**Geoscientific
Model Development**

Assimilation of OMI NO$_2$ retrievals into the limited-area chemistry-transport model DEHM (V2009.0) with a 3-D OI algorithm

J. D. Silver[1], **J. Brandt**[1], **M. Hvidberg**[2], **J. Frydendall**[3], and **J. H. Christensen**[1]

[1]Department of Environmental Science, Aarhus University, Frederiksborgvej 399, 4000 Roskilde, Denmark
[2]National Survey and Cadastre, Danish Ministry of the Environment, Rentemestervej 8, 2400 Copenhagen NV, Denmark
[3]Institute of Mathematical Modelling, Danish Technical University, Richard Petersens Plads, Building 321, 2800 Lyngby, Denmark

Correspondence to: J. D. Silver (jds@dmu.dk)

Abstract. Data assimilation is the process of combining real-world observations with a modelled geophysical field. The increasing abundance of satellite retrievals of atmospheric trace gases makes chemical data assimilation an increasingly viable method for deriving more accurate analysed fields and initial conditions for air quality forecasts.

We implemented a three-dimensional optimal interpolation (OI) scheme to assimilate retrievals of NO$_2$ tropospheric columns from the Ozone Monitoring Instrument into the Danish Eulerian Hemispheric Model (DEHM, version V2009.0), a three-dimensional, regional-scale, offline chemistry-transport model. The background error covariance matrix, **B**, was estimated based on differences in the NO$_2$ concentration field between paired simulations using different meteorological inputs. Background error correlations were modelled as non-separable, horizontally homogeneous and isotropic. Parameters were estimated for each month and for each hour to allow for seasonal and diurnal patterns in NO$_2$ concentrations.

Three experiments were run to compare the effects of observation thinning and the choice of observation errors. Model performance was assessed by comparing the analysed fields to an independent set of observations: ground-based measurements from European air-quality monitoring stations. The analysed NO$_2$ and O$_3$ concentrations were more accurate than those from a reference simulation without assimilation, with increased temporal correlation for both species. Thinning of satellite data and the use of constant observation errors yielded a better balance between the observed increments and the prescribed error covariances, with no appreciable degradation in the surface concentrations due to the observation thinning. Forecasts were also considered and these showed rather limited influence from the initial conditions once the effects of the diurnal cycle are accounted for.

The simple OI scheme was effective and computationally feasible in this context, where only a single species was assimilated, adjusting the three-dimensional field for this compound. Limitations of the assimilation scheme are discussed.

1 Introduction

Chemistry-transport models (CTMs) are widely used for forecasting air pollution, evaluating proposed emission reductions, studying chemical or physical processes, and assessing climate-scale effects and forcings related to atmospheric components (Jacobson, 2005). Modelled concentrations are often highly uncertain, and can be improved in several ways, such as better parameterisation of sub-grid scale processes, more accurate estimates of forcings at the lateral and lower boundary conditions, higher spatial resolution, higher order numerical methods, and more accurate initial conditions. Data assimilation (DA) involves estimating initial conditions by combining previous forecasts with recent observations (Kalnay, 2003).

In recent decades, satellite retrievals of atmospheric constituents have complemented observations from ground-based monitoring stations (Martin, 2008). Satellite retrievals provide concentration estimates for the total vertical column, for a partial column (e.g. the troposphere) or at a range of vertical levels, and they cover a far greater geographical range across the planet compared to ground-based measurements of surface concentrations. They therefore present great potential for use in "chemical DA" (i.e. DA for CTMs). For a comprehensive review of chemical DA, see Carmichael et al. (2008).

Optimal interpolation (OI) is the one of simplest DA algorithms currently applied to CTMs; it is based on a least-squares formulation of the DA problem. While the assumptions underpinning OI are relatively crude, this algorithm is simple to implement and may be computationally cheaper than other more sophisticated DA methods, provided that neither the number of observations nor the number of model variables is too large. In meteorology, OI has long been surpassed by variational or Kalman filtering methods (Kalnay, 2003), however it is still in use with chemistry-transport models. For example, Mok et al. (2008) used OI with a Gaussian puff model for sulfur dioxide over Lisbon, Portugal. Adhikary et al. (2008) and Matsui et al. (2004) applied OI to assimilate satellite retrievals of aerosol optical depth when modelling aerosol concentrations over Southeast Asia and the eastern United States, respectively.

In the case of off-line CTMs, a small perturbation in the initial conditions will typically decay as the simulation proceeds, mainly due to forcing from sources and sinks such as chemistry and emissions (Carmichael et al., 2008; Wu et al., 2008). Thus the quality of the initial conditions is less critical in air quality modelling than for numerical weather prediction (NWP) models, where perturbations tend to grow with time. In the case of short-lived chemical species, the duration of the initial perturbation may be quite brief (e.g. one day) and this limits the extent to which better initial conditions can improve forecasts. Chemical DA can, nonetheless, be used for historical re-analysis.

The conceptual and practical simplicity of OI makes the algorithm a reasonable starting point for use of DA in CTMs. Wu et al. (2008) compared four different DA methods (OI, two types of Kalman filter, and four-dimensional variational assimilation) applied to ozone forecasting. They demonstrated that OI, although a relatively simple method, was comparable in performance to the more advanced and computationally intensive variational and Kalman filter methods.

This study concerns the assimilation of satellite-derived estimates of tropospheric concentrations of nitrogen dioxide (NO$_2$). NO$_2$ plays an important role in atmospheric chemistry. In the stratosphere, it is involved in catalytic cycles that destroy ozone (O$_3$); in the troposphere NO$_2$ is a key O$_3$ precursor, especially in polluted urban environments (Seinfeld and Pandis, 2006, Chapters 5 and 6). NO and NO$_2$ interconvert and atmospheric lifetime of NO$_x$ varies from hours

to days at the surface to a couple of weeks in the upper troposphere (Seinfeld and Pandis, 2006, p. 224). There is also substantial seasonal variation; Schaub et al. (2007) estimated the lifetime of NO$_x$ to be around 3 h during summer and 13 h during winter. Thus NO$_2$ is a relatively "local" pollutant. It is also of interest from a human health perspective; for example, exposure to NO$_2$ has been linked to reduced lung function, asthma and increased mortality (reviewed by Searl, 2004).

In this study we make use of tropospheric NO$_2$ column concentrations, derived from measurements by the Ozone Monitoring Instrument (OMI) aboard the NASA satellite AURA (see Sect. 2.1). These NO$_2$ retrievals have been studied in a number of contexts. They have been used to re-estimate NO$_x$ emission rates (Zhao and Wang, 2009). They have been validated against ground-based measurements (Lamsal et al., 2008), spectrometers (Ionov et al., 2006), aircraft campaigns (Boersma et al., 2008). They are a resource for validation of air quality models (Huijnen et al., 2010) or comparison with retrievals from other satellites (Boersma et al., 2008). Furthermore, they have be used to study particular pollution or emission reduction events (Wang et al., 2007).

In this article, we assimilated retrieved tropospheric NO$_2$ columns from OMI into a limited-area CTM with a three-dimensional OI scheme. We describe how the background error covariance matrix was parameterised based using the difference between paired simulations (the "NMC method" of Parrish and Derber, 1992). We ran a number of simulations relating to different treatment of the observation error statistics. The analysed concentration fields are compared to ground-based observations of NO$_2$ concentrations (see Sect. 3.2). In Sect. 4, we discuss these results in the broader context of forecasting and chemical DA.

2 Assimilation and modelling framework

2.1 OMI retrievals

Tropospheric NO$_2$ concentrations were retrieved from radiances measured by the Dutch–Finnish Ozone Monitoring Instrument (OMI) aboard the NASA satellite Aura. Aura's orbit is sun-synchronous, crossing the equator between 13:30 and 14:00 local time, passing over Europe shortly after. The retrieval scheme is described in Boersma et al. (2002, 2007). Retrievals from pixels with a cloud radiance fraction in excess of 50 % were excluded or with a surface albedo greater than 0.3, as recommended by Boersma et al. (2011). We used the DOMINO version 2.0 (produced September 2010) of the level 2 retrieved tropospheric column NO$_2$ concentrations. The retrieval process yielded a measure of the estimate's uncertainty. The retrieved total column was found to be highly correlated with the associated uncertainty measure ($R^2 = 0.83$ for the year 2005 for retrievals within the DEHM

domain). Around 20 % of total column estimates were negative, and these were also included despite the non-physical nature of the quantity; excluding them would lead to a positive bias over regions with low NO$_x$ concentrations, and also because they represent uncertainty intrinsic in the retrieval process.

The OMI data represent a different temporal and spatial resolution compared to that of the CTM used in this study (see Sect. 2.2). The model domain covered Europe, and satellite readings in this area were available several times a day, usually few hours before and after 12:00 UTC. Multiple images are produced per satellite overpass. The spatial resolution of the CTM used in this study is approximately 50 km × 50 km across the model domain. Resolution of the OMI images varies across the camera's swath. Nadir pixels are 13 km × 24 km, while pixels furthest from nadir are 13 km × 128 km.

Finally, OMI data from the outermost (i.e. first and last) row and column of the CTM's 96 × 96 grid were not assimilated, thus providing a buffer for potential boundary effects (e.g. see Fig. 13).

2.2 Chemistry-transport model: DEHM

The Danish Eulerian Hemispheric Model (DEHM) is an off-line, Eulerian, three-dimensional, long-range CTM (Christensen, 1997; Frohn et al., 2002; Brandt et al., 2012). The model simulates atmospheric transport and diffusion, chemical transformations, wet and dry deposition, and emissions from a range of biogenic and anthropogenic sources.

Version V2009.0 of DEHM was used in this study: this is the version developed for the 2009 annual report for the Danish Air Quality Monitoring Programme (NOVANA; Ellermann et al., 2010). In the present configuration of the model, the horizontal domain was spatially discretised with a 96 × 96 grid using a polar stereographic projection. In the vertical, the model extends from the surface to 100 hPa in 20 vertical layers using terrain-following σ-coordinates. This version of the model describes a total of 58 gaseous chemical species and 9 classes of particulate matter. The chemistry scheme is similar to that used in the European Monitoring and Evaluation Programme (EMEP) model (Simpson et al., 2003).

For all but one of the CTM simulations presented here, meteorological parameters (e.g. wind speed, temperature, pressure) were calculated by the Eta mesoscale NWP model (Janjic, 1994). In the remaining simulation, the MM5 (V3.7) NWP model (Grell et al., 1995) was run to provide meteorological inputs. In both cases, the meteorological initial conditions were taken from the NCEP FNL (Final) Operational Global Analysis dataset (available at http://dss.ucar.edu). The simulations with MM5 meteorology involved a hemispheric domain, with two-way nesting over Europe; the horizontal grid-spacing for the hemispheric and European domains were 150 km and 50 km, respectively, at 60° N (Fig. 1a). The
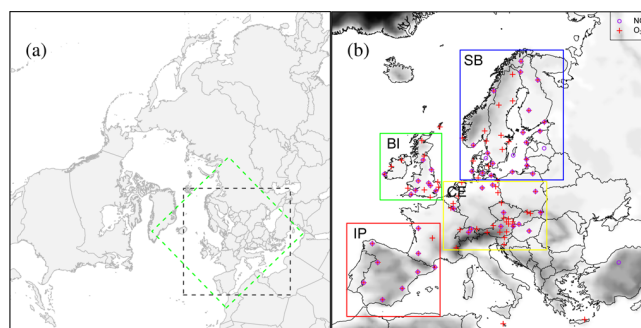

Fig. 1. Left panel: simulations with MM5 meteorology were run with a hemispheric domain (in the region shown), with a nested European domain (indicated by a black dashed box). Simulations with Eta meteorology used only a single nest (indicated by a green dashed box). Right panel: Locations of the NO$_2$ and O$_3$ EMEP monitoring stations are indicated within the Eta domain, and the terrain height used by DEHM is given by the density of the shading. The coloured boxes outline the four sub-regions examined separately within this article.

simulations with Eta meteorology did not include nesting, and the domain was rotated with respect to the European MM5 nest. The horizontal grid-spacing for the Eta-based simulations was also 50 km at 60° N.

Over Europe anthropogenic emissions were based on the EMEP emission inventory (Vestreng and Klein, 2002), and elsewhere RCP2.6 emissions (Lamarque et al., 2010) were assumed. Natural emissions were based on the GEIA inventory (Benkovitz et al., 1996), including NO$_x$ emissions from lightning and soil. For wildfire emissions, the RETRO database was used (Schultz et al., 2007). Aircraft emissions are not accounted for in DEHM.

The extended continuity equation is split into several sub-equations, which are in turn solved sequentially (Lanser and Verwer, 1999). Horizontal advection is solved via "accurate space derivatives" (Dardub and Seinfeld, 1994), and by applying Forester and Bartnicki filters to resolve, respectively, spurious oscillations and negative mass (Forester, 1977; Bartnicki, 1989). Finite elements with linear shape functions are applied to vertical advection. Time integration for the advection is solved using a third-order Taylor series expansion. Diffusion is solved using a combination of the finite elements method and the θ-method (e.g. Morton and Mayers, 2005, Sect. 2.10). The chemistry solver involved a combination of a second-order, two-step, variable step-size backwards differentiation formula (Verwer et al., 1996) and the Euler backward iterative method (Hertel et al., 1993). Lateral boundary conditions are either free or fixed, depending on the wind direction at the boundaries – see Frohn et al. (2002) for further references and details.
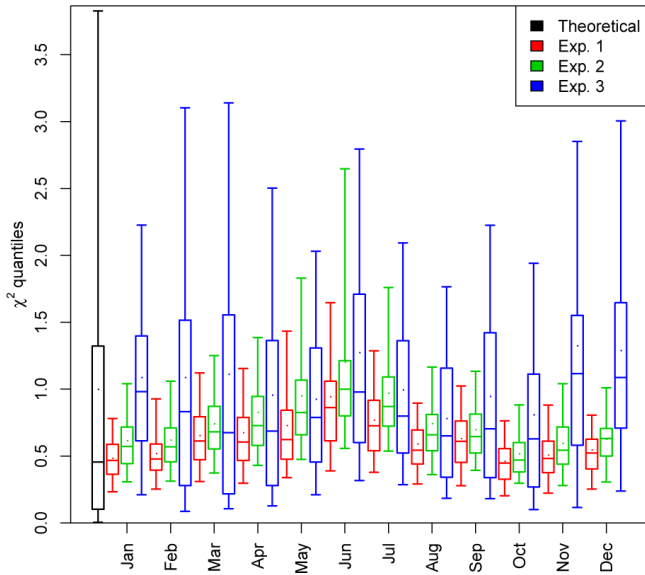
Fig. 2. Box-plots for the χ^2 statistic for Exp. 1–3, separated by each month, as well as for the χ_1^2 distribution. The lower and upper whiskers show the 0.05 and 0.95 quantiles, respectively. The lower bound, centre line and upper bound of the box correspond to the 0.25, 0.5 and 0.75 quantiles, respectively. The dot within each box denotes the mean of the sample.

2.3 Assimilation scheme

The algorithm presented here will be referred to as V1.0 of the three-dimensional OI scheme for DEHM (as distinct from the two-dimensional scheme described in Frydendall et al., 2009).

The assimilation was performed once an hour, if any retrievals were available in the model domain in the previous hour. No special treatment was used to account for the time discrepancy between the model and OMI data, due to short interval between assimilation cycles.

Let x_b be the NO$_2$ concentration field (expressed in units of 10^{15} molecules per cm^2 per model level) and y be the retrieved OMI NO$_2$ tropospheric column concentrations (expressed in units of 10^{15} molecules per cm^2). Let $\mathbf{B}$ and $\mathbf{R}$ be the error covariance matrices for x_b and y, respectively. Let $\mathbf{H}$ be the linear transformation from the model space to the observation space. Then the analysed field x_a is estimated (Kalnay, 2003, pp.150–156) by

$$x_a = x_b + \mathbf{BH}^\top \left(\mathbf{HBH}^\top + \mathbf{R} \right)^{-1} (y - \mathbf{H}x_b). \qquad (1)$$

Given available computational resources it was not possible to solve Eq. (1) as posed. Indeed, the background covariance $\mathbf{B}$ has dimension $n_x n_y n_z = 184\,320$, and would thus require 126 GB of memory at single precision. The following steps were taken to obtain analysis increments. First, $\mathbf{B}$ was implemented algorithmically (i.e. calculating only values required for a given operation). Second, $\mathbf{H}$ was represented as

a sparse matrix, and multiplication by $\mathbf{H}$ exploited this. As described below, $\mathbf{H}$ involved interpolation to the nearest grid point, and this increased the sparsity the matrix (compared to bi-linear interpolation, for example). Third, $\mathbf{R}$ was treated as diagonal, and thus it was only necessary to add to the diagonal elements of $\mathbf{HBH}^\top$. Fourth, the system $\mathbf{HBH}^\top + \mathbf{R}$ was symmetric positive definite, thus the Cholesky decomposition could be used. Fifth, the analysis increment $(x_a - x_b)$ was calculated by multiplying each term by the corresponding vector on the right. In other words:

$$r_1 := \left(\mathbf{HBH}^\top + \mathbf{R} \right)^{-1} (y - \mathbf{H}x_b) \qquad (2)$$

$$r_2 := \mathbf{H}^\top r_1 \qquad (3)$$

$$x_a - x_b = \mathbf{B}r_2. \qquad (4)$$

The most computationally demanding step was Eq. (4), which would require $(n_x n_y n_z)(n_z n_{\text{obs}})$ operations if we only exploited the sparsity of r_2. The product $\mathbf{B}r_2$ was approximated by truncating covariances to zero if the corresponding correlation was below a threshold of 10^{-4}, which greatly reduced the work required for this calculation. A final approximation was to process observations in batches of no more than 1000 at a time, and the analysis from assimilating one batch of observations was used as the background for the next batch (as in Houtekamer and Mitchell, 2001). This is an approximation since sequential assimilation of observations is only strictly valid when data with correlated observation errors are processed in the same batch (Houtekamer and Mitchell, 2001) – this point is discussed further in Sect. 4. Each step of the assimilation scheme was parallelised where possible, and a year-long simulation with assimilation and observation thinning (discussed later) took about 8 % longer than the corresponding reference run without assimilation.

The assimilation step involves an additive increment to the concentration field, and in a large fraction of assimilation cycles, negative values were present in the analysis. Such values are not only unphysical (representing negative concentrations), but may cause further problems in other components of the CTM. Consequently, after the assimilation step any negative values in a given layer were (somewhat arbitrarily) set to the lowest non-zero concentration present in the background field at this layer. This problem is distinct from the negative concentrations due to numerical artifacts of certain higher-order advection schemes, where mass-conservation is required (Bartnicki, 1989).

The background error covariance matrix is parameterised as

$$\mathbf{B}_{i,j;m,h} = \sigma_{i;m,h} \sigma_{j;m,h} r_{l_i,l_j;m,h} \psi(d_{i,j}, L_{l_i,l_j;m,h}), \qquad (5)$$

where subscripts i and j refer to two grid points in the three-dimensional model domain, and subscripts m and h are the indices for the hour of day and month of the year. All parameters were estimated separately for each month of the year and hour of the day, to allow for diurnal and seasonal

patterns in the concentration field, and the associated error statistics. The background error standard deviation for grid point i is termed $\sigma_{i;m,h}$. The subscript l_i denotes the model level of grid point i, and $r_{l,l';m,h}$ the correlation between pairs of grid points at levels (l, l') with zero horizontal separation. The term $\psi(d_{i,j}, L_{l_i,l_j;m,h})$ represents the horizontal correlation between two grid points, and is modelled based on the assumption of horizontally homogeneous and isotropic correlations. This models correlations with a second-order autoregressive function (Balgovind et al., 1983),

$$\psi(d_{i,j}, L_{l_i,l_j;m,h}) = (1 + d_{i,j}/L_{l_i,l_j;m,h})e^{-d_{i,j}/L_{l_i,l_j;m,h}} \quad (6)$$

where $d_{i,j}$ is the horizontal separation between points i and j, and $L_{l,l';m,h}$ is the correlation length scale for pairs of grid points at levels (l, l'). The estimation of the parameters $\sigma_{j;m,h}$, $r_{l_i,l_j;m,h}$ and $L_{l_i,l_j;m,h}$ is presented in Sect. 2.4. For the decomposition of **B** into its variance and correlation components, see Kalnay (2003, p. 151).

The linear observation operator, **H**, calculates the model-equivalent of the OMI tropospheric retrievals. For simplicity only the DEHM column closest to the OMI pixel was used, thus each column of **H** is non-zero for at most $n_z = 20$ values. The DOMINO product includes averaging kernels, which provide the estimated sensitivity of the retrieval to NO$_2$ at different layers of the atmosphere (Eskes and Boersma, 2003). The averaging kernels tend to increase with altitude; thus while NO$_2$ concentrations are highest at the surface (over populated regions, at least), the tropospheric retrievals are far more informative for the free troposphere. The ratio of the tropospheric and total-column air-mass factors was used to convert from total-column to tropospheric averaging kernels (Boersma et al., 2011). The DOMINO averaging kernels are valid at model levels of the chemistry-transport model TM4 (Meirink et al., 2006), which provides the a priori profiles for the retrieval, and thus it was necessary to calculate a weighted average, for each DEHM level, of the averaging kernels (weighted by the proportion of the corresponding TM4 layers overlapping the DEHM level).

The observation matrix, **R**, was treated as a diagonal (i.e. assuming that observation errors are uncorrelated). The choice of observation error variances is discussed in Sect. 3.1.

2.4 Parameter estimation

Parameters for the background error covariance matrix were estimated using differences between paired simulations, also known as the NMC method (Parrish and Derber, 1992). Two DEHM simulations were run, forced with meteorology from the Eta (Janjic, 1994) and MM5 (Grell et al., 1995), respectively; these two simulations will be respectively referred to as Ref. Eta and Ref. MM5. The simulation ran for 2005 and the three-dimensional modelled NO$_2$ concentration field was stored each hour. We will denote the concentration fields
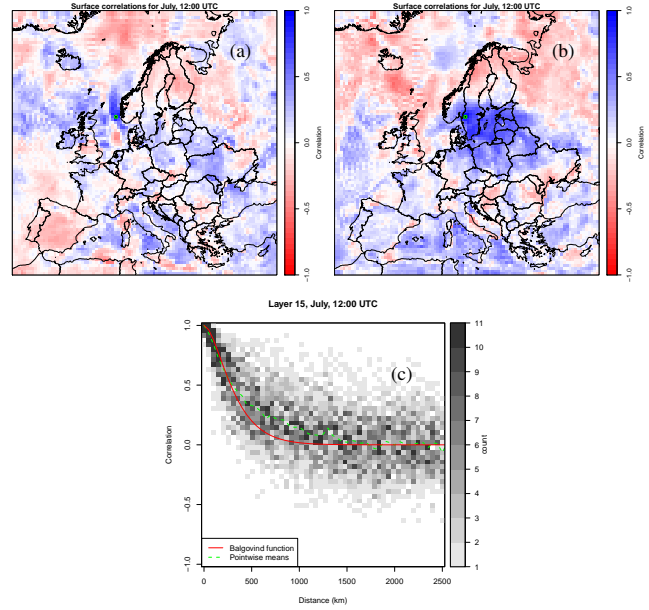
Fig. 3. Top row: sample correlations between a selected point (highlighted with a green circle) and all other grid points, shown for the surface layer. Bottom panel: An example of background error correlations as a function of horizontal separation between pairs of grid points on the same vertical layer. The fitted Balgovind correlation function (Eq. 6) is shown, as well as the moving average value (averaged over 50 km windows). The colour scale shows the number of points inside each of the cells.

from the two simulations as C and $\tilde{C}$. Estimation of parameters was handled separately for each month and each hour of the day, due to the large annual and diurnal cycles in atmospheric concentrations and lifetimes of NO$_2$.

For each time point, t, differences between the paired simulations, were centred at zero and used to calculate the standard deviations (Eqs. 7–9). The $\langle \cdot \rangle$ denotes the averaging operator, and $T(t)$ the set of time points in for the same month and hour as time point t (i.e. the number of elements of $T(t)$ is equal to the number of days in the month). The correlation between points i and j can then be calculated by Eq. 10.

$$e_{i,t} = C_{i,t} - \tilde{C}_{i,t} \quad (7)$$

$$\hat{e}_{i,t} = e_{i,t} - \langle e_{i,t'} \rangle_{t' \in T(t)} \quad (8)$$

$$\sigma_{i;m,h} = \left(\langle \hat{e}_{i,t'} \rangle_{t' \in T(t)} \right)^{1/2} \quad (9)$$

$$\hat{\psi}_{i,j;m,h} = \frac{\langle \hat{e}_{i,t'} \hat{e}_{j,t'} \rangle_{t' \in T(t)}}{\sigma_{i;m,h} \sigma_{j;m,h}}. \quad (10)$$

The above is based on part of the error covariance estimation procedure described in Kahnert (2008).

Estimation of the remaining parameters required in Eq. (6) (the horizontal length scale and inter-level correlation) was based on a sample of 5000 randomly chosen pairs of grid points in the horizontal domain with separation distance range from 0 km to 2500 km. Pairs of grid points were chosen

to ensure a roughly even distribution of horizontal separation distances. For each pair of points, sample correlations were calculated for each pair of vertical levels (l, l'). The inter-level correlation $r_{l,l';m,h}$ was estimated as the mean correlation between the points with zero separation distance; we note that $r_{l,l';m,h} = 1$ by definition if $l = l'$.

For each pair of vertical levels (l, l'), the length scale was estimated by fitting Eq. (6) to the sample of the estimated correlations. The fitting involved minimising the objective function

$$g(L_{l,l';m,h}) = \qquad\qquad\qquad\qquad\qquad (11)$$

$$\sum_{k=1}^{K} \frac{1}{(d_{i_k,j_k} + 1)^2} \left(\hat{\psi}_{i_k,j_k;m,h} - r_{l,l';m,h} \psi(d_{i_k,j_k}, L_{m,h}) \right)^2$$

where subscript k refers to the k-th the pair of grid points i_k, j_k (with points i_k, j_k positioned at levels l, l', respectively), and $K = 5000$ is the total number of sample correlations (illustrated in Fig. 3c). In many cases, the apparent correlation decay length scale was quite short (e.g. less than 100 km), due to the short life-time and consequent high-resolution spatial heterogeneity in NO$_2$ concentrations. In these cases, correlations corresponding to short length scales were of most interest, and hence pairs were down-weighted by distance (using the $(d_{i_k,j_k} + 1)^2$ normalisation factor).

In most cases (i.e. different months, hours, pairs of levels), the correlation model fitted the available data well and the estimation procedure yielded consistent results, in the sense that realistic patterns were observed in the parameter values. These patterns are discussed below and illustrated in Figs. 4–6.

However in some cases, the parameter estimation gave apparently unrealistic results. In particular, it was difficult to estimate correlation decay length scales for pairs of levels if there was also very low vertical correlation between the layers. Also, the upper-most model layer (and to some extent, the layer below this) appeared to show artificially low standard deviations and negative correlations with other layers, which was due to strong damping from the upper boundary condition. Three steps were taken to correct for such artifacts. First, background standard deviations for the upper-most layer were replaced with a copy of those from the layer below. Second, if $r_{l,l';m,h} < 0$ and $l < l'$ then set $r_{l,l';m,h} := r_{l-1,l';m,h}$. Third, correlation length scales between two levels (which showed the highest number of unrealistic artifacts) were modelled as a weighted sum of the original estimated length scale and the sum of the length scales for the individual levels, weighted by the square of the inter-level correlation.

While the fitted correlation captures the overall behaviour of the sample correlations, there is evidence that the covariance model described by Eqs. (5) and (6) was not optimal. In particular, the sample correlations showed evidence of horizontal heterogeneity and anisotropy. These features can be seen in sample correlations between a particular grid point
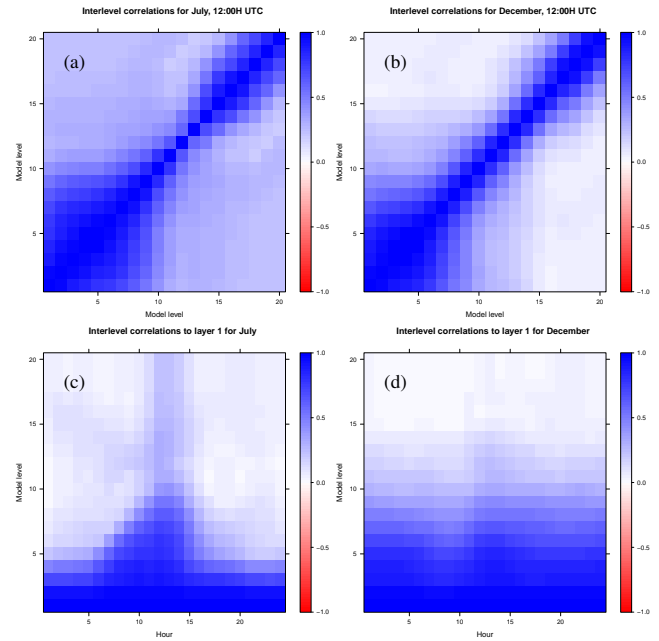

Fig. 4. Correlations between model variables with zero separation distance, $r_{l,l';m,h}$, at different vertical layers (upper row), and diurnal variation in correlation between level 1 and other levels (lower row). The left and right columns present results for July and December, respectively.

and all other grid points (Fig. 3a and b). More sophisticated covariance modelling or use of a stochastic assimilation technique, such as the ensemble Kalman filter, would be required to properly account for this.

The use of time-varying length scales and background covariances appears to be further justified by evidence of diurnal and seasonal variation in these parameters. Figure 4 illustrates variation in the parameter $r_{l,l';m,h}$. Vertical correlations are strongest in the lowest model layers (i.e. within the boundary layer). During the summer, the diurnal cycle is far more prominent, due to strong daytime vertical mixing from surface convection and to the effects of the photochemistry. In winter, the boundary layer is more often stably stratified, resulting in weaker correlations with the upper model levels and stronger vertical correlations between the lower model levels.

In general, estimated correlation length scales ($L_{l,l';m,h}$) tend to increase with altitude and also with the separation between vertical layers (Fig. 5); while this latter point was to some extent enforced by the correction to the estimated length scales, the trend was also seen in the uncorrected estimates. This increase in length scale with height reflects the longer atmospheric lifetime of NO$_x$ and the increase in wind speed. During the summer months, horizontal length scales can be seen to decrease somewhat around layer 10, due to substantial diurnal variation in the boundary layer height, which is diagnosed differently by the Eta and MM5
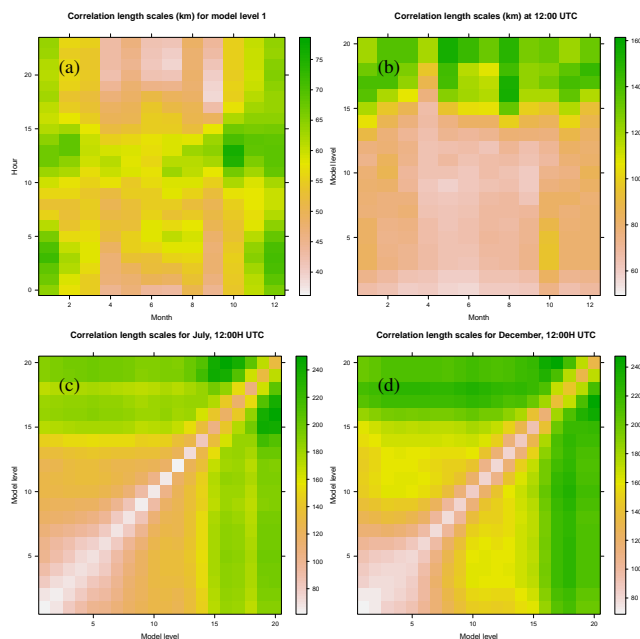
Fig. 5. The estimated correlation length scale, $L_{l,l';m,h}$, as a function of month, hour of the day and vertical level. **(a)** shows diurnal and seasonal variation in the lowest model level, while **(b)** presents seasonal variation in $L_{l,l;m,h}$ (i.e. length scales valid for two points on the same vertical level) at different vertical layers for 12:00 UTC only. The relationship between length scale and pairs of vertical layers in July and December are shown in **(c)** and **(d)**. Note that different colour scales were used for each panel to highlight variation within each graph.

models (discussed in Sect. 3.2). Longer length scales at the surface are seen during the winter months. A diurnal pattern in $L_{1,1;m,h}$ is most pronounced during the summer months, as NO$_2$ can accumulate at night whereas it is dissipated during the daylight hours due to O$_3$ formation and an increasing boundary layer.

The background error standard deviations ($\sigma_{i;m,h}$) are partly determined by the magnitude of the concentrations observed in each grid-cell (Fig. 6). This can be seen in higher estimated error variances in areas of high NO$_x$ emission density (e.g. the Benelux region). During the winter months, NO$_2$ concentrations exhibit substantial temporal variation (Fig. 7), and this is reflected in the background errors (as they were estimated by the standard deviation of a time series of values with one value per day). The vertical profile of $\sigma_{i;m,h}$ tends to increase over the first 6–10 model levels and then decay, following the trend seen in the modelled concentrations (in terms of mass NO$_2$ per model level) at each level; this is in spite of the general decay of modelled NO$_2$ mixing ratio with altitude (Fig. 12a). This is due to the fact that the mass of NO$_2$ per unit area per model level is the product of the mixing ratio (which decays with height), the density of air (which decays exponentially with altitude) and the thickness
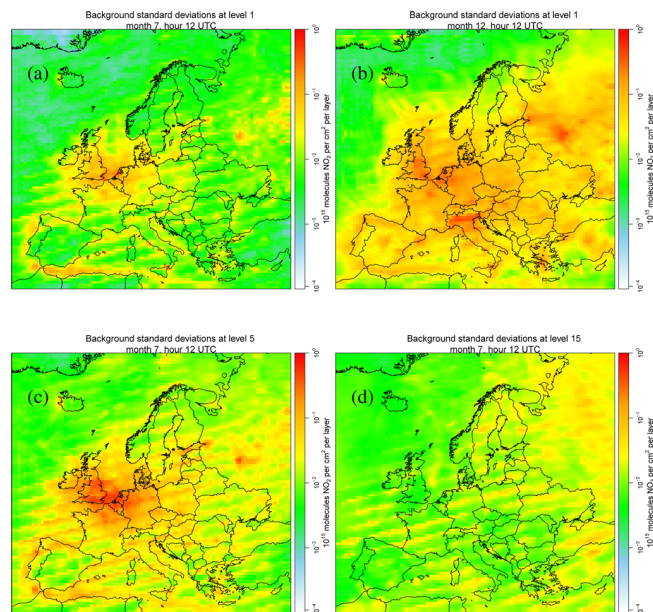
Fig. 6. Background error standard deviations, $\sigma_{i;m,h}$ in July and December at the lowest model layer (top row), and for layers 5 and 15 in July (bottom row). Note that the colour scale denotes is logarithmic.

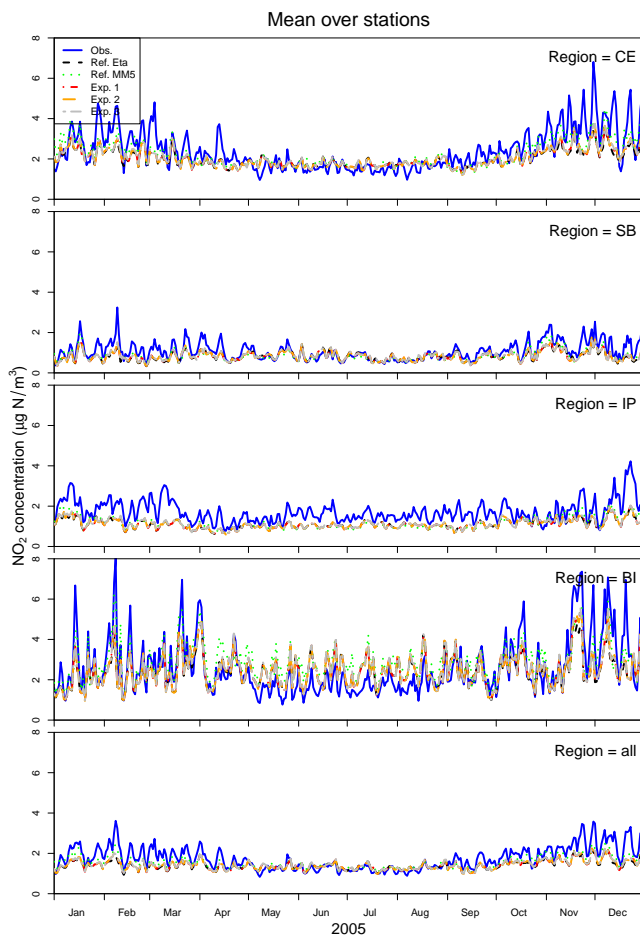of the model layer (levels are thinnest at the surface and increase).

The estimated background errors exhibit artifacts (appearing as ripples in Fig. 6a, c and d) from the interpolation from the MM5 grid to the Eta grid (Fig. 1a). These are most clearly visible when shown on a logarithmic scale (very little spatial structure is visible when the same data are plotted with a linear scale). No filtering was applied to correct for this.

3 Verification

3.1 Experiments

Along with the two simulations used for the background covariance estimation, which are denoted Ref. Eta and Ref. MM5, three simulations were run using the assimilation procedure described above (summarised in Table 1). They were designed to explore the treatment of observation errors and the problem of correlated observation errors.

Satellite data can be classified as level-1 products (calibrated and collocated radiances) and level-2 products (derived geophysical quantities, such as temperature, humidity, concentration of trace gases). The process of retrieving level-2 information from level-1 data involves an inverse-modelling framework similar to that used in data assimilation, requiring an a priori estimate of the state of the atmosphere. This results in correlations between observation errors; the assumption of independent errors in level-1 data fits much better than for level-2 products (Kunzi et al., 2011).

Fig. 7. Time series of daily-averaged NO$_2$ concentrations for the different regions indicated in Fig. 1b. Each time series, for observations and the modelled values from the five simulations, is shown with a distinct colour and line type.

Table 1. A summary of the differences between the experiments. Abbreviations used: DA – whether data assimilation was applied, Thin – thin satellite data to a given resolution, Const. **R** – use constant observation errors for each month, NWP – the numerical weather prediction model used to provide the meteorological inputs.

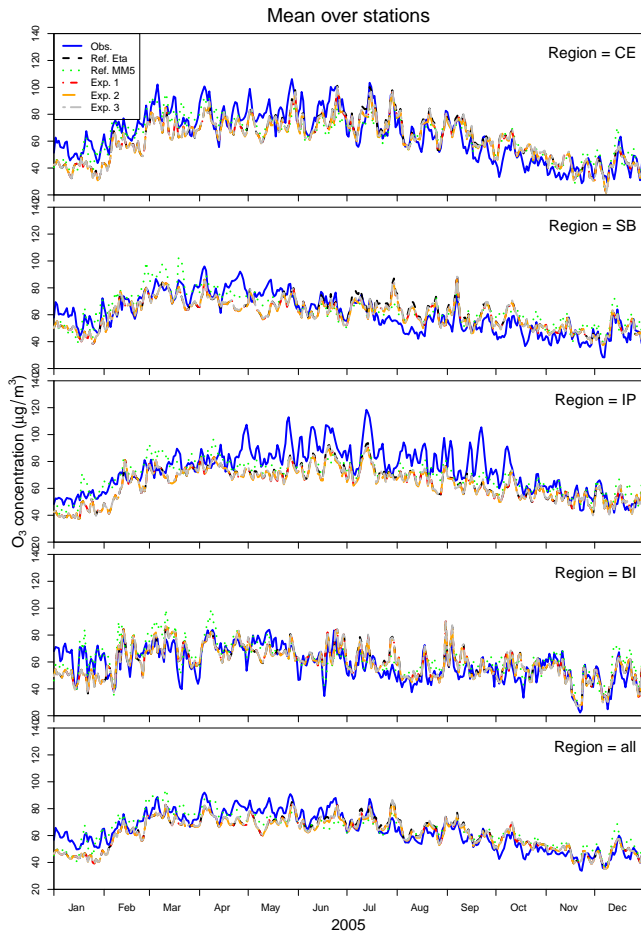| Name	DA	Thin	Const. $\mathbf{R}_{ii}$	NWP	Outer nest
Ref. Eta	No	–	–	Eta	No
Ref. MM5	No	–	–	MM5	Yes
Exp. 1	Yes	No	No	Eta	No
Exp. 2	Yes	Yes	No	Eta	No
Exp. 3	Yes	Yes	Yes	Eta	No

as $13\,km \times 24\,km$. Thus model and observations are representative on different scales, which contributes to the representativity component of the observation errors. In experiments Exp. 2 and Exp. 3 this issue was addressed by thinning observations so that they were at least $50\,km$ apart, thereby in principle reducing the observation error correlation. In Exp. 1, all available observations were used that passed the quality-control criteria described in Sect. 2.1.

Figure 13 illustrates the spatial distribution of the total column concentration, as well as the effect of the assimilation. Panels a and b present, respectively, the background and analysis (for simulation Exp. 2) projected into observation space, while panel c shows the OMI retrievals. The grid-averages of these data were calculated for the month of July 2005, averaging over all assimilation cycles in this period. Only observations used in the assimilation (i.e. passing the quality control and thinning) were included in the averages presented, thus no data are available over areas of persistent ice cover or in the outermost rows and columns of the model domain.

The OMI field shows a greater contrast between "clean" and "polluted" regions (n.b. retrievals in some remote regions, such as over oceans, may be negative even after averaging). The grid-averaged retrievals also show greater spatial heterogeneity than either the averaged $\mathbf{H}x_b$ or $\mathbf{H}x_a$ fields. As we would expect, the analysis appears as a merger of the background and the observations. However one can also see the results of the ripple-like interpolation artifacts from the background error correlations (see Sect. 2.4); these are most apparent in relatively unpolluted areas.

The observation errors used in the assimilation experiments were based on the reported errors in the DOMINO product. As stated in Sect. 2.1, the retrieved tropospheric column concentration was highly correlated with the associated error. This is may be an appropriate description from a measurement perspective, however it may be problematic in an assimilation context. If the higher retrieved values are assigned higher observation errors, then these data will have less influence on the analysis, resulting in negatively biased analysis increments. In simulations Exp. 1 and Exp. 2, the
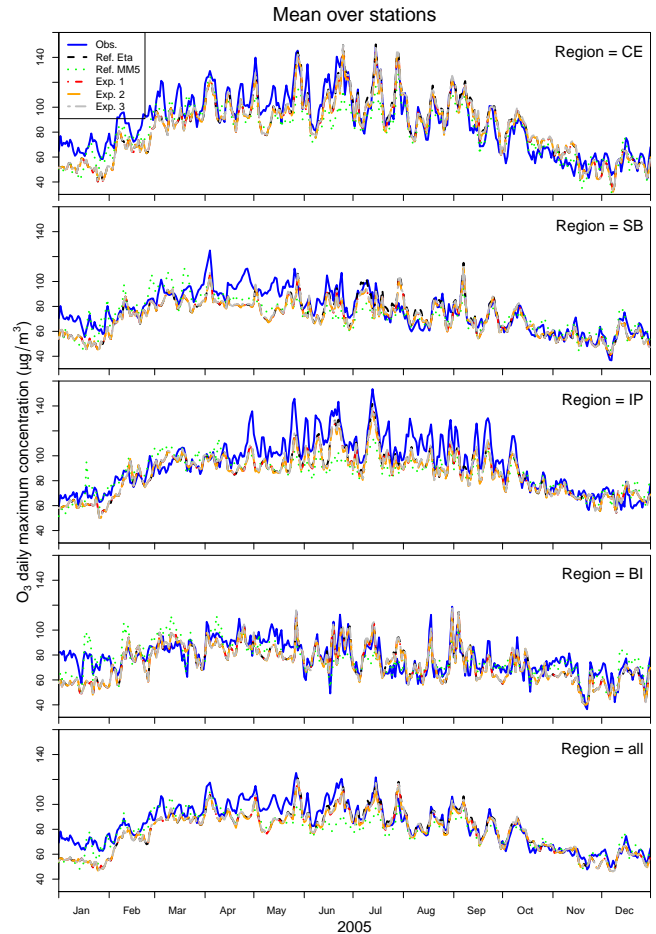
The best treatment of correlated observation errors for data assimilation is still an open research question. The most common approaches used in operational meteorology are the assumption of uncorrelated errors in conjunction with "thinning" (discarding observations to reduce the data density), "superobbing" (assimilating the average of a group of nearby observations) or observation error variance inflation (Stewart et al., 2008). If observation errors are indeed correlated but are treated as independent in the assimilation scheme, then to achieve an appropriate balance between observation and background errors the observation error variances must be artificially inflated. A final and promising approach is to model the observation error correlations explicitly (Stewart et al., 2008), however this is beyond the scope of the present study.

The assimilation experiments addressed two different aspects of the problem of correlated errors. First, the OMI observations were at a much higher spatial resolution than the model's resolution: a DEHM grid-cell represents a roughly $50\,km \times 50\,km$ region, whereas the OMI pixels were as small

Fig. 8. Time series of daily-averaged O₃ concentrations for the different regions indicated in Fig. 1b.

Fig. 9. Time series of daily-maximum O₃ concentrations for the different regions indicated in Fig. 1b.

observation errors used were those reported in the DOMINO product, whereas in Exp. 3, the observation error standard deviation was fixed at a constant value, chosen to be the median of the tropospheric column errors reported in the DOMINO product for the given month.

To assess the balance of the background and observation errors, we conducted an a posteriori validation of the assimilation using the $\hat{\chi}^2$ metric (Ménard and Chang, 2000), defined in Eq. (12), where p is the number of observations available at the assimilation step.

$$\hat{\chi}^2 = (\boldsymbol{y} - \mathbf{H}\boldsymbol{x}_b)^\top \left(\mathbf{H}\mathbf{B}\mathbf{H}^\top + \mathbf{R} \right)^{-1} (\boldsymbol{y} - \mathbf{H}\boldsymbol{x}_b)/p \qquad (12)$$

If the standard assumptions hold (unbiased background and observations, Gaussian errors) and if the background and observation error covariances adequately describe the residuals (i.e. the difference $\boldsymbol{y} - \mathbf{H}\boldsymbol{x}_b$), then the statistic $\hat{\chi}^2$ should follow a χ_1^2 distribution with 1 degree of freedom. It then follows from the properties of the χ_1^2 distribution that $\langle \hat{\chi}^2 \rangle \approx 1.0$. This provides an a posteriori check of whether the observation and background errors are appropriately specified.

The $\hat{\chi}^2$ metric was calculated for simulations Exp. 1–3 for each month, and are shown alongside the theoretical distribution statistic (Fig. 2). The background errors were determined, as described above, by the difference between paired forecasts. Assuming that these have been appropriately specified, then the $\hat{\chi}^2$ results reflect on the observation errors. Exp. 1 and 2 yield $\hat{\chi}^2$ values consistently less than 1.0, indicating that the magnitude of observed residuals was smaller than that prescribed by the error variances. The observation errors would need to be inflated substantially (as to compensate for correlated observation errors) to correct this. The $\hat{\chi}^2$ values for Exp. 1 were lower than for Exp. 2, suggesting that the impact of observation error correlations was less severe for the Exp. 2, where observations were thinned. For Exp. 3, which used fixed observation error variances as well as thinned observations, the $\hat{\chi}^2$ values show a much closer fit to the χ_1^2 distribution compared to Exp. 1 and 2. This is partly attributable to the fact that none of the prescribed observation errors for this experiment were relatively small (i.e. from the lower tail of the reported DOMINO errors).
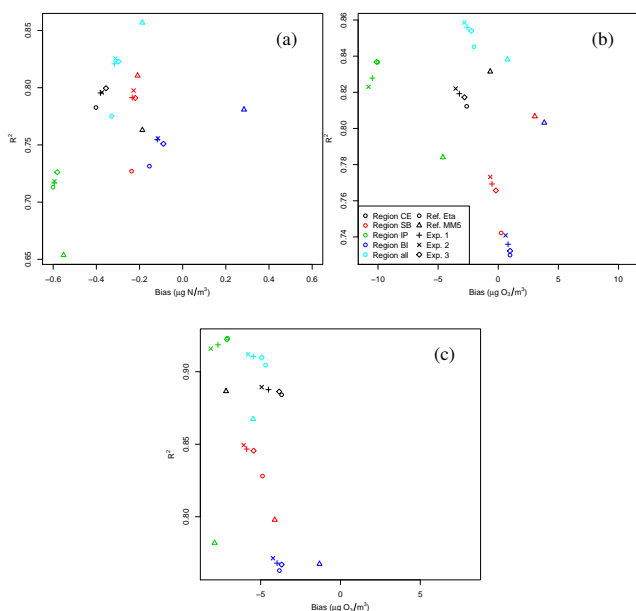
Fig. 10. Verification statistics for daily-averaged NO$_2$ **(a)**, daily-averaged O$_3$ **(b)** and daily-maximum O$_3$ concentrations **(c)**. These were based on the time series of the mean over stations, averaged over different regions (Figs. 7, 8 and 9). Results for the different simulations and regions are denoted by different point types and colours. Note that the ranges of the x- and y-axes differ between the three panels.

Thinning observations clearly results in a loss of information, which is not the aim of this exercise. In order to assess the impact of the loss of information, we now turn our attention to verification with ground-based monitoring data.

3.2   Verification with EMEP data

The DEHM describes a total of 67 atmospheric components, however we examine results for only two: NO$_2$ and O$_3$. Our main interest is in how the accuracy of NO$_2$ estimates varies when assimilating this species. We also chose to examine O$_3$, as it has a close chemical relationship with NO$_2$ and has important consequences for human health.

Observations of NO$_2$ and O$_3$ concentrations for the year 2005 were obtained for European air-quality monitoring stations in the EMEP network (Aas, 2008). Stations were selected for inclusion based on four quality-control criteria: at most 50 % data were missing, the station altitude was below 2500 m above sea level, the difference between the station altitude and the land-surface height represented by DEHM differed was below 200 m, and (for NO$_2$ only) the measurement technique was reported. Four areas within the Eta domain were defined: Scandinavia and the Baltic region (SB), central Europe (CE), the Iberian Peninsula (IP) and the British Isles (BI), as shown in Fig. 1b. The numbers of monitoring stations for each species in each region are reported in Table 2. Hourly O$_3$ measurements were available for all stations, and

Table 2. The number of monitoring stations satisfying the inclusion criteria in each of the regions considered (Fig. 1b).

Name	Acronym	# NO$_2$ stations	# O$_3$ stations
Eta European domain	–	53	99
Central Europe	CE	12	36
Scandinavian and Baltic region	SB	22	31
Iberian Peninsula	IP	9	10
British Isles	BI	9	18

daily-average NO$_2$ data were provided for all stations, with 25 of those passing the quality-control criteria also reporting NO$_2$ at hourly frequency.

The in situ NO$_2$ data were measured with either chemiluminescence or aqueous phase techniques (variants on the Griess-Saltzman reaction). The molybdenum oxide catalysts used in chemiluminescence monitors have been shown to reduce a range of nitrates, and thus the resulting measurements are subject to interference from such species (Dunlea et al., 2007). It was thus necessary to compare such measurements with NO$_2$ plus the sum of modelled nitrates ($2N_2O_5 + HNO_3 + HO_2NO_2 + CH_3COOONO_2 + NO_3^-$). For stations that used an aqueous-phase measurement, it was possible to compare directly to the modelled NO$_2$. In both cases, modelled and measured values are expressed in units of µg N m^{-3}. The details of which measurement technique was used at the individual monitoring sites was obtained from AirBase station configuration files (Mol and de Leeuw, 2005).

The EMEP monitoring sites are located to measure regional background concentrations in Europe. Yet the concentrations measured at these sites are necessarily subject to local factors, and the sites differ in their proximity to emission sources. The model's 50 km × 50 km grid resolution will not capture such local variation, yet by averaging geographically, the influence of these sub-grid scale effects is diminished. Thus for each day, the average over all station (in the Eta domain as well for the four sub-regions) was calculated for the observed and calculated concentrations (Figs. 7–9). Given the availability of hourly O$_3$ measurements, it was also possible to assess how well the model captures diurnal variation as expressed in the daily maximum O$_3$ concentration (Fig. 9). Based on the observed and calculated time series, we computed the correlation coefficient (R^2) and bias (modelled minus observed). Figure 10 presents the verification statistics for the 5 simulations.

These time series show only very small differences between the five simulations. The largest differences being due to the choice of meteorological input, and Ref. MM5 was clearly distinct from the other simulations (which used meteorology from the Eta NWP model). For NO$_2$ (Fig. 7), the differences were most apparent during the winter months, which may be related to differences in the mixing heights diagnosed from the two meteorological datasets. The Eta
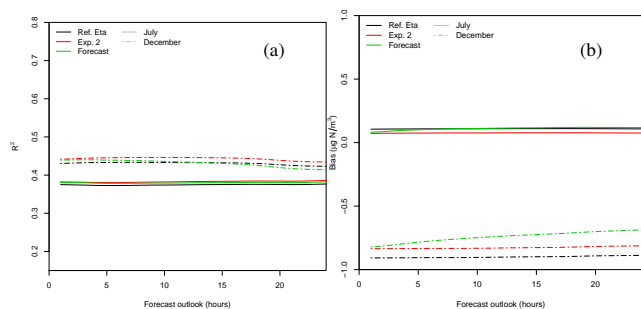
Fig. 11. Verification statistics for forecasts of NO$_2$ (compared to Exp. 2, which was used to provide the initial conditions, and to Ref . Eta) in terms of correlation **(a)** and bias **(b)**. The line colour indicates the simulation, the line type denotes the month in question.

Fig. 12. Modelled concentration profiles, averaged over region CE (Fig. 1b), for July and December and for each of the five simulations.

model uses the η (step-mountain) vertical coordinate and was configured with 32 vertical levels, whereas both MM5 and DEHM use the same terrain-following σ vertical coordinates with 20 vertical levels. The σ vertical coordinate allows for a higher vertical resolution in the planetary boundary layer (PBL), yet the η vertical coordinate places a limit to the vertical resolution in the PBL and thus on the minimum height of the mixed-layer height that can be resolved.

During the winter months, the height of the mixed layer can be quite low, leading to higher concentrations of directly emitted compounds (such as NO$_x$), particularly in cases of nocturnal inversion layers. This results in higher variability in the observed NO$_2$ concentrations, and the wintertime variability is resolved better by Ref. MM5 than the Eta-driven simulations. This is most apparent in region CE, which has a high density of NO$_x$ emissions and is subject to low inversion layers during the winter months. Low mixed layers heights are also common in the region SB, especially in northern areas, but there is much lower wintertime variability in NO$_2$ due to the lower NO$_x$ emission density.

The region BI showed high-amplitude temporal fluctuations in NO$_2$ concentrations, partly due to the small number of monitoring stations to average over. Differences between Ref. Eta and Exp. 1–3 can be seen for this region during the highly variable concentrations in November and December of the study period. In region BI, the model does not capture seasonal variation very well, tending to under-predict NO$_2$ concentrations during the winter months and over-predict during the summer months (leading to a relatively low overall bias).

Of all the areas considered, the DEHM shows least skill for the region IP, where NO$_2$ levels are underestimated throughout the year and the model does not appear to capture the observed temporal fluctuations; these comments apply both to the MM5- and Eta-driven simulations. The region IP is at a lower latitude to the other regions considered, thus receiving more solar radiation and a shorter atmospheric lifetime of NO$_x$ (due to photolysis). Hence local emission sources are
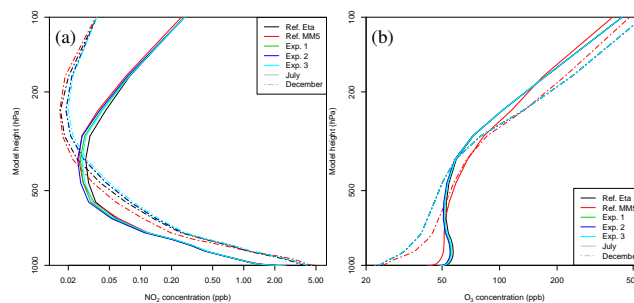
relatively more important, and the coarseness of the DEHM grid-cells poorly describe local emission patterns.

The verification statistics for NO$_2$ (Fig. 10a) show that for all regions considered, simulations Exp. 1–3 had a higher correlation with observations than Ref. Eta, although there was little change to the bias. The Ref. MM5 simulation had the highest correlation of the five simulations for regions BI, SB and the entire Eta domain, and the lowest correlation for the other regions. In region BI, the Ref. MM5 run was also far more biased than the other simulations.

For O$_3$, the Ref. MM5 simulation again has the most distinct results (Fig. 8). Ozone has a much longer atmospheric lifetime than NO$_2$ and thus the nesting within the hemispheric domain (Fig. 1a) ensures that the European nest is affected by episodes of long-range transport (unlike for the Eta domain, which uses fixed inflow concentrations based on monthly averages from hemispheric simulations). The different meteorological inputs will also influence the photochemistry (e.g. due to cloud cover), affecting O$_3$ production during the summer months. Differences between Ref. Eta and Exp. 1–3 are most prominent during the summer months (e.g. during peak episodes in regions CE and SB in June and July) due to the higher photochemical production and to the changes to O$_3$ precursors by the assimilation procedure. In the region IP, the model tends to underestimate O$_3$ concentrations during the summer months. This may be partly due to the general underestimation of NO$_2$ in this region (Fig. 7), and possibly related to poor specification of emissions of biogenic volatile organic compounds (Zare et al., 2012).

Verification statistics for daily average O$_3$ (top-right panel of Fig. 10b) show little change in bias in the simulations using assimilation (with respect to Ref. Eta), but an increase in correlation for these simulations in all regions except IP. The performance of Ref. MM5 stands out, and has the highest correlation of the five simulations for region SB, BI and CE, and the lowest for regions IP and for the Eta domain.

Many of the same comments apply to the daily maximum O$_3$ concentration (Fig. 9) as for daily average O$_3$. Peak O$_3$ concentrations typically occur around 14:00–15:00 local time, thus shortly after any OMI overpass. Thus the extra
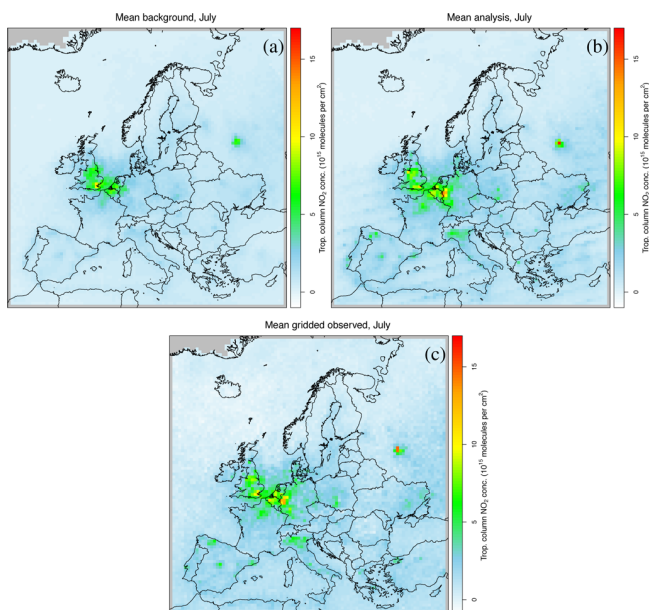
Fig. 13. Tropospheric column concentration averaged over all assimilation cycles for a month. **(a)** and **(b)** show grid-averaged fields of $\mathbf{H}x_b$ and $\mathbf{H}x_a$, respectively, for Exp. 2. **(c)** presents the grid-averaged OMI retrievals. Grid cells with no observations are shown in grey. Only observations passing the quality-control criteria described in Sect. 2.1 were considered, and this is reflected in the grey areas shown.

information provided by the assimilation of tropospheric NO$_2$ retrievals should be most prominent around this time of day. However, there were only small differences in the verification scores for simulations using Eta meteorology. Exp. 1–3 were slightly more biased than Ref. Eta, and there was little difference in the correlation (apart from region SB, where Exp. 1–3 had slightly higher correlation). The Ref. MM5 simulation showed much lower correlations with observed daily maximum O$_3$ concentrations in regions IP, SB and the Eta domain. Finally, we note that on average the model tends to under-predict daily maximum O$_3$ concentrations in all regions, particularly from March to June.

3.3 Forecasts

The model results in Sect. 3.2 are based on daily or hourly averages, and the assimilation procedure was performed once an hour for each hour when observations were available. Thus these are effectively a mixture of background (input for an assimilation), analysis (result of an assimilation) and forecast (the time-integration using an analysis as input), and can be thought of as reanalyses. It is also of interest to assess the impact of the initial conditions on CTM forecasts.

From Exp. 2, instantaneous concentration fields were stored for every hour for July and December, and these were used as initial conditions for 48-h forecasts. It was necessary to start forecasts at each hour of the 24-h cycle in order to

smooth out diurnal effects (which dominate the results otherwise). Hourly NO$_2$ observations were available at 25 EMEP stations. For these stations, for each hour of the 48-h forecast outlook we calculated the bias and correlation between modelled and observed values. The bias and correlation were calculated separately for each station and these were summarised as a weighted average over the per-station values, with weights based on the number of non-missing observations for the station in the period considered. For comparison, the time-matched verification scores for simulations Ref. Eta and Exp. 2 are shown together with the results of the forecasts. The first 24 h are shown in Fig. 11. The forecasts were run with the same meteorological inputs as for the other Eta-driven simulations.

The bias in NO$_2$ shows a rapid adjustment in July (over 3–4 h), and somewhat slower in December (over 10–15 h), reflecting the seasonal variation in NO$_x$ lifetimes. While in July the forecasts adjust towards the Ref. Eta as expected, the December forecasts tend away from the corresponding results from the free run. The reasons for this are unclear, and it may suggest that either the free run (Ref. Eta) is not the only equilibrium solution when the assimilation is turned off, or that other factors in the initial conditions (influenced indirectly by the changes from the assimilation of NO$_2$) have a longer relaxation time. In July, the correlation was relatively constant over the forecast outlook considered, whereas in December the improvement in correlation appears to decay after around 15 h.

3.4 Modelled profiles

Due to vertical variation in both $\mathbf{H}$ (due to the averaging kernels) and $\mathbf{B}$, the increment is not expected to be evenly distributed throughout the vertical profile. When averaged temporally over the months of July and December and spatially over the region CE (Fig. 1b), the differences between the modelled NO$_2$ profiles for the different experiments are very small and only discernable in the free troposphere and lower stratosphere (Fig. 12a). In December, the NO$_2$ mixing ratios in Exp. 1–3 were slightly higher than for Ref. Eta, whereas in July they were marginally lower. For Ref. MM5, NO$_2$ mixing ratios were higher at the surface in December than was seen for the simulations with Eta meteorology.

The differences in modelled O$_3$ concentrations between Ref. Eta and Exp. 1–3 were only visible for the simulations in July, where the assimilation tended to result in slightly lower O$_3$ concentrations in the boundary layer and free troposphere. The differences between the Ref. MM5 and the Eta-driven simulations are more pronounced for O$_3$ than for NO$_2$, due to combination of the longer atmospheric lifetime of O$_3$ and the dynamic boundary conditions provided by the hemispheric nest.

4 Discussion

The assimilation of the OMI retrieved tropospheric NO$_2$ column concentrations led to increased temporal correlation with surface measurements of NO$_2$, and somewhat surprisingly with surface O$_3$. However the differences between the free run (Ref. Eta) and the simulations with assimilation (Exp. 1–3) were rather small and were difficult to see in the time series of the modelled concentrations. The short forecasts showed only minor differences from the other simulations once diurnal effects were averaged out (by starting forecasts at each hour of the day), and did not necessarily relax toward the free run as expected.

While the impact of assimilation of OMI retrievals may be positive, at the surface at least, it appears rather small; there may be several factors at play. First, due to the short atmospheric lifetime of NO$_2$, any improvement due to extra information from an observation has a relatively short duration and is rather localised. Second, in a given day only a fraction of the model domain benefits from observations. This is partly due to the nature of the satellite overpasses, and partly due to the quality control procedure, which excluded a large fraction of the retrievals due to high cloud fraction or surface albedo.

In spite of the preceding considerations, Wang et al. (2011) report a substantial improvement in surface NO$_2$ concentrations due to assimilation of OMI NO$_2$ column concentrations. This suggests that the relatively small improvement found here was due to a non-optimal usage of the data rather than, for example, high errors in the retrievals themselves. Yet Wang et al. (2011) estimated parameters for their **B** in order to optimise performance at surface monitoring sites. This is likely to have contributed to the greater performance gains than may be achieved without such tuning; the NMC method used here to parameterise **B** did not allow for "tuning" (by its nature, and also due to the large number of parameters involved).

The assimilation experiments were all run with meteorology from the Eta NWP model, however we also examined results from the free run with meteorology from MM5, which was used in the construction of the background covariances. It was seen that for three of the regions considered, Ref. MM5 showed much higher correlation with surface NO$_2$ measurements than any of the Eta-driven simulations. This highlights the importance of considering the influence of model inputs in chemical data assimilation research.

The three-dimensional OI scheme presented here is a development from an earlier two-dimensional scheme (Frydendall et al., 2009). The three-dimensional version makes appropriate use of the averaging kernels, thus providing extra vertical structure in the analysis increment. Despite the extra dimension involved, a number of algorithmic speed-ups and approximations were implemented and the OI scheme to limit the extra computational burden.

Background covariances were estimated from the difference between paired CTM simulations, run with different meteorology and a different domain structure. The background covariances were calibrated for each month and for each hour, to allow for seasonal and diurnal patterns in the modelled NO$_2$ field. The resulting parameter estimates illustrated not only these temporal patterns but also vertical structure in the NO$_2$ error correlations.

This construction of the background covariances incorporates uncertainties due to meteorological inputs only. While the meteorology is an important source of uncertainty in the modelled concentration field (Vautard et al., 2012), other key factors (such as emission rates, the choice of physical and chemical parameterisations, grid resolution) should also be accounted for (Mallet and Sportisse, 2006).

Observation errors for the OMI retrievals can be assumed to be correlated due to the use of level-2 satellite data. The effects were investigated by thinning observations to a spatial resolution comparable to the model grid, as well as assuming fixed error variances for all observations (during a given month). The combination of these yielded a better balance between the magnitude of the observed residuals and the specified error variances. The loss of information resulting from the observation thinning did not appear to systematically degrade the quality of the modelled NO$_2$ or O$_3$ surface concentrations (based on the verification statistics).

The assumption of independent observation errors was not only applied in specification of the observation error matrix, **R**, but also in the use of sequential processing of observations. Such batch processing is only strictly valid if data with correlated observation errors are assimilated in the same batch (Houtekamer and Mitchell, 2001). This could have been avoided if a shorter assimilation cycle were used (as in Wang et al., 2011), thus reducing the number of observations per cycle. However such a treatment ignores correlations in observation errors between successive assimilation cycles (Dee, 2005). Appropriate treatment of observation error correlations requires further investigation.

The year-long span of the simulations allowed us to consider seasonal variation. For example, the assimilation showed the most pronounced effects on modelled NO$_2$ concentrations in the winter, and had the biggest impact on O$_3$ during the summer months.

Barbu (2010, pp. 65–78) assimilated the OMI tropospheric NO$_2$ retrievals into the LOTOS-EUROS model, using a bias-aware ensemble Kalman filter. Ensemble spread was generated by perturbing the emissions of NO$_x$ and VOC emissions. While the bias-aware assimilation led to much more accurate estimates of surface NO$_2$, the simulation period covered only one month and verification statistics for O$_3$ were not reported. This significance of model bias is emphasised by Dee and da Silva (1998), who showed that a biased background field will result in a biased increment. Biases in the background or observations were not accounted for in the work presented here, and could be addressed as an extension

of this system, such as with the sequential bias correction scheme given in Eqs. (17)–(18) of Dee (2005).

The OI assimilation scheme is both conceptually simple and relatively straightforward to implement. If both the length of control vector (i.e. those modelled variables adjusted by the assimilation) and the number of observations are sufficiently small, then it may be less computationally demanding than other assimilation schemes. However the framework of OI is, in several ways, rather limited (Kalnay, 2003). First, it does not scale well as the number of observations or the size of state increases. Second, it requires the preparation of a climatological **B** matrix, which must be recalibrated if extra variables are added to the state vector. Third, unlike variational techniques it cannot be used for parameter estimation or quality control of observations. Far more flexible approaches to chemical DA are offered by variational or Kalman filter-based schemes (see Lahoz et al., 2007, and references therein).

For historical re-analysis and for single-component assimilation, then OI may be a sufficiently competent scheme, beyond which any gains are marginal and require a significantly more complicated and computationally demanding DA procedure (Wu et al., 2008). In a forecasting context, however, the strong forcings from chemistry and emissions limit the potential benefit from more accurate initial conditions, although the rate depends on the atmospheric lifetime of the species in question and whether they are emitted directly or formed chemically. The framework of DA can be used to re-estimate highly uncertain model parameters (e.g. emission rates, as were examined by Elbern et al., 2007), and this appears to be a promising means of addressing forecast accuracy for directly emitted, short-lived atmospheric components such as NO$_x$; however, this is not possible with OI.

The DA scheme described here could be developed in several ways. For example, the observations were originally provided as pixels, thereby corresponding to a two-dimensional region. However for simplicity they were summarised at the pixel's midpoint. A more accurate version of the observation operator would account for the fact that pixels may span multiple grid-cells. Furthermore, the observation operator was constructed by considering only the vertical column above each OMI pixel, thus not accounting for horizontal variation in NO$_2$ concentrations resulting from non-nadir scan angles.

5 Conclusions

In this study, we have presented an algorithm used to assimilate remotely-sensed tropospheric NO$_2$ column concentrations. Results were compared to observations from regional background monitoring stations. The analysed concentration fields showed increased temporal correlation for both NO$_2$ and O$_3$ across the different sub-regions considered.

Non-separable, horizontally homogeneous and isotropic background covariances were estimated from the difference between paired simulations, effectively attributing all model uncertainties as arising from the meteorological inputs. Despite the incomplete survey of sources of model uncertainty, the parameters of the three-dimensional background covariances reflect temporal patterns and vertical structures that fit with the modelled variation in NO$_2$ concentrations.

Despite its limitations, the optimal interpolation algorithm is conceptually simple and relatively straightforward to implement, and proved it to be useful in this context. We have demonstrated that the effects of chemical DA are not limited to the assimilated species, and can be seen in chemically related compounds. Finally, the effectiveness of chemical DA in a forecasting context must be considered in conjunction with the atmospheric lifetime of the species in question.

References

Aas, W. (Ed.): Data quality 2006, quality assurance, and field comparisons, EMEP/CCC-Report 3/2008, EMEP Co-operative Programme for Monitoring and Evaluation of the Long-range Transmission of Air Pollutants in Europe, Norwegian Institute for Air Research, Kjeller, Norway, available at: www.emep.int (last access: 21 December 2012), 2008.

Adhikary, B., Kulkarni, S., Dallura, A., Tang, Y., Chai, T., Leung, L. R., Qian, Y., Chung, C. E., Ramanathan, V., and Carmichael, G. R.: A regional scale chemical transport modeling of Asian aerosols with data assimilation of AOD observations using optimal interpolation technique, Atmos. Environ., 42, 8600–8615, 2008.

Balgovind, R., Dalcher, A., Ghil, M., and Kalnay, E.: A stochastic-dynamic model for the spatial structure of forecast error statistics, Mon. Weather Rev., 111, 701–722, 1983.

Barbu, A. L.: Ensemble-based data assimilation schemes for atmospheric chemistry models, Ph.D. thesis, Delft University of Technology, 2010.

Bartnicki, J.: A simple filtering procedure for removing negative values from numerical solutions of the advection equation, Environ. Softw., 4, 187–201, 1989.

Benkovitz, C. M., Scholtz, M. T., Pacyna, J., Tarrasón, L., Dignon, J., Voldner, E. C., Spiro, P. A., Logan, J. A., and Graedel, T. E.: Global gridded inventories of anthropogenic emissions of sulfur and nitrogen, J. Geophys. Res., 101, 29239–29254, 1996.

Boersma, F., Bucsela, E., Brinksma, E., and Gleason, J. F.: Chapter 2: NO$_2$, in: OMI algorithm theoretical basis document, OMI trace gas algorithms, edited by: Chance, K., IV, NASA technical report, 2002.

Boersma, K. F., Eskes, H. J., Veefkind, J. P., Brinksma, E. J., van der A, R. J., Sneep, M., van den Oord, G. H. J., Levelt, P. F., Stammes, P., Gleason, J. F., and Bucsela, E. J.: Near-real time retrieval of tropospheric NO$_2$ from OMI, Atmos. Chem. Phys., 7, 2103–2118, doi:10.5194/acp-7-2103-2007, 2007.

Boersma, K. F., Jacob, D., Bucsela, E., Perring, A., Dirksen, R., van der A, R., Yantosca, R., Park, R., Wenig, M., Bertram, T., and Cohen, R.: Validation of OMI tropospheric NO$_2$ observations during INTEX-B and application to constrain NO$_x$ emissions over the eastern United States and Mexico, Atmos. Environ., 42, 4480–4497, 2008.

Boersma, K. F., Braak, R., and van der A, R. J.: Dutch OMI NO$_2$ (DOMINO) data product v2.0: HE5 data file user manual, available at: http://www.temis.nl/docs/OMI_NO2_HE5_2.0_2011.pdf (last access: 21 December 2012), 2011.

Brandt, J., Silver, J., Frohn, L., Geels, C., Gross, A., Hansen, A., Hansen, K., Hedegaard, G., Skjøth, C., Villadsen, H., Zare, A., and Christensen, J.: An integrated model study for Europe and North America using the Danish Eulerian Hemispheric Model with focus on intercontinental transport of air pollution, Atmos. Environ., 53, 156–176, 2012.

Carmichael, G. R., Sandu, A., Chai, T. F., Daescu, D. B., Constantinescu, E. M., and Tang, Y. H.: Predicting air quality: Improvements through advanced methods to integrate models and measurements, J. Comput. Phys., 227, 3540–3571, 2008.

Christensen, J. H.: The Danish Eulerian hemispheric model – a three-dimensional air pollution model used for the arctic, Atmos. Environ., 31, 4169–4191, 1997.

Dardub, D. and Seinfeld, J. H.: Numerical advective schemes used in air quality models – sequential and parallel implementation, Atmos. Environ., 28, 3369–3385, 1994.

Dee, D. P.: Bias and data assimilation, Q. J. Roy. Meteorol. Soc., 131, 3323–3343, 2005.

Dee, D. P. and da Silva, A.: Data assimilation in the presence of forecast bias, Q. J. Roy. Meteorol. Soc., 124, 269–295, 1998.

Dunlea, E. J., Herndon, S. C., Nelson, D. D., Volkamer, R. M., San Martini, F., Sheehy, P. M., Zahniser, M. S., Shorter, J. H., Wormhoudt, J. C., Lamb, B. K., Allwine, E. J., Gaffney, J. S., Marley, N. A., Grutter, M., Marquez, C., Blanco, S., Cardenas, B., Retama, A., Ramos Villegas, C. R., Kolb, C. E., Molina, L. T., and Molina, M. J.: Evaluation of nitrogen dioxide chemiluminescence monitors in a polluted urban environment, Atmos. Chem. Phys., 7, 2691–2704, doi:10.5194/acp-7-2691-2007, 2007.

Elbern, H., Strunk, A., Schmidt, H., and Talagrand, O.: Emission rate and chemical state estimation by 4-dimensional variational inversion, Atmospheric Chemistry and Physics, 7, 3749–3769, doi:10.5194/acp-7-3749-2007, 2007.

Ellermann, T., Nordstrøm, C., Brandt, J., Christensen, J. H., Ketzel, M., and Jensen, S. S.: The Danish Air Quality Monitoring Programme, Annual Summary for 2009, NERI Technical Report 799, National Environmental Research Institute, Aarhus University, available at: http://www2.dmu.dk/Pub/FR799.pdf (last access: 21 December 2012), 2010.

Eskes, H. J. and Boersma, K. F.: Averaging kernels for DOAS total-column satellite retrievals, Atmos. Chem. Phys., 3, 1285–1291, doi:10.5194/acp-3-1285-2003, 2003.

Forester, C. K.: Higher order monotonic convective difference schemes, J. Comput. Phys., 23, 1–22, 1977.

Frohn, L. M., Christensen, J. H., and Brandt, J.: Development of a high-resolution nested air pollution model, J. Comput. Phys., 179, 68–94, 2002.

Frydendall, J., Brandt, J., and Christensen, J. H.: Implementation and testing of a simple data assimilation algorithm in the regional air pollution forecast model, DEOM, Atmos. Chem. Phys., 9, 5475–5488, doi:10.5194/acp-9-5475-2009, 2009.

Grell, G. A., Dudhia, J., and Stauffer, D. R.: A description of the fifth-generation Penn State/ NCAR mesoscale model (MM5), NCAR Technical Note NCAR/TN-398+STR, National Center for Atmospheric Research, Boulder, Colorado, USA, 1995.

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

Houtekamer, P. L. and Mitchell, H. L.: A sequential ensemble Kalman filter for atmospheric data assimilation, Mon. Weather Rev., 129, 123–137, 2001.

Huijnen, V., Eskes, H. J., Poupkou, A., Elbern, H., Boersma, K. F., Foret, G., Sofiev, M., Valdebenito, A., Flemming, J., Stein, O., Gross, A., Robertson, L., D'Isidoro, M., Kioutsioukis, I., Friese, E., Amstrup, B., Bergstrom, R., Strunk, A., Vira, J., Zyryanov, D., Maurizi, A., Melas, D., Peuch, V.-H., and Zerefos, C.: Comparison of OMI NO$_2$ tropospheric columns with an ensemble of global and European regional air quality models, Atmos. Chem. Phys., 10, 3273–3296, doi:10.5194/acp-10-3273-2010, 2010.

Ionov, D., Goutail, F., Pommereau, J.-P., Bazureau, A., Kyro, E., Portafaix, T., Held, G., Ericksen, P., and Dorokhov, V.: Ten Years of NO$_2$ Comparisons Between Ground-Based SAOZ and Satellite Instruments (GOME, Sciamachy, OMI), in: Proceedings of Atmospheric Science Conference, edited by Lacoste, H. and Ouwehand, L., no. ESA SP-628 in European Space Agency, ESRIN, Frascati Italy, 2006.

Jacobson, M. Z.: Fundamentals of atmospheric modelling, Cambridge University Press, Cambridge, UK, 2005.

Janjic, Z. I.: The step-mountain Eta coordinate model: Further developments of the convection, viscous sublayer, and turbulence closure schemes, Mon. Weather Rev., 122, 927–945, 1994.

Kahnert, M.: Variational data analysis of aerosol species in a regional CTM: background error covariance constraint and aerosol optical observation operators, Tellus, 60B, 753–770, 2008.

Kalnay, E.: Atmospheric modelling, data assimilation and predictability, Cambridge University Press, Cambridge, United Kingdom, 2003.

Kunzi, K., Bauer, P., Eresmaa, R., Eriksson, P., Healy, S. B., Mugnai, A., Livesey, N., Priget, C., Smith, E. A., and Stephens, G.: Microwave absorption, emission and scattering: trace grases and meteorological parameters, in: The remote sensing of tropospheric composition from space, edited by: Burrows, J. P., Platt, U., and Borell, P., Springer, Heidelberg, Germany, Chapter 4, 153–230, 2011.

Lahoz, W. A., Errera, Q., Swinbank, R., and Fonteyn, D.: Data assimilation of stratospheric constituents: a review, Atmos. Chem.

Phys., 7, 5745–5773, doi:10.5194/acp-7-5745-2007, 2007.

Lamarque, J.-F., Bond, T. C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C., Mieville, A., Owen, B., Schultz, M. G., Shindell, D., Smith, S. J., Stehfest, E., Van Aardenne, J., Cooper, O. R., Kainuma, M., Mahowald, N., McConnell, J. R., Naik, V., Riahi, K., and van Vuuren, D. P.: Historical (1850–2000) gridded anthropogenic and biomass burning emissions of reactive gases and aerosols: methodology and application, Atmos. Chem. Phys., 10, 7017–7039, doi:10.5194/acp-10-7017-2010, 2010.

Lamsal, L. N., Martin, R. V., van Donkelaar, A., Steinbacher, M., Celarier, E. A., Bucsela, E., Dunlea, E. J., and Pinto, J. P.: Ground-level nitrogen dioxide concentrations inferred from the satellite-borne Ozone Monitoring Instrument, J. Geophys. Res., 113, doi:10.1029/2007JD009235, 2008.

Lanser, D. and Verwer, J.: Analysis of operator splitting for advection-diffusion-reaction problems from air pollution modelling, J. Comput. Appl. Math., 111, 201–216, 1999.

Mallet, V. and Sportisse, B.: Uncertainty in a chemistry-transport model due to physical parameterizations and numerical approximations: An ensemble approach applied to ozone modeling, J. Geophys. Res. Atmos., 111, 01302, doi:10.1029/2005JD006149, 2006.

Martin, R. V.: Satellite remote sensing of surface air quality, Atmos. Environ., 42, 7823–7843, 2008.

Matsui, T., Kreidenweis, S. M., Pielke Sr., R. A., Schichtel, B., Yu, H., Chin, M., Chu, D. A., and Niyogi, D.: Regional comparison and assimilation of GOCART and MODIS aerosol optical depth across the eastern U.S., Geophys. Res. Lett., 31, L21101, doi:10.1029/2004GL021017, 2004.

Meirink, J. F., Eskes, H. J., and Goede, A. P. H.: Sensitivity analysis of methane emissions derived from SCIAMACHY observations through inverse modelling, Atmos. Chem. Phys., 6, 1275–1292, doi:10.5194/acp-6-1275-2006, 2006.

Ménard, R. and Chang, L.-P.: Assimilation of stratospheric chemical tracer observations using a Kalman filter, Part II: χ^2-validated results and analysis of variance and correlation dynamics, Mon. Weather Rev., 128, 2672–2686, 2000.

Mok, K. M., Miranda, A. I., Leong, K. U., and Borrego, C.: A Gaussian puff model with optimal interpolation for air pollution modelling assessment, Int. J. Environ. Pollut., 35, 111–137, 2008.

Mol, W. J. A. and de Leeuw, F. A. A. M.: AirBase: a valuable tool in air quality assessments, in: Proceedings of the 5th International Conference on Urban Air Quality, edited by: Sokhi, R. S., Millán, M. M., and Moussiopoulos, N., Valencia, Spain, 2005.

Morton, K. W. and Mayers, D. F.: Numerical solution of partial differential equations, Cambridge University Press, 2nd Edn., 2005.

Parrish, D. and Derber, J. C.: The National Meteorological Center's spectral statistical interpolation analysis system, Mon. Weather Rev., 120, 1747–1763, 1992.

Schaub, D., Brunner, D., Boersma, K. F., Keller, J., Folini, D., Buchmann, B., Berresheim, H., and Staehelin, J.: SCIAMACHY tropospheric NO$_2$ over Switzerland: estimates of NO$_x$ lifetimes and impact of the complex Alpine topography on the retrieval, Atmos. Chem. Phys., 7, 5971–5987, doi:10.5194/acp-7-5971-2007, 2007.

Schultz, M., Rast, S., van het Bolscher, M., Pulles, T., Brand, R., Pereira, J., Mota, B., Spessa, A., Dalsøren, S., van Nojie, T., and Szopa, S.: Emission data sets and methodologies for estimating emissions, RETRO project report D1-6, 2007.

Searl, A.: A review of the acute and long term impacts of exposure to nitrogen dioxide in the United Kingdom, Research Report TM/04/03, Institute of Occupational Medicine, Edinburgh, United Kingdon, 2004.

Seinfeld, J. H. and Pandis, S. N.: Atmospheric chemistry and physics: from air pollution to climate change, Wiley, 2nd Edn., 2006.

Simpson, D., Fagerli, H., Jonson, J. E., Tsyro, S., Wind, P., and Tuovinen, J.-P.: The EMEP Unified Eulerian Model, Model Description,, EMEP MSC-W Technical Report 1/2003, Norwegian Meteorological Institute, Oslo, Norway, 2003.

Stewart, L. M., Dance, S. L., and Nichols, N. K.: Correlated observation errors in data assimilation, Int. J. Num. Method. Fluids, 56, 1521–1527, 2008.

Vautard, R., Moran, M. D., Solazzo, E., Gilliam, R. C., Matthias, V., Bianconi, R., Chemel, C., Ferreira, J., Geyer, B., Hansen, A. B., Jericevic, A., Prank, M., Segers, A., Silver, J. D., Werhahn, J., Wolke, R., Rao, S., and Galmarini, S.: Evaluation of the meteorological forcing used for the Air Quality Model Evaluation International Initiative (AQMEII) air quality simulations, Atmos. Environ., 53, 15–37, doi:10.1016/j.atmosenv.2011.10.065, 2012.

Verwer, J. G., Blom, J. G., van Loon, M., and Spee, E. J.: A comparison of stiff ODE solvers for atmospheric chemistry problems, Atmos. Environ., 30, 49–58, 1996.

Vestreng, V. and Klein, H.: Emission data reported to UNECE/EMEP: Quality assurance and trend analysis & Presentation of WebDab, Msc-w status report, Norwegian Meteorological Institute, Oslo Norway, 2002.

Wang, X., Mallet, V., Berroir, J.-P., and Herlin, I.: Assimilation of OMI NO2 retrievals into a regional chemistry-transport model for improving air quality forecasts over Europe, Atmos. Environ., 45, 485–492, 2011.

Wang, Y., McElroy, M. B., Boersma, K. F., Eskes, H. J., and Veefkind, J. P.: Traffic restrictions associated with the Sino-African summit: Reductions of NO$_x$ detected from space , Geophys. Res. Lett., 34, L08814, doi:10.1029/2007GL029326, 2007.

Wu, L., Mallet, V., Bocquet, M., and Sportisse, B.: A comparison study of data assimilation algorithms for ozone forecasts, J. Geophys. Res., 113, D20310, doi:10.1029/2008JD009991, 2008.

Zare, A., Christensen, J. H., Irannejad, P., and Brandt, J.: Evaluation of two isoprene emission models for use in a long-range air pollution model, Atmos. Chem. Phys., 12, 7399–7412, doi:10.5194/acp-12-7399-2012, 2012.

Zhao, C. and Wang, Y.: Assimilated inversion of NO$_x$ emissions over east Asia using OMI NO$_2$ column measurements, Geophys. Res. Lett., 36, L06805, doi:10.1029/2008GL037123, 2009.