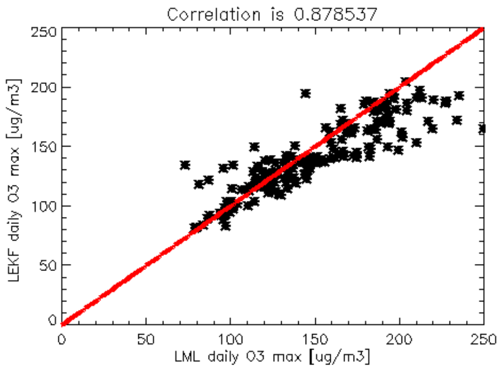
Figure 16. Time series of hourly ozone concentrations at 3 rural LML stations and average over all LML rural stations.
*: LML; lines: LOTOS-EUROS (blue line: σ = 0.2, red line: σ = 0.4, green line: σ = 0.5, magenta line: σ = 0.8, orange line: σ = 0.9).



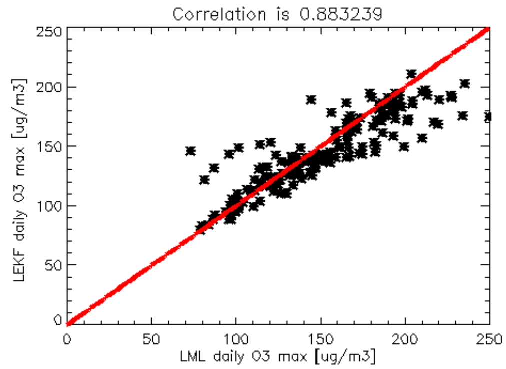
σ = 0.2



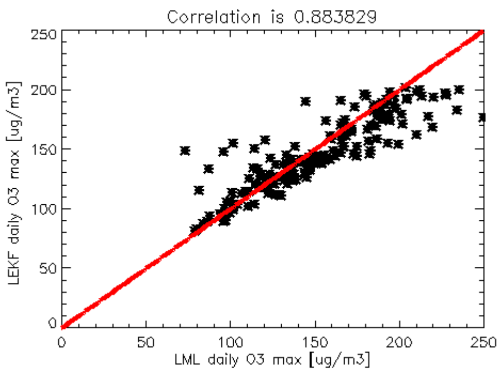
σ = 0.4



σ = 0.5

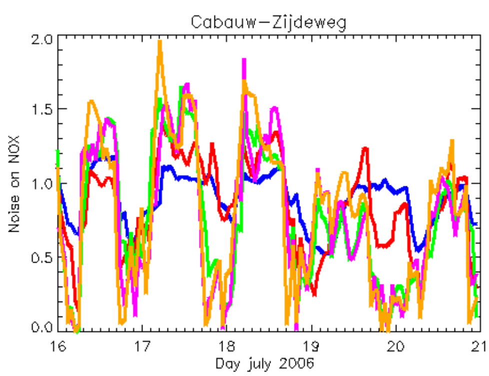


σ = 0.6

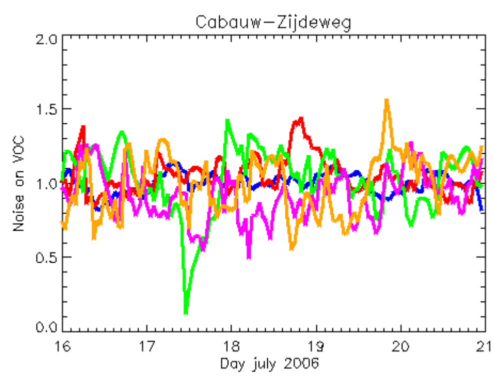


σ = 0.9

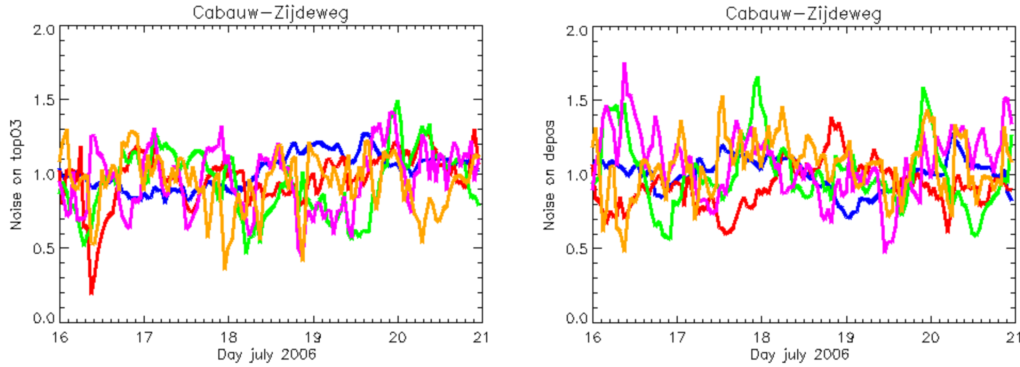
Figure 17. Scatter plots of modelled ozone maxima against LML measurements for different values of σ.



NO_x noise factors



VOC noise factors



top O_3 b.c. noise factors　　　　deposition noise factors

Figure 18. Noise factors on NO_x and VOC emissions, top O_3 boundary conditions and deposition at the station Cabauw-Zijdeweg. LOTOS-EUROS (blue: σ = 0.2, red: σ = 0.4, green: σ = 0.5, magenta: σ = 0.8, orange: σ = 0.9).

Both plots on stations and scatter plots of ozone maxima indicate that a choice of σ = 0.9 gives the best results: the scatter plots show highest correlation and the time series for stations shows the best approximation of ozone peaks. On the other hand the noise factors of the run with σ = 0.9 has values far above and below the indicated 0.5 and 1.5. For this reason it is not suitable.

The optimum at the unrealistic σ value of 0.9 is caused by the fact that not only errors in the NO_x and VOC emissions, top O_3 boundary conditions and deposition, but also model errors are taken into account in the noise factors. We can also see this back in the noise factors: they are for example unrealistically high for NO_x , in order to be able to obtain the (high) ozone peaks. In order to obtain noise factors between 0.5 and 1.5, we chose σ to be 0.5 as setting for our evaluation of the analysis / forecast runs.

3.3.5 Conclusion of Kalman filter settings

From the runs evaluated it was found that the optimal settings for the Kalman filter are as follows:

1) 15 ensemble modes
2) a standard deviation on the noise σ = 0.5
3) a correlation time τ = 0.25 day, following the daily ozone cycle
4) a correlation length of 25 km and 61 Dutch LML, German and Belgian stations.

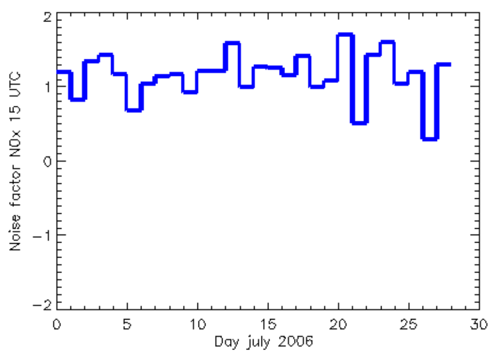
3.4 From assimilation to forecast

In all forecast runs, LOTOS-EUROS is driven by meteorological forecasts. Analysed measurements are used up to the moment that the forecast starts; after that the forecast is run for 3 days. This forecast should be better than the LOTOS-EUROS forecast without data assimilation, as the run starts from an assimilated concentration field. The choice of the noise factor we use in the forecast may be critical to the success of the forecast. Therefore we evaluated a number of choices for this *inheritance* process of the noise:

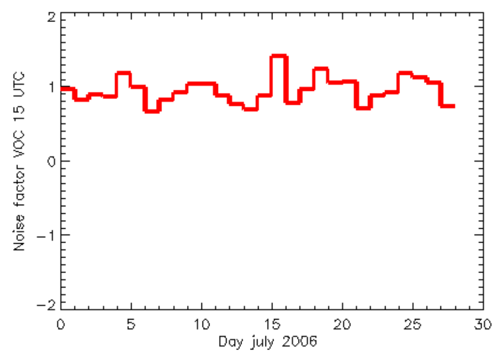
a) Noise from 15 h. UTC of the last assimilation day (i.e. around the ozone peak)
b) Noise of the last 24 hours of the last assimilation day (daily cycle)
c) Average noise between 6 and 18 h. UTC of the last assimilation day
d) Default noise of 1.

In the following figures, the noise factors for these different inheritance settings are shown for the measuring station Cabauw.

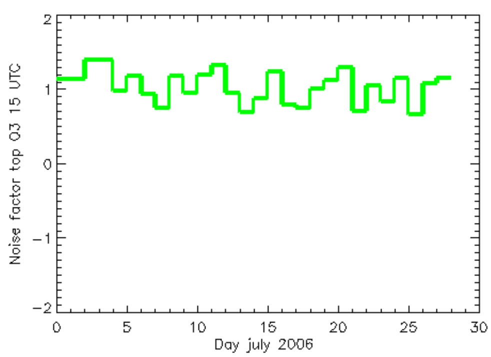
3.4.1 Noise at 15 h. UTC



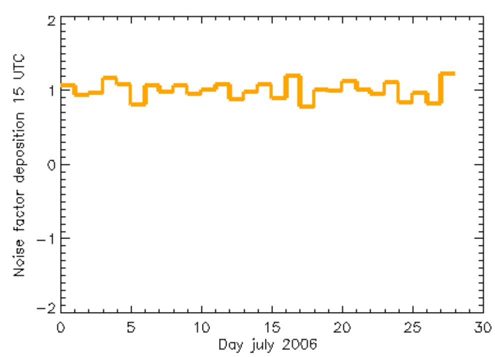
NO_x noise factors



VOC noise factors



top O_3 b.c. noise factors

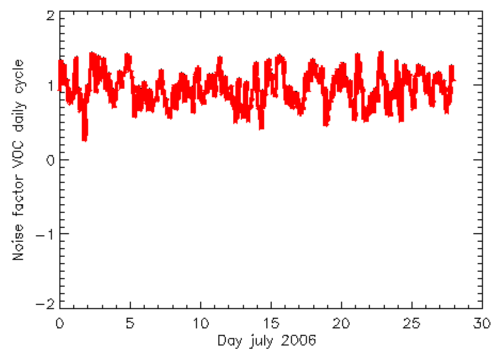


deposition noise factors

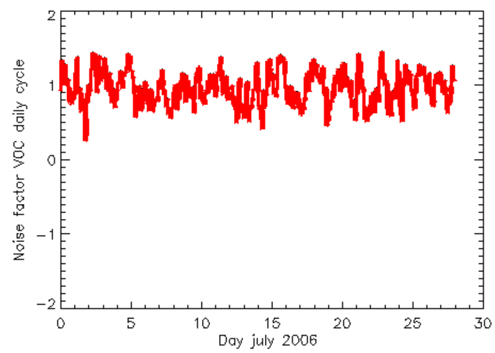
Figure 19. Noise factors from 15h. UTC

Compared to the other settings, this setting gives lowest day to day changes, making it most suitable as a noise inheritance setting.

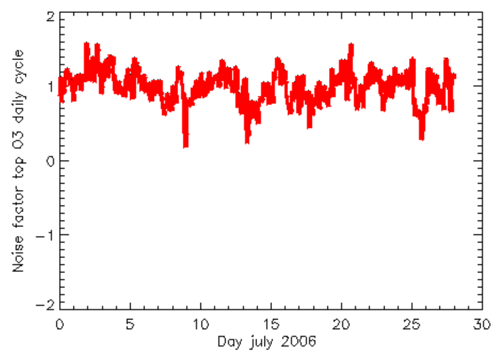
3.4.2 Daily cycle



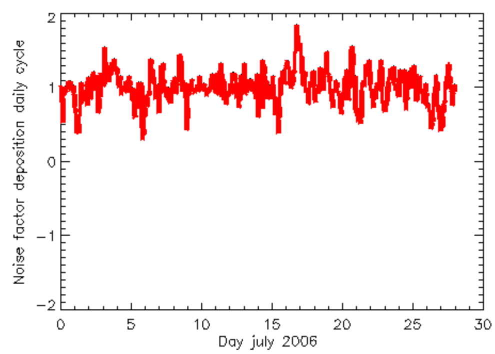
NO_x noise factors



VOC noise factors



top O_3 b.c. noise factors

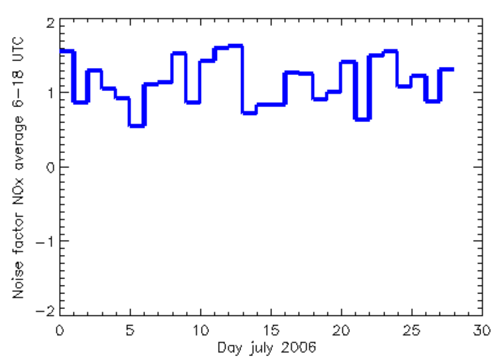


deposition noise factors

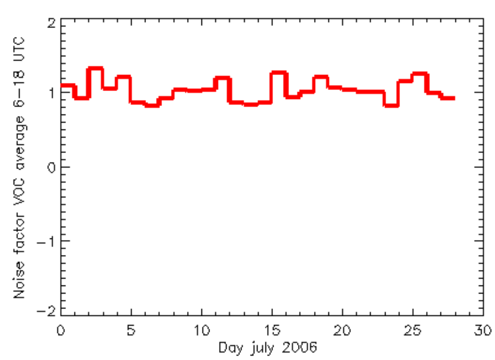
Figure 20. Noise factors, hourly values.

Figure 20 shows a relatively large day to day change in noise pattern, making this inheritance setting not really suitable for a forecast.

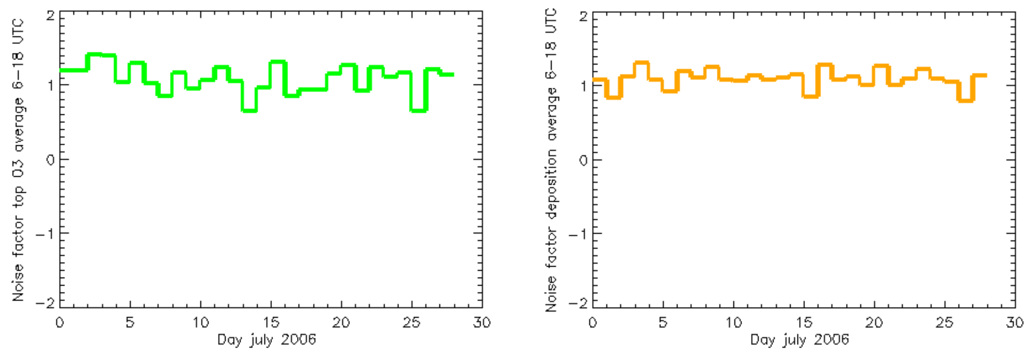
3.4.3 Average noise from 6-18 h. UTC



NO_x noise factors



VOC noise factors



top O_3 b.c. noise factors
deposition noise factors

Figure 21. Noise factors, average between 6 – 18 h. UTC

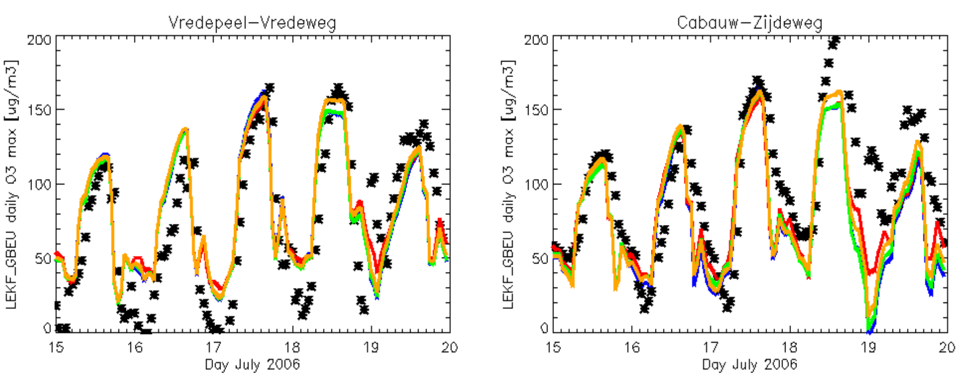
Using this setting, noise factors change a little more than the noise from 15 h. UTC scenario, making it a little less suitable.

3.4.4 Default noise of 1

We see that for each choice of inheritance, noise factors change relatively strong from day to day. This means that there are always errors due to a wrong noise factor in the forecasting days. Therefore a constant noise factor of 1 (i.e. using the default input parameters of LOTOS-EUROS) might be a useful choice.

3.4.5 Test runs for inheritance

In the following figures, we show results of test runs with these four inheritance settings.



1-day forecast

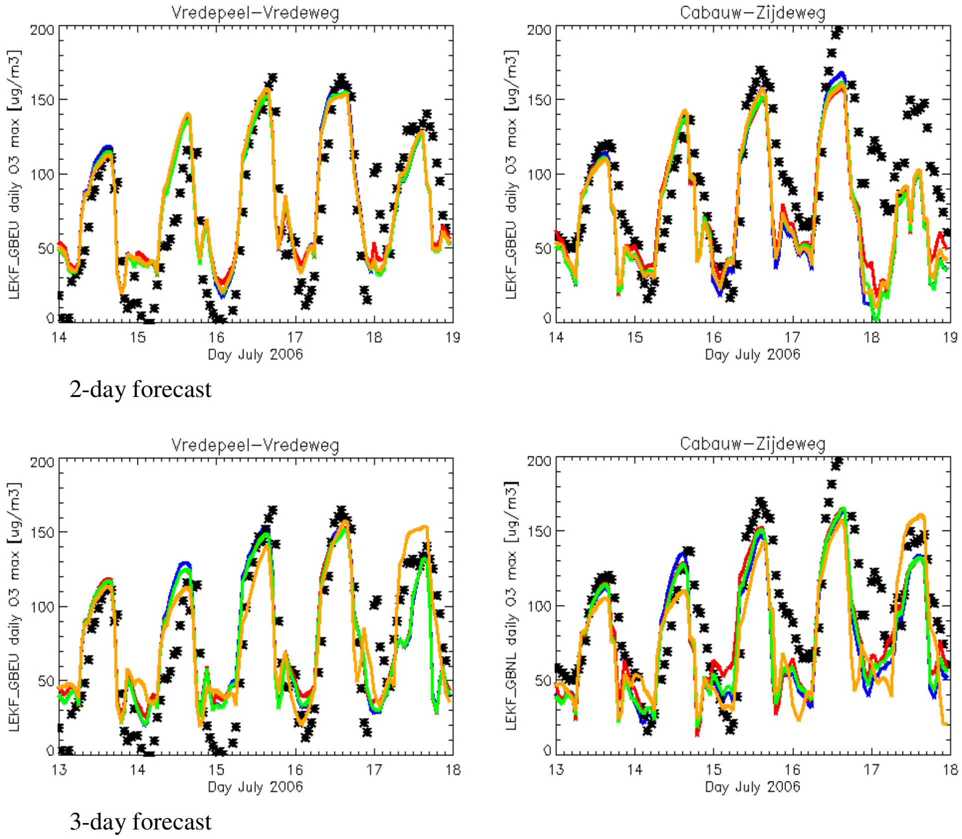


Figure 22. Time series of hourly ozone concentrations at Vredepeel and Cabauw. Assimilation of GBEU (Dutch, Belgian and German) stations. *: LML; lines: LOTOS-EUROS (blue line :noise from 15h. , red line : daily cycle, green line : average noise 6 – 18 h, orange line : noise =1).

These figures suggest that there is not a large difference between the different noise inheritance implementations. Only the default noise differs at some times significantly from the other implementations.

The implementation with the noise from 15 h. UTC gives good results and we took this setting as a default for the runs in chapter 4.

3.5 OMI assimilation

For the assimilation of OMI data it was decided to use the LOTOS-EUROS model extended with an implementation of the ensemble Kalman filter assimilation technique. The advantage of this technique is the possibility to relate differences between observed and modelled concentrations to specific uncertain model processes, and emission strengths in particular.

During the SMOGPROG project a new software module was developed to link the OMI data to the LOTOS-EUROS and Ensemble Kalman Filter model. This module implements the so-called observation operator: it generates a model prediction of the OMI measurement, based on the relevant LOTOS-EUROS model data and information from the measurement such as the geolocation information and averaging kernel. The tropospheric column of NO_2 reported by OMI is compared with the inner product of the LOTOS-EUROS NO_2 vertical profile at the OMI footprint, and the averaging kernel vertical profile. The averaging kernel represents the sensitivity of the OMI instrument to NO_2 at different heights above the surface, which depends strongly on aspects like cloud coverage and surface reflectivity. Figure 23 shows one example of an averaging kernel in near cloud-free conditions.

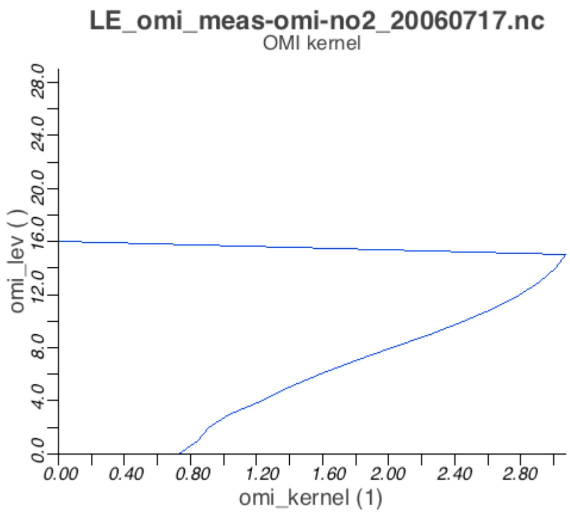


Figure 23. Example of an averaging kernel provided in the OMI data product. The kernel is plotted as a function of the level index (levels provided in the OMI data file). This kernel corresponds to a nearly cloud free situation. The cut-off at level 16 corresponds to the tropopause: the stratospheric NO_2 column has been subtracted and only the tropospheric column is reported. We can see that kernel has a sensitivity near the surface which is typically a factor 3 smaller than the sensitivity to NO_2 at the tropopause. However, in regions like Western Europe the tropospheric NO_2 profile is dominated by NO_2 in the boundary layer, and despite the larger sensitivity the contribution of the free troposphere to the OMI signal is often quite small (order 10%).

The following two images show one day of OMI NO_2 measurements over the European domain. Typically there are about 14000 observations available to the assimilation. The OMI overpass is in the early afternoon, and this is the time when these observations are assimilated.

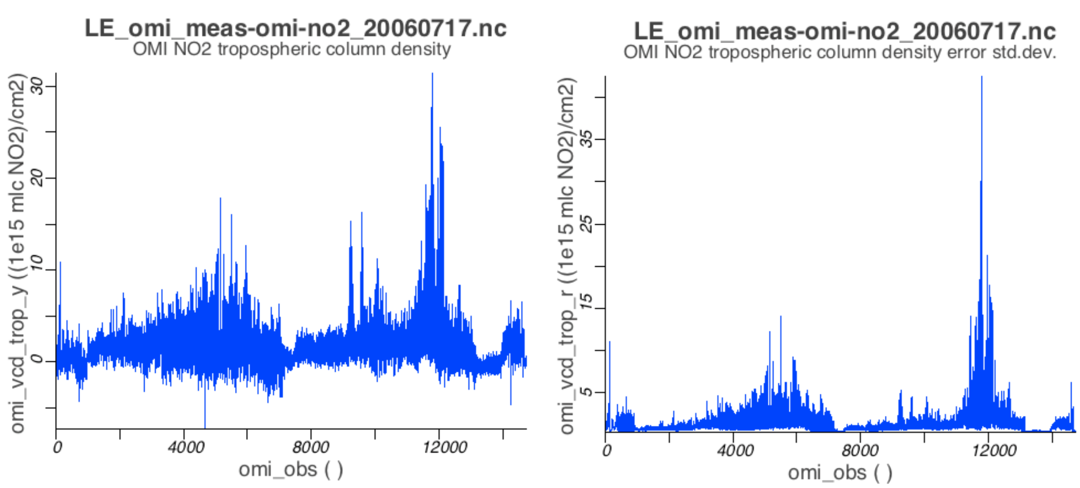


Figure 24. One day of OMI measurements over Europe (July 17, 2006) (left plot) and the corresponding error bars provided in the OMI data product (right plot). The high values, e.g. around observation nr. 12000, are related to areas with elevated NO_x concentrations (Netherlands, Belgium, Ruhr area, England, Po valley etc.).

Negative tropospheric columns occur in the data product, related to the noise in the retrievals and the fact that NO_2 concentrations vary over several orders of magnitude, with values below the detection limit over clean areas, e.g. the ocean. These values are complemented by error bars of the order of 100%. A large portion of the error is proportional to the actual retrieval. This explains the clear correlation between the NO_2 column (left plot) and the corresponding error (right plot).

The next two plots show the result of applying the observation operator. The LOTOS-EUROS (LE) predicted OMI observations show a good overall correspondence with the OMI measurements, demonstrating that the distribution of the major NO_x areas in Europe, as observed by OMI is reproduced by the model. However there are also differences. One major issue is the overall bias that seems to exist between model and observation, OMI being roughly 2 times as large. This is a major problem: the cause of this bias should be identified and remedied before a successful assimilation can be done. This is now the first issue to study in the coming months. It is interesting to note here that the NO_2 concentrations from LOTOS_EUROS at ground level compare well with ground based measurements. When looking at the details in the OMI observation and LOTOS-EUROS model results there are qualitative small-scale differences which may indicate errors in bottom-up estimates of local emissions. This is where we hope OMI will bring important new information.

When comparing the LOTOS-EUROS values after the analysis (right plot) with the LOTOS-EUROS values of the free model run x_B (left plot) we see an overall increase of the values by about 20%. The model is corrected in the right direction, but the adjustment is not very large.

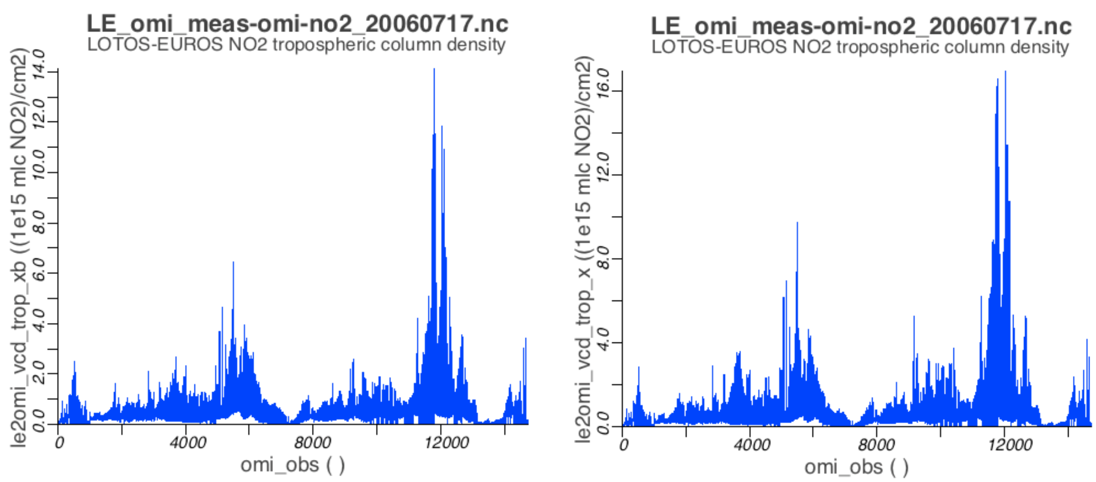


Figure 25. LOTOS-EUROS predicted tropospheric NO_2 columns for the free model run (left) and the model state after assimilation (right).

The next plot shows the OMI observations displayed in a latitude-longitude map on July 17, 2006. This was an exceptionally clear day, with almost no clouds over Western Europe. This plot may be compared with LOTOS-EUROS modelled profiles (same colour scale). Again, the most prominent feature is the much smaller column values seen in the model over land. Over sea things are the opposite: here the model predicts high NO_2 values, especially over the North Sea and the Channel. These signals come from ship emissions. The OMI plot seems to show similar features, but in this case the concentrations are smaller. Apart from the overall scaling factor over land one can observe that the patterns of NO_2 show a good overall correspondence.

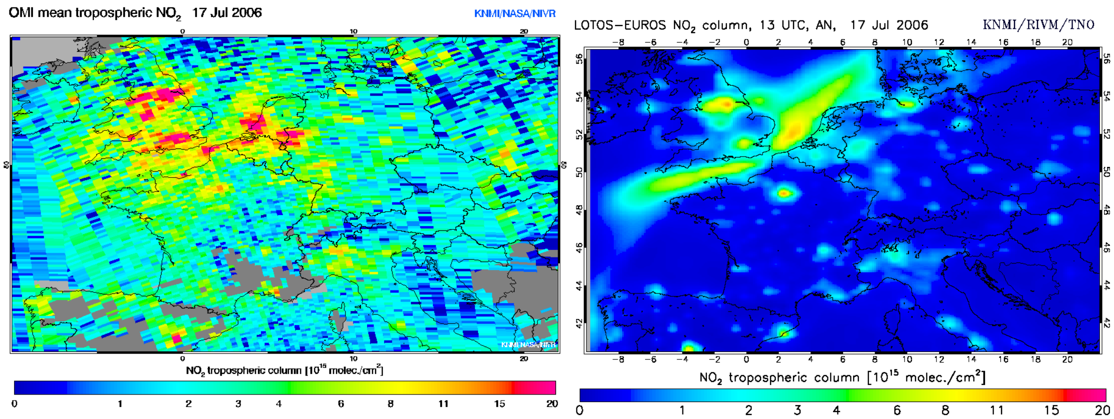
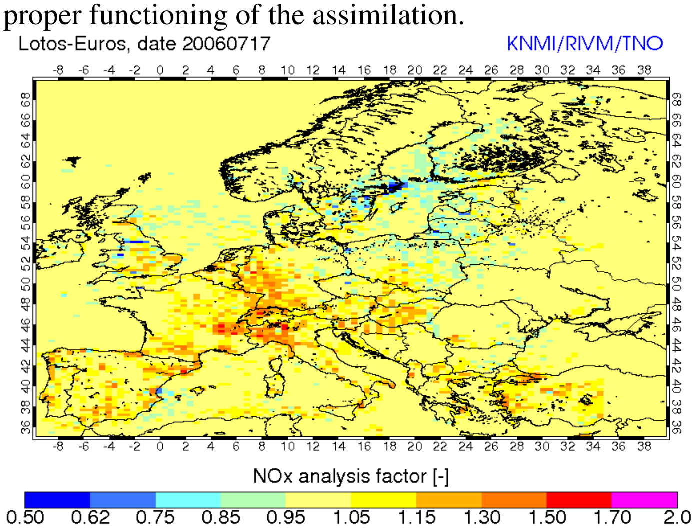


Figure 26. OMI NO2 tropospheric column measurements over Europe on 17 July 2006 (left panel) and LOTOS-EUROS modelled NO2 columns for the same time, 13 UTC. Measurements with a substantial cloud cover are removed from the plot (grey areas).

The settings of the Kalman filter are based on the settings used in the ozone surface observation assimilation described above. There are two notable changes:

- A time correlation of two days has been introduced. In this way the OMI observations - which are available only early afternoon - will impact the emission strengths also during the rest of the day.
- The automatic quality control implemented was causing problems, and has been switched off. There is no clear indication that there exist individual “bad” OMI retrievals which need to be filtered out, and removing the quality control seems justified.

The next three plots shows the adjustments made by the ensemble Kalman filter to the NO_x emissions, the VOC emissions and the ozone boundary condition. From this it is clear that the OMI NO_2 assimilation has a substantial impact on the NO_x emissions. As mentioned before, these adjustments are in the order of 20%. Smaller adjustments are observed for VOC and ozone boundary conditions. The OMI observations lead to an increase in NO_x emissions over western and central Europe. This behaviour is in agreement with expectations and indicates a proper functioning of the assimilation.



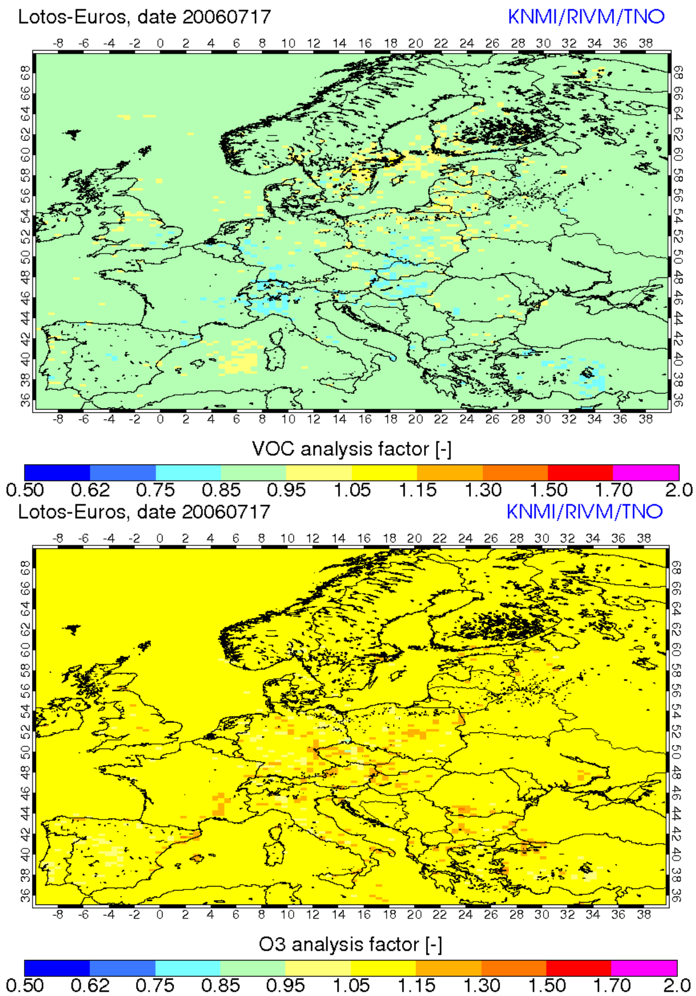


Figure 27. Analysis factors ("noise factors") for NOx emissions (top), VOC emissions (middle) and ozone boundary conditions (bottom), resulting from the assimilation of OMI NO_2 observations. The analysis factor is the ratio between the analysed and *a-priori* values. The uniform background reflects the uncertainty in the *a-priori* and is not related to OMI.

4 Evaluation of a historical smog episode

4.1 Introduction

In the previous chapter, the development of LOTOS-EUROS with data assimilation was described. Based on a test episode of 5 days the optimal settings of the system were derived experimentally. In this chapter, we evaluate the overall performance of PROZON and the LOTOS-EUROS versions without and with data assimilation for the whole month of July 2006. By evaluating a full month, more reliable statistics are obtained than for the five day run. Although a month is still too short for a solid evaluation, the special character of the test month enables the evaluation under the most difficult and most relevant conditions. The models were run in a semi-operational setting to mimic the results which would be used in a daily forecasting practice. In the runs we used noise on NO_x and VOC emissions, top ozone boundary conditions and ozone deposition velocity.

4.2 Description of the runs

We have selected the month of July in 2006 for evaluation, as described in Chapter 1. In this period, several smog episodes with ozone values higher than 180 $\mu g/m^3$ occurred, which made it very suitable for our model evaluation. Models are evaluated in order of increasing number of assimilated data, as listed below:

1) the statistical model PROZON
2) the chemical transport model LOTOS-EUROS
3) the chemical transport model LOTOS-EUROS using Ensemble Kalman filter data assimilation of ozone ground measurements from Dutch LML stations
4) the chemical transport model LOTOS-EUROS using Ensemble Kalman filter data assimilation of ozone ground measurements from Dutch, German and Belgian stations.
5) the chemical transport model LOTOS-EUROS using Ensemble Kalman filter data assimilation of OMI NO_2 column data

PROZON only gives a forecast, starting from observed ozone maxima at the ‘analysis’ day. The results presented here are a reconstruction of the daily forecasts and are based on forecasted temperature, for both the ‘analysis’ day and the forecast days. The reconstructed forecasts are not identical to the operational forecasts due to differences in availability of information for PROZON. These differences are not considered as essential.

All LOTOS-EUROS forecasts are driven by ECMWF forecast meteorology. For the assimilation of ground based ozone measurements, in the Ensemble Kalman filter data are assimilated up to the moment that the forecast starts. After that the forecast is run with fixed parameter settings for emissions, boundary conditions and deposition velocity. In the runs described in this chapter, the parameter values at 15 UTC of the analysis day were taken, since this gave the best forecast performance over the 5 days test period.

Data assimilation of OMI NO_2 column observations is technically possible, but the analysis is not good enough to start forecasts from it. Therefore only analysis results will be shown.

4.3 Results

In this section, results of the runs described in paragraph 7.1 are presented. Model performance is evaluated by showing pictures of:

1) ozone daily maximum values. Results of the individual LML stations Balk (north), Cabauw (middle) and Vredepeel (southeast) are shown to indicate regional differences. The average of 22 LML rural stations was used to indicate the overall performance. Shown are analysis, 1-day, 2-day and 3-day forecast and LML measurements. These pictures are used to evaluate the beginning and the end of the episode and the performance of the ozone maximum modelling. For all graphs the same legend is used (Table 3).
2) maps of daily ozone maximum for an analysis run and a 1-day forecast. The maps indicate the spatial distribution of ozone and ozone gradients. At the right page, we show a map of the Netherlands, at the left page a larger area around the Netherlands.
3) scatter plots of daily maximum ozone values of model against LML measurements. These give an indication of a model under- and overestimation.

Forecasts are started from 30 June at earliest, which implies that for the 2-day forecast data for July 1st and for the 3-day forecast data for July 1st and 2nd are not available. A discussion of the results and conclusions are given in the next section.

Table 3: legend entries for graphs

marker/color	used for
*	LML observation
blue line	model analysis
red line	forecast 1 day
green line	forecast 2 days
orange line	forecast 3 days

PROZON

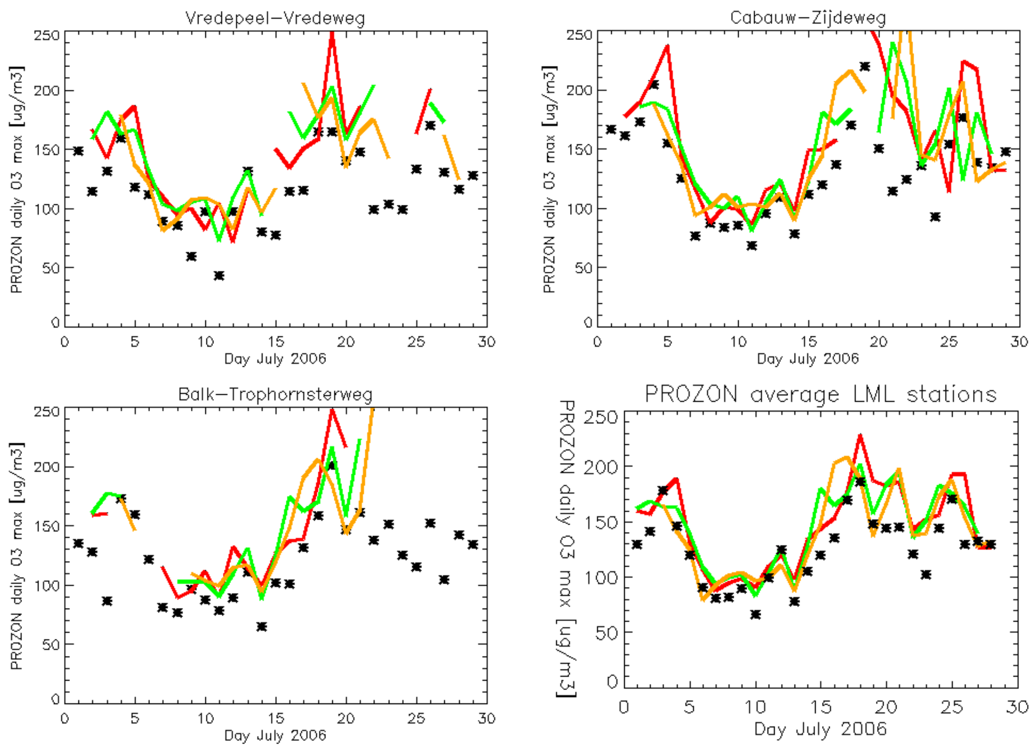


Figure 28. Daily maximum ozone concentrations (µg/m^3) for July 2006 as modelled by PROZON. See Table 3 for legend.

Maps of larger area around The Netherlands not applicable

PROZON

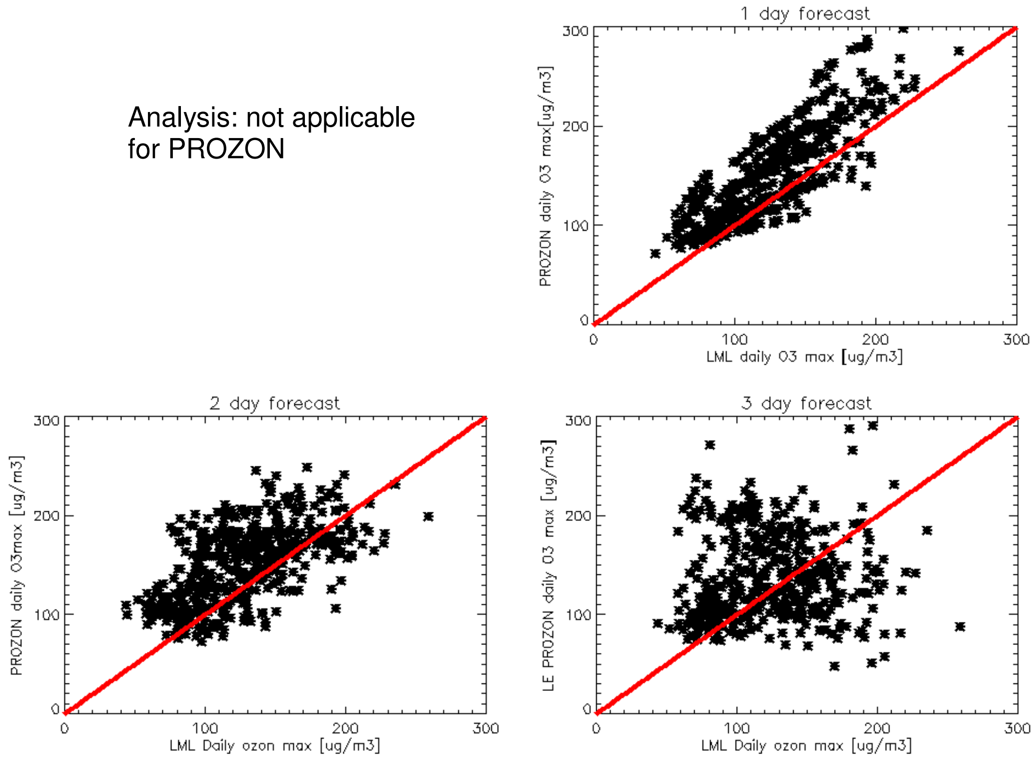


Figure 29. Daily maximum ozone concentrations (μg/m^3) for July 2006; scatter plot PROZON versus LML measurements.

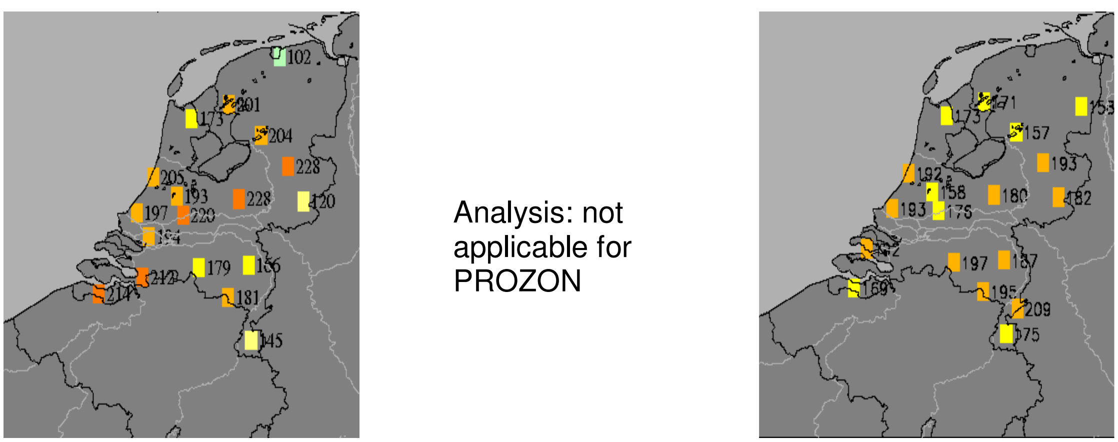
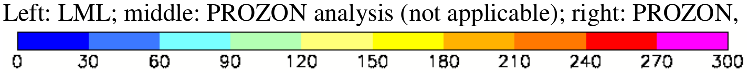


Figure 30. Daily maximum ozone concentrations (μg/m^3) for July 19th, 2006 in The Netherlands, as modelled by PROZON.

Left: LML; middle: PROZON analysis (not applicable); right: PROZON, 1-day forecast



4.3.1 LOTOS-EUROS without data-assimilation

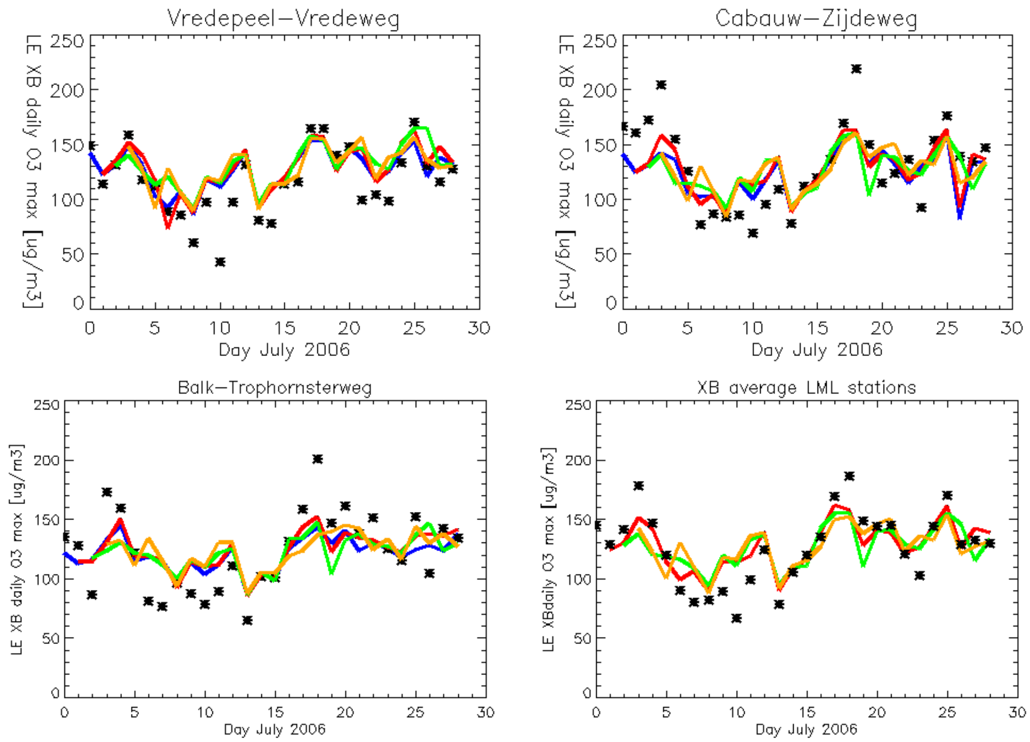


Figure 31. Daily maximum ozone concentrations (µg/m^3) for July 2006 as modelled by LOTOS-EUROS without data-assimilation. See Table 3 for legend.

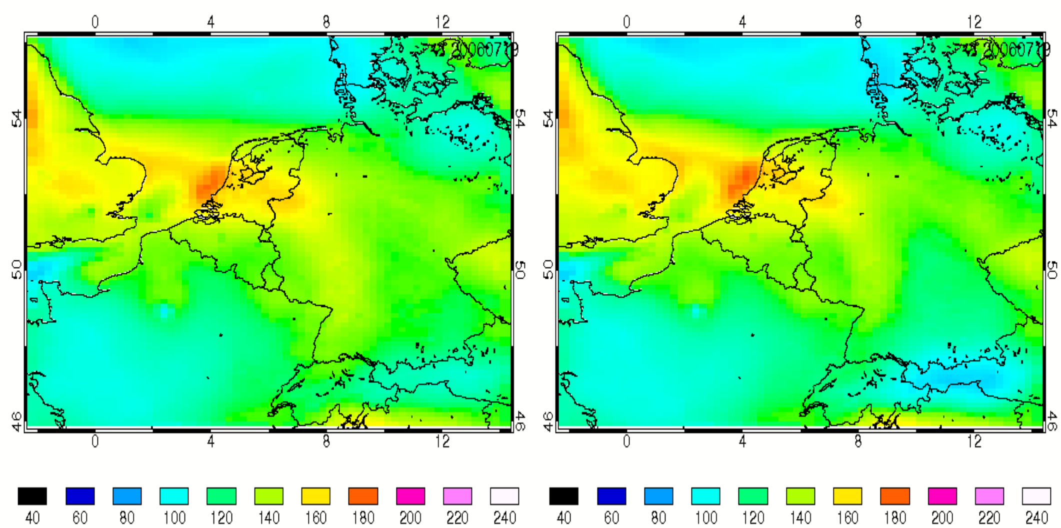


Figure 32: . Daily maximum ozone concentrations (µg/m^3) for July 19th, 2006 as modelled by LOTOS-EUROS without data assimilation. Left: analysis; right: 1-day forecast. In this case the only difference is the analysed or forecast meteo used.

LOTOS-EUROS without data-assimilation

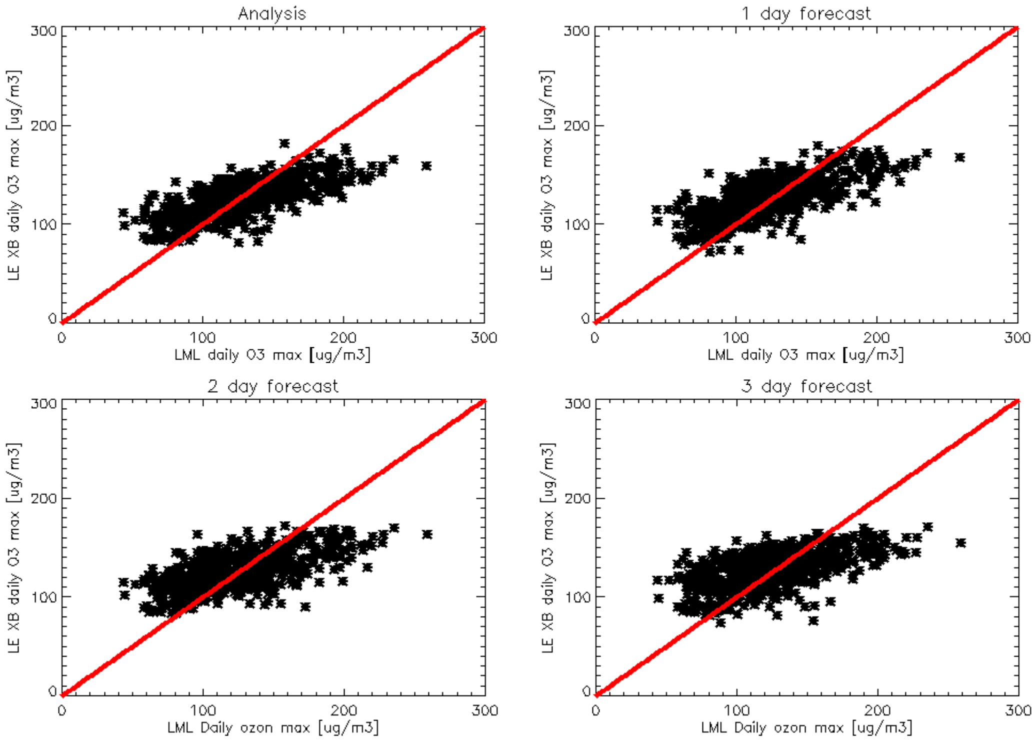


Figure 33. Daily maximum ozone concentrations (μg/m^3) for July 2006; scatter plot LOTOS-EUROS without data-assimilation versus LML measurements.

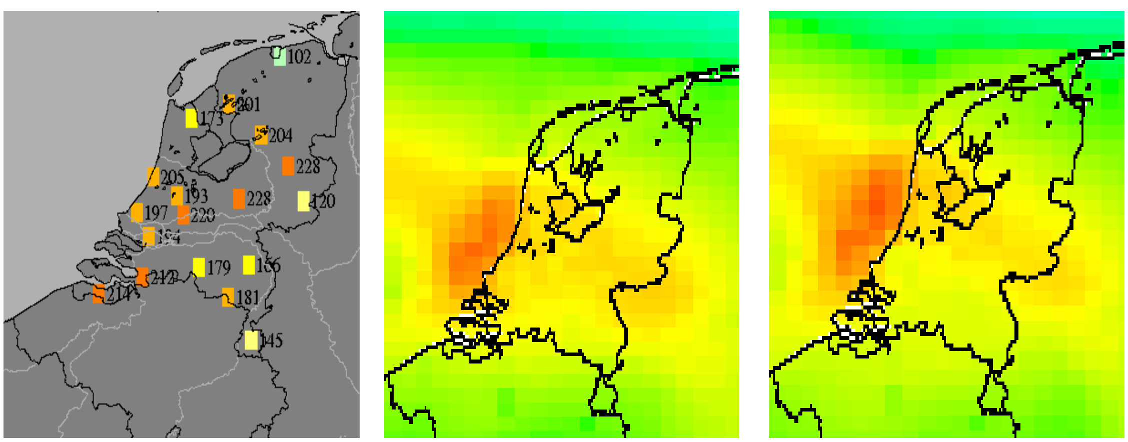
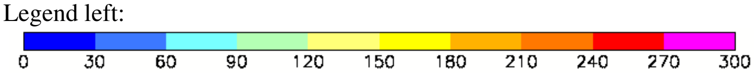


Figure 34. Daily maximum ozone concentrations (μg/m^3) for July 19th, 2006 in The Netherlands, as modelled by LOTOS-EUROS without data-assimilation.
Left: LML; middle: LOTOS-EUROS, analysis; right: LOTOS-EUROS , 1-day forecast
Legend middle, right: the same as Figure 32.
Legend left:



4.3.2 LOTOS-EUROS with data-assimilation of rural stations in The Netherlands

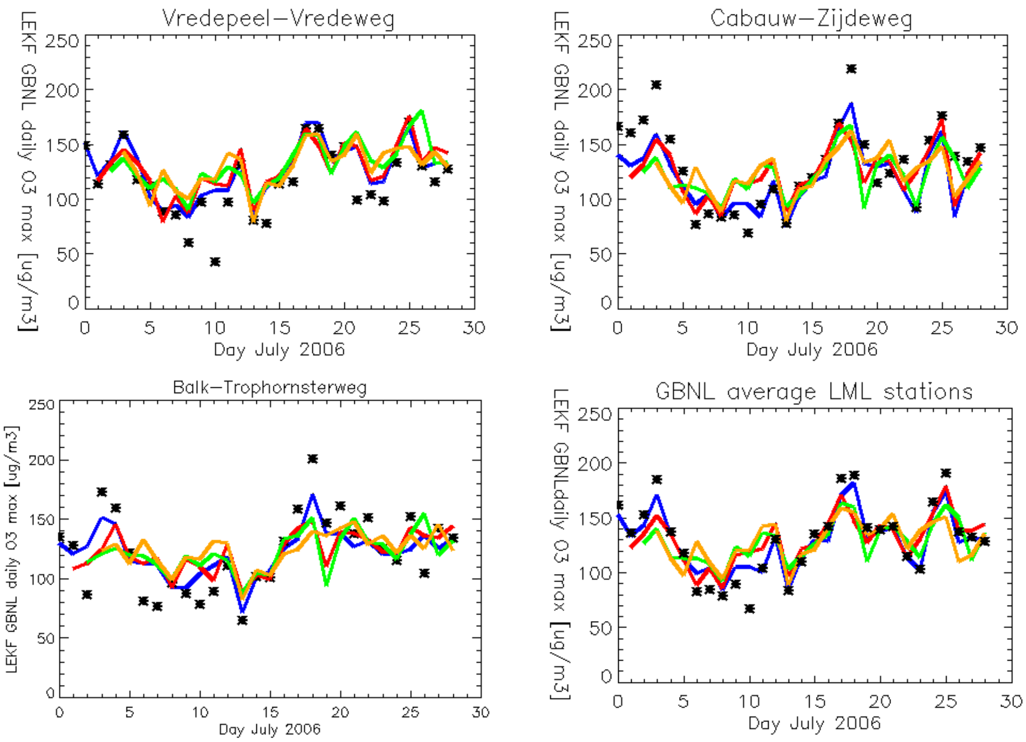


Figure 35. Daily maximum ozone concentrations (µg/m^3) for July 2006 as modelled by LOTOS-EUROS using EnKF data assimilation, assimilating rural stations from The Netherlands. See Table 3 for legend.

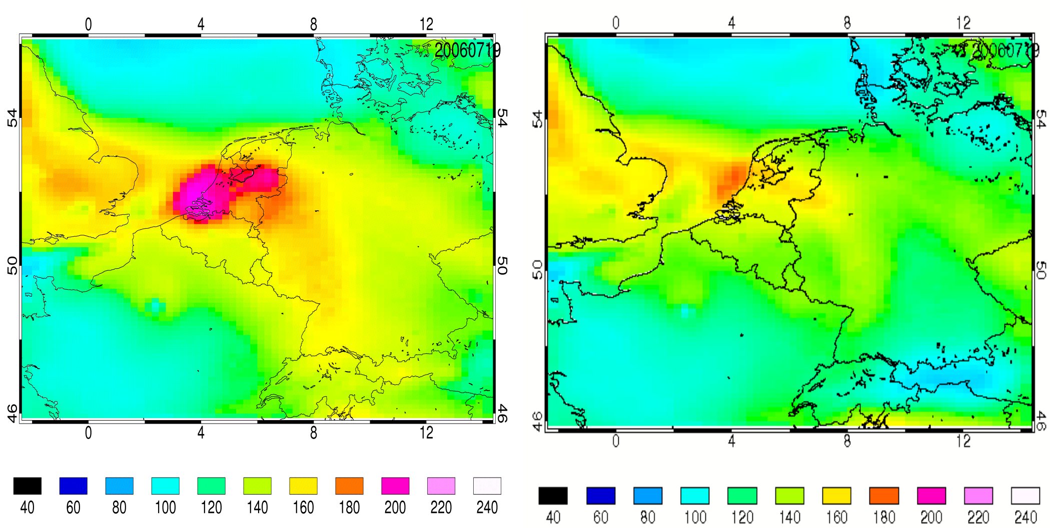


Figure 36. Daily maximum ozone concentrations (µg/m^3) for July 19th, 2006 as modelled by LOTOS-EUROS using EnKF data assimilation, assimilating rural stations from The Netherlands.
Left: analysis; right: 1-day forecast

LOTOS-EUROS with data-assimilation of rural stations in The Netherlands

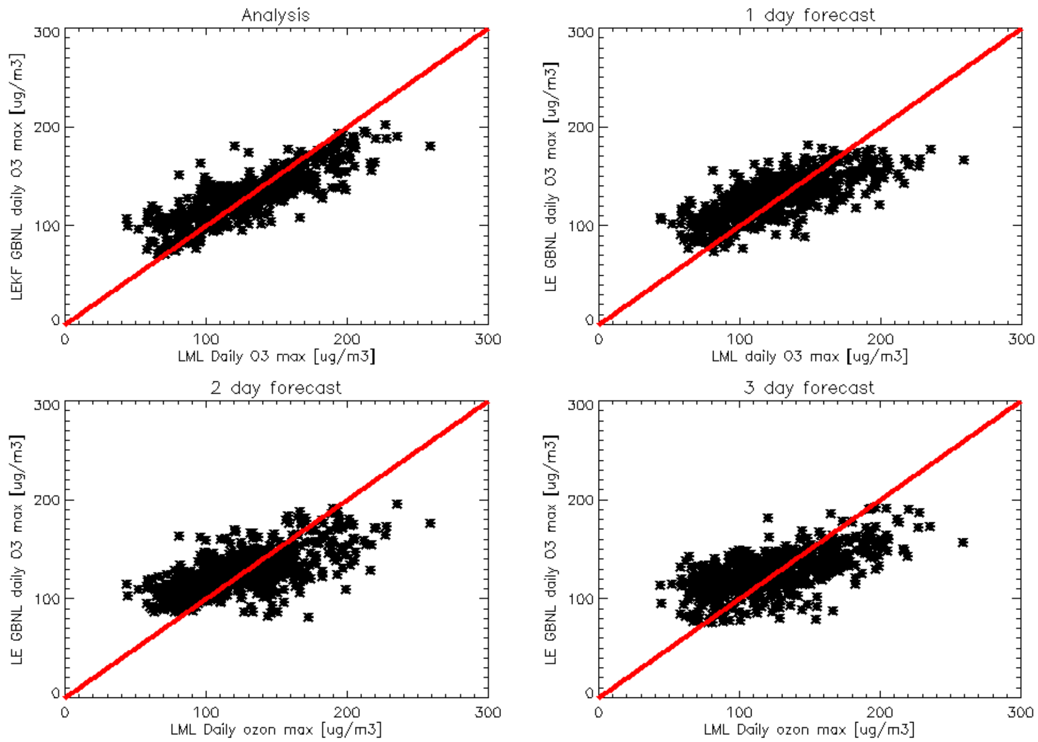


Figure 37. Daily maximum ozone concentrations (μg/m^3) for July 2006; scatter plot LOTOS-EUROS using EnKF data assimilation, assimilating rural stations from The Netherlands versus LML measurements.

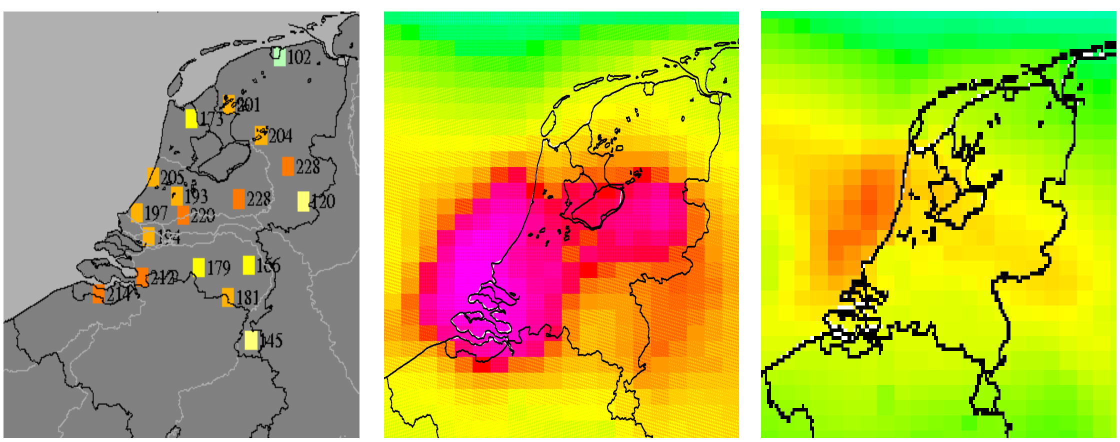
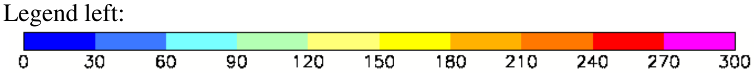


Figure 38. Daily maximum ozone concentrations (μg/m^3) for July 19th, 2006 in The Netherlands, as modelled by LOTOS-EUROS using EnKF data assimilation, assimilating rural stations from The Netherlands.
Left: LML; middle: LOTOS-EUROS, analysis; right: LOTOS-EUROS , 1-day forecast.
Legend middle, right: the same as Figure 36
Legend left:



4.3.3 LOTOS-EUROS with data-assimilation of rural stations in The Netherlands, Belgium and Germany

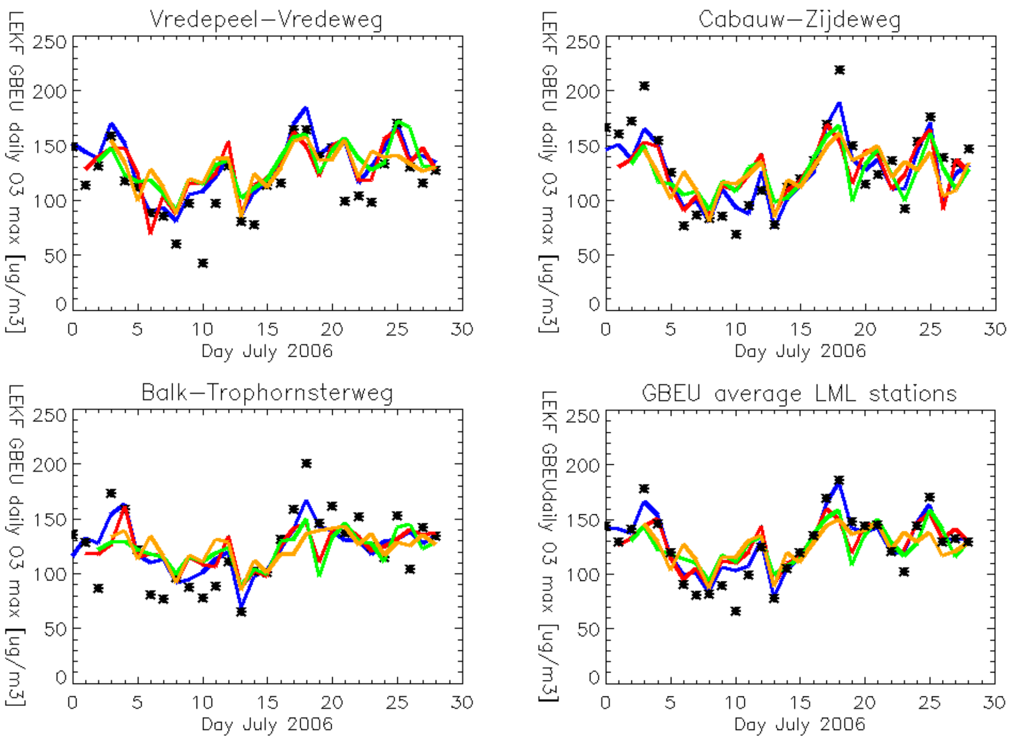


Figure 39. Daily maximum ozone concentrations (µg/m^3) for July 2006 as modelled by LOTOS-EUROS using EnKF data assimilation, assimilating rural stations in The Netherlands, Belgium and Germany. See Table 3 for legend.

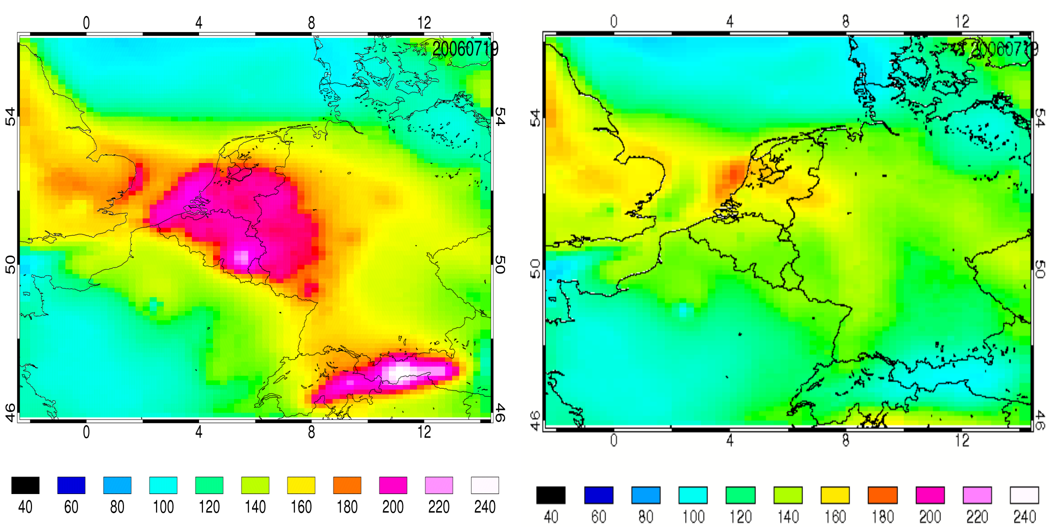


Figure 40. Daily maximum ozone concentrations (µg/m^3) for July 19th, 2006 as modelled by LOTOS-EUROS using EnKF data assimilation, assimilating rural stations in The Netherlands, Belgium and Germany.
Left: analysis; right: 1-day forecast

4.3.4 LOTOS-EUROS with data-assimilation of rural stations in The Netherlands, Belgium and Germany

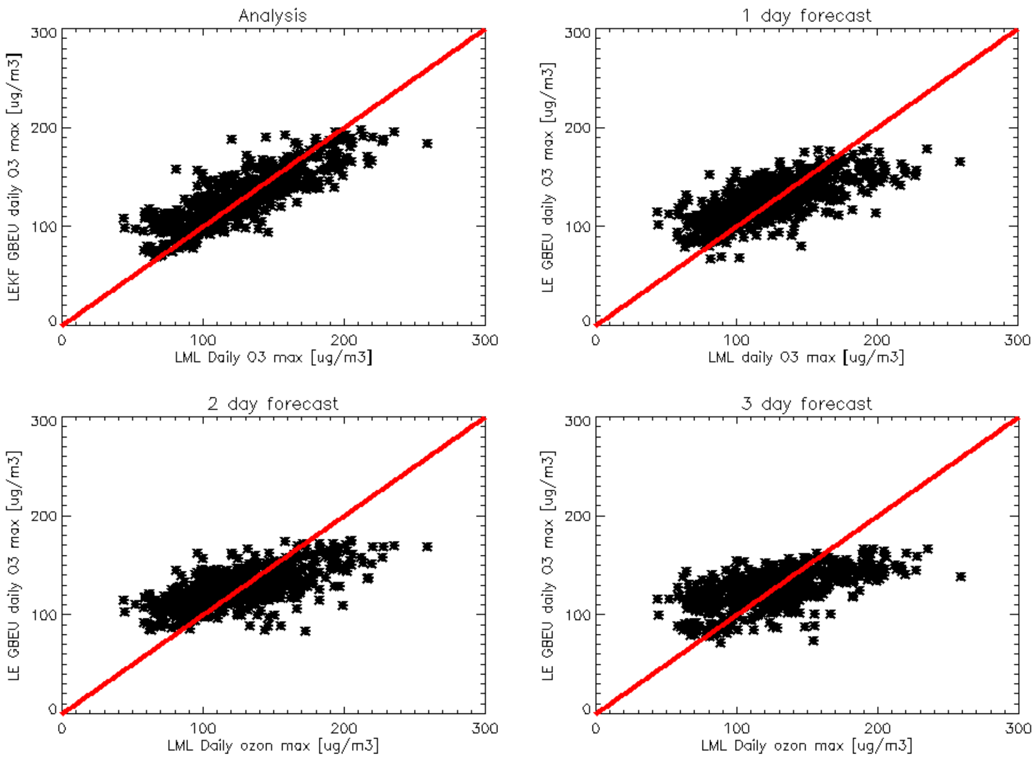


Figure 41. Daily maximum ozone concentrations (μg/m^3) for July 2006; scatter plot LOTOS-EUROS using EnKF data assimilation, assimilating rural stations in The Netherlands, Belgium and Germany versus LML measurements.

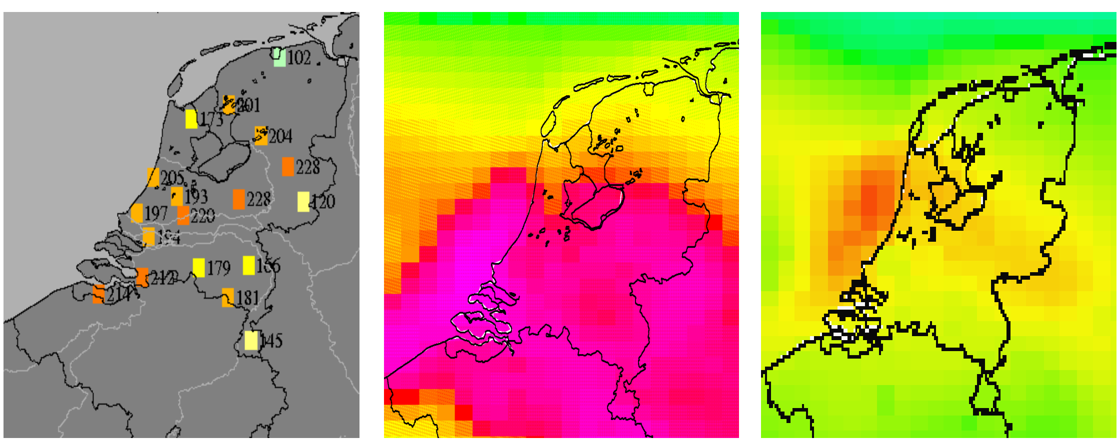
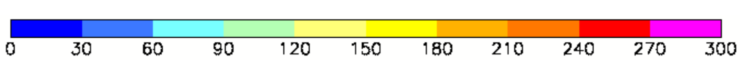


Figure 42. Daily maximum ozone concentrations (μg/m^3) for July 19th, 2006 in The Netherlands, as modelled by LOTOS-EUROS using EnKF data assimilation, assimilating rural stations in The Netherlands, Belgium and Germany. Left: LML; middle: LOTOS-EUROS, analysis; right: LOTOS-EUROS, 1-day forecast.

Legend middle, right: the same as in Figure 40

Legend left:



4.3.5 LOTOS-EUROS with data-assimilation of OMI NO_2 data

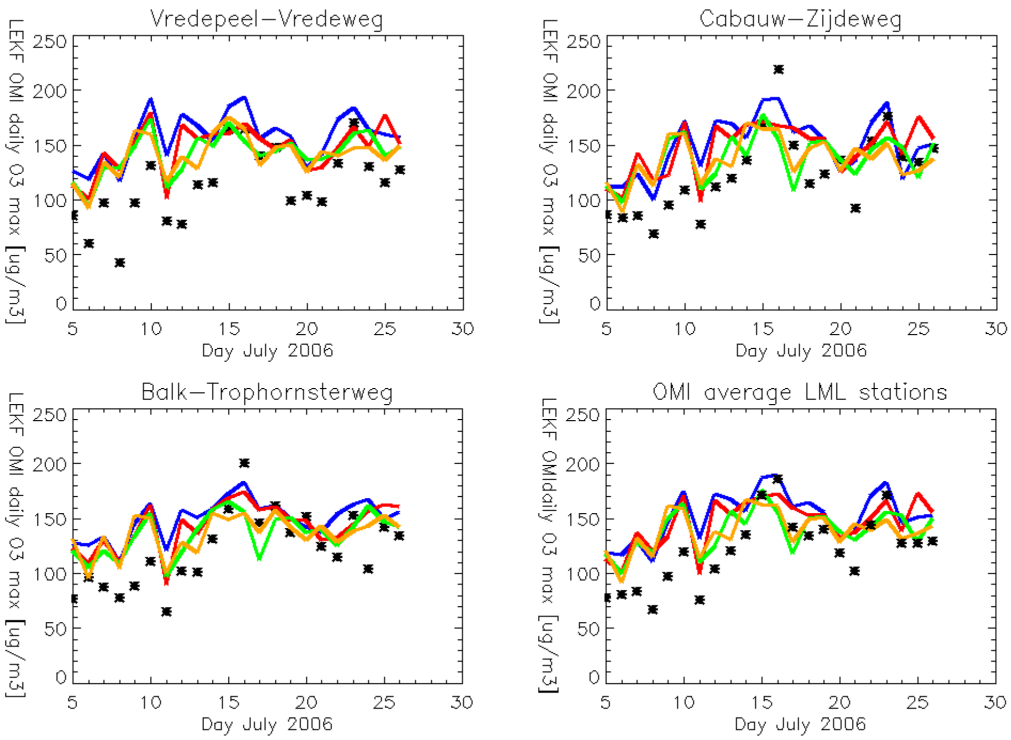


Figure 43. Daily maximum ozone concentrations (µg/m^3) for July 2006 as modelled by LOTOS-EUROS using EnKF data assimilation, assimilating OMI NO_2 data. See Table 3 for legend.

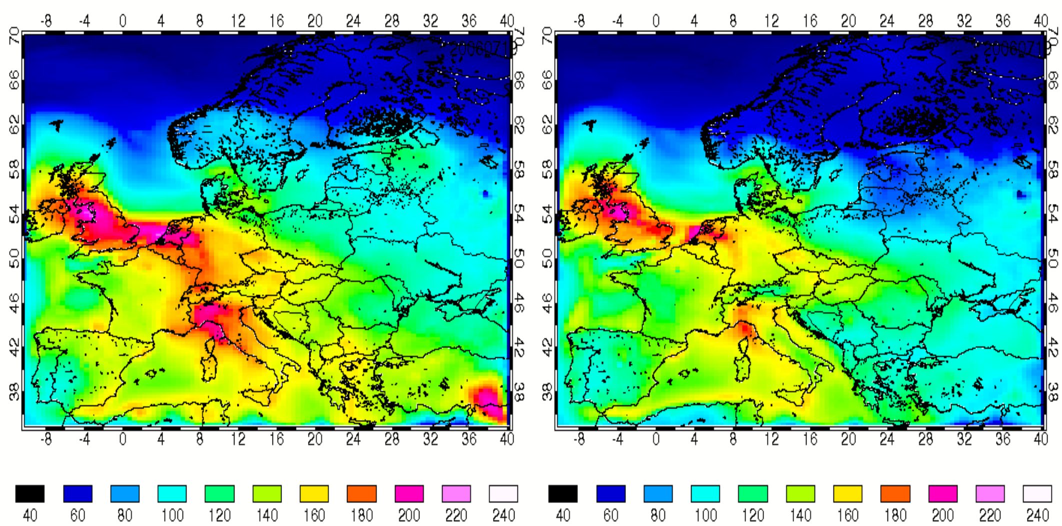


Figure 44. Daily maximum ozone concentrations (µg/m^3) for July 19th, 2006 as modelled by LOTOS-EUROS using EnKF data assimilation, assimilating OMI NO_2 data.
Left: analysis,; right: 1-day forecast.

4.3.6 LOTOS-EUROS with data-assimilation of OMI NO_2 data

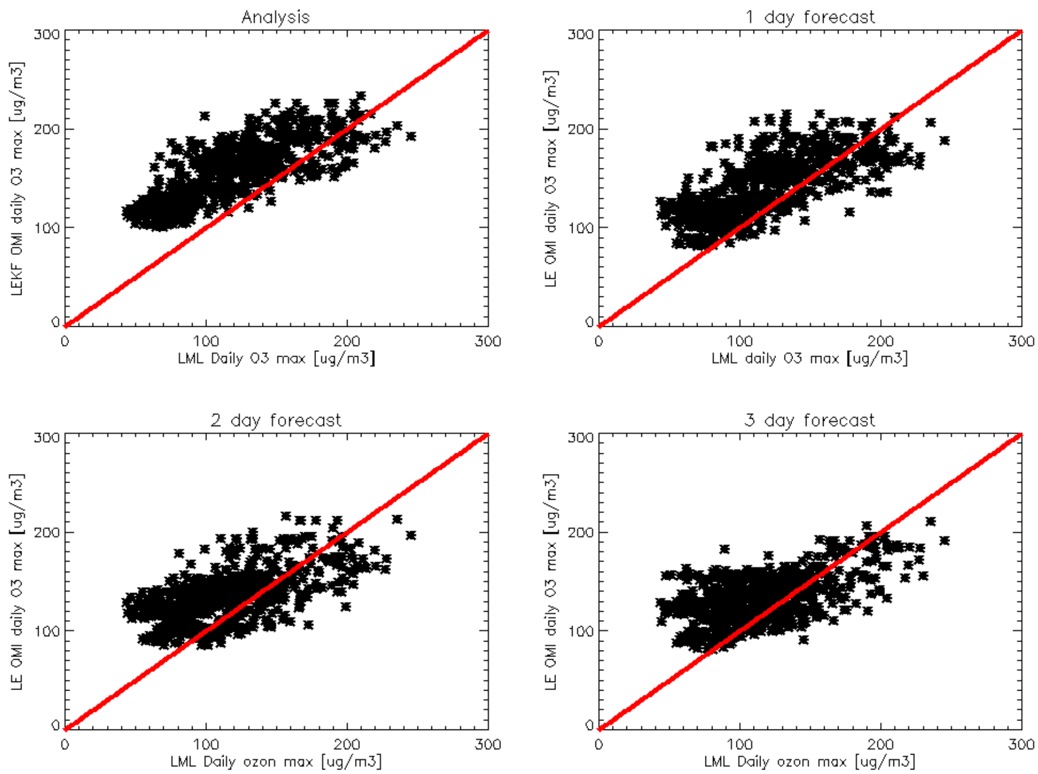


Figure 45. Daily maximum ozone concentrations (µg/m^3) for July 2006; scatter plot LOTOS-EUROS using EnKF data assimilation, assimilating OMI NO_2 data versus LML measurements.

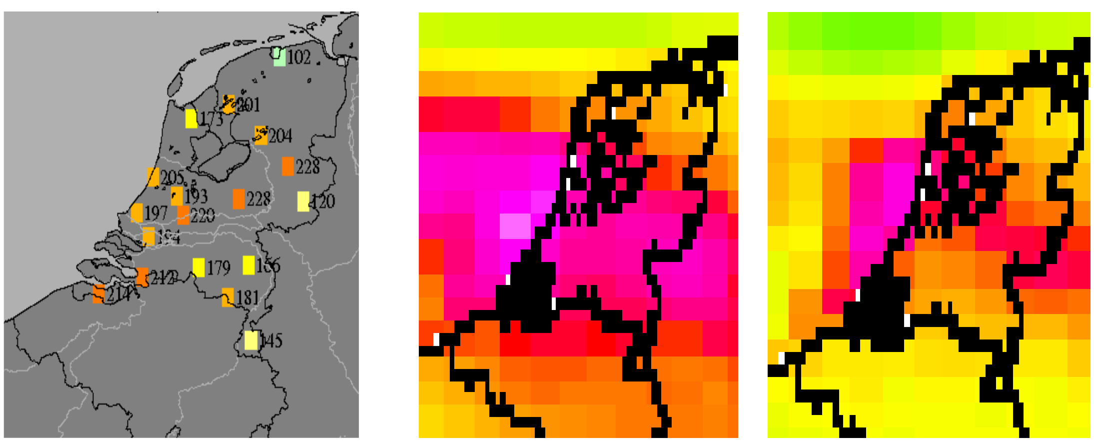


Figure 46. Daily maximum ozone concentrations (µg/m^3) for July 19th, 2006 in The Netherlands, as modelled by LOTOS-EUROS using EnKF data assimilation, assimilating OMI NO_2 data.
Left: LML; middle: LOTOS-EUROS, analysis; right: LOTOS-EUROS, 1-day forecast.

Legend the same as in Figure 44.

4.4 Discussion

In this discussion, we refer to a day maximum with the term *maximum*, the term *peak* is used for a high daily maximum as compared with moderate daily maxima.

4.4.1 PROZON

Figure 28 - Figure 30

In the time series of the individual stations (Figure 28), some PROZON forecasts are missing. When some hourly observations are missing for a particular station, PROZON does not give a forecast for this station.

PROZON has the tendency to overestimate the ozone peak values, which is most prominently visible for the 1-day forecast (Figure 28 and Figure 29). From the picture showing the PROZON average over all LML stations it can be concluded that the peaks at the 3^{rd}, 19^{th} and the 25^{th}, are often forecasted 1 day too late. The timing is sometimes even 1 day too early (19^{th}). Also the end of the episode (6^{th}, 13^{th}, 23^{rd}) is not well predicted. The 1-day forecast gives the best performance, the 2-day forecast gives the same overestimation but with a much larger spread and lower concentrations and for the 3-day forecast the spread is so large that the forecast cannot be used in practice. Part of the decreasing performance with increasing forecast time can probably be attributed to uncertainties in the meteorological forecast, which increases in time, and the unability of the model to take the gradual development of weather systems into account.

4.4.2 LOTOS-EUROS without data-assimilation

Figure 31 - Figure 34

The free-running LOTOS-EUROS model (without data-assimilation) predicts high ozone peaks at the correct days (Figure 31). The free-running model with ECMWF analysis performs best, the forecasts become slightly less accurate due to the increasing uncertainty in the meteo forecast, as with PROZON. The difference between the 1-, 2- and 3-day forecasts is however much smaller than for PROZON. The time series and especially the scatter plots (Figure 33) indicate that, in spite of the reasonable time correlation, the absolute values of the daily maximum do not represent the low and high values well. LOTOS-EUROS tends to overestimate low ozone peaks and to underestimate high ozone peaks. This lack of extreme values is a shortcoming of the model. The 3-day forecast is clearly better than for PROZON, which illustrates that the model is driven by meteorological developments in a realistic way.

4.4.3 LOTOS-EUROS with data-assimilation of rural stations in The Netherlands

Figure 35 - Figure 38

Including the LML observations in LOTOS-EUROS clearly improves the analysis (Figure 35) in the sense that LOTOS-EUROS comes closer to the observed high peaks. Low daily maxima are still overestimated. Also the 1-day forecast is slightly improved, with peaks at the right day except for the peak at the 18^{th} that should be at the 19^{th}. With increasing forecast time we see that data-assimilation has only minor impact on the forecast.

Comparing the figures showing the average over all LML stations of LOTOS-EUROS with and without data assimilation (Figure 35 and Figure 31), it can be concluded that the end of the episode is better predicted by the LOTOS-EUROS run with data assimilation. Also the absolute value of the ozone peak is better predicted, even though with increasing forecast day it gets worse.
Compared to the maps of the run without data-assimilation (compare Figure 32 and Figure 34 to Figure 36 and Figure 38 "analysis"), we see that by assimilating the Dutch ozone ground measurements the modelled peak ozone concentrations become significantly better in the Netherlands. Both low and high ozone concentrations are better represented by the model. Comparing the maps of the 1-day forecast, we see that the data-assimilation does not improve the model results significantly. Scatter plots (Figure 9) indicate that the forecasts were slightly improved with respect to the case without data-assimilation, especially the 2- and 3-day forecast.

Ozone formation in a CTM is a conservative system, which means that changes in initial conditions do no strongly affect future concentrations. This is unlike weather systems, where slight changes in initial conditions have large impact on future weather. Therefore, it was already anticipated that also the settings of the noise obtained by the data assimilation should be kept in some way in the forecast. As we saw in Chapter 3, noise factors are changing relatively strongly from day to day, introducing errors if they are kept constant in the forecast. It is difficult to distinguish the effect of improved initial conditions and the effect of adjusted emissions and boundary conditions. Especially further ahead, the noise from the Kalman filter may not apply, the correlation time may not be long enough. In the present setting, these effects are intertwined.

4.4.4 LOTOS-EUROS with data-assimilation of rural stations in The Netherlands, Belgium and Germany

Figure 39 - Figure 42
Since air is transported away and replaced by air from surrounding countries, it may be important to adjust emissions upstream from The Netherlands. In case of ozone smog over the Netherlands, this air generally comes from the south and east, from Belgium and Germany. Therefore, ozone observations from rural stations in Belgium and the western part of Germany were assimilated. Figure 39 shows that both the analysis and the forecast runs in almost all cases predict the ozone peak at the correct day. When comparing the results of the average over all stations with the results from the other model runs, the end of the episode is best predicted by LOTOS-EUROS assimilating stations from The Netherlands, Belgium and Germany. With increasing forecast time, there is less variability of the maximum values when Belgian and German stations are assimilated, which can be observed when comparing figures Figure 35 and Figure 39. The 1-day forecast is slightly better than the LOTOS-EUROS run without assimilation and comparable to the run of LOTOS-EUROS assimilating only Dutch LML stations. The performance of the 2-day forecast is similar to LOTOS-EUROS assimilating only Dutch LML measurements; the 3-day forecast tends to be worse.

Comparing the maps in Figure 40 and Figure 42 to those of the runs without assimilation and assimilation of Dutch stations only, we see that by assimilating ozone ground measurements of Dutch, German and Belgian stations, the ozone concentration becomes significantly higher over the Netherlands, Belgium and Germany. We also see that ozone is transported from

Belgium and south-west Germany to the Netherlands. In the 1-day forecast, the effect of the improved initial conditions (analysis) and the noise settings seems nearly lost.

4.4.5 LOTOS-EUROS with data-assimilation of OMI NO_2 data

Figure 43 - Figure 46

On average in July 2006, the daily ozone maximum is increased by the assimilation of OMI NO_2. The increase is between 0 and 50 µg/m^3. On average, this increase leads to a positive bias with respect to the LML ground observations. Also in the forecasts, ozone concentrations tend to be too high but they decrease with increasing forecast time.

The dynamical range in ozone values has increased. In particular, the occurrence of peaks and the relative increase at high ozone episodes seem to be better captured in the assimilation run with OMI observations. This is a promising result, showing how OMI NO_2 observations may have a positive impact on the analysis/forecast of surface ozone.

The main problem at this point is the large discrepancy between the modelled NO_2 and the OMI observed NO_2 values. A meaningful bias correction scheme for both OMI and LOTOS-EUROS needs to be introduced to bring the model and observations closer together and to better obey the bias-free assumption underlying the Kalman assimilation approach. Such bias estimates should be based on comparisons with surface observations and other independent validation measurements.

4.5 Conclusions

4.5.1 Beginning and end of an episode

In the above comparison of all models, it is shown that LOTOS-EUROS is better in predicting high peak values of ozone at the correct day than PROZON, even without data assimilation, especially at the beginning and end of a smog episode.

4.5.2 Absolute value of peak

PROZON is better in forecasting the absolute value of a peak one day ahead with the tendency to overestimate the peaks, whereas LOTOS-EUROS often underestimates the peaks significantly, even after data assimilation. LOTOS-EUROS overestimates low ozone daily maxima. Overall, LOTOS-EUROS does not reach the extremes, unless forced by the data assimilation. The beneficial effect of data-assimilation decays quickly in the forecast.

4.5.3 Analysis

In the **analysis**, the best model performance in modelling ozone values (peaks) is obtained when both the LML and the Belgian and German measurements are assimilated. Both low and high values are better captured. LOTOS-EUROS without data assimilation performs worst. It can be concluded that assimilating ground based data for the analysis significantly improves the initial conditions. The analysis with OMI NO_2 columns results in too high ozone concentrations. A bias correction or model improvement is necessary before the assimilation of NO_2 column can be used for the forecast.

4.5.4 Forecasts

For the **1-, 2- and 3-day forecast** of LOTOS-EUROS with data assimilation performs slightly better than the free running model. The 1-day forecast benefits most from the data assimilation. This can be explained by the better starting situation and by using adjusted emissions and boundary conditions. The effect of improved initial conditions decays rapidly with increasing forecast time and for longer forecast times the model without adjusted emissions and boundary conditions may even perform better. Added assimilation of Belgian and German stations performs a little better for the high ozone peaks. PROZON is better in predicting one day ahead, for two days ahead the spread becomes larger than for LOTOS-EUROS and LOTOS-EUROS becomes better in the long-term forecast. PROZON is overall better in predicting absolute values, LOTOS-EUROS is better in predicting the timing of the ozone maxima.

5 Operational forecasting

The central aim of the SmogProg project is to provide improved information and forecasts of air quality in the Netherlands. This is achieved by introducing prognostic models and by the assimilation of both surface observations and OMI NO_2 measurements from space.
The operational smog forecast is issued on a daily basis by RIVM. The project therefore aims to improve the instrumentation of the RIVM smogteam for this task. Indirectly, clients of the RIVM smog forecasts will benefit from the project. They will receive more accurate and more detailed information. These other clients include the Dutch provinces, the municipal health agencies (GGD's), and the Dutch public.

For the purpose of communicating this information two websites have been developed where daily forecasts, validation information as well as archived model results are made available. These are:

1. The public website of the project, (www.lml.rivm.nl/data/smogprog/index.html), is accessible through a link on the RIVM LML website with surface observations.
2. A password-protected website (www.temis.nl/luchtkwaliteit) has been set up for the project members and other interested parties. This website provides additional information to the air quality team at RIVM, with forecasts and model output from different models and different assimilation experiments.

Two phases can be distinguished within the SmogProg project, namely a demonstration phase (summer 2007- summer 2008) and operational phase (summer 2008 onwards).

5.1 SmogProg demonstration phase

In the summer of 2007 a demonstration air quality forecast has become available together with the SmogProg project web site. This demonstration service was limited to ozone only. Daily three-day forecasts of surface ozone are provided for the Netherlands. This information is based on forecast runs (without data assimilation) with the CHIMERE model (version 200606a) run on the European domain with 0.5 degree resolution, driven by ECMWF meteorological data.

The SmogProg website provides the following information:

- The predicted ozone peak value for today, tomorrow and the day after tomorrow.
- The predicted maximum 8-hourly running mean ozone concentration for today, tomorrow and the day after tomorrow.
- A movie of the development of ozone during the forecast period, based on hourly model output.
- An operational verification page which presents a comparison of the ozone peak and ozone daily mean model results compared with the Landelijk Meetnet Luchtkwaliteit (LML) measurements of yesterday.
- Apart from this dynamical content the website also provides background information on air quality thresholds, on air quality modelling and on the SmogProg project.

The password-protected project website provides more extensive information based on the following operation sources:

- LOTOS-EUROS forecasts. LOTOS-EUROS version 1.3 of July 2007 is providing daily 3-day forecasts of ozone and related trace gases and aerosols. The model (without assimilation) is run on the European domain with a resolution of 0.5 by 0.25 degree, driven by meteorological forecasts of the ECMWF.

- CHIMERE forecasts. CHIMERE version 200606a is providing daily 3-day forecasts of ozone and related trace gases and aerosols. The model (without assimilation) is run on the European domain with a resolution of 0.5 by 0.5 degree, driven by meteorological forecasts of the ECMWF.
- Within SmogProg a routine transfer of LML surface data from RIVM to KNMI was established. This data is converted to images showing the daily peak, mean, or 8-hourly mean for ozone, NO_2 and PM_{10} for the rural stations.
- OMI and SCIAMACHY nitrogen dioxide measurements are available in near-real time, within three hours after measurement.

The password-protected project website presents this information in a couple of different ways:

- There are dedicated pages that show ozone, nitrogen dioxide and PM_{10} for the two models, for the LML measurements and for OMI and SCIAMACHY NO_2. These results are presented side-by-side for a direct comparison. Web pages are available for: 1. The NO_2 column over Europe; 2. Ozone peak values over the Netherlands; 3. Ozone peak values over Europe; 4. Ozone daily-mean values over the Netherlands; 5. Ozone daily-mean values over Europe; 6. The maximum 8-hour mean ozone concentration over the Netherlands; 7. Same, but over Europe; 8. NO_2 peak values over the Netherlands; 9. NO_2 peak values over Europe; 10. NO_2 daily-mean values over the Netherlands; 11. NO_2 daily-mean values over Europe; 12. PM_{10} over the Netherlands; 13. PM_{10} over Europe.
- An archive of all these images for the different model and observational information sources.

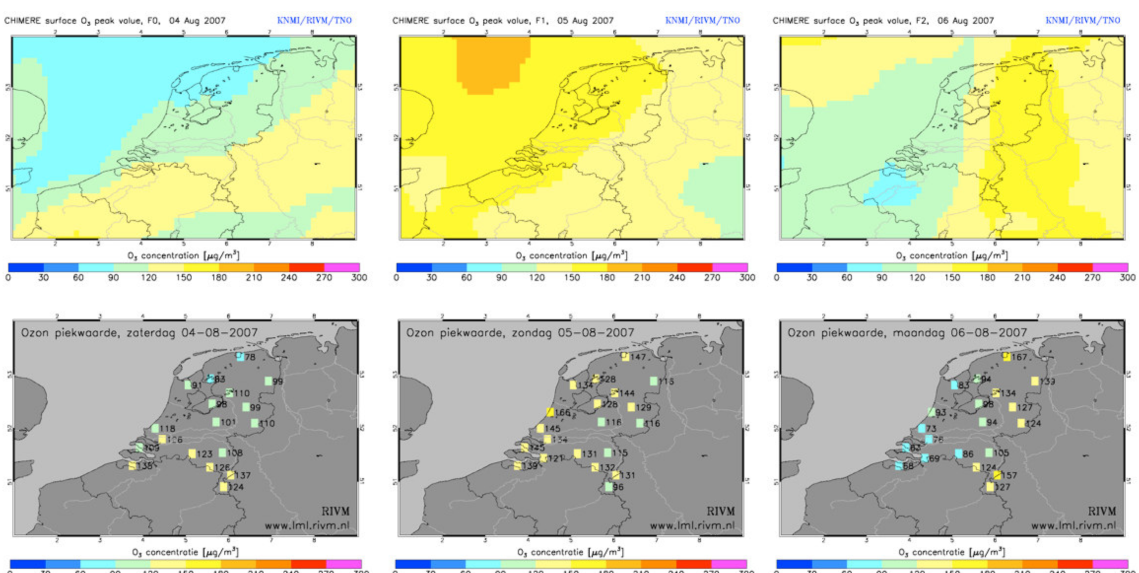


Figure 47. A forecast of ozone compared with measurements from rural LML-stations. The top three panels show the 1-day (4 Aug), 2-day (5 Aug) and 3-day forecast of surface ozone. The results are compared with surface measurements that became available 1-3 days later. The forecast predicts rapid changes in the ozone distribution related to changes in the weather, with the highest ozone in the South, North-West or East, for 3, 4 or 5 August respectively. This change in the ozone gradient over the Netherlands is very similar to what was observed.

The figure shows one interesting ozone forecast made during the demonstration phase, where changes in the weather result in pronounced gradients of ozone over the Netherlands. Such a detailed change in ozone distributions can not be predicted by the statistical forecast that is now in use at RIVM for providing information to the public. This shows the added value of using prognostic air quality models driven by state-of-the art weather forecasts.

5.2 SmogProg operational phase

In summer 2008 the demonstration service described above is further upgraded to include in particular the results from the assimilation. The final SmogProg operational service will consist of the following configurations:

1. Free-running LOTOS-EUROS and CHIMERE models, complemented by Dutch near-real ground based air quality data from the RIVM-LML network, as well as near-real time NO_2 data from the OMI satellite instrument on EOS-Aura.
2. Analysis and forecast runs of LOTOS-EUROS with an assimilation of surface observations over the Netherlands from the LML network. This model will replace the operational model Chimere on the public website of SmogProg.
3. Analysis and forecast runs of LOTOS-EUROS with an assimilation of OMI NO_2 observations over Europe.
4. Analysis and forecast runs of LOTOS-EUROS with an assimilation of surface observations over the Netherlands from the LML network extended with surface observations from Germany and Belgium.
5. A combined assimilation of surface observations and OMI NO_2 observations.

A few additional comments:

- Configuration 1: The LOTOS-EUROS model is run in a nested setup: First the model is run on the European domain with a resolution of 0.5 by 0.25 degree. Secondly, a nested zoom forecast run is performed with a resolution of 0.25 by 0.125 degree, forced by boundary conditions from the European run. Maps of ozone, NO_2 and PM_{10} are made on the European domain based on the European model forecast, and over the Netherlands based on the zoom forecast run. These runs are operationally available since the end of July 2008.
- Configuration 2: This will become available in August 2008.
- Configuration 3: This will become available in August/September 2008 on the password-protected site. Additional work is planned to optimise the assimilation and to provide a good bias-correction scheme.
- Configuration 4: Depends on the availability of near-real time surface data, ideally provided through EEA.
- Configuration 5: The combined assimilation of both OMI and surface observations requires more work. The bias problem needs to be dealt with and the impact of OMI data on in particular surface ozone needs to be quantified.

The SmogProg operational service has the main aim to provide customised information to the RIVM to support the daily smog forecasts. In particular information will be provided in a form as requested in the “SmogProg Programma van Eisen”, see Appendix B.

The following information is demanded:

- Maps for the Netherlands with the expected maximum ozone concentration for today and similar maps for the coming 24 hours near the surface. This information is already provided.

- Animation of the hourly concentrations of ozone. This information is already provided.
- A file with the maximum hourly ozone concentration per model pixel, for today and for the coming 24 hours. This information is available, but is not used yet. The SmogProg and LML teams will further discuss a user-friendly way of providing this data.
- The resolution should be of the order of 10 km. This is covered by the zoom version of Lotos-Euros which provides information with a resolution of about 13-15 km.
- Two updates per day of the forecast should be provided. This is not yet available, but will be implemented soon.

Apart from this there is a wish-list in the "SmogProg Programma van Eisen":

- Forecasts of the 8-hourly means. The maximum 8-hourly mean is already provided. An animation of the 8-houly means will be provided soon.
- Provide information for the individual provinces. This is not yet implemented.
- The choice of stations used in the assimilation should be flexible, and the SmogProg system should be able to rapidly adopt the assimilation in case of for instance changes in quality and availability of stations. This is already foreseen: The list of stations that are actively assimilated can be easily adopted within one working day.
- There should be a validation report: The results presented in this report partly cover this. The content of the validation report will be further discussed in the SmogProg/LML teams.

It should be noted that the SmogProg project team will continue to exist and continue to work on the project for at least 5 months after this report. Many if not all of the indicated open issues will be resolved in this period.

5.3 Review from the LML Smogteam

In addition to the current used statistical ozone model PROZON the members of the Smogteam started to use the model output on the password protected website in June 2008. In the summer months, Smogteam members focus on ozone measurements and forecast. Less attention was given to other pollutants.
Smog episodes due to ozone exist especially in periods of warm and sunny weather. Two short periods of moderate smog were observed. The first one was in the beginning of July and the second on the last day of July. The available output of the different models in the mentioned periods was compared with the statistical PROZON output. The extra information proved to be a very helpful tool to make a better judgement of the smog situation. The Smogteam experienced that the extra information also provides an indicator of the reliability of the separate models including the currently used model PROZON.
Based on all the available data the Smogteam decided twice to warn the provinces that moderate smog by ozone was expected. In both cases the warning was justified. Moderate smog levels were measured.

The RIVM Smogteam is looking forward to the full implementation of the SmogProg model including the data assimilation part, and to the full implementation of the web based interface to the Smogteam as defined in the Smogteam definition study. It will work closely with the SmogProg project team to tailor the final product to the Smogteam's operational needs.

6 Cost-Benefit analysis

As described in the previous chapters, a system was developed which assimilates ozone ground measurements and OMI NO_2 column observations in a chemistry transport model and produces an ozone forecast. This system is more advanced and has more possibilities than PROZON, the smog forecasting system that is currently available. Apart from the costs of developing the system, the system will be more expensive to maintain than PROZON. However, ozone has adverse health effects and accurate prediction of smog episodes may help to reduce these effects by reducing exposure of the population. Preventing smog episodes by ad hoc measures is very difficult, but public awareness can help in reducing long-term emissions and push back the ozone peaks. The savings on public health can outweigh the costs of building and maintaining SmogProg. We expect that the fully operational model will be used for a period of at least 7 year. Costs and benefits are summarised in Table 4.

6.1 Costs

6.1.1 Development of SmogProg

In the development of SmogProg, the following steps were taken. We have

- adapted the LOTOS-EUROS data assimilation scheme, as developed for PM_{10}, to meet specific requirements for ground-based ozone observations (daily cycle, nonlinearity, bias, uncertainties)
- developed the LOTOS-EUROS data assimilation module to assimilation of column NO_2 satellite observations
- set up data streams (use of ground observations LML + stations from Belgium and Germany through EEA, OMI satellite data, ECMWF NWP data)
- evaluated the SmogProg forecasts
- implemented SmogProg in a semi-operational setting (automatic run every night)
- set up data stream for output, publication to a public and a password protected website

The total project costs were k€ 413. These are non-recurrent costs.

6.1.2 Operationalisation

Some effort has still to be made to make the SmogProg system operational.

- the data stream with near real time data from EEA has to be set up
- the data stream to an official RIVM website and videotext has to be set up, including the possibility for a smog team member to provide comment on the website
- data assimilation of satellite data has to be improved before it can contribute to an improvement of the forecast

Converting the SmogProg software to a real 24/7 operational application in the future will involve additional costs (initial development, hardware extensions and process monitoring). These costs are non-recurrent and estimated to k€ 25.

6.1.3 Maintenance

Producing and communicating a daily forecast to the public is a legal task of RIVM. In the operational practice, daily forecasts were made and published to videotext and the internet.

The maintenance of the data stream from a model run to stakeholders and the public, including a quality check by a smog team member is a core task of RIVM which has to be carried out in any case. Therefore this part of the process will be left out of the cost benefit analysis. We focus on the costs of running and the preceding data streams.

A single run with PROZON (one day forecast) only takes a few seconds. The costs of running PROZON at RIVM are k€ 0.2-0.3 per year, consisting of computer system costs (50%) + personnel costs (50%) including occasional updates of the statistics. However, in the near future the computer system on which it is running will be replaced. The computer system costs of PROZON may increase, alternatively a possible migration to a different platform will result in non-recurrent personnel costs of about 3 k€.

A single run with SmogProg takes about 3 hours (analysis + three days forecast). SmogProg is running on KNMI's computing facility. Because no additional hardware investments were made, there was no additional cost during the SmogProg project duration. The costs of running SmogProg are largely determined by personnel costs (5-10 k€ per year).

6.1.4 Observations

Although running SmogProg is more expensive than PROZON, modelling is still far cheaper than monitoring air quality. Observations of ground-level ozone and NO_2 concentrations are mandatory (European legislation) and do not contribute to the project costs. However, the use in the SmogProg project gives them added value and it is worthwhile to look at the costs of the observations for a feeling for the order of magnitude. The costs of observations are not included in the cost-benefit analysis itself.

Observations of ground-level ozone concentrations are a legal obligation. The cost of ozone measurements in LML are k€ 300 per year. Apart from LML data, also data from surrounding countries, which were made freely available through the EEA are used. Ozone observations are in fact more crucial for the functioning of PROZON than for LOTOS-EUROS (SmogProg). The capability of LOTOS-EUROS to assimilate not only Dutch observations but also observations from the surrounding countries puts the observations in a better perspective, especially in the border regions.

The total costs of the OMI satellite are estimated to be € 100 million, with an expected lifetime of 10 years this is k€10000 per year. The OMI satellite yields global observations for a number of gases and was designed in particular for measurements on the ozone layer. The SmogProg project exploits only a fraction of the OMI data, but by doing so it gives added value to the OMI mission.

6.2 Benefits

The costs of PROZON are far lower than those of SmogProg and the forecasts are at least as good regarding absolute values, with the exception of the correlation with observations at the beginning and end of a smog episode. A further improvement of SmogProg is foreseen, since the underlying LOTOS-EUROS model is under continuous development, but this cannot be taken into account in the present cost-benefit analysis. However, public awareness is a key parameter and the possibities of SmogProg to inform the general public are clearly stronger. Here we estimate the positive effect of SmogProg on public health with respect to the existing situation.

6.2.1 Direct health related costs in the Netherlands

Although absolute values within a smog episode are still forecasted better by PROZON, the new SmogProg tool has several advantages:

- a better spatial coverage
- a better time resolution, forecast of time and duration of high concentrations
- it takes better account of regional differences in air pollution levels
- it gives a forecast on a larger area than only The Netherlands
- the beginning and end of a smog episode are forecasted better

These advantages yield better possibilities to inform the public. This is important since ozone has adverse health effects (e.g. WHO 2003). High concentrations can induce asthma attacks. Also a relationship has been found between ozone and *Minor restricted activity days*, which means that people have to adapt their activities because of health impacts, but without staying home from work. More dramatically, high ozone concentrations were found to causes premature deaths. During the summers of 2001, 2003 and 2006 heat waves resulted in relatively high ozone concentrations. In 2006, 1 to 2 % of the premature deaths in The Netherlands (which means 1000-2000 people) were attributed to exposure to high ozone concentrations (in excess to the victims that are attributed to the heat itself, MNC 2007). For the other periods similar numbers were obtained, and even in summers without such extreme conditions several hundreds of premature deaths were attributed to ozone (Fischer et al 2004). Also about 200 emergency admissions to the hospital during these summers were related to ozone.

It is difficult to estimate the costs of the health effects (De Hollander, 2004, Chapter 5). We adopted the estimates taken in the Belgian study by Torfs et al 2004. If we assume the cost of premature death to be k€166 and of hospitalisation to be k€ 4, the net costs are 1000x166+200x4= k€ 166800 per year. In years with less severe smog, these costs will be lower. But on average, the order of magnitude will be k€100000. The impact of asthma attacks and minor restricted activity days was not taken into account. In a study by De Hollander (2004) the annual total health costs were estimated to k€150000. As a conservative order of magnitude estimate we will stick to k€100000.

Exposure can be avoided by staying inside and limiting physical exertion outside during peak hours. Some people with lung problems already inform themselves using the near real time data and the operational PROZON forecast at the RIVM website and videotext. An alert is given when the alarm threshold is reached. The advantage of the new SmogProg system is that regional differences and timing of the peak is also forecasted so that people can plan their activities. Restrictions to normal behaviour are thus reduced to a minimum. Animations of the analysed and forecasted ozone concentrations, like the weather forecast, will make the information easier to understand for the general public. If 1% of the people would adapt its behaviour based on these forecasts, we could reduce the annual costs of health impacts due to ozone exposure by 1%. This would save k€1000 per year. These benefits are clearly larger than the costs of maintaining SmogProg.

When the public is better aware of the situation, it may be not only be motivated to adapt its behaviour during smog episodes to reduce exposure, but may contribute to a more structural reduction of the emission of ozone precursors. The latter will reduce the need to avoid exposure and contributes in this way to the benefits of the project.

When the peak prediction becomes more accurate and more timely, it becomes possible to plan emission reductions for specific days to prevent high concentrations. However, substantial emission reductions (overall in the order of 25%, e.g. Palacios et al 2002) are necessary to yield a substantial improvement and the reductions must be spread over a large area (at least at national scale) including all sectors (traffic, power plants etcetera). Short-term abatement measures like local traffic restrictions are generally regarded as inefficient in reducing local ozone concentrations (e.g. Smeets and Beck 2001). The larger scale measures may be effective but are difficult to realise.

6.2.2 European perspective

The forecast produced by SmogProg is not restricted to the Netherlands but covers a substantial part of Western Europe. By comparing the SmogProg results with other European forecasts (e.g. CHIMERE, Eurad), as is currently done in the European PROMOTE project, the strengths and weaknesses of the individual models can be compared and used to improve the models. Furthermore a multi-model approach (model ensemble prediction) in general leads to a better forecast as uncertainties and model weaknesses of individual models are averaged out. At the same time, the ensemble gives an indication of the uncertainty of the forecast. By actively taking part in such a European ensemble system, one can both improve the ozone forecast on the European and Dutch scale. At the same time, participating in such a project helps to keep up with the state of the art. Such knowledge is important to be able to fully interpret the SmogProg results and to improve them in the future.

6.3 Conclusion

The forecasts obtained by SmogProg are not always better than those obtained by PROZON, since peak values are often too low, but the beginning and end of an episode can be forecasted better. More importantly, SmogProg offers better means of communicating smog forecasts to the general public by means of animations and maps. By doing so, the public awareness is increased. Improved public awareness may lead to changes in behaviour, resulting in less exposure to ozone and inherently to a decrease of adverse health effects. Since these adverse health effects are very costly, the benefits easily outweigh the estimated costs. Even if only 0.1% of the people would adapt its behaviour, the benefits of SmogProg are larger than the costs.

Table 4. Yearly costs and benefits, assuming that SmogProg will be used for 7 years. Numbers are rounded to obtain order of magnitude estimates. PROZON has no additional benefts as it represents the benchmark situation.

	Costs SmogProg	Costs Prozon	Benefits SmogProg
Development (k€)	60	n.a.	
Operationalisation(k€)	5	n.a.	
Maintenance (k€)	15	1	
Health (k€)			1000
Total (k€)	80	1	1000

7 Outreach

The main end-users of smog forecast products are civil authorities and the general public. As part of this project we aim to inform these users about the project and its results, including the new use of satellite data in the forecast.

7.1 SmogProg flyer

A compact brochure/flyer was produced at the beginning of the project (see Appendix C) and distributed among civil authorities, especially the Dutch Ministry of Environment VROM, the Netherlands Environmental Assessment Agency PBL, users at the National Institute of Public Health and the Environment RIVM, and all users of the current smog forecast, including all Dutch provinces and all municipal health services (GGD's) in the Netherlands.

Figure 48. The SmogProg flyer

7.2 SmogProg webpage

A webpage with similar contents as the flyer was created to inform the general public. This webpage was hosted at RIVM and linked to the existing operational smog pages of RIVM. In the course of the project data from the new SmogProg machine were included and show air quality development from day to day, including a movie of the development of ozone concentrations. See: http://www.lml.rivm.nl/data/smogprog/index.html.

7.3 SmogProg email address

A special email address, SmogProg@rivm.nl, was opened for reactions to the SmogProg team. Very positive reactions were received, some requesting special features to be added, or commenting on existing features. We anticipate that feedback through this address will help us to further improve the SmogProg products when SmogProg replaces PROZON as operational tool in the future.

7.4 Final presentation

At the end of the project, results will be communicated in a presentation and demonstration to stakeholders at VROM, RIVM, PBL, local authorities and health organisations. We also aim to show some of the results in an attractive way to the general public. These activities will be scheduled when the project reaches a mature state. This moment is anticipated at the end of the formal project duration, approximately 5 month from now.

8 Conclusions and Outlook

8.1 Conclusions

A new system has been developed to diagnose and forecast air quality over the Netherlands. Novel aspects are the real time assimilation of surface ozone measurements from the Netherlands and neighbouring countries, and of tropospheric nitrogen dioxide column measurements from space.

Tested on a historical dataset, but in a procedure completely similar to a real forecast, the system has shown its capability to generate hourly prognostic maps of air quality over the Netherlands, providing the spatial and temporal resolution desired.

Compared to the current operational forecast tool PROZON, the new system provides better temporal behaviour, especially at the beginning and end of smog episodes, which are hard to predict by PROZON.

Compared to the current operational forecast tool PROZON, the new system in its current status of development is less able to predict high concentration levels correctly. Concentration peaks larger than 150 $\mu g/m^3$ are at this moment substantially underestimated by the new system. This underestimation is confirmed by comparison to actual surface ozone observations. As the relevance of smog forecasting is closely linked to predicting correctly the crossing of the European information and alarm thresholds of 180 and 240 $\mu g/m^3$, this is a major flaw.

In the current assimilation scheme, four model parameters were adjusted by assimilation: emissions of NO_x, emissions of VOC's, ozone concentration at the model's top boundary, and ozone dry deposition. Only the NO_x adjustments turn out to have substantial impact on the model output. For the other three parameters sensitivity is substantially less. This is unexpected, since VOC's and NO_x are both drivers in ozone formation. Indications are that this is caused by the current chemistry scheme.

The new system but without its assimilation part has shown its capability to run on a quasi-operational basis for many months. It provides a daily data stream to the RIVM smog team and the web, including one and two day forecasts, hindsight comparison to surface measurements, and a movie showing hourly maps of ozone from yesterday till the day after tomorrow.

The new system *including assimilation* has not been tested yet in a real time forecast situation. For the assimilation of surface ozone measurements, only the data from the Dutch monitoring network are available in real time on a solid basis. Some modifications at the EEA database are necessary to make available data from other countries in real time. This limitation is expected to be resolved in the coming months.

For the assimilation of OMI tropospheric columns of NO_2, the data stream is in place and the first forecasts based on historic data have been generated. These results are very recent and need to be analysed further. Indications are, that some further work is needed on the model before a meaningful forecast can be generated. Assimilations *combining* groundbased and space data in one assimilation scheme have not been performed yet, but are planned.

A cost benefit analysis has shown that the additional costs of implementation and operation of the new system are small with respect to the annual smog-related healthcare costs. On the benefit side, the system is much better suited to communicate potential adverse health conditions to non-experts and the general public. Only a minor fraction of the population needs to change its exposure by changed behaviour to reduce healthcare costs by a similar amount. It is therefore anticipated that from a societal point of view implementation of SmogProg will be cost-effective financially even without capitalising the gain in health and well-being.

8.2 Outlook

Conclusions above clearly show that, while results are promising, the development is not complete yet. The SmogProg proposal was originally written for a two-year period. This report shows the progress made in the first 19 months (Jan 2007-July 2008). The project will continue according to its original planning for an additional 5 months, and the project partners will continue their work on the following issues:

- The implementation of a solid real time data flow from EU-wide surface ozone measurements. This will be facilitated by the European Environmental Agency. Member states will be obliged to provide these data in real time under a new EU-directive.
- Improvements to the SmogProg model LOTOS-EUROS. In any case the model will be adapted to better facilitate the assimilation of OMI data. Other options will be studied to improve the model's absolute response to peak values: a better chemistry scheme, improved biogenic emissions and deposition modules. Other options are: a bias correction for surface ozone in the assimilation of ground based monitoring data, or the combination of statistical and physical modelling.
- Further development of the (web) interface towards the smog team, other civil authorities and the public.
- When time allows: assimilation of both groundbased and space data in one assimilation scheme.

This report will be updated in approximately six months time, and report on this additional work. When the steps above can be completed successfully, the resulting system will be compliant with the user requirements as set out by the RIVM Smogteam. It will then be presented to the stakeholders VROM, RIVM, PBL, other civil authorities like provinces, and NGO's.

In anticipation of these positive results, project partners have already started the implementation of the SmogProg forecast model in the operational model environment of KNMI, and the development of the operational dataflow from this model towards the smog team at RIVM. The current schedule aims for an operational SmogProg forecast to run starting June 2009.

Long term perspective: GMES, MACC, TROPOMI

In a broader scope and somewhat longer timescale, the efforts to combine modelling with ground based and space borne observations will be further developed in the European

framework of the GMES[8] program, in the MACC[9] project that is the implementation of the GMES Atmosphere core service for Europe, and in the GMES downstream services for atmosphere. A very active role for the Netherlands is anticipated here, given its national focus on atmosphere in GMES, and its leading role in the satellite instruments OMI and TROPOMI. Data from these satellite instruments form a cornerstone of this approach, and vice versa this approach is an essential element in the interpretation and harvesting of satellite observations of air quality.
RIVM, TNO and KNMI all plan to contribute to these developments in the coming years, and have set aside strategic research funds to participate. KNMI and TNO already participate in MACC, while RIVM is very active in GMES coordinating bodies. Several research proposals are currently being formulated from the SmogProg research team, aiming for funding through Dutch and European Science programs. It is therefore likely that also after the 5 month period mentioned above, the further development of SmogProg will continue.

---

[8] The EU program "Global Monitoring for Environment and Security", recently rebaptised to "Kopernikus"
[9] The EU 7[th] framework project "Monitoring Atmospheric Composition and Climate"

9 References

Acarreta, J. R., J. F. De Haan, and P. Stammes, Cloud pressure retrieval using the O2−O2 absorption band at 477 nm, J. Geophys. Res., 109, D05204, doi:10.1029/2003JD003915

De Hollander, A.E.M., Assessing and evaluating the health impact of environmental exposures. PhD Thesis, Utrecht University, 2004

Dirksen, R., H. Eskes, F. Boersma, P. Levelt, P. Veefkind, R. van der A, Derivation of Ozone Monitoring Instrument tropospheric NO2 in near-real time (DOMINO). NIVR final report, 2008

Evensen, G., Advanced data assimilation for strongly nonlinear dynamics, *Mon. Weather Rev.*, 125, 1342-1354, 1997

Fischer, P.H., B. Brunekreef and E. Lebret, Air pollution related deaths during the 2003 heat wave in the Netherlands Atmos. Environment, 38, 2004, doi:10.1016/j.atmosenv.2003.11.010

Logan, J., An analysis of ozonesonde data for the troposphere, recommendations for testing 3-D models and development of a gridded climatology for tropospheric ozone, *J. Geophys. Res. 104, 16,* 1998.

Manders, A.M.M, L. Nguyen and R. Hoogerbrugge, Evaluation of the RIVM forecasting models for ozone and particulate matter, PROZON and PROPART, (in Dutch) RIVM report no 680704004, 2008

MNP, CBS and WUR, Healt effects of particulate matter and ozone, (Gezondheidseffecten van fijn stof en ozon, in Dutch). www.milieuennatuurcompendium.nl (v07, 19 nov 2007). MNP, Bilthoven, CBS, Voorburg en WUR, Wageningen.

Noordijk, H., PROZON en PROPART; statistische modellen voor smorgprognose, RIVM report no 724301012, 2003

Palacios, M., F. Kirchner, A. Martilli, A. Clappier, F. Martín, M.E. Rodríguez Summer ozone episodes in the Greater Madrid area. Analyzing the ozone response to abatement strategies by modelling,. Atmos. Env. 36, 2002

Schaap, M., M. Roemer, F. Sauter, G. Boersen, R. Timmermans, P.J.H. Builtjes, LOTOS-EUROS: Documentation, TNO-report, B&O-A R 2005/297, 2005.

Schaap, M., Timmermans, R.M.A., Roemer, M., Boersen, G.A.C., Builtjes, P.J.H. Sauter, F.J., Velders, G.J.M. and Beck, J.P., 'The LOTOS–EUROS model: description, validation and latest developments', *Int. J. Environment and Pollution*, Vol. 32, No. 2, pp.270–290, 2008.

SmogProg project website, http://www.lml.rivm.nl/data/smogprog/index.html.

SmogProg development website, http://www.temis.nl/luchtkwaliteit/.

Smogregeling 2001 Government Gazette 11, June 2001 no. 109 / page 16, 1996/62/EC Official Journal of the European Union, 27 September 1996, no. L 296/ page 55; 1999/30/EC Official Journal of the European Union, 22 April 1999, no. L 163/ page 41; 2002/3/EC Official Journal of the European Union, 9 March 2002, no. L 67/ page 14

Smeets, C.J.P.P. and J. P. Beck, Effects of short-term abatement measures on peak ozone concentrations during summer smog episodes in the Netherlands, RIVM report 725501004, 2001

Torfs,R., L. Int Panis, L. De Nocker and S. Vermoote, External costs of ozone concentrations and abatements (in Dutch),proceedings van de studiedag ozon op leefniveau en ozonprecursoren, wetenschappelijke instrumenten en beleid, Brussels, June 10 2004.

Van Loon, M., R. Vautard, M. Schaap, R. Bergström, B. Bessagnet, J. Brandt, P.J.H. Builtjes, J. H. Christensen, K. Cuvelier, A. Graf, J.E. Jonson, M. Krol, J. Langner, P. Roberts, L. Rouil, R. Stern, L. Tarrasón, P. Thunis, E. Vignati, L. White, P. Wind, Evaluation of long-term ozone simulations from seven regional air quality models and their ensemble average, Atmospheric Environment, 41 (10), 2083-2097, 2007.

Vautard, R., Builtjes, P. H. J., Thunis, P., Cuvelier, K., Bedogni, M., Bessagnet, B., Honoré, C., Moussiopoulos, N., Pirovano G., Schaap, M., Stern, R., Tarrason, L., Van Loon, M., Evaluation and intercomparison of Ozone and PM10 simulations by several chemistry-transport models over 4 european cities within the City-Delta project, Atmospheric Environment*, 41, 173-188, 2007.*

Visschedijk, A.J.H. and Denier van der Gon, H.A.C., Gridded European anthropogenic emission data for NOx, SOx, NMVOC, NH_3, CO, PPM10, PPM2.5 and CH_4 for the year 2000, TNO-Report B&O-A R 2005/106, 2005.

WHO(2003), Health Aspects of Air Pollution with Particulate Matter, Ozone and Nitrogen Dioxide. WHO, Bonn, Germany, http://www.euro.who.int/document/e79097.pdf

Appendix A. Measuring stations in The Netherlands, Belgium and Germany

In this appendix a map of the measuring stations that are used for data-assimilation in the validation study in Chapter 4 is shown in Figure 49.

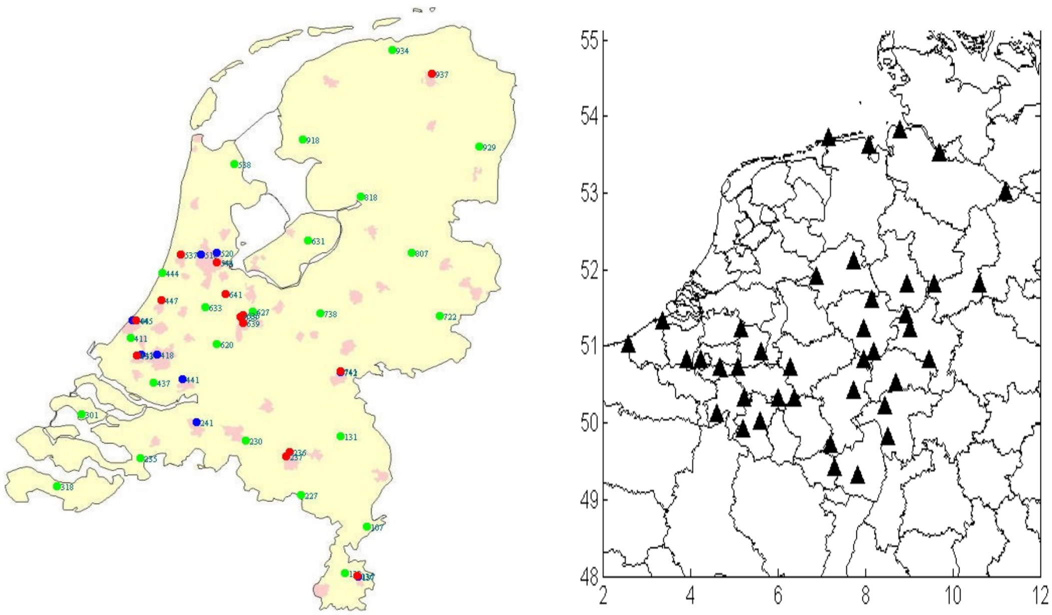


Figure 49. Assimilated rural stations in The Netherlands (left, green dots), Belgium and Germany (right).

Furthermore a list of these stations is given and an indication of which stations are used for assimilation and which are used for monitoring, during the test runs of Chapter 3.

GBNL: ρ=25 km
Assimilated
Posterholt-Vlodropperweg
Vredepeel-Vredeweg
Wijnandsrade-Opfergeltstraat
Budel-Toom
Biest Houtakker-Biestsestraat
Huijbergen-Vennekenstraat
Zierikzee-Lange Slikweg
Philippine-Stelleweg
Schipluiden-Groeneveld
Westmaas-Groeneweg
De Zilk-Vogelaarsdreef
Wieringerwerf-Medemblikkerweg
Cabauw-Zijdeweg
Biddinghuizen-Hoekwantweg
Zegveld-Oude Meije
Eibergen-Lintveldseweg
Wekerom-Riemterdijk
Hellendoorn-Luttenbergerweg
Barsbeek-De Veenen
Balk-Trophornsterweg
Valthermond-Noorderdiep
Kollumerwaard-Hooge Zuidwal

GBNL: ρ=50 km
Assimilated
Posterholt-Vlodropperweg
Philippine-Stelleweg
De Zilk-Vogelaarsdreef
Wieringerwerf-Medemblikkerweg
Cabauw-Zijdeweg
Eibergen-Lintveldseweg
Wekerom-Riemterdijk
Hellendoorn-Luttenbergerweg
Balk-Trophornsterweg
Valthermond-Noorderdiep
Kollumerwaard-Hooge Zuidwal
Philippine-Stelleweg

Monitored
Vredepeel-Vredeweg
Wijnandsrade-Opfergeltstraat
Budel-Toom
Biest Houtakker-Biestsestraat
Huijbergen-Vennekenstraat
Zierikzee-Lange Slikweg
Schipluiden-Groeneveld
Westmaas-Groeneweg
Biddinghuizen-Hoekwantweg
Zegveld-Oude Meije
Barsbeek-De Veenen

GBNL: ρ=100 km
Assimilated
Posterholt-Vlodropperweg
Philippine-Stelleweg
De Zilk-Vogelaarsdreef
Eibergen-Lintveldseweg
Kollumerwaard-Hooge Zuidwal

Monitored
Vredepeel-Vredeweg
Wijnandsrade-Opfergeltstraat
Budel-Toom
Biest Houtakker-Biestsestraat
Huijbergen-Vennekenstraat
Zierikzee-Lange Slikweg
Schipluiden-Groeneveld
Westmaas-Groeneweg
Wieringerwerf-Medemblikkerweg
Cabauw-Zijdeweg
Biddinghuizen-Hoekwantweg
Zegveld-Oude Meije
Wekerom-Riemterdijk
Hellendoorn-Luttenbergerweg
Barsbeek-De Veenen
Balk-Trophornsterweg
Valthermond-Noorderdiep

GBEU: 25 km
Assimilated
Posterholt-Vlodropperweg
Vredepeel-Vredeweg
Wijnandsrade-Opfergeltstraat
Budel-Toom
Biest Houtakker-Biestsestraat
Huijbergen-Vennekenstraat
Zierikzee-Lange Slikweg
Philippine-Stelleweg
Schipluiden-Groeneveld
Westmaas-Groeneweg
De Zilk-Vogelaarsdreef
Wieringerwerf-Medemblikkerweg
Cabauw-Zijdeweg
Biddinghuizen-Hoekwantweg
Zegveld-Oude Meije
Eibergen-Lintveldseweg
Wekerom-Riemterdijk
Hellendoorn-Luttenbergerweg
Barsbeek-De Veenen
Balk-Trophornsterweg
Valthermond-Noorderdiep
Kollumerwaard-Hooge Zuidwal
DE0735A
DE1303A
DE1004A
DE1305A
DE0674A
DE0874A
DE0680A
DE0732A
DE0692A
DE0996A
DE1294A
DE1295A
DE1296A
DE1299A
DE0685A
DE0688A
DE0737A
DE0686A
DE1122A
DE0687A
DE1172A
DE1392A
DE1634A
DE1253A
DE0649A
DE0422A
BE0238A
BE0300A
BE0295A
BE0298A
BE0211A
BE0294A
BE0033R
BE0311A
BE0345A
BE0301A
BE0198A
BE0291A
BE0304A

GBEU 50 km

Assimilated

Posterholt-Vlodropperweg
Philippine-Stelleweg
De Zilk-Vogelaarsdreef
Wieringerwerf-Medemblikkerweg
Cabauw-Zijdeweg
Eibergen-Lintveldseweg
Wekerom-Riemterdijk
Hellendoorn-Luttenbergerweg
Balk-Trophornsterweg
Valthermond-Noorderdiep
Kollumerwaard-Hooge Zuidwal
Philippine-Stelleweg
DE0735A
DE1303A
DE1004A
DE1305A
DE0674A
DE0874A
DE0680A
DE0732A
DE0692A
DE0996A
DE1294A
DE1295A
DE1296A
DE1299A
DE0685A
DE0688A
DE0737A
DE0686A
DE1122A
DE0687A
DE1172A
DE1392A
DE1634A
DE1253A
DE0649A
DE0422A
BE0238A
BE0300A
BE0295A
BE0298A
BE0211A
BE0294A
BE0033R
BE0311A
BE0345A
BE0301A
BE0198A
BE0291A
BE0304A

Monitored

Vredepeel-Vredeweg
Wijnandsrade-Opfergeltstraat
Budel-Toom
Biest Houtakker-Biestsestraat
Huijbergen-Vennekenstraat
Zierikzee-Lange Slikweg
Schipluiden-Groeneveld
Westmaas-Groeneweg
Biddinghuizen-Hoekwantweg
Zegveld-Oude Meije
Barsbeek-De Veenen

GBEU 100 km

Assimilated

Posterholt-Vlodropperweg
Philippine-Stelleweg
De Zilk-Vogelaarsdreef
Eibergen-Lintveldseweg
Kollumerwaard-Hooge Zuidwal
Posterholt-Vlodropperweg
Philippine-Stelleweg
De Zilk-Vogelaarsdreef
Eibergen-Lintveldseweg
Kollumerwaard-Hooge Zuidwal
DE0735A
DE1303A
DE1004A
DE1305A
DE0674A
DE0874A
DE0680A
DE0732A
DE0692A
DE0996A
DE1294A
DE1295A
DE1296A
DE1299A
DE0685A
DE0688A
DE0737A
DE0686A
DE1122A
DE0687A
DE1172A
DE1392A
DE1634A
DE1253A
DE0649A
DE0422A
BE0238A
BE0300A
BE0295A
BE0298A
BE0211A
BE0294A
BE0033R
BE0311A
BE0345A
BE0301A
BE0198A
BE0291A
BE0304A

Monitored:

Vredepeel-Vredeweg
Wijnandsrade-Opfergeltstraat
Budel-Toom
Biest Houtakker-Biestsestraat
Huijbergen-Vennekenstraat
Zierikzee-Lange Slikweg
Schipluiden-Groeneveld
Westmaas-Groeneweg
Wieringerwerf-Medemblikkerweg
Cabauw-Zijdeweg
Biddinghuizen-Hoekwantweg
Zegveld-Oude Meije
Wekerom-Riemterdijk
Hellendoorn-Luttenbergerweg
Barsbeek-De Veenen
Balk-Trophornsterweg
Valthermond-Noorderdiep

Appendix B. Programma van eisen SmogProg

11 februari 2008

9.1.1 Inleiding

Het project SmogProg heeft als doel een continu operationeel product op te leveren, dat kwalitatief beter is dan de huidige statistische smogverwachting en welke de huidige smogverwachting kan vervangen. Verbeteringen worden verwacht op het gebied van (i) het voorspellen van het begin en eind van een smog-episode, (ii) geografisch detail en (iii) communicatie van smog-episodes naar niet-experts en het algemene publiek. Dit document beschrijft de eisen die de gebruikers stellen aan het product.
De eisen vloeien voort uit wetgeving, praktische invulling en mogelijke toekomstige veranderingen. De wetgeving is vastgelegd in de smogregeling 2001 (Staatcourant 11 juni 2001 nr. 109 / pag. 16). De smogregeling 2001 is een uitwerking van regels over de informatievoorziening van periodes van smog aan het publiek welke vastgelegd zijn in de Europese richtlijnen. De praktische zaken komen voort uit de ervaringen van medewerkers met de huidige smogmodellen.

9.1.2 Belangrijke punten uit de smogregeling en overige wetgeving voor SmogProg

In de smogregeling 2001 is vastgelegd wat verschillende instanties en overheden moeten uitvoeren in geval van smog. Hieruit is af te leiden wat de minimum eisen zijn voor SmogProg. Echter, de smogregeling is gemaakt vooruitlopend op de definitieve Europese en de daaruit voortvloeiende Nederlandse wetgeving. Als de regels naast elkaar worden gelegd kan er een verschil van mening zijn over de te volgen koers. Ook heeft VROM aangegeven om samen met de partners de smogregeling in 2008 te evalueren[10]. Dit kan tot gevolg hebben dat de smogregeling op punten gewijzigd gaat worden.
In de volgende paragrafen worden eerst de belangrijke punten uit de smogregeling behandeld. Hierna volgt aanvullende informatie met betrekking tot overige wetgeving, waaronder knelpunten. Vervolgens zal er worden geprobeerd om alvast inzicht te geven in de vragen die mogelijk aan de orde komen in de evaluatie in 2008 en die van invloed zijn op SmogProg. Tot slot wordt er een overzicht gegeven van alle eisen en wensen die de eindgebruiker (het RIVM smogteam) stelt aan het eindproduct.

Indicatoren
Er zijn vier indicatoren voor smog: ozon, fijn stof, zwaveldioxide en stikstofdioxide. De onderstaande tabel geeft de smogklassen voor deze stoffen weer. In Nederland zijn drie smogklassen gedefinieerd: geen of geringe smog, matige smog en ernstige smog. De grenzen van de smogklassen zijn vastgesteld aan de hand van grenswaarden, streefwaarden en informatie- en alarmdrempels uit de Europese richtlijnen. Een uitzondering hierop is de alarmdrempel voor PM10 van 200 $\mu g/m^3$. Deze waarde bestaat niet in de Europese regels en is gevormd om een uniforme smogregeling te maken, waarbij voor elke smogindicator een min of meer gelijk criterium geldt. Binnen het SmogProg project wordt slechts gewerkt aan de voorspelling van ozon en stikstofdioxide.

[10] Deze evaluatie is vertraagd en staat inmiddels gepland voor 2009.

Tabel 1: Indeling smog in drie klassen, concentraties in µg/m^3

	geen of geringe smog	**matige smog**	**ernstige smog**
Ozon (uurgemiddelde)	< 180	180 - 240	> 240
Fijn stof (daggemiddelde)	< 50	50 - 200	> 200
Zwaveldioxide (uurgemiddelde)	< 350	350 - 500	> 500 *
Stikstofdioxide (uurgemiddelde)	< 200	200 - 400	> 400 *

* overschrijding van de uurgemiddelde concentratie gedurende drie opeenvolgende uren

Eisen gesteld aan de verwachting van ozon in de smogregeling
Voor ozon zijn de grenzen voor matige en ernstige smog gekoppeld aan de informatiedrempel (180 µg/m^3) en de alarmdrempel (240 µg/m^3). In de smogregeling 2001 is besloten om voor smog door ozon alleen actieve informatievoorziening te starten bij ernstige smog.
Tijdens smogepisodes moet er **twee** maal per dag een update worden gegeven van de smogsituatie. Hierbij wordt gevraagd om ook een prognose te geven. De volgende gegevens moeten hierbij worden gegeven:

- De ontwikkeling van de concentraties;
- Het betrokken gebied;
- Duur van de overschrijding.

In de smogregeling (en overige wetgeving) is **niets** vastgelegd over het tijdstip van de update. Op dit moment worden rond 9:30u en 16:30u de smogverwachtingen richting het publiek ververst. De modeluitdraai is echter een uur eerder bekend bij het RIVM smogteam. Zo kunnen zij alvast inspelen op de situatie en eventueel de prognose bijstellen.

Teletekst
In de smogregeling is vastgelegd dat de smogverwachting gepresenteerd moet worden op teletekst. Voor het nieuw te ontwikkelen product houdt dit twee dingen in. Ten eerste zal het product geschikt moeten zijn om te presenteren op teletekst. Ten tweede zal het nieuwe product om een andere presentatie vragen dan het huidige. Dit houdt in dat de teletekst pagina gewijzigd moet worden. Hiervoor is toestemming nodig van de NOS. Voor deze toestemming worden vooralsnog geen problemen voorzien.

Zaken uit overige wetgeving.
Naast de smogregeling 2001 kent de Nederlandse wet ook de Regeling luchtkwaliteit ozon (Staatscourant 19 november 2004, nr. 224 / pag. 18). Dit is de Nederlandse uitwerking van de Europese 3^e dochterrichtlijn 2002/3/EG betreffende ozon in lucht. In de Regeling luchtkwaliteit ozon staat duidelijker dan in de smogregeling 2001 aangegeven dat de waarschuwing aan het publiek gebeurt op basis van uurgemiddelde concentraties. Hierbij moet er minimaal inzicht kunnen worden gegeven over de concentraties van de volgende middag, maar liever meer dagen vooruit. Naast een uurgemiddelde concentratie wordt er in de Regeling luchtkwaliteit ozon ook gesproken over een 8-uursgemiddelde. Het 8-uursgemiddelde is gedefinieerd als een voortschrijdend gemiddelde. Hierbij is het eerste 8-uursgemiddelde van dag x het gemiddelde over 17:00u van dag x-1 tot 1:00u van dag x. Het laatste 8-uursgemiddelde is van 16:00u tot 24:00u van dag x. Deze 8-uursgemiddelde is een belangrijke parameter bij de beoordeling van de bescherming van de volksgezondheid. De 8-uursgemiddelde concentratie mag in 2010 maximaal vijfentwintig dagen per kalender jaar gemiddeld over drie jaar hoger zijn dan 120 µg/m^3. De langetermijndoelstelling van 2020 schrijft voor dat een concentratie van 120 µg/m^3 helemaal niet meer overschreden mag

worden. Tijdens een smogepisode moet in een bericht de actuele 8-uursgemiddelde concentratie worden opgenomen. Over dit gemiddelde hoeft echter geen prognose te worden gegeven. Toch is het wenselijk om ook hier inzicht in te hebben.
Daarnaast worden er eisen gesteld aan de representativiteit van een meetstation. Voor een regionaal station geldt dat de representativiteit van het station zich moet uitstrekken tot enkele honderden vierkante kilometers. Bij voorstadstations is dit enkele tientallen vierkante kilometers. Voor de smogverwachting is een grid van 10 bij 10 km afdoende om bij deze eisen aan te sluiten. De aanzuighoogte van de opstelling moet zich bevinden tussen 1,5 en 4 meter. De aanzuighoogte van de meethutten van het Landelijk Meetnet Luchtkwaliteit zijn – uitzonderingen daargelaten – ingesteld op 3,5 meter hoogte. De verwachte maximale ozonconcentratie zou op die hoogte bepaald moeten worden. Tot slot bevat de 3^{e} dochterrichtlijn een minimum kwaliteitseis van 50% onzekerheid in uurgemiddelde concentraties indien er gebruik gemaakt wordt van modellen.

Toekomstige acties
In 2008 zullen in ieder geval twee zaken worden bekeken.

- De huidige indeling van zones en agglomeraties zal worden geëvalueerd en mogelijk worden aangepast.
- De smogregeling zal worden geëvalueerd.

Dat de smogregeling wordt geëvalueerd zal voor de component ozon weinig gevolgen hebben. De protocollen liggen hierbij redelijk vast. Hooguit kan het tijdstip van melding veranderen. Als er in 2008 besloten wordt voor een andere indeling in zones en agglomeraties dan zal dat voor de modelberekeningen zelf weinig invloed hebben. Het kan echter wel leiden tot aanpassing van het eindproduct.

9.1.3 Eisen

Verwachte maximale ozonconcentratie

a. een kaart van Nederland met daarop de verwachte maximale uurconcentratie ozon van die dag per pixel op 3,5m hoogte.
b. een kaart van Nederland met daarop de verwachte maximale uurconcentratie ozon van de dag volgend op de dag van de modelrun per pixel op 3,5m hoogte.
c. Een animatie met het verwachte verloop in uurconcentraties ozon voor de komende 24 uur.
d. Een bestand met daarin de maximale uurconcentratie ozon per pixel voor de dag van de modelrun.
e. Een bestand met daarin de maximale uurconcentratie ozon per pixel voor de dag volgend op de dag van de modelrun.

De **resolutie** van de resultaten moet van de orde 10 x 10 km zijn. Voor de vertaling naar teletekst is eventueel een grover grid nodig.
In de 3^{e} dochterrichtlijn wordt een minimale eis aan de onzekerheid van 50% gesteld. Een verdere uitwerking van de nauwkeurigheid zal op een later tijdstip volgen in gezamenlijk overleg tussen het RIVM smogteam en het SmogProg consortium.

Smog houdt niet op bij de grens van Nederland. In de kaarten en animatie van de resultaten moet ook een groot gebied van Europa zichtbaar zijn. Te denken valt daarbij aan een gebied van 2000 bij 2000 km met Nederland als middelpunt. Deze grenzen zijn louter indicatief. Daarnaast moet er de mogelijkheid zijn om ook een kaart te tonen ingezoomd op Nederland.

Tijdsgemiddelde waarden
Voor ozon moet er voor de producten a t/m e een verwachting van de uurwaarde komen

Animatie met uurverwachtingen
Er moet een animatie kunnen worden getoond, waarbij de ontwikkeling in de komende 24 uur kan worden getoond.

Kleurovergangen grenswaarden
De kleurovergangen voor de verschillende grenswaarden moeten duidelijk zichtbaar zijn. De grenswaarden voor de ozonpiekwaarde zijn: 180 $\mu g/m^3$ voor matige smog en 240 $\mu g/m^3$ voor ernstige smog. De grenswaarde voor het ozon 8-uurs gemiddelde is 120 $\mu g/m^3$.

Update van de prognose
Er moet minimaal twee maal op een dag een update komen van de prognose.

Vertaling resultaat naar teletekst, persberichten, e-mails en SMS (producten d en e)
Op dit moment krijgen de gebruikers een SMS-bericht over de maximale waarden per zone en een e-mailbericht over de maximale waarde per dag in zones en agglomeraties. Deze maximale waarden worden ook gebruikt in de presentatie naar teletekst. Er moet een file komen waaruit de SMS-, e-mail- en teletekst-berichten samengesteld kunnen worden. Daarnaast moet het product geschikt zijn om te verwerken in een persbericht.

9.1.4 Wensen

8-uursgemiddelden
Voor ozon is een verwachting van de 8-uursgemiddelde waarde wenselijk voor de producten a t/m c. Daarnaast is het prettig als de verwachting de ontwikkeling in de komende 48 uur toont en als deze kan worden gecombineerd met animatie van actuele concentraties. Dus actuele concentraties gaan over in een verwachting.

Opdeling naar provincies
Een aantal provincies is erg geïnteresseerd in de smogverwachting voor hun provincie. Het is daarmee wenselijk om, naast een kaart van Nederland, ook een opdeling te kunnen maken naar provincies.

Tijdstip van draaien
Bij voorkeur komt er een update om 8.30u en om 15.30u. De presentatie naar internet vindt een uur later plaats. Dit geeft het smogteam de gelegenheid om de uitkomsten te beoordelen op betrouwbaarheid (zie ook het volgende punt). Een update om 5.00u UTC (dus om 6.00u of 7.00u lokale tijd) en een update om 11.00u UTC (dus om 12.00u of 13.00u lokale tijd) sluit hier voldoende bij aan.

Controle door medewerkers van het RIVM smogteam
Bij de prognose moet de medewerker van het RIVM smogteam de mogelijkheid hebben om commentaar te plaatsen.

Wijzigen smoggrenzen
Er is een kans dat de informatie- en alarmdrempels, die de overgang van de ene naar de andere smogklasse definiëren, in de toekomst gaan veranderen. Concrete voorbeelden liggen niet bij ozon, maar denk hierbij aan PM10 (elk land gebruikt een andere grens om de ernst van de smog weer te geven) en aan SO_2 (de WHO kwam met een voorstel om de grenswaarden voor SO_2 naar beneden bij te stellen). De beheerders van het model moeten de grenswaarden eenvoudig kunnen wijzigen.

Station wel/niet meenemen in data-assimilatie
Het model moet in staat zijn om uitbijters in een dataset te kunnen detecteren en deze niet mee te nemen in de data-assimilatie. Het criterium hiervoor wordt vastgesteld in overleg met het RIVM smogteam. Daarnaast moet het mogelijk zijn om data van buitenlandse stations op te nemen in de data-assimilatie.

Betrouwbaarheidsfactor
Er is de wens om de betrouwbaarheid van een modelrun weer te geven met een parameter. Zo kan de betrouwbaar van een modelrun worden ingeschat. Deze parameter zal samengesteld worden door het RIVM smogteam en het SmogProg consortium in gezamenlijk overleg. Hierbij wordt opgemerkt dat het opnemen van zo'n betrouwbaarheidsfactor erg veel werk met zich mee brengt en waarschijnlijk niet gerealiseerd kan worden binnen het SmogProg project.

Validatierapport
Het product dient getest te worden aan de eisen in dit verslag. De belangrijkste criteria hierbij zijn het correct voorspellen van het begintijdstip en de duur van een smogperiode, evenals de maximumconcentratie per regio gedurende een smogperiode. De resultaten van deze tests worden gepresenteerd in een zogeheten validatierapport.

Beheer
Voor de duur van het project ligt het beheer van het model bij het SmogProg consortium. Wanneer de in dit project ontwikkelde smogverwachting de huidige RIVM smogverwachting vervangt zullen er nadere afspraken over beheer gemaakt worden.

Handleiding
Bij het model moet een handleiding komen waarin het gebruik van het model wordt uitgelegd.

Toekomst
Het is de wens om het product in de toekomst uit te breiden naar de smogindicatoren SO_2 en PM10 en wellicht ook PM2.5. De resolutie voor deze nieuwe componenten moet in de buurt van bronnen (b.v. wegen) misschien wel omhoog (verdichting). Het kan zijn dat de afstanden dan kleiner dan een kilometer worden.
Voor NO_2 en SO_2 moet een verwachting komen van uurgemiddelde waarden. Voor PM10 en PM2.5 gaat het om een verwachting van 24-uurgemiddelde waarden.

Appendix C. Flyer

Verwacht eindresultaat

Het verwachte eindresultaat is een operationeel product dat de huidige statistische smogverwachting kan vervangen. Het nieuwe product zal beter in staat zijn het begin en eind van een smog-episode te voorspellen, geeft meer detail in ruimte en tijd en is intuïtief inzichtelijk voor het algemene publiek.

Bredere toepassing

Het SmogProg product heeft een bredere toepassing dan alleen de smogverwachtingen. Het product levert diagnostische en prognostische informatie over de luchtkwaliteit in Nederland, ook bij geringe en matige smog. Deze informatie kan gebruikt worden voor milieuvraagstukken op het gebied van klimaat en luchtkwaliteitsonderzoek. Het SmogProg product kan een belangrijke toevoeging zijn aan deze gebieden van onderzoek. Het levert nieuwe informatie door informatie van verschillende bronnen te combineren

Samenhang met Europese ontwikkelingen

De ontwikkeling van de smogverwachting zal aansluiten bij vergelijkbare ontwikkelingen in Europa, in het bijzonder in het ESA-GMES project PROMOTE, het EU FP6 GMES project GEMS en het EU FP7 GMES project MACC. De ontwikkeling van het SmogProg product en de beoordeling van de kwaliteit zal deels gebeuren in de context van het PROMOTE project.

Website:
http://www.lml.rivm.nl/data/smogprog/index.html

Contact: smogprog@rivm.nl

Het SmogProg project (Jan 2007 - Dec 2008) wordt gefinancierd door het NIVR, in het kader van het Nationaal Programma Gebruikersondersteuning (GO) 2005 (Project Nr. GO-2005/84).

Foto voorkant: J.P. van Geuns

RIVM
Rijksinstituut voor Volksgezondheid en Milieu

Laboratorium voor Milieumetingen

Postbus 1
3720 BA Bilthoven
www.rivm.nl

SmogProg:
Naar een operationele smogverwachting gebaseerd op grondmetingen en satellietmetingen

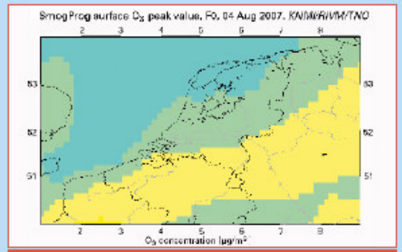
Wat is SmogProg?

SmogProg is een gezamenlijk project van RIVM, KNMI en TNO. In dit project worden grondmetingen en satellietmetingen gecombineerd met een computermodel om te komen tot een smogverwachting. Doel is om de huidige -statistische- smogverwachting van het RIVM te verbeteren.

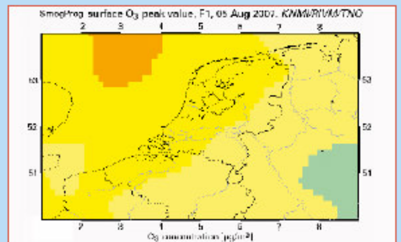
Achtergrond

Het RIVM heeft de wettelijke taak om dagelijks –en gedurende smog-episodes twee maal per dag- een smogverwachting te geven. De huidige smogverwachting is gebaseerd op een statistisch model. Dit model werkt goed. Echter, de huidige satellietmetingen, gecombineerd met nieuwe modellen geven aanleiding tot verbetering op een drietal punten. Ten eerste het voorspellen van het begin en eind van een smog-episode. Ten tweede het verbeteren van het geografische detail en tot slot het beter voorspellen van het verloop in de tijd. De laatste twee punten maken een betere presentatie naar het algemene publiek mogelijk.

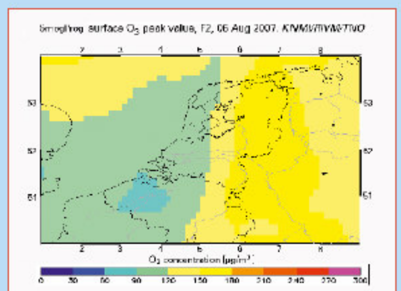
Verwachting ozon piekwaarde



Verwachting vandaag



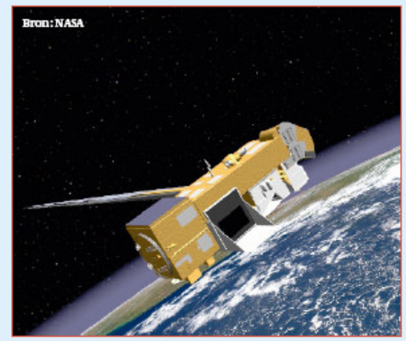
Verwachting morgen



Verwachting overmorgen

Hoe?

Voor het SmogProg product worden grondmetingen en satellietmetingen gecombineerd (geassimileerd) met een regionaal chemisch transportmodel. Het model gebruikt verder meteorologische gegevens zoals wind, temperatuur en neerslag en bevat daarnaast ook informatie over vervuilingsbronnen. Het SmogProg product levert vervolgens tweedaagse verwachtingen van de luchtkwaliteit in Nederland op uurbasis.



Producten

Binnen het huidige project wordt gekeken naar de smogindicatoren ozon en stikstofdioxide. Het SmogProg product zal kaarten leveren met de uurlijkse verwachting voor de verdelingen van ozon- en stikstofdioxideconcentraties boven Nederland. Deze gegevens zullen gebruikt worden voor smogwaarschuwingen en het dagelijks communiceren van de verwachtingen naar overheden en het publiek.