



Regional scale ozone data assimilation using an ensemble Kalman filter and the CHIMERE chemical transport model

B. Gaubert¹, A. Coman¹, G. Foret¹, F. Meleux², A. Ung², L. Rouil², A. Ionescu³, Y. Candau³, and M. Beekmann¹

¹LISA, CNRS/INSU UMR7583, Université Paris-Est Créteil et Université Paris Diderot, Institut Pierre Simon Laplace, 94010 Créteil, France

²Institut National de l'Environnement Industriel et des Risques, INERIS, Parc Technologique ALATA-B.P. No. 2, 60550 Verneuil en Halatte, France

³Centre d'Études et de Recherche en Thermique, Environnement et Systèmes EA 3481, Université Paris-Est Créteil, 94010 Créteil, France

Correspondence to: B. Gaubert (benjamin.gaubert@lisa.u-pec.fr)

Received: 22 April 2013 – Published in Geosci. Model Dev. Discuss.: 30 May 2013

Revised: 25 December 2013 – Accepted: 7 January 2014 – Published: 13 February 2014

Abstract. An ensemble Kalman filter (EnKF) has been coupled to the CHIMERE chemical transport model in order to assimilate ozone ground-based measurements on a regional scale. The number of ensembles is reduced to 20, which allows for future operational use of the system for air quality analysis and forecast. Observation sites of the European ozone monitoring network have been classified using criteria on ozone temporal variability, based on previous work by Flemming et al. (2005). This leads to the choice of specific subsets of suburban, rural and remote sites for data assimilation and for evaluation of the reference run and the assimilation system. For a 10-day experiment during an ozone pollution event over Western Europe, data assimilation allows for a significant improvement in ozone fields: the RMSE is reduced by about a third with respect to the reference run, and the hourly correlation coefficient is increased from 0.75 to 0.87. Several sensitivity tests focus on an a posteriori diagnostic estimation of errors associated with the background estimate and with the spatial representativeness of observations. A strong diurnal cycle of both these errors with an amplitude up to a factor of 2 is made evident. Therefore, the hourly ozone background error and the observation error variances are corrected online in separate assimilation experiments. These adjusted background and observational error variances provide a better uncertainty estimate, as verified by using statistics based on the reduced centered random variable. Over the studied 10-day period the overall EnKF performance over evaluation stations is found relatively unaffected by different formulations of observation

and simulation errors, probably due to the large density of observation sites. From these sensitivity tests, an optimal configuration was chosen for an assimilation experiment extended over a three-month summer period. It shows a similarly good performance as the 10-day experiment.

1 Introduction

Tropospheric ozone plays a major role in air pollution due to its impact on human health and vegetation growth (WHO, 2003; Felzer et al., 2004). Ozone as a strong oxidant affects the human respiratory system and is associated with a risk of premature mortality (Bell et al., 2005). Cumulative ozone uptake through leaf stomata over a given threshold causes injury to vegetation (Fowler et al., 2009). Among other tasks, the GMES (Global Monitoring for Environment and Security) programs foster the development of environmental monitoring of ozone and other pollutants using a combination of state-of-the-art numerical models and in situ and space-borne observations. In this framework, the European project GEMS (Global and regional Earth-system Monitoring using Satellite and in-situ data) and the follow-up projects MACC and MACC II (Monitoring Atmospheric Composition and Climate, <http://www.gmes-atmosphere.eu/>) promoted and continue to promote monitoring of atmospheric constituents from global to regional scale at a high spatio-temporal resolution (Hollingsworth et al., 2008).

Operational forecasting systems of regional air quality are generally based on modeling platforms in synergy with observations for evaluation, model post-processing and/or data assimilation. A review of such platforms has recently been performed by Kukkonen et al. (2012). They are often based on regional chemical transport models (rCTMs) whose deterministic forecasts are driven in real time by an offline or online numerical weather forecast model providing meteorological fields and by global CTMs for chemical boundary conditions. Among these air quality forecasting systems, the PREV' AIR platform (www.prevoir.org, Rouil et al., 2009) delivers daily air quality forecasts of pollutant concentrations such as ozone, nitrogen oxides and particulate matter. This computational chain involves the rCTM CHIMERE (Schmidt et al., 2001; Bessagnet et al., 2004; Menut et al., 2013) on the continental scale (Europe) and over a French domain. The regional scale simulations based on the CHIMERE model have been widely evaluated against measurements and give satisfactory results, particularly to simulate ozone peaks (Honoré et al., 2008). In addition to model simulations, analyses resulting from data assimilation of observations in near real time provide a better representation of the surface pollutant concentrations.

One of the challenges of these air quality modeling chains is to provide uncertainties and errors associated with the modeling results. These uncertainties can be estimated by the comparison of simulations obtained by different models and the same sets of observations. For instance, model simulations have been compared against ozone surface observations in Europe (Vautard et al., 2007, 2009) and in the United States (Solazzo et al., 2012b) in the context of the AQMEII (Air Quality Model Evaluation International Initiative). Ensembles of model simulations and their statistical combinations have also been evaluated against satellite observations for tropospheric NO₂ (Huijnen et al., 2010) and O₃ (Zyryanov et al., 2012). These studies helped to identify if uncertainties given by the ensemble spread can represent the error distribution for instance in terms of geographical and temporal patterns.

Data assimilation methods, which consist in the integration of chemical observations in the simulations, are now recognized as crucial in the air quality community (Carmichael et al., 2008; Zhang et al., 2012). After pioneering work in numerical weather predictions such as the 4D-Var (Talagrand and Courtier, 1987; Courtier and Talagrand, 1987), assimilation methods already developed for meteorology have been successfully applied to air quality simulations and were focused in particular on ozone pollution. These are optimal interpolation (OI, Blond et al., 2003), the 4D-Var (Elbern et al., 1997), the ensemble Kalman filter and the reduced rank square root filter (RRSQRT, Van Loon et al., 2000; Hanea et al., 2004). These data assimilation methods can also be used to improve the short-term forecast (Blond and Vautard, 2004; Elbern and Schmidt, 2001). The two key points of the data assimilation process are the representation on the one

hand of the background error covariance matrix (BECM) and on the other hand of the observation error covariance matrix (OECM). The relative values of the BECM and of the OECM in the observation space allow weighting of the confidence between observations and background estimate. In addition, the BECM matrix will serve to propagate the innovations (observed value minus background one) spatially. The principal origin of uncertainties in rCTMs are model parameterizations (chemistry, transport, deposition) and input data, which include emissions, meteorological fields and chemical boundary conditions (Beekmann and Derognat, 2003; Mallet and Sportisse, 2006). Thus, the sensitivity to these factors and the fact that some species are transported out of the limited model domain reduce the importance of initial conditions in the overall error budget (Blond and Vautard, 2004; Sandu and Chai, 2011). In order to obtain an accurate 4-D analysis, a strategy consists in the correction of indirect or possibly unobserved quantities such as emissions rates or even wind fields using variational (Elbern et al., 2007; Semane et al., 2009) or sequential methods (Brunner et al., 2012; Miyazaki et al., 2012) such as Kalman filters. However, if surface ozone observations are assimilated into a one-hour frequency, this reduces the importance of the error growth during the forecast.

In this study, an ensemble Kalman filter (EnKF) method is employed to correct the near surface ozone fields directly. To do this, background and observation errors need to be estimated. An ensemble of simulations following a Monte Carlo approach is used to determine a flow-dependent BECM matrix. In former studies, improving the background error representation greatly helped in increasing the analysis accuracy and the forecast performance of the EnKF (Constantinescu et al., 2007a, b; Wu et al., 2008; Agudelo et al., 2011; Tang et al., 2011; Curier et al., 2012). One approach employed to construct a representative background ensemble was to perturb the main model parameters affecting the ozone error variance and correlation (Hanea et al., 2004). Another strategy employed in data assimilation is to adjust the BECM iteratively using diagnostics such as Desroziers diagnostics, which derive background errors intrinsically from the assimilation procedure (Desroziers et al., 2005; Schwinger and Elbern, 2010).

The rCTM CHIMERE that is used in this study has already been successfully coupled to a local EnKF square root scheme (Evensen, 2004) in order to assimilate tropospheric ozone columns derived from the IASI instrument (Coman et al., 2012). In this paper, we assess the setup of the CHIMERE-ENKF in the context of the assimilation of surface ozone data. We have built a consistent ensemble of assimilation and evaluation stations for a summertime episode (Fig. 1). According to the model resolution, the observation error is mainly due to the site representativeness for ozone, which depends on space and time. The observations data set can be thinned a priori according to a rigorous classification (Curier et al., 2012) and/or estimated in the assimilation

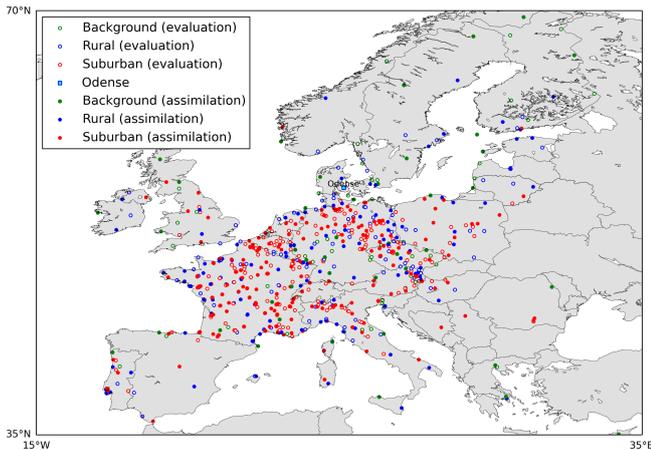


Fig. 1. Modeled domain with sites retained for assimilation (filled circles) and validation (filled in white circles). Colors indicate the station type with red for suburban, blue for rural and green for background/remote stations. The rural evaluation station for which the ozone concentration time series is plotted in Fig. 6 (located in the city of Odense in Denmark) is shown as a square filled in cyan.

procedure following a parameterization dependent on station types (Elbern et al., 2007). Then a first focus of this paper is to assess the impact of the observation representativeness on the assimilation performance. Following Fleming et al. (2005) (hereafter denoted FLEM05), we have performed a classification of the ozone stations that give a qualitative estimation of spatial representativeness. A second focus is to investigate different formulations and diagnostics of the model and observation error and their impact on the assimilation skills. In particular, we study the diurnal variation of these errors. We investigate sensitivities to the BECM properties by perturbing both model parameters and the ozone state and by using the Desroziers diagnostic to estimate and tune the covariance inflation factor (Li et al., 2009b). We also diagnose the observations error variance to evaluate its temporal variation as a function of the observation types defined by the classification. In addition, we have tested an alternative way of the prescription of the OEMC by using the Desroziers diagnostic. We compare the performance of the assimilation system using these different errors formulations for a 10-day simulation over Europe including a photochemical episode. The ensemble Kalman filter assimilation setup, including a description of the rCTM CHIMERE and the a posteriori diagnostics is described in Sect. 2. We present the classification of the surface ozone observations in Sect. 3. Section 4 shows the evaluation of the CHIMERE reference run. Different assimilation experiments of a short period are presented in Sect. 5, while an ozone analysis for a longer summer period is evaluated in Sect. 6. The conclusion and future directions are given in Sect. 7.

2 The CHIMERE–EnKF data assimilation setup and diagnostics

2.1 The EnKF algorithm

We implemented the EnKF that was first introduced by Evensen (1994). This sequential filter allows a relatively simple implementation of a sophisticated data assimilation scheme appropriate for a large three-dimensional model such as CHIMERE. We created an ensemble of $N = 20$ perturbed model states using a Monte Carlo method. They evolve forward in time in order to obtain a forecast $\mathbf{x}_{i,k}^f$ from the time step $(k - 1)$ to the time step k .

$$\mathbf{x}_{i,k}^f = M(\mathbf{x}_{i,k-1}^a + \mathbf{q}_{i,k-1}\mathbf{x}_{i,k-1}^a) \quad (1)$$

where $\mathbf{q}_{i,k-1} = sd * \boldsymbol{\eta}_{i,k-1}$.

Here $\mathbf{x}_{i,k}^f$ represents the vector of forecasted ozone concentrations where the subscript i indicates the ensemble number. Ozone concentrations at the next time step are simulated using the CHIMERE model M . The noise $\mathbf{q}_{i,k-1}$ is the product of spatially correlated fields $\boldsymbol{\eta}_{i,k-1}$ and a tunable coefficient with a relative standard deviation (SD). Pseudo-random fields ($\boldsymbol{\eta}_{i,k-1}$) are derived from a two-dimensional Gaussian distribution with some fixed characteristics, namely zero mean and unitary variance (Evensen, 1994, 2003) and a fixed horizontal decorrelation length of 200 km (Boynard et al., 2011; Coman et al., 2012). This parameter is close to the value of 270 km used in several other studies (Chai et al., 2007; Constantinescu et al., 2007c; Frydendall et al., 2009) and in any case our results are similar to both values. These perturbations are added to the ozone fields after the analysis step. As suggested in Sandu and Chai (2011), the same noise is applied for all vertical layers inside the calculated boundary layer and thus induces a vertical correlation in the background error. The ensemble mean value over the N ensemble members is defined in Eq. (2):

$$\bar{\mathbf{x}}_k^f = \frac{1}{N} \sum_{i=1}^N \mathbf{x}_{i,k}^f. \quad (2)$$

At the analysis step, the BECM is approximated by the ensemble spread over the N realizations of the model at a given time:

$$\mathbf{P}_k^f = \frac{1}{N} \sum_{i=1}^N (\mathbf{x}_{i,k}^f - \bar{\mathbf{x}}_k^f)(\mathbf{x}_{i,k}^f - \bar{\mathbf{x}}_k^f)^T. \quad (3)$$

Then, measurements \mathbf{y}_k are available along with the OEMC-noted \mathbf{R} ; each ensemble member is updated following Eq. (4):

$$\mathbf{x}_{i,k}^a = \mathbf{x}_{i,k}^f + \mathbf{P}_k^f H_k^T (H_k \mathbf{P}_k^f H_k^T + \mathbf{R}_k)^{-1} (\mathbf{y}_k - H \mathbf{x}_{i,k}^f), \quad (4)$$

where H is a projection operator from the model space to the observation space. In our case, it is a bi-linear interpolation of the closest model grid cells values onto the observation location. Equation (5) yields an analysis model state

$\mathbf{x}_{i,k}^a$, which is used for the initialization of the next forecast. The background error covariance matrix is also updated at the analysis step using the following formula:

$$\mathbf{P}_k^a = \mathbf{P}_k^f - \mathbf{P}_k^f \mathbf{H}_k^T \left(\mathbf{H}_k \mathbf{P}_k^f \mathbf{H}_k^T + \mathbf{R}_k \right)^{-1} \mathbf{H}_k \mathbf{P}_k^f. \quad (5)$$

In an ensemble Kalman filter algorithm, the analysis error statistics are given through the analysis procedure. One of the drawbacks of this method is the introduction of sampling errors due to the limited ensemble size. This leads to an underestimation of the analysis errors caused by an artificial decrease of the ensemble variance in Eq. (5) and to spurious correlation in the analyzed fields. This makes it necessary to localize spatially the analysis, in order to prevent long-range spurious correlation (Houtekamer and Mitchell, 2001; Hamill et al., 2001). The first approach to achieve this covariance localisation is based on the introduction of a distance correlation in the BECM or in the observation error covariance matrix (OECM) that smoothes the gain progressively to zero when distances between observations and model grid cells increase. In our study we use a second method, which is known as local analysis. Only observations within a fixed area around the analyzed cell are assimilated. Despite their different algorithms, covariance localization and local analysis yield similar results (Sakov and Bertino, 2010). Thus, at each analysis step and for each grid cell, we perform a local analysis with a horizontal cut-off radius of 250 km around the analyzed cell. In the vertical, levels completely or partially located within the boundary layer are included. These choices are made to avoid spurious correlation, as the ensemble size is quite low. Only three-dimensional fields for ozone are included in the state vector at the analysis step.

There are two types of algorithms to solve Eqs. (4) and (5): the stochastic EnKF, where observations are treated as random variables (Burgers et al., 1998; Houtekamer and Mitchell, 1998; Evensen, 2003) and several forms of deterministic EnKF that use a square root decomposition of the background error covariance matrix (Whitaker and Hamill, 2002; Tippett et al., 2003; Evensen, 2004; Hunt et al., 2007, and reference therein). These square root filters avoid the need to apply measurement perturbations and thus have a lower analysis error by reducing this additional source of sampling errors.

2.2 A posteriori diagnostics and error modeling

In an EnKF, the background error is sampled by an ensemble of model realizations, which has the advantage of evolving in time (in contrast to a static BECM associated with OI methods). As mentioned above, with finite and generally small ensemble sizes, and due to significant model errors, the EnKF generally underestimates the analysis error covariance matrix. Also, errors are increased due to the model error during the forecast step. Thus, a particular strategy must be employed to inflate the ozone perturbations

(Eq. 1) by treating model and analysis error contributions in the same framework. Finally, the ensemble design must reflect the background error and generate adequate error correlations. This goal is difficult to achieve, as processes affecting model error sources are often not well known and/or estimated. Because of the chaotic structure of the atmosphere and the ocean dynamics, the errors were first naturally represented by a set of perturbed initial conditions (Evensen, 1994). One way to build the ensemble in the EnKF framework consists of applying a perturbation to the state vector, here the ozone concentration fields. Furthermore, the model error can be approximated by the perturbation of physical parameters and uncertain inputs of the model to set up a more physically sound model ensemble, as initiated by the work of Hanea et al. (2004). This approach is also tested in our study, but the noise is not updated during the analysis step. For the CHIMERE model, following the work of Boynard et al. (2011), we add stochastic perturbations to the uncertain inputs, namely anthropogenic and biogenic emissions, the boundary conditions, the land use for dry deposition, and meteorological variables (Table 1).

Concerning ozone assimilation with an EnKF, Constantinescu et al. (2007b) have shown that the most efficient way to maintain a sufficiently dispersive ensemble was to apply an additive perturbation for the covariance inflation. Rather than just evaluating the consistency of the background error weight in the analysis, we propose here also to use the Desroziers et al. (2005) diagnostics to derive the background error variance. In practice, diagnostics are computed in the observation space over p observations following Eq. (6) for the background error:

$$(\sigma_b)^2 = \frac{1}{p} \sum_{i=1}^p \left(y_i^a - y_i^f \right) \left(y_i^o - y_i^f \right). \quad (6)$$

Because they require the analyzed ensemble mean $\mathbf{y}^a = \mathbf{H}\bar{\mathbf{x}}^a$ (in the observation space), and the observations y_i^o and the forecasted ensemble mean $\mathbf{y}^f = \mathbf{H}\bar{\mathbf{x}}^f$ (in the observation space), these quantities have to be computed a posteriori. The observation error can also be diagnosed following Eq. (7):

$$(\sigma_o)^2 = \frac{1}{p} \sum_{i=1}^p \left(y_i^o - y_i^a \right) \left(y_i^o - y_i^f \right). \quad (7)$$

The difference in these two equations is represented by the first term of the product. Diagnosed background or observation errors will be small if the analysis estimate is close to the background or observations, respectively. The background error variance has been diagnosed in a variational framework for the assimilation of total ozone columns in global CTMs (Massart et al., 2012; Schwinger and Elbern, 2010) and surface ozone observations (Jaumouillé et al., 2012). In Li et al. (2009b), the suggested tuning strategy of the covariance inflation factor in EnKF consists in adaptively adjusting the model error standard deviation prescription. The ability of

Table 1. Description of the model input field perturbation. Uncertainties are modeled by a log-normal perturbation with a prescribed standard deviation and a fixed spatial–temporal (τ) decorrelation length.

Parameter	Standard deviation	Decorrelation length	T
Boundary conditions			
Top	30 %	800 km	6 h
Lateral	30 %	3 km \times 500 km	6 h
Dry deposition (9 land-use types)	25 %	4000 km	6 h
Anthropogenic emissions	50 %	800 km	6 h
Vertical diffusion coefficient	40 %	800 km	6 h
Attenuation coefficient	20 % to 100 %	800 km	6 h
Photolytic rate	20 % (no clouds)		
Biogenic emissions	30 %	800 km	24 h
Temperature	1 %	800 km	6 h

the tuned ensemble to represent more accurate error statistics has been demonstrated in assimilation exercises that take into account different ranges of true background errors (Li et al., 2009b). However, this error estimation requires (as the whole assimilation procedure) that the forecast model mean bias is small in practise. Besides, an accurate prior estimate of these observations and background error variances allows their simultaneous estimation, although the solution is not unique (Li et al., 2009b; Schwinger and Elbern, 2010).

In addition, we propose to use an appropriate diagnostic of the ensemble design, useful to compare the different experiments, which is the reduced centered random variable (RCRV) diagnostic (Candille et al., 2007). For each observation i of the system, the RCRV is defined by the ratio between the innovation and its associated error:

$$\text{RCRV} = \frac{y_i^o - H\bar{x}^f}{\sqrt{\sigma_o^2 + H\sigma_b^2}}. \quad (8)$$

Innovations are the differences between the observations and the forecasted ensemble mean at the observation location. Errors are estimated by the square root of the sum of observation error variance (σ_o^2) and background error variance (represented by the ensemble variance, σ_b^2). Here the observation error variance (σ_o^2) is set to 25 ppb² for all experiments (cf. Sect. 5.1); background ensemble mean and variance (σ_b^2) are calculated in the observation space. For a representative ensemble (i.e., when errors variances are comparable to the innovations), the RCRV should be normally distributed with a zero mean and a standard deviation of 1. The mean of the RCRV indicates the weighted bias of the one-hour forecast. It is important in an EnKF framework, since systematic errors are not taken into account in the analysis formulation. However, the analysis allows globally an initialization of the ensemble forecast from an unbiased ozone state.

2.3 The CHIMERE regional chemistry transport model

CHIMERE is a state-of-the-art rCTM (www.lmd.polytechnique.fr/chimere/; Menut et al., 2013). The general formulation and the first evaluation on a regional scale were presented by Schmidt et al. (2001). Our simulations cover the European continental domain (Fig. 1) for eight hybrid (σ , p) vertical levels from 995 hPa to 500 hPa (the height at the top of each box is on average: 42 m, 115 m, 240 m, 455 m, 838 m, 1520 m, 2820 m and 5500 m). To reduce computational time, we worked on the $0.5^\circ \times 0.5^\circ$ horizontal grid spacing; also, aerosols are not included in the simulation. Meteorological variables are obtained at $0.25^\circ \times 0.25^\circ$ resolution from the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF). Analyses at 00:00 UT and 12:00 UT and three-hourly output forecasts are linearly interpolated on an hourly basis and bi-linearly interpolated on the CHIMERE spatial domain. The PBL height parameterization is described in Menut et al. (2013); in the stable case, it is diagnosed following a K diffusion approach (Troen and Mahrt, 1986) and a thermal plume approach is used in an unstable case (Cheinet and Teixeira, 2003). Biogenic emissions are calculated using the MEGAN model (Guenther et al., 2006; Curci et al., 2009) and hourly anthropogenic emission fluxes are derived from the TNO (<http://www.tno.nl/>) inventory (Visschedijk et al., 2007). The MOZART 3.5 (Kinnison et al., 2007) global chemical transport model running at $1.875^\circ \times 1.875^\circ$ has been coupled to IFS (Flemming et al., 2009) and provides global trace gas compositions on an hourly basis for O₃, CO, HCHO, NO_x and SO₂. In this way, we obtain three-hourly boundary conditions that are temporally and spatially interpolated on the CHIMERE grid.

3 Characterization of the surface observations

Ozone observations used in this study are operated by national and regional networks across Europe and collected through the European Environment Agency (EEA) in the Airbase database (<http://acm.eionet.europa.eu/databases/airbase/>). A classification of ozone measurement sites corresponding to their representativeness is crucial for their use within the data assimilation procedure (cf. Sect. 5.1), either for assimilating observed information (assimilation sites) or for evaluating analyzed fields (validation sites). Rather than relying on a station type classification of EEA (see for instance the European Union directive 2008/50/CE), we derive here a classification defined by a specific criterion related to ozone concentrations. The area of representativeness associated with a station depends on the chemical and physical sources and sinks of the compound of interest. This area can be estimated knowing the surrounding emission intensity and atmospheric and surface fluxes such as deposition rates and vertical mixing (Henne et al., 2010). This spatial representativeness can be characterized using metadata such as population density, land cover, emission inventories, topography and transport model results (Tarasova et al., 2007; Spangl et al., 2007; Henne et al., 2010). Another, much simpler approach is based on the identification of the pollution regime following statistics of the ozone concentrations themselves (FLEM05, Joly and Peuch, 2012). In this paper, we choose this approach because it has the advantage of being directly targeted on the pollutant of interest (ozone) and because it does not require any other metadata. The FLEM05 classification is based on only two criteria, which are the yearly median (P50) of the daily average (P50DA) and of the daily variability (P50DV), which is the difference between the daily maximum and minimum divided by the daily average (DA). Based only on the summertime P50DV statistics (Table 1 in FLEM05), we derive four station types from the entire set of available background Airbase stations with an altitude below 800 m above sea level (a.s.l.). These are remote/mountain stations ($P50DV < 0.68$), rural stations ($0.68 < P50DV < 1.07$), suburban stations (namely “U1” in FLEM05, $1.07 < P50DV < 1.45$) and urban stations (corresponding to “U2”, “U3” and street (“S”) classes in FLEM05, $P50DV > 1.45$). According to this method, we find that the ozone hourly mean over each station type decreases when variability increases (Table 2). In the FLEM05 classification, generally high levels of NO₂ and low levels of ozone are found in densely urbanized areas and the inverse is true in rural and remote areas. This feature will be important later in the paper when we will use different types of stations for the validation of the assimilation procedure. Urban stations that have the highest daily variability are considered as not representative of the model grid spacing used for this study ($0.5^\circ \times 0.5^\circ$); we will neither use these data for evaluation nor for assimilation. The general behavior of the ozone mean and variability is close to the findings of FLEM05 based

on a set of German stations (Fig. 2 and Table 2). We plotted in Fig. 2 the daily ozone profile obtained for the summer of 2009 averaged over each type of observation. Compared to the yearly average, the summertime suburban profile shows a higher ozone concentration in the daytime while keeping the lowest nighttime ozone level, which is in accordance with the variability criterion. Thus, differences in average ozone concentrations between station types are more important in the early morning and are weakest at 15:00 UTC, when all observation types reach similar daily maxima. Observations associated with the lowest daily variability are mostly representative of a remote environment instead of a mountainous one because we reject stations located above 800 m height. Remote stations located between 300 m a.s.l. and 800 m a.s.l. exhibit a higher baseline in ozone concentrations probably associated with the tropospheric ozone vertical gradient (Chevalier et al., 2007). These stations are therefore discriminated from the remaining remote ones. The latter are shared between remote continental and coastal stations under the influence of generally rather clean marine air masses such as the Mace Head station in Ireland. Finally, the use of the P50DV statistics can be extended to the whole of Europe with some exceptions; for instance, we can notice that remote stations with low ozone variability that are located in Scandinavia do not have the highest daily mean, as usually observed for remote stations. Furthermore, we note that the geographical distribution of the stations is coherent with their attributed environment (Fig. 1); for example, suburban stations (red circles) are often located in high emission regions close to urban areas such as Paris, Berlin or in the Po Valley. Also, as seen from Table 2, results obtained with the FLEM05 approach are often consistent with the Airbase classification: around 90 % of the remote stations (FLEM05) are rural (Airbase) and more than 80 % of the suburban stations (FLEM05) are also “suburban” or “urban” in Airbase. However, discrepancies are found in the ozone categorization with respect to the standard classification. For instance, the effect of the urban environment on ozone concentrations for many “urban” Airbase sites is small enough so that these sites are still representative of a larger environment: these sites constitute around 40 % of our (FLEM05) total rural sites. Thus, the classification procedure applied here allows us to obtain a significantly larger observational database than the initial classification proposed by Airbase.

4 Selection of the simulation period and evaluation of the reference

Using the model configuration described above, a reference simulation is performed for the period running from mid-May to the end of August 2009. We evaluate the simulated ozone fields against the selected set of observations for June/July/August (JJA) 2009 including the period for which assimilation experiments are performed from 14 August to

Table 2. Number of stations, daily mean and median of the daily variability (P50DV) average over the different station types (lines) of the FLEM05 classification (remote, rural, suburban). The number of corresponding stations in the initial Airbase (right columns) classification (rural, suburban, and urban) is also indicated with the relative contribution (percentage) of the new station type.

	<i>N</i>	mean (ppb) (period: JJA 2009)	P50DV (period: JJA 2009)	Rural (<i>N</i> = 405)	Suburban (<i>N</i> = 256)	Urban (<i>N</i> = 391)
remote (> 300 m)	47	38.8	0.56	42 (89 %)	4 (9 %)	1 (2 %)
remote (< 300 m)	42	36.5	0.58	32 (76 %)	3 (7 %)	7 (17 %)
rural	228	33.2	0.9	98 (43 %)	40 (18 %)	90 (39 %)
suburban	376	31.6	1.17	63 (17 %)	70 (19 %)	243 (65 %)
urban	359	30.1	1.46	170 (47 %)	139 (39 %)	50 (14 %)

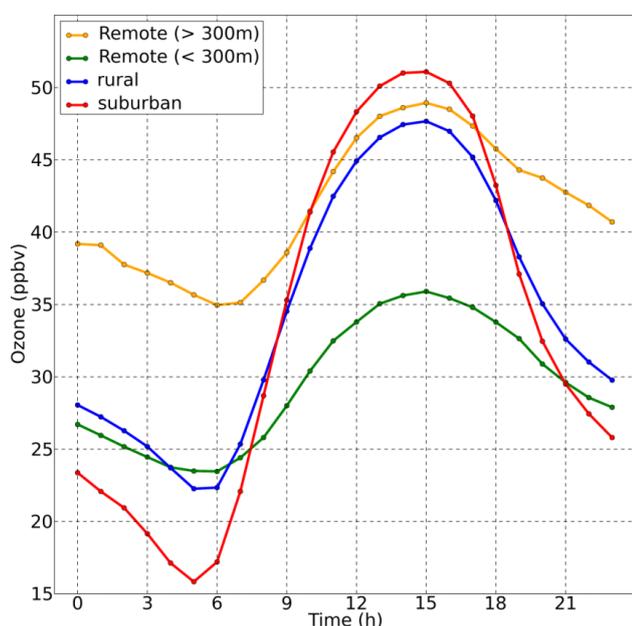


Fig. 2. Average daily ozone profile during JJA 2009 for each type of observation chosen for assimilation and validation (see ozone statistics in Table 2). Colors indicate the station type, with red for suburban, blue for rural, green for background stations under 300 m a.s.l. and orange for background stations between 300 m and 800 m a.s.l.

23 August 2009. During the latter period, anticyclonic conditions over central Europe followed by a low-pressure system led to a short ozone pollution episode between 19 and 21 August (see observations and the CHIMERE reference simulation in Fig. 5). Generally, the modeled ozone fields appropriately represent the synoptic pattern: lower values are simulated for westerly regimes when marine air masses flow into Europe and higher levels for periods of stagnation under anticyclonic conditions. In order to characterize the simulation accuracy in different environments, the evaluation is conducted separately for each station type using the classification presented above.

Figure 3 shows the average daily profile of the ozone observations, the simulations, and the associated root mean

square error (RMSE) for the JJA period. Statistics over the whole summer demonstrate that daily values are on average well reproduced by the model simulations. There is a minimum of the RMSE of about 8 ppb for all station types in the afternoon. The decreasing amplitude of the diurnal cycle of ozone concentrations from the suburban to the remote stations is also captured by the model. At suburban and rural stations, comparisons between observations and simulations show a good temporal correlation of the hourly values around 0.7–0.8 (Table 4); on average, observations are overestimated all day long, except mid-afternoon. These errors can be typically attributed to the model resolution in both the horizontal and vertical directions. It does not allow a good estimation of the subgrid processes such as vertical turbulent transport and spatial variability of anthropogenic emissions (Valari and Menut, 2010). This leads to an uncertain representation of night-time and early morning chemistry, especially ozone titration, and probably also dry deposition. The background ozone level (i.e., at the remote stations) is also well captured; average errors range between 8 and 10 ppb, but the correlation is lower (0.58). Simulation of remote stations above 300 m a.s.l. exhibits a much stronger positive bias during nighttime, as these stations are generally more representative of higher model layers which are less affected by dry deposition and NO emissions. Therefore, we also plot in Fig. 3 the average simulated ozone values and RMSE for the second level, which corresponds to a height varying between 115 m and 240 m. It shows that nighttime values are better simulated at the second level for both the ozone mean and RMSE.

5 Setup and results of assimilation experiments

In this section, we shall present the setup and the results of the assimilation experiments applied to the 10-day summertime period in August 2009 including an ozone pollution episode. A first setup will be presented and evaluated. Then, several sensitivity experiments will be presented with a focus on the background error (Sect. 5.3) and on the observation error (Sect. 5.4). A review of the different assimilation experiments is shown in Table 3.

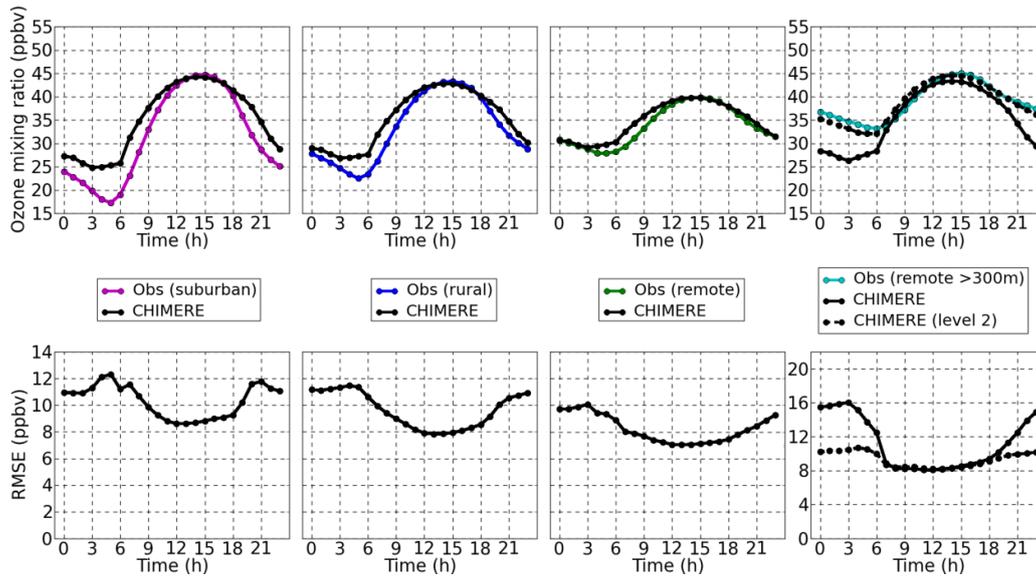


Fig. 3. Simulated (CHIMERE) and observed average diurnal ozone cycle and their associated RMSE during the 2009 summer. Results are shown for the suburban stations (left), rural stations (middle left), and for background stations located below (middle right) and above 300 m a.s.l. (right). The corresponding number of stations and statistics are indicated in Table 4.

Table 3. List of assimilation experiments with different formulations of OECM and BECM. For the OECM, the range of standard deviations is indicated; see Figs. 4 and 8 for an example of the noise and background error standard deviation profile.

	OECM	BECM
REF_ASSIM	5 ppb	fixed hourly noise profile
MOD_DESR	5 ppb	diagnosed hourly noise profile
NEWPAR	5 ppb	fixed noise profile and parameter perturbations
NEWPAR_MOD_DESR	5 ppb	diagnosed noise and parameter perturbations
TUNNING_OBS_TYPE	$f(\text{station type})$ (3–7 ppb)	fixed hourly noise profile

5.1 Setup of assimilation experiments

In order to increase the available data set and the spatial coverage of observations, we keep observations from all the three station types (i.e., remote, rural and suburban) both for assimilation and evaluation. Then, for each observation type, we randomly divide the station set into two subsets (Fig. 1). In this way, we get one set for the assimilation with 350 stations (remote: $N = 44$; rural: $N = 117$; suburban: $N = 189$) and another one for evaluation with 344 stations (remote: $N = 45$; rural: $N = 112$; suburban: $N = 187$). Finally the average nearest distance between two stations is around 37 km for the entire set of selected observations (assimilation + evaluation) and around 61 km for the subset of assimilated observations. The assimilation period extends from 14 August at midnight to 23 August at 23:00 UTC, with an hourly assimilation step.

The inspection of relative errors of the CHIMERE reference run shows that these errors exhibit a clear diurnal cycle (right panel of Fig. 4). This is due to the fact that on

the one hand the photochemical buildup of ozone and the associated mid-day maximum is quite well reproduced by the model (relative RMSE < 20 %); on the other hand, the nighttime minimum of ozone associated with the lowest daily PBL height is partly missed (relative RMSE up to 40 %). We have chosen to prescribe relative perturbations that represent this diurnal error cycle (light blue curve in the left panel of Fig. 4). Finally, following the cycle of ensemble analysis, perturbations and forecasts, the obtained absolute forecasted ensemble standard deviation ranges between 4 and 6 ppb (Fig. 8).

The prescription of the OECM (R) is also a key point of the assimilation procedure. Assuming no observational error correlation, the OECM matrix is defined by the observation error variance (diagonal terms). As highlighted above (Sect. 2), the main observational error source is the representativeness error (and thus is dependent on the model resolution). The random part of the observation error can be defined as an average value of the standard deviation of observations within a grid cell. For the city of Paris, Blond

Table 4. Comparison of the accuracy measures (hourly bias, RMSE and correlation coefficient, daily maximum and maximum of the daily eight-hour ozone mean RMSE) for the reference run over suburban stations ($N = 376$), rural stations ($N = 228$), and for remote stations that have altitude below 300 m a.s.l. ($N = 42$) and above ($N = 47$). These statistics are computed from 1 June to 30 August 2009.

Accuracy measures	Total	Background	Rural	Suburban	Background (> 300 m)
Bias	-2.86	-0.89	-2.0	-3.6	3.45
RMSE (hourly)	10.19	8.43	9.88	10.57	11.76
Correlation	0.72	0.58	0.69	0.76	0.51
RMSE (8 h max)	7.36	6.61	7.17	7.55	9.57
RMSE (daily max)	9.5	8.64	9.5	9.6	13.92

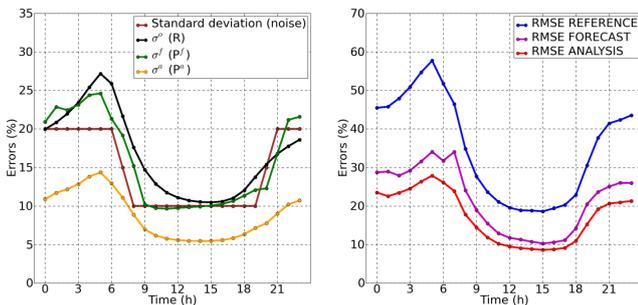


Fig. 4. Relative errors of the assimilated stations ($N = 295$) for the REF_ASSIM experiment. In the left panel, we plot the diurnal profile of the standard deviation of the spatially correlated noise added to the analyzed state after each assimilation step (brown curve) and the observation error standard deviation (black curve). The forecast ensemble standard deviation is in green and the analyzed standard deviation in orange. The RMSE (right panel) is calculated for the reference run (blue curve), the ensemble mean analysis (red curve) and the one-hour ensemble forecast (purple curve).

et al. (2003) determined an observation error of 5 ppb at 15:00 UTC for a 6 km model resolution. Using observations from the AIRNOW database in the USA, Chai et al. (2006) evaluated the ozone standard deviation inside a model grid cell (60 km horizontal resolution) and got on average a daily range between 5 ppb and 13 ppb (at night). Finally, in their assimilation framework, they assumed an observation error of 8 ppb. Using the Hollingsworth and Lönnberg method (Hollingsworth and Lönnberg, 1986), Flemming et al. (2004) got on average an absolute standard deviation of 5 ppb independently of the station type of the FLEM05 classification. Following this last study, an observation error standard deviation of 5 ppb is used a priori. This value is also consistent with those typically used in other data assimilation systems (Hanea et al., 2004; Wu et al., 2008). It corresponds to a lower relative error for the background stations as their mean concentration is higher and also a lower relative error during the afternoon ozone maximum (Fig. 4). In the following, the setups of the model and observational errors will be referred to as the REF_ASSIM experiment; they correspond to the reference case to which the other assimilation experiments will be compared.

5.2 Evaluation of the base case assimilation experiment

In order to analyze the performance of this REF_ASSIM assimilation experiment, we compare results from the CHIMERE reference run and the analysis (i.e., the mean of the analyzed ensemble) against observations from evaluation stations (which are not assimilated). Figure 5a shows the simulated surface ozone on 20 August 2009 at 15:00 UTC along with the observations. The reference run correctly displays spatial patterns associated with the episode in which a cold front separates Europe into two parts: a western part with marine air masses associated with low ozone values, and an eastern part associated with higher ozone concentrations expanding from Italy up to Norway. Compared to the reference, the analyzed field exhibits higher values and more pronounced spatial ozone gradients that seem to be more realistic, as confirmed by observations (Fig. 5b); for instance, the general underestimation of the peaks over a large region of Denmark, eastern Germany and Austria is corrected. As an example, the time series of surface ozone observation at the Odense stations (10.4° E, 55.4° N) in Denmark along with the REF_ASSIM analysis and the CHIMERE reference run is plotted in Fig. 6. It shows a clear improvement of the analysis which, contrary to the reference run, captures the high temporal variability of the ozone measurements including the ozone peak of 20 August. At this station, the temporal correlation coefficient over the period is increased from 0.8 for the reference run to 0.9 for the analysis. In locations where only few measurements are available, observations assimilated from a single station are propagated over large areas as in Spain or Greece (see Fig. 1 for the location of assimilated stations). Over these regions, often no validation stations are available to verify if these changes are realistic. However, the spatial shape of the corrections, for instance over the North Sea, illustrates the ability of the sequential assimilation to extend innovations along with the ozone flow (in the northwest direction) during the forecast step. We plot the average diurnal profile of ozone concentrations and for the RMSE during the period between 14 August and 23 August 2009 (Fig. 7a). We notice an increase in the ozone baseline of 5 ppb and in the amplitude of the diurnal cycle for suburban, rural and background stations located above 300 m a.s.l. For this period, the maxima over the suburban and rural stations are

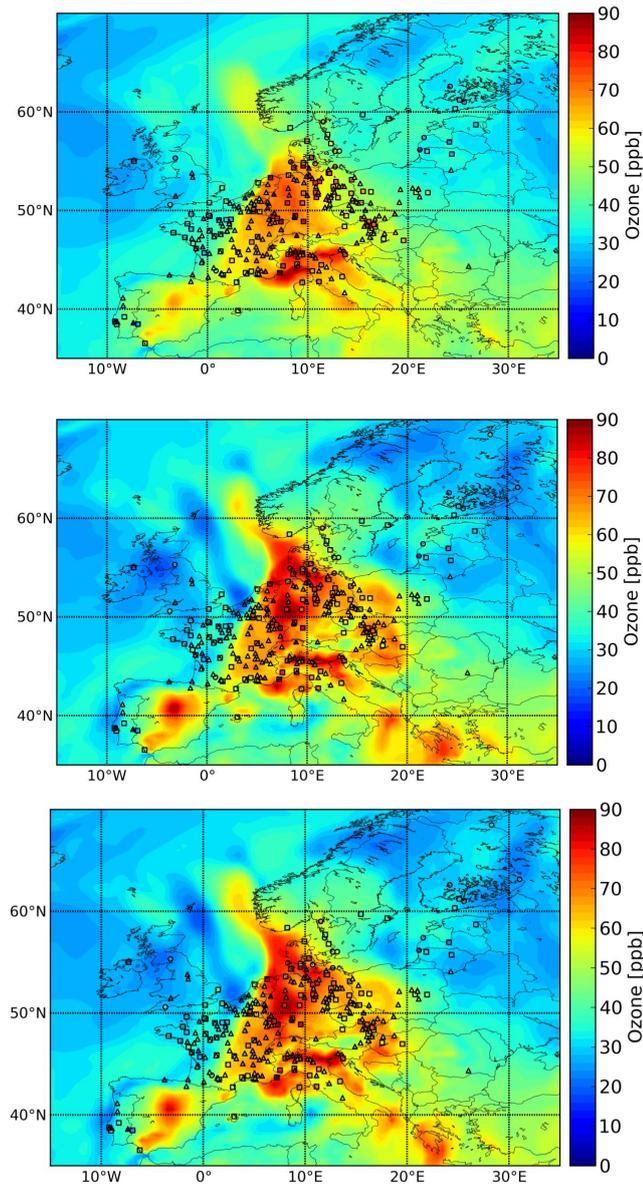


Fig. 5. Simulated ozone fields for (a) the CHIMERE reference run (top), (b) the REF_ASSIM analysis (middle), and (c) the MOD_DESR analysis (bottom) on 20 August at 15:00 UTC. Ozone measurements at validation stations are plotted by circles, squares and triangles for, respectively, background, rural and suburban stations.

underestimated by the CHIMERE simulation. Remote stations do not show the same behavior, and the ozone level is even reduced with respect to summer average values. This is because some stations are located in Scandinavia or in the United Kingdom, where weather conditions were not favorable for ozone formation (Fig. 5a). The diurnal cycle of errors shows overall reduced values and still shows a minimum of 6 ppb during daytime for all station types. Contrasting results are obtained for nighttime regarding the station types;

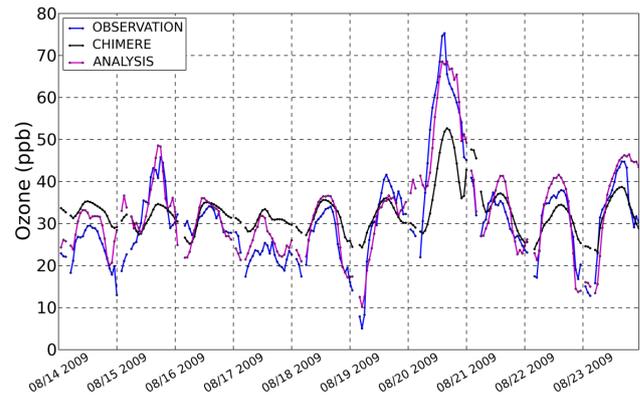


Fig. 6. Ozone measured (blue line) at the Odense evaluation site (not assimilated) in Denmark (10.4° E, 55.4° N) compared with the CHIMERE reference run (black line) and the REF_ASSIM analysis (purple line).

for instance, in the morning, suburban ozone levels are overestimated by the analysis while rural as well as mountain ozone concentrations are underestimated. Some of the errors can also be the result of reduced spatial representativeness of the observations early in the morning.

In addition to these particular elements, we present the statistics for the whole set of sites and for each station type of the validation set on a daily basis for the hourly profile, daily peaks and the daily maximum of the running eight-hour mean, which is an indicator for the impact of ozone exposure to human health (REF_ASSIM statistics in Table 5). The RMSE of the analysis is largely reduced: by 30 % as compared to the reference run; around 95 % of sites retained for evaluation show a decrease of RMSE. The hourly correlation average over each day and over all the stations increased to 0.87 against 0.75 for the reference run; it is improved for all station types (Table 5). Both the analysis and the reference run only show a small bias (below 5 ppb for around 75 % of the evaluation stations). The bias is not corrected (slightly more positive) for the rural stations, but is reduced for suburban and background ones. In order to evaluate an indicator of the impact of ozone exposure to human health, we also calculate the daily maximum of the running eight-hour mean; its RMSE decreases from about 8 ppb in the reference run to 5 ppb in the analysis. The assimilation allows an improvement in the reproduction of the daily maxima, with errors around 6–7 ppb (Table 5) while reducing the negative bias for rural and suburban stations.

5.3 Sensitivity to the background error covariance description

Since the representation of BECM can be a crucial point in assimilation systems, we have evaluated the sensitivity to its formulation. Especially, we have built different BECM following approaches recently used in the field: either by

Table 5. Comparison of the accuracy measures (hourly bias, RMSE and correlation coefficient, daily maximum and maximum of the daily eight-hour ozone mean RMSE) for the reference run and for each assimilation experiment over the evaluation set for suburban stations ($N = 187$), rural stations ($N = 112$), and for remote stations that have altitude below 300 m a.s.l. ($N = 18$) and above ($N = 27$). These statistics are computed over the whole assimilation period (i.e., from 14 August to 23 August 2009).

Accuracy measures	Simulation name	Evaluation set	Background	Rural	Suburban	Background (> 300 m)
Bias	REFERENCE	−0.42	−1.91	0.78	−0.99	7.61
	REF_ASSIM	0.18	0.67	1.43	−0.62	7.95
	MOD_DESR	0.26	0.6	1.49	−0.52	8.12
	NEWPAR	0.12	0.45	1.39	−0.67	7.89
	NEWPAR_MOD_DESR	0.15	0.58	1.39	−0.63	8.01
	TUNNING_OBS_TYPE	0.08	0.58	1.28	−0.68	7.77
RMSE (hourly)	REFERENCE	10.79	8.5	11.04	10.86	15.77
	REF_ASSIM	7.76	7.23	8.16	7.58	13.06
	MOD_DESR	7.74	7.11	8.13	7.58	13.08
	NEWPAR	7.69	7.13	8.07	7.52	12.93
	NEWPAR_MOD_DESR	7.69	6.99	8.04	7.54	13.07
	TUNNING_OBS_TYPE	7.8	7.29	8.14	7.65	12.96
Correlation	REFERENCE	0.75	0.53	0.71	0.8	0.43
	REF_ASSIM	0.87	0.69	0.85	0.9	0.67
	MOD_DESR	0.87	0.69	0.85	0.91	0.67
	NEWPAR	0.87	0.69	0.85	0.91	0.68
	NEWPAR_MOD_DESR	0.87	0.7	0.85	0.91	0.67
	TUNNING_OBS_TYPE	0.87	0.68	0.85	0.9	0.67
RMSE (8 h max)	REFERENCE	7.97	6.57	8.07	8.05	11.87
	REF_ASSIM	4.86	5.25	4.95	4.77	9.63
	MOD_DESR	4.85	4.91	4.99	4.76	9.61
	NEWPAR	4.79	5.04	4.89	4.71	9.45
	NEWPAR_MOD_DESR	4.83	5.01	4.94	4.75	9.56
	TUNNING_OBS_TYPE	4.86	5.25	4.95	4.77	9.55
RMSE (daily max)	REFERENCE	10.83	7.85	11.19	10.9	17.16
	REF_ASSIM	7.22	7.16	7.6	7.0	14.66
	MOD_DESR	7.3	6.83	7.67	7.13	14.58
	NEWPAR	7.13	6.93	7.5	6.93	14.44
	NEWPAR_MOD_DESR	7.21	6.96	7.59	7.01	14.64
	TUNNING_OBS_TYPE	7.24	7.11	7.58	7.05	14.58

adjusting the covariance inflation factor based on diagnostics of previous assimilations, or by perturbing physical parameters that control the background error to get a more physical ensemble and in this way a better error representation. We show here the results of the three sensitivity tests made in this respect: first we present the online tuning of the covariance inflation factor (MOD_DESR), then we present the combined perturbation of ozone and models parameters (NEWPAR), and finally the combination of the model parameter perturbation with the online tuning (NEWPAR_MOD_DESR). In the case of the MOD_DESR experiment, the diurnal cycle of the BECM profile is adjusted from an online diagnostic (Desroziers et al., 2005). This adjustment is performed by calculating, at each analysis step, the product (Eq. 6) of the differences between analysis and background and between observation and background state in the

observation space. This product is divided by the analyzed ozone average over assimilated observations. At the end of the day, this diurnal profile is used for the prescription of the relative noise standard deviation for the next day. This treatment is important because background errors strongly vary with the time of the day, for instance dispersion and photochemical processes, and their associated model errors show a strong diurnal cycle (as discussed in Sect. 5.2).

In an EnKF, the spread of the ensemble (i.e., square root of Eq. 3) is employed to estimate background errors, therefore quantitatively it has to be compared with the background error diagnostic (Eq. 6). For the REF_ASSIM experiment, these diagnosed background errors (light blue curve in the left panel of Fig. 8) show lower values than the ensemble standard deviation (green curve). Thus, the diagnostic suggests an overestimation of the background error in the

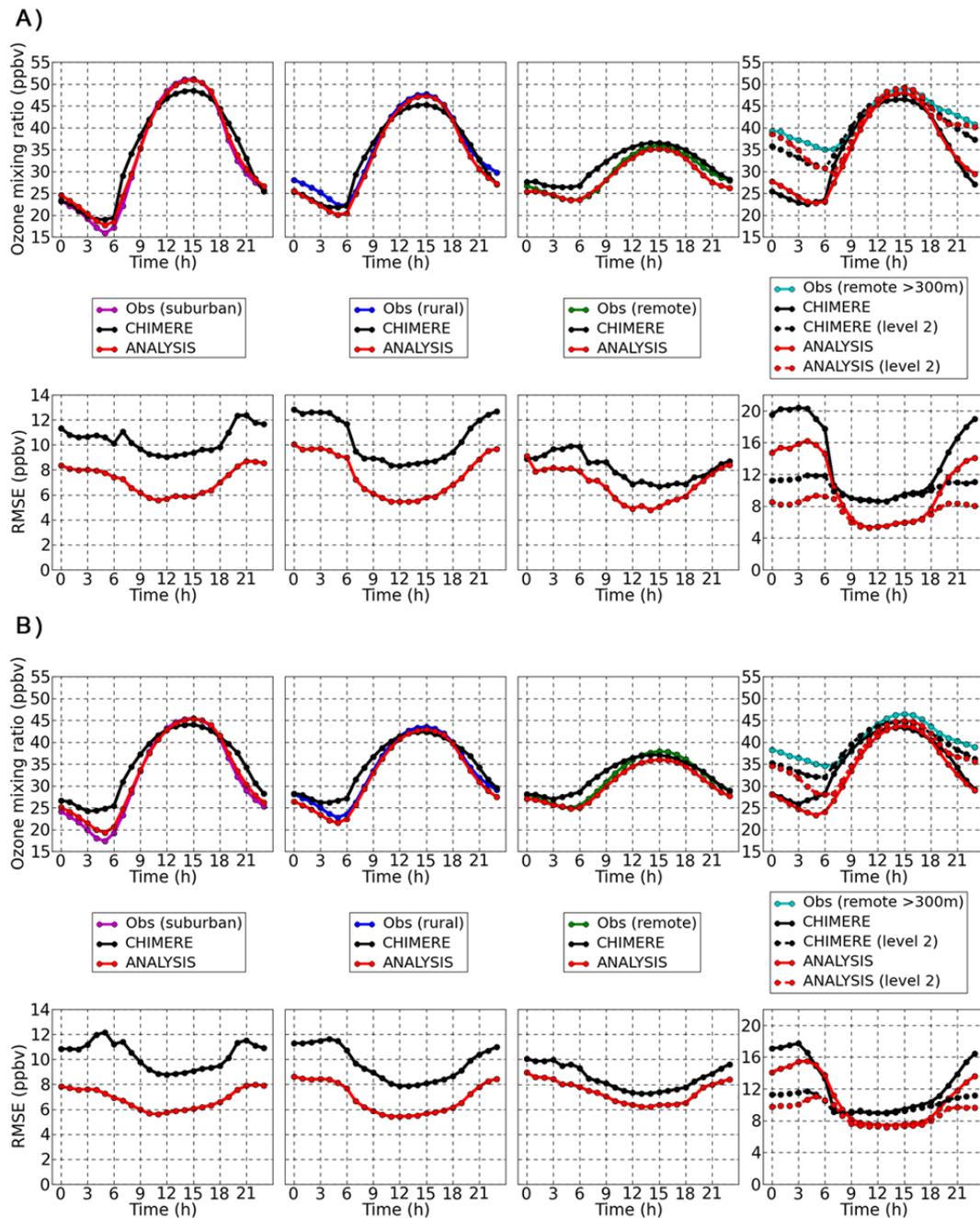


Fig. 7. Average diurnal cycle of ozone and the associated RMSE for the reference run and the REF_ASSIM analysis for the assimilation period at validation stations (upper panel, **a**) and for the summer period (lower panel, **b**). Separated panels are shown for the suburban stations (left, $N = 172$), rural stations (middle left, $N = 102$), and for remote stations with an altitude below 300 m a.s.l. (middle right, $N = 17$) and above ($N = 27$).

REF_ASSIM experiment. The comparison with the forecast RMSE indicates that the daily variation of the background error diagnostic is more realistic than that of the initial background error. In particular, the morning peak in RMSE corresponds to a peak in the background error diagnostic. In the MOD_DESR experiment (Fig. 8, right panel), the ensemble standard deviation is close to the diagnostic of the first

experiment. This demonstrates a rapid convergence towards a stable background error through this procedure. Thus, the ensemble standard deviation globally represents much better the shape of the diurnal cycle of the forecast RMSE. Particularly, the morning (07:00 UTC) and evening peaks (19:00 and 20:00 UTC) of the RMSE are well diagnosed. However, the nighttime ensemble standard deviations are almost at the

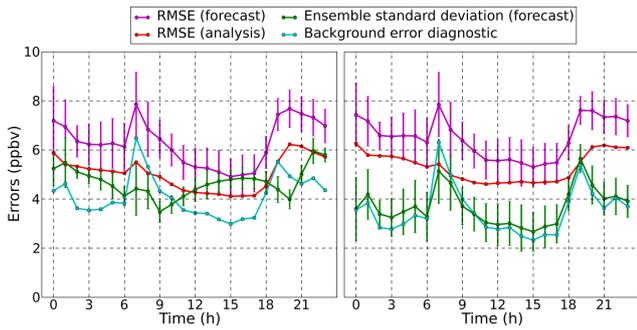


Fig. 8. Mean hourly value (over all assimilated observations) of the RMSE for the ensemble mean analysis (red curve), the one-hour ensemble forecast (purple curve), the ensemble standard deviation (green) and the background error diagnostic (light blue). The left panel displays these values for the base case assimilation run (REF_ASSIM) and the right panel for the simulation using the online tuning Desroziers diagnostics (MOD_DESR). Error bars display the temporal standard deviation.

same level as those obtained during daytime, suggesting that the higher nighttime RMSE are not due to errors in model formulation, but rather to observational (i.e., representativeness) errors. The reduction of the ensemble variance gives more weight to the model (compared to the observations) and therefore increases the RMSE of the analysis against assimilated stations. Thus, the lower background error variances in the MOD_DESR experiment reduce the magnitude and the spatial extent of the analysis increment. This is illustrated by the analyzed maps (Fig. 5b and c): over large areas, for instance in Spain (one station near Madrid) or in Greece (one station near Athens), the ozone field is controlled by only a few assimilated observations. This is an advantage when errors are spatially correlated; it leads to a reduced RMSE against evaluation stations for daily peaks. However, 8 h mean average and hourly statistics are not substantially modified.

A further sensitivity study is performed by creating different ensemble members using perturbations of the uncertain input data of the model, instead of the ozone fields directly. Choices of the perturbed parameter, their standard deviation and spatio-temporal correlation (Table 1) are inspired from several previous studies (Hanna et al., 2001; Beekmann and Derognat, 2003; Hanea et al., 2004; Wu et al., 2008; Boynard et al., 2010). The scheme of the perturbations is similar to the one applied to the ozone fields with fixed horizontal spatial decorrelation (Evensen et al., 2003), but this time the distribution is log-normal and we also include a temporal correlation. The standard deviation of the 20 ensemble members of the free run (without assimilation) with a perturbed parameter reaches a maximum of 8 ppb at some locations after four days of simulation, but is only about 1 ppb on average, so it generally strongly underestimates the background error. It turns out that this ensemble cannot create enough variability

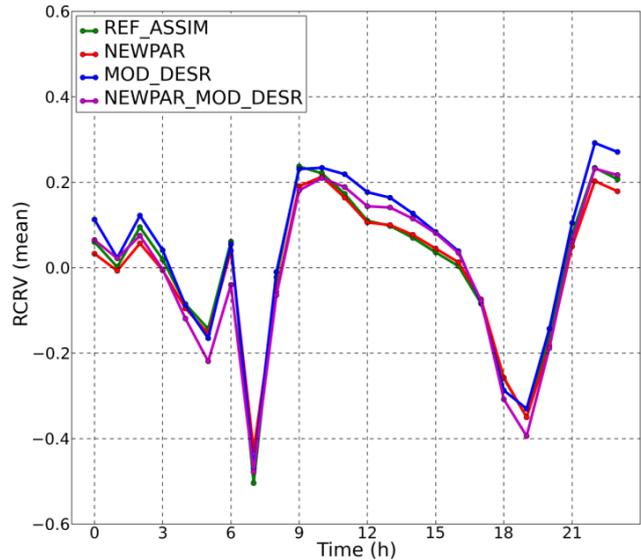


Fig. 9. Mean of the reduced centered random variable (RCRV) averaged over all assimilation steps for the REF_ASSIM, NEWPAR, MOD_DESR, and the NEWPAR_MOD_DESR assimilation experiments.

(ensemble spread) between two analyses, i.e., during a one-hour time step, to avoid a more and more reduced ensemble spread. Thus we choose to combine this perturbation of the model parameters with the classical perturbation of the ozone field itself, in this way creating a hybrid ensemble. We then reproduce the two previous experiments with the addition of the parameter perturbations, namely NEWPAR associated with the REF_ASSIM one and NEWPAR_MOD_DESR for the online tuning experiment (MOD_DESR). Note that in these experiments, the model parameters are not included in the state vector.

In order to check and compare the ensemble dispersion from these simulations, we calculate the RCRV (Eq. 8). Figure 9 displays the spatially and hourly averaged mean of the RCRV. It indicates that the ensemble predictions have a weak positive bias in the afternoon and in the evening and a negative bias at 07:00 UTC and around 19:00 UTC. In the morning and in the early evening, the bias is larger and is formed rapidly. At that time, the model simulation can be very uncertain, due to increasing or still large emissions and a developing (in the morning) or fading (in the evening) of the PBL height. These processes are locally variable, and thus lead to a reduced site representativeness that is not resolved at the models' horizontal and vertical resolution. An evaluation of the time evolution of the spatial pattern of the ensemble bias indicates a large variability in space and time, but the morning bias is more pronounced at suburban stations and especially at those located in the Po Valley and in southeastern Europe (not shown). Globally, a slightly lower bias is found for REF_ASSIM and NEWPAR, for which analyses

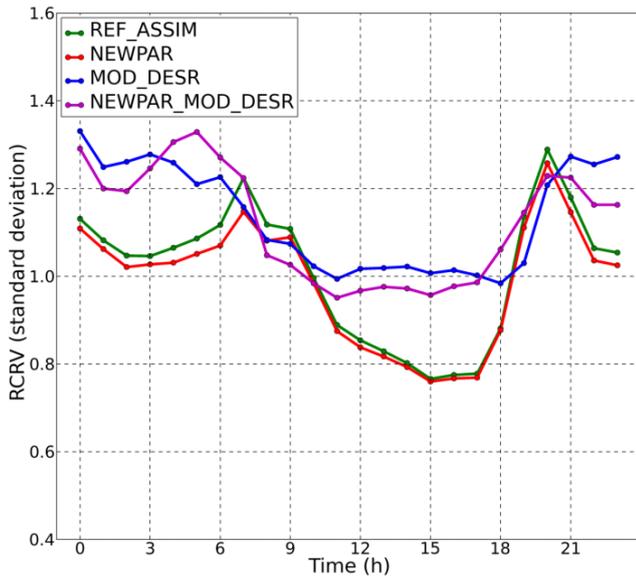


Fig. 10. Standard deviation of the RCRV averaged over all assimilation steps for REF_ASSIM, NEWPAR, MOD_DESR, and the NEWPAR_MOD_DESR assimilation experiment.

are closer to the observations, but only small differences are found between the different experiments. In fact, the analysis field is not biased because it is removed by the assimilation procedure. The use of the diagnostic leads to an increase in the additive noise; thus, it allows a prediction of the increase in the ensemble spread due to the higher bias (Figs. 8 and 9). A step further would be the use of a bias correction procedure or to add these perturbed parameters in the state vector in order to improve the ensemble forecast and subsequently the assimilation performance.

The standard deviation of the RCRV (Fig. 10) provides a framework for the verification of the ensemble spread, namely, if it is greater/lower than one. This means that the ensemble is underdispersive/overdispersive. The tuned simulations (MOD_DESR and NEWPAR_MOD_DESR) display a correct behavior of error statistics during daytime, with a standard deviation of the RCRV close to 1. By contrast, the ensemble is under dispersive for all simulations during nighttime. The addition of the perturbation for the model parameters slightly improves the error representation during nighttime. As the observational error is considered constant in the calculation of the RCRV statistics, the results from Desroziers diagnostic could also suggest that observation errors are underestimated in the night time (see Sect. 5.4).

Although we obtain different ensemble spread profiles that change weights between model and observation in the assimilation procedure, only small changes are found in the comparison against the observations used for validation. The ozone mean, the average bias and RMSE are not significantly modified in the different simulations. However, it is preferable to reduce the amount of perturbation because the

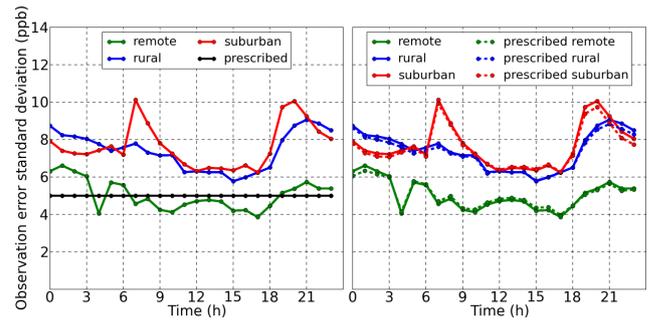


Fig. 11. Prescribed observational error variance and corresponding diagnostics averaged over all assimilation steps for background sites below 300 m a.b.l. (23 stations), rural (106 stations) and suburban (166 stations) stations for the REF_ASSIM (left panel) and the TUNING_OBS_TYPE (right panel) experiments.

ensemble standard deviation can become important in large areas where observations are scarce or even absent; this may lead to unrealistic spatial correlation of the error fields, see for instance the differences in the ozone fields in Spain or in Greece when only one or two observations are assimilated (Fig. 5b and c). Furthermore, we can notice that the model parameter perturbations (slightly) improve the results in all cases.

5.4 Sensitivity to the observation error formulation

Similarly to the previous section (i.e., Sect. 5.3) concerning the BECM, we propose to examine the system sensitivity to the formulation of the OEM. In the case of surface ozone observations, the measurement errors are generally weak, but the representativeness errors can be important. This latter error depends on the horizontal and vertical model resolution. In this section, we investigate different approaches to estimate the observation error variance, which should lead to a more realistic and detailed error estimation, rather than fixing it to a constant value as done in the previous sections. Besides this modification, the ensemble configuration is the same as for the base case experiment (REF_ASSIM). Figure 11 (left panel) shows the 5 ppb error (black line) prescribed for the REF_ASSIM experiment, and in addition the Desroziers error diagnostic estimated for each subset of stations following the FLEM05 classification (cf. Sect. 3). Firstly, it shows that the diagnosed errors are rather similar to the prescribed error for remote stations but larger for rural and suburban station types. For the suburban type, it is probably linked to the coarse horizontal resolution of our model (0.5°), which is not representative enough for urbanized environments. This result is close to the one of Chai et al. (2006) for similar grid spacing. Secondly, the diagnosed error shows a diurnal cycle with lowest values at noon (about ~ 6 ppb for suburban and rural sites), close to the prescribed value, but which are higher at nighttime and in the morning. More specifically, suburban stations show an error maximum when the

boundary layer develops and fades. In these cases, ozone titration by NO emissions is most effective, and representativeness errors are expected to be larger.

Next, we use the above results in the assimilation scheme. As for the tuning of the background error, we apply the diagnostic (Eq. 7) of the observational error diurnal profile from the previous day to the OECM, as a function of the station type (TUNING_OBS_TYPE experiment). Figure 11 (right panel) shows that the mean error diurnal cycle as a function of station type is well captured, because prescribed and newly diagnosed errors coincide (convergence). Nevertheless, the RMSE of the analysis does not show a significant improvement or modification with respect to the reference experiment (Table 3); only a slight reduction in errors is found for the mountain stations due to their reduced prescribed error.

Following Eqs. (6) and (7) of the diagnostics for each time step, one can see that the sum of the diagnosed observational and the background error variances must be equal to the mean square error of the ensemble forecast. Therefore, the use of the diagnostics suggests that the representativeness error is higher than expected and would contribute to a major part of the total error. However, it should be noted that the diagnostic efficiency can be reduced by a model bias; this can lead to an overestimation of observation error variances (Li et al., 2009b). Although observational errors strongly depend on the site representativeness and type, the sensitivity to the observational error is globally small. Therefore, accuracy measures (at evaluation stations) do not show a substantial sensitivity to the modification of observational error variance.

6 Evaluation of the summer analysis

Different assimilation experiments have been conducted for a period of 10 days only (in August 2009). Then, in order to assess the performance statistics over different meteorological conditions, an analysis of 90 days has been evaluated. The period covers three months of the summer of 2009 (JJA, 2009), which allows the evaluation of rather different meteorological conditions. The configuration of the data assimilation system is similar to the PAR_MOD_DESR experiment. An ensemble of 20 members is used; the observation error variance is set to 25 ppb²; the covariance inflation factor is tuned every day from the hourly diagnostics of the background error and several model parameters are perturbed (Table 1) but not corrected in the assimilation process. The set of assimilated as well as evaluation stations is unchanged with respect to the previous experiments. We employed the same evaluation method based on accuracy measures (Table 6) and on analysis of the diurnal ozone and error cycle (Fig. 7b). A similar improvement to the analysis with regards to the reference run is found for the whole summer as for the 10-day period: the averaged hourly correlation coefficient is 0.87 and

the RMSE is around 7 ppb (with respect to 0.72 and 10.2 ppb for the reference simulation). These errors exhibit a similar diurnal cycle with an average value of 8 ppb during the night and 6 ppb in the daytime (Fig. 7b). These results indicate that the data assimilation system is suitable for analysis or reanalysis of longer periods. However, it should still be tested in other seasons even if ozone pollution is most important during the summer season.

7 Conclusions

In this paper we present a data assimilation system based on the rCTM CHIMERE in an EnKF framework and using surface ozone observations provided by the European Airbase database. These stations were classified based on their diurnal ozone variability. 350 stations labelled as remote, rural and suburban were selected for assimilation with an average closest distance of 61 km as compared to 37 km for the entire set of almost 700 stations. The system is based on a local analysis computed using a deterministic square root filter and is applied to an ensemble of 20 perturbed CHIMERE rCTM simulations. The assimilation algorithm is well suited to an operational framework; an analysis of one day (including the control run) takes 5 h where only CHIMERE simulations are parallelized (and not yet the assimilation procedure). For a 10-day period in summer 2009, the reference run shows good performance, with a root mean square error (RMSE) around 10 ppb and an average correlation coefficient of 0.75. However, model simulations generally overestimate nighttime and morning concentrations, while they underestimate mid-day ozone peaks in particular during the regional ozone pollution episode that occurs during this period. This error cycle is caused by physical processes during nighttime such as vertical mixing or NO emissions titrating ozone that are not resolved at the chosen model resolution of 0.5 degrees. The statistical evaluation of the analyzed fields using a substantial set of unassimilated observations indicates a 30 % reduction of RMSE, which reaches on average a minimum around 6 ppb in the afternoon, regardless of the station type. Similar performance can be found for analyzed fields for a larger period covering JJA 2009. These improvements are similar to EnKF analyses performed in Hanea et al. (2004), where background errors were estimated by the correction of the LOTOS-EUROS model parameter. The analysis increment can be propagated on a synoptic scale during the forecast step, thus allowing corrections downwind to the continental observations, as shown for example in the northern part of Europe. More than 95 % of the evaluation set shows a decrease in RMSE. The ensemble mean of the resulting 1 h forecast also shows a better performance than the reference run. However, it has more limited success at transitions between day and night and vice versa. This is mainly due to the formation of a positive bias, which is more pronounced for suburban stations. Therefore, the background error reduction

Table 6. Comparison of the accuracy measures (hourly bias, RMSE and correlation coefficient, daily maximum and maximum of the daily eight-hour ozone mean RMSE) for the reference run and for the analysis over the evaluation set for suburban stations ($N = 187$), rural stations ($N = 112$), and for background stations that have altitude below 300 m a.s.l. ($N = 18$) and above ($N = 27$). These statistics are computed at evaluation stations for the period from 1 June to 30 August 2009.

Accuracy measures	Simulation name	Evaluation set	Background	Rural	Suburban	Background (> 300 m)
Bias (ppb)	REFERENCE	-2.15	-0.82	-1.27	-2.81	4.9
	ANALYSIS	0.0	1.14	0.94	-0.67	6.64
RMSE (ppb)	REFERENCE	10.2	8.73	9.89	10.53	12.92
	ANALYSIS	7.09	7.55	7.16	7.0	11.51
Correlation	REFERENCE	0.72	0.55	0.68	0.76	0.47
	ANALYSIS	0.87	0.7	0.84	0.89	0.7
RMSE (8 h max, ppb)	REFERENCE	7.7	6.97	7.47	7.9	10.86
	ANALYSIS	4.92	6.09	4.9	4.81	9.45
RMSE (daily max, ppb)	REFERENCE	10.16	8.97	10.19	10.26	15.64
	ANALYSIS	7.12	8.31	7.4	6.83	13.93

is in this case not sensitive to the improvement of ozone initial conditions at each forecast cycle. As suggested by other studies, a step further would be the simultaneous adjustment of precursor initial conditions and emissions rate (Tang et al., 2011). Also, the application of a bias correction procedure combined with the data assimilation scheme (Li et al., 2009a) could improve the EnKF performance.

A series of different assimilation experiments was performed with different formulations of the OECM and BECM matrices. We focused on the estimation of the hourly variation of the model and observational errors using Desroziers diagnostics. We used the reduced centered random variable (RCRV) standard deviation as a tool to evaluate the results of the tuning of the model error using the Desroziers diagnostic. By the comparison of the innovations statistics and prescribed errors, RCRV statistics allow the evaluation of both weighted bias and error prescription. The ensemble dispersion is more consistent with the tuned background error during daytime, while an underestimation of the background error is found during nighttime; it is attributed to the larger and unaccounted observational error at that time. Generally, the diagnostic indicates a large contribution of the observational error, higher than expected, especially for rural stations. As pointed out by Li et al. (2009b), an overestimation of the observation error can be found when the model forecast shows large bias. However, the use of the FLEM05 classification of the measurement sites allowed diagnosing of observational errors for specific types of sites. As expected, larger representativeness errors were estimated for suburban and rural than for remote sites. These errors show as expected a daily variation with a minimum in the daytime and a nighttime increase according to the increase of the ozone precursor level close to the sources. This underlies the need to determine a priori the observation representativeness such as done in the FLEM05 classification. As a perspective for future work, the diagnostic could provide a determination of observation data suitable

for data assimilation (observation thinning), since both observation error variances/correlations and analysis sensitivity can be estimated by using the a posteriori diagnostics. Indeed, stations used for assimilation or evaluation are spatially close, where the ozone observations network is spatially dense (over Western and Central Europe). Thus many observation sites are present with respect to a given model grid cell within the localisation radius (250 km) and within the correlation length. Then, in the assimilation procedure, the weight of the observations will be large independently of “not too big” variations of the background and observation errors; the weight of the background will remain small. Thus, in terms of accuracy measures, only small changes in performance statistics are found among experiments even for substantial changes (up to a factor of 2) in the model and observation errors. However, for a dense network correlations between assimilated observations cannot be excluded. A step further would be to perform an a posteriori diagnostic of the observation error correlation and if necessary to take it into account in the assimilation procedure.

For future work, the evaluation of the system performance should be made over a longer time period covering different seasons. Also, ideally, to be physically consistent, the BECM should rely more on the perturbation of uncertain physical parameters instead of the ozone concentrations. This would also allow assessing of the impact of an ozone correlation on other chemical variables by including and updating them in the state vector (Brunner et al., 2012; Miyazaki et al., 2012). In addition to the model parameter perturbations already implemented, or by replacing them, a new ensemble should be designed taking the model inputs as emissions, meteorological forcing, and chemical boundary conditions from different sources (models). Finally, due to its robustness, the present system already appears suitable for implementation in operational systems such as the one supported by the European FP7 MACC project.

Acknowledgements. This work has been supported by the MACC and MACC-II (Monitoring Atmosphere and Climate Change) projects from the European Community's Seventh Framework Programme (FP7 THEME [SPA.2011.1.5-02]) under grant agreement no. 283576 and by the French LEFE project ADOMOCA (Assimilation de Données pour les Modèles de Chimie Atmosphérique). This work was performed using HPC resources from GENCI-CCRT (Grant 2013- t2013016695). This study has been also supported by CNES in the framework of the IASI-TOSCA project. Most of the charts have been made using the Python plotting library Matplotlib (Hunter 2007). B. Gaubert's PhD was supported by a grant from the French Ministry of Higher Education and Research.

Edited by: H. Garny



The publication of this article is financed by CNRS-INSU.

References

- Agudelo, O. M., Barrero, O., Peter, V., and De Moor, B.: Assimilation of ozone measurements in the air quality model AURORA by using the Ensemble Kalman Filter, Decision and Control and European Control Conference (CDC-ECC), 2011 50th IEEE 12–15 December 2011, 4430–4435, doi:10.1109/CDC.2011.6160444, 2011.
- Beekmann, M. and Derognat, C.: Monte Carlo uncertainty analysis of a regional-scale transport chemistry model constrained by measurements from the Atmospheric Pollution Over the Paris Area (ESQUIF) campaign, *J. Geophys. Res.*, 108, 8559–8559, 2003.
- Bell, M. L., Dominici, F., and Samet, J. M.: A Meta-Analysis of Time-Series Studies of Ozone and Mortality With Comparison to the National Morbidity, Mortality, and Air Pollution Study, *Epidemiology*, 16, 436–445, 2005.
- Bessagnet, B., Hodzic, A., Vautard, R., Beekmann, M., Cheinet, S., Honoré, C., Liousse, C., and Rouil, L.: Aerosol modeling with CHIMERE – preliminary evaluation at the continental scale, *Atmos. Environ.*, 38, 2803–2817, 2004.
- Blond, N. and Vautard, R.: Three-dimensional ozone analyses and their use for short-term ozone forecasts, *J. Geophys. Res.*, 109, D17303, doi:10.1029/2004JD004515, 2004.
- Blond, N., Bel, L., and Vautard, R.: Three-dimensional ozone data analysis with an air quality model over the Paris area, *J. Geophys. Res.*, 108, 4744, doi:10.1029/2003JD003679, 2003.
- Boynard, A., Beekmann, M., Foret, G., Ung, A., Szopa, S., Schmechtig, C., and Coman, A.: An ensemble assessment of regional ozone model uncertainty with an explicit error representation, *Atmos. Environ.*, 45, 784–793, 2011.
- Brunner, D., Henne, S., Keller, C. A., Reimann, S., Vollmer, M. K., O'Doherty, S., and Maione, M.: An extended Kalman-filter for regional scale inverse emission estimation, *Atmos. Chem. Phys.*, 12, 3455–3478, doi:10.5194/acp-12-3455-2012, 2012.
- Burgers, G., Van Leeuwen, P. J. and Evensen, G.: Analysis scheme in the ensemble Kalman filter, *Mon. Weather Rev.*, 126, 1719–1724, 1998.
- Candille, G., Côté, C., Houtekamer, P. L., and Pellerin, G.: Verification of an Ensemble Prediction System against Observations, *Mon. Weather Rev.*, 135, 2688–2699, doi:10.1175/MWR3414.1, 2007.
- Carmichael, G. R., Chai, T., Sandu, A., Constantinescu, E. M., and Daescu, D.: Predicting air quality: Improvements through advanced methods to integrate models and measurements, *J. Comp. Phys.*, 227, 3540–3571, 2008.
- Chai, T., Carmichael, G. R., Tang, Y., Sandu, A., Hardesty, M., Pilewskie, P., Whitlow, S., Browell, E. V., Avery, M. A., Nédélec, P., Merrill, J. T., Thompson, A. M., and Williams, E.: Four-dimensional data assimilation experiments with International Consortium for Atmospheric Research on Transport and Transformation ozone measurements, *J. Geophys. Res.*, 112, D12S15, doi:10.1029/2006JD007763, 2007.
- Cheinet, S. and Teixeira, J.: A simple formulation for the eddy diffusivity parameterization of cloud-topped boundary layers, *Geophys. Res. Lett.*, 30, 1930, doi:10.1029/2003GL017377, 2003.
- Chevalier, A., Gheusi, F., Delmas, R., Ordóñez, C., Sarrat, C., Zbinden, R., Thouret, V., Athier, G., and Cousin, J.-M.: Influence of altitude on ozone levels and variability in the lower troposphere: a ground-based study for western Europe over the period 2001–2004, *Atmos. Chem. Phys.*, 7, 4311–4326, doi:10.5194/acp-7-4311-2007, 2007.
- Coman, A., Foret, G., Beekmann, M., Eremenko, M., Dufour, G., Gaubert, B., Ung, A., Schmechtig, C., Flaud, J.-M., and Bergametti, G.: Assimilation of IASI partial tropospheric columns with an Ensemble Kalman Filter over Europe, *Atmos. Chem. Phys.*, 12, 2513–2532, doi:10.5194/acp-12-2513-2012, 2012.
- Constantinescu, E. M., Sandu, A., Chai, T., and Carmichael, G.: Ensemble-based chemical data assimilation I: General approach, *Q. J. Roy. Meteorol. Soc.*, 133, 1229–1243, doi:10.1002/qj.76, 2007a.
- Constantinescu, E. M., Sandu, A., Chai, T., and Carmichael, G. R.: Ensemble-based Chemical Data Assimilation II: Covariance Localization, *Q. J. Roy. Meteorol. Soc.*, 133, 1245–1256, 2007b.
- Courtier, P. and Talagrand, O.: Variational assimilation of meteorological observations with the adjoint vorticity equation, II: Numerical results, *Q. J. Roy. Meteorol. Soc.*, 113, 1329–1347, 1987.
- Curci, G., Beekmann, M., Vautard, R., Smiatek, G., Steinbrecher, R., Theloke, J., and Friedrich, R.: Modelling study of the impact of isoprene and terpene biogenic emissions on European ozone levels, *Atmos. Environ.*, 43, 1444–1455, doi:10.1016/j.atmosenv.2008.02.070, 2009.
- Curier, R. L., Timmermans, R., Calabretta-Jongen, S., Eskes, H., Segers, A., Swart, D., and Schaap, M.: Improving ozone forecasts over Europe by synergistic use of the LOTOS-EUROS chemical transport model and in-situ measurements, *Atmos. Environ.*, 60, 217–226, 2012.
- Desroziers, G., Berre, L., Chapnik, B., and Poli P.: Diagnosis of observation, background and analysis-error statistics in observation space, *Q. J. Roy. Meteorol. Soc.*, 131, 3385–3396, 2005.
- Elbern, H. and Schmidt, H.: Ozone episode analysis by four-dimensional variational chemistry data assimilation, *J. Geophys. Res.*, 106, 3569–3590, 2001.

- Elbern, H., Schmidt, H., and Ebel, A.: Variational data assimilation for tropospheric chemistry modeling, *J. Geophys. Res.*, 102, 15967–15985, 1997.
- Elbern, H., Strunk, A., Schmidt, H., and Talagrand, O.: Emission rate and chemical state estimation by 4-dimensional variational inversion, *Atmos. Chem. Phys.*, 7, 3749–3769, doi:10.5194/acp-7-3749-2007, 2007.
- European Parliament: DIRECTIVE 2008/50/EC of the European Parliament and of the council of 21 May 2008 on ambient air quality and cleaner air for Europe.
- Evensen, G.: Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte Carlo methods to forecast error statistics, *J. Geophys. Res.*, 99, 10143–10162, doi:10.1029/94JC00572, 1994.
- Evensen, G.: The Ensemble Kalman Filter: Theoretical Formulation and Practical Implementation, *Ocean Dynam.*, 53, 343–367, 2003.
- Evensen G.: Sampling strategies and square root analysis schemes for the EnKF, *Ocean Dynam.*, 54, 539–560, 2004.
- Felzer, B., Kicklighter, D., Melillo, J., Wang, C., Zhuang, Q., and Prinn, R.: Effects of ozone on net primary production and carbon sequestration in the conterminous United States using a biogeochemistry model, *Tellus B*, 56, 230–248, 2004.
- Flemming, J., Van Loon, M., and Stern, R.: Data assimilation for CTM based on optimum interpolation and Kalman filter, in: *Air Pollution Modelling and its Application XVI*, edited by: Borrego, C. and Incecik, S., Kluwer Academic/Plenum Publishers, New York, 373–383, 2004.
- Flemming, J., Stern, R., and Yamartino, R. J.: A new air quality regime classification scheme for O₃, NO₂, SO₂ and PM₁₀ observations sites, *Atmos. Environ.*, 39, 6121–6129, 2005.
- Flemming, J., Inness, A., Flentje, H., Huijnen, V., Moinat, P., Schultz, M. G., and Stein, O.: Coupling global chemistry transport models to ECMWF's integrated forecast system, *Geosci. Model Dev.*, 2, 253–265, doi:10.5194/gmd-2-253-2009, 2009.
- Fowler, D., Pilegaard, K., Sutton, M., Ambus, P., Raivonen, M., Duyzer, J., Simpson, D., Fagerli, H., Fuzzi, S., Schjoerring, J., Granier, C., Neftel, A., Isaksen, I., Laj, P., Maione, M., Monks, P., Burkhardt, J., Daemmgen, U., Neiryneck, J., Personne, E., Wichink-Kruit, R., Butterbach-Bahl, K., Flechard, C., Tuovinen, J., Coyle, M., Gerosa, G., Loubet, B., Altimir, N., Gruenhage, L., Ammann, C., Cieslik, S., Paoletti, E., Mikkelsen, T., Ro-Poulsen, H., Cellier, P., Cape, J., Horvath, L., Loreto, F., Niinemets, U., Palmer, P., Rinne, J., Misztal, P., Nemitz, E., Nilsson, D., Pryor, S., Gallagher, M., Vesala, T., Skiba, U., Brüggemann, N., Zechmeister-Boltenstern, S., Williams, J., O'Dowd, C., Facchini, M., de Leeuw, G., Flossman, A., Chaumerliac, N., and Erisman, J.: Atmospheric composition change: Ecosystems-Atmosphere interactions, *Atmos. Environ.*, 43, 5193–5267, 2009.
- 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.
- Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P. I., and Geron, C.: Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), *Atmos. Chem. Phys.*, 6, 3181–3210, doi:10.5194/acp-6-3181-2006, 2006.
- Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter, *Mon. Weather Rev.*, 129, 2776–2790, 2001.
- Hanea, R. G., Velders, G. J. M., and Heemink, A.: Data assimilation of ground-level ozone in Europe with a Kalman filter and chemistry transport, *J. Geophys. Res.*, 109, D10302, doi:10.1029/2003JD004283, 2004.
- Hanna, S. R., Zhigang, L., Frey, C. H., Wheeler, N., Vukovich, J., Arunachalam, S., Fernau, M., and Hansen, A. D.: Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain, *Atmos. Environ.*, 35, 891–903, 2001.
- Henne, S., Brunner, D., Folini, D., Solberg, S., Klausen, J., and Buchmann, B.: Assessment of parameters describing representativeness of air quality in-situ measurement sites, *Atmos. Chem. Phys.*, 10, 3561–3581, doi:10.5194/acp-10-3561-2010, 2010.
- Hollingsworth, A. and Lönnberg, P.: The statistical structure of short-range forecast errors as determined from radiosonde data, Part I: The wind field, *Tellus A*, 38A, 111–136, doi:10.1111/j.1600-0870.1986.tb00460.x, 1986.
- Honoré, C., Rouil, L., Vautard, R., Beekmann, M., Bessagnet, B., Dufour, A., Elichegaray, C., Flaud, J.-M., Malherbe, L., Meleux, F., Menut, L., Martin, D., Peuch, A., Peuch, V.-H., and Poisson, N.: Predictability of European air quality : Assessment of 3 years of operational forecasts and analyses by the PREV'AIR system, *J. Geophys. Res.*, D113, D04301, doi:10.1029/2007JD008761, 2008.
- Hollingsworth, A., Engelen, R. J., Textor, C., Benedetti, A., Boucher, O., Chevallier, F., Dethof, A., Elbern, H., Eskes, H., Flemming, J., Granier, C., Kaiser, J. W., Morcrette, J.-J., Rayner, P., Peuch, V.-H., Rouil, L., Schultz, M. G., Simmons, A. J., and the GEMS Consortium: Toward a monitoring and forecasting system for atmospheric composition: The GEMS Project", *B. Am. Meteorol. Soc.*, 89, 1147–1164, 2008.
- Houtekamer, P. L. and Mitchell, H. L.: Data assimilation using an Ensemble Kalman Filter technique, *Mon. Weather Rev.*, 126, 796–811, 1998.
- 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 NO₂ 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.
- Hunt, B. R., Kostelich, E. J., and Szunyogh, I.: Efficient data assimilation for spatiotemporal chaos: A local ensemble transform Kalman filter, *Physica D: Nonlinear Phenomena*, 230, 112–126, 2007.
- Hunter, J. D.: Matplotlib: A 2D Graphics Environment, *Comput. Sci. Eng.*, 9, 90–95, doi:10.1109/MCSE.2007.55, 2007.
- Jaumouillé, E., Massart, S., Piacentini, A., Cariolle, D., and Peuch, V.-H.: Impact of a time-dependent background error covariance matrix on air quality analysis, *Geosci. Model Dev.*, 5, 1075–1090, doi:10.5194/gmd-5-1075-2012, 2012.

- Joly, M. and Peuch, V.-H.: Objective classification of air quality monitoring sites over Europe, *Atmos. Environ.*, 47, 111–123, 2012.
- Kinnison, D. E., Brasseur, G. P., Walters, S., Garcia, R. R., Marsh, D. R., Sassi, F., Harvey, V. L., Randall, C. E., Emmons, L., Lamarque, J. F., Hess, P., Orlando, J. J., Tie, X. X., Randel, W., Pan, L. L., Gettelman, A., Granier, C., Diehl, T., Niemeier, U., and Simmons, A. J.: Sensitivity of chemical tracers to meteorological parameters in the MOZART-3 chemical transport model, *J. Geophys. Res.*, 112, D20302, doi:10.1029/2006JD007879, 2007.
- Kukkonen, J., Olsson, T., Schultz, D. M., Baklanov, A., Klein, T., Miranda, A. I., Monteiro, A., Hirtl, M., Tarvainen, V., Boy, M., Peuch, V.-H., Poupkou, A., Kioutsioukis, I., Finardi, S., Sofiev, M., Sokhi, R., Lehtinen, K. E. J., Karatzas, K., San José, R., Astitha, M., Kallos, G., Schaap, M., Reimer, E., Jakobs, H., and Eben, K.: A review of operational, regional-scale, chemical weather forecasting models in Europe, *Atmos. Chem. Phys.*, 12, 1–87, doi:10.5194/acp-12-1-2012, 2012.
- Li, H., Kalnay, E., Miyoshi, T. and Danforth, C. M.: Accounting for model errors in ensemble data assimilation, *Mon. Weather Rev.*, 137, 3407–3419, 2009a.
- Li, H., Kalnay, E., and Miyoshi, T.: Simultaneous estimation of covariance inflation and observation errors within an ensemble Kalman filter, *Q. J. Roy. Meteorol. Soc.*, 135, 523–533, 2009b.
- Mallet, V. and Sportisse, B.: Ensemble-based air quality forecasts: A multimodel approach applied to ozone, *J. Geophys. Res.*, 111, D18302, doi:10.1029/2005JD006675, 2006.
- Massart, S., Piacentini, A., and Pannekoucke, O.: Importance of using ensemble estimated background error covariances for the quality of atmospheric ozone analyses, *Q. J. Roy. Meteorol. Soc.*, 138, 889–905, 2012.
- Menut, L., Bessagnet, B., Khvorostyanov, D., Beekmann, M., Blond, N., Colette, A., Coll, I., Curci, G., Foret, G., Hodzic, A., Mailler, S., Meleux, F., Monge, J.-L., Pison, I., Siour, G., Turquety, S., Valari, M., Vautard, R., and Vivanco, M. G.: CHIMERE 2013: a model for regional atmospheric composition modelling, *Geosci. Model Dev.*, 6, 981–1028, doi:10.5194/gmd-6-981-2013, 2013.
- Miyazaki, K., Eskes, H. J., Sudo, K., Takigawa, M., van Weele, M., and Boersma, K. F.: Simultaneous assimilation of satellite NO₂, O₃, CO, and HNO₃ data for the analysis of tropospheric chemical composition and emissions, *Atmos. Chem. Phys.*, 12, 9545–9579, doi:10.5194/acp-12-9545-2012, 2012.
- Rouil, L., Honore, C., Vautard, R., Beekmann, M., Bessagnet, B., Malherbe, L., Meleux, F., Dufour, A., Elichegaray, C., Flaud, J.-M., Menut, L., Martin, D., Peuch, A., Peuch, V.-H., and Poisson, N.: PREV'AIR: An Operational Forecasting and Mapping System for Air Quality in Europe, *B. Am. Meteorol. Soc.*, 90, 73–83, 2009.
- Sakov, P. and Bertino, L.: Relation between two common localisation methods for the EnKF, *Comput. Geosci.*, 15, 225–237, 2010.
- Sandu, A. and Chai, T. Chemical Data Assimilation – An Overview, *Atmos. Environ.*, 2, 426–463, 2011.
- Schmidt, H., Derognat, C., Vautard, R., and Beekmann, M.: A comparison of simulated and observed ozone mixing ratios for the summer of 1998 in Western Europe, *Atmos. Environ.*, 35, 6277–6297, 2001.
- Schwinger, J. and Elbern, H.: Chemical state estimation for the middle atmosphere by four-dimensional variational data assimilation: A posteriori validation of error statistics in observation space, *J. Geophys. Res.*, 115, D18307–D18307, 2010.
- Semane, N., Peuch, V.-H., Pradier, S., Desroziers, G., El Amraoui, L., Brousseau, P., Massart, S., Chapnik, B., and Peuch, A.: On the extraction of wind information from the assimilation of ozone profiles in Météo-France 4-D-Var operational NWP suite, *Atmos. Chem. Phys.*, 9, 4855–4867, doi:10.5194/acp-9-4855-2009, 2009.
- Solazzo, E., Bianconi, R., Vautard, R., Appel, K. W., Moran, M. D., Hogrefe, C., Bessagnet, B., Brandt, J., Christensen, J. H., Chemel, C., Coll, I., van der Gon, H. D., Ferreira, J., Forkel, R., Francis, X. V., Grell, G., Grossi, P., Hansen, A. B., Jercevic, A., Kraljevic, L., Miranda, A. I., Nopmongcol, U., Pirovano, G., Prank, M., Riccio, A., Sartelet, K. N., Schaap, M., Silver, J. D., Sokhi, R. S., Vira, J., Werhahn, J., Wolke, R., Yarwood, G., Zhang, J., Rao, S., and Galmarini, S.: Model evaluation and ensemble modelling of surface-level ozone in Europe and North America in the context of AQMEII, *Atmos. Environ.*, 53, 60–74, doi:10.1016/j.atmosenv.2012.01.003, 2012b.
- Spangl, W., Schneider, J., Moosmann, L., and Nagl, C.: Representativeness and classification of air quality monitoring stations, Umweltbundesamt report, 2007.
- Talagrand, O. and Courtier, P.: Variational assimilation of meteorological observations with the adjoint vorticity equation. I: Theory, *Q. J. Roy. Meteorol. Soc.*, 113, 1311–1328, 1987.
- Tang, X., Zhu, J., Wang, Z. F., and Gbaguidi, A.: Improvement of ozone forecast over Beijing based on ensemble Kalman filter with simultaneous adjustment of initial conditions and emissions, *Atmos. Chem. Phys.*, 11, 12901–12916, doi:10.5194/acp-11-12901-2011, 2011.
- Tarasova, O. A., Brenninkmeijer, C. A. M., Jöckel, P., Zvyagintsev, A. M., and Kuznetsov, G. I.: A climatology of surface ozone in the extra tropics: cluster analysis of observations and model results, *Atmos. Chem. Phys.*, 7, 6099–6117, doi:10.5194/acp-7-6099-2007, 2007.
- Tippett, M. K., Anderson, J. L., Bishop, C. H., Hamill, T. M., and Whitaker, J. S.: Ensemble Square Root Filters*, *Mon. Weather Rev.*, 131, 1485–1490, 2003.
- Troen, I. and Mahrt, L.: A simple model of the atmospheric boundary layer: Sensitivity to surface evaporation, *Bound.-Lay. Meteorol.*, 37, 129–148, 1986.
- Valari, M. and Menut, L.: Does an Increase in Air Quality Models Resolution Bring Surface Ozone Concentrations Closer to Reality?, *J. Atmos. Ocean. Technol.*, 25, 1955–1968, 2008.
- Van Loon, M., Builtjes, P. J. H., and Segers, A. J.: Data assimilation of ozone in the atmospheric transport chemistry model LOTOS, *Environ. Model. Softw.*, 15, 603–609, 2000.
- Vautard, R., Builtjes, P., Thunis, P., Cuvelier, K., Bedogni, M., Bessagnet, B., Honoré, C., Moussiopoulos, N., Schaap, M., Stern, R., Tarrason, L., and 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, *Atmos. Environ.*, 41, 173–188, 2007.
- Vautard, R., Schaap, M., Bergström, R., Bessagnet, B., Brandt, J., Builtjes, P., Christensen, J., Cuvelier, C., Foltescu, V., Graff, A., Kerschbaumer, A., Krol, M., Roberts, P., Rouil, L., Stern, R., Tarrason, L., Thunis, P., Vignati, E., and Wind, P.: Skill and uncer-

- tainty of a regional air quality model ensemble, *Atmos. Environ.*, 43, 4822–4832, 2009.
- Visschedijk, A. J. H., Zandveld, P. Y. J., and Denier Van Der Gon, H. A. C.: A high resolution gridded European emission database for the EU Integrate Project GEMS, TNO, Apeldoorn, Netherlands, TNO report, 2007.
- Whitaker, J. S. and Hamill, T. M.: Ensemble Data Assimilation without Perturbed Observations, *Mon. Weather Rev.*, 130, 1913–1924, 2002.
- World Health Organization (WHO): Health aspects of air pollution with particulate matter, ozone and nitrogen dioxide, Report on a WHO Working Group, Regional Office for Europe; Bonn, Germany, EUR/03/5042688, 13–15 January 2003.
- 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.
- Zhang, Y., Bocquet, M., Mallet, V., Seigneur, C., and Baklanov, A.: Real-time air quality forecasting, part II: State of the science, current research needs, and future prospects, *Atmos. Environ.*, 60, 656–676, doi:10.1016/j.atmosenv.2012.02.041, 2012.
- Zyryanov, D., Foret, G., Eremenko, M., Beekmann, M., Cammas, J.-P., D’Isidoro, M., Elbern, H., Flemming, J., Friese, E., Kioutsioutkis, I., Maurizi, A., Melas, D., Meleux, F., Menut, L., Moinat, P., Peuch, V.-H., Poupkou, A., Razinger, M., Schultz, M., Stein, O., Suttie, A. M., Valdebenito, A., Zerefos, C., Dufour, G., Bergametti, G., and Flaud, J.-M.: 3-D evaluation of tropospheric ozone simulations by an ensemble of regional Chemistry Transport Model, *Atmos. Chem. Phys.*, 12, 3219–3240, doi:10.5194/acp-12-3219-2012, 2012.