



This discussion paper is/has been under review for the journal Geoscientific Model Development (GMD). Please refer to the corresponding final paper in GMD if available.

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. Beekmann²

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

Received: 22 April 2013 – Accepted: 14 May 2013 – Published: 30 May 2013

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

Published by Copernicus Publications on behalf of the European Geosciences Union.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Abstract

The Ensemble Kalman Filter is an efficient algorithm for data assimilation; it allows for an estimation of forecast and analysis error by updating the model error covariance matrices at the analysis step. This algorithm has been coupled to the CHIMERE chemical transport model in order to assimilate ozone ground measurements at the regional scale. The analyzed ozone field is evaluated using a consistent set of observations and shows a reduction of the quadratic error by about a third and an improvement of the hourly correlation coefficient despite of a low ensemble size designed for operational purposes. A classification of the European observation network is derived from the ozone temporal variability in order to qualitatively determine the observation spatial representativeness. Then, an estimation of the temporal behavior of both model and observations error variances of the assimilated stations is checked using a posteriori Desroziers diagnostics. The amplitude of the additive noise applied to the ozone fields can be diagnosed and tuned online. The evaluation of the obtained background error variance distribution through the Reduced Centered Random Variable standard deviation shows improved statistics. The use of the diagnostics indicates a strong diurnal cycle of both the model and the representativeness errors. Another design of the ensemble is constructed by perturbing model parameter, but does not allow creating enough variability if used solely. Finally, the overall filter performance over evaluation stations is found to be relatively unaffected by different formulations of observation and simulation errors.

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 to a risk of premature mortality which increases with its concentration (Bell et al. 2005). Integrated programs

GMDD

6, 3033–3083, 2013

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



such as GMES (Global Monitoring for Environment and Security) foster the development of environmental monitoring of ozone and other pollutants using a combination of state-of-the-art numerical models, 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/>) promote the 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 need well-defined modelling platforms which perform direct model simulation and forecast in synergy with observations. A review of such platforms has recently been performed by Kukkonen et al. (2012). They are 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.prevail.org, Rouil et al., 2009) delivers daily analysis and air quality forecasts of pollutant concentrations such as ozone, nitrogen oxides and particulate matter. This computational chain involves the rCTM CHIMERE at different spatial scales (Schmidt et al., 2001; Bessagnet et al., 2004). 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). The analysis product is a fundamental result which depicts the best representation of the surface pollutant concentrations. All the reporting activities, including calculations of air quality indicators (threshold exceedances ...) derived from PREV’AIR are based on the analysis.

One of the challenges of these air quality modelling chains is to provide the uncertainty associated to the modelling results. An ensemble of independent models has been evaluated 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

**Regional scale
ozone data
assimilation**

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



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.

A step further that is now recognized as crucial in the air quality community is the integration of chemical observations in the simulations by data assimilation methods (Carmichael et al., 2008; Zhang et al., 2012). After pioneering work in numerical weather predictions (Courtier et al., 1998; Houtekamer and Mitchell, 1998), methods already developed in this field have been successfully applied to air quality simulations: 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 studies were intended to obtain more accurate surface ozone analyzed fields as well as to improve the short-term forecast (Blond et al., 2004; Elbern and Schmidt, 2001). One key point of the data assimilation process is the representation of the Background Error Covariance Matrix (BECM) which needs to include simulation errors. Another key point is the construction of the Observation Error Covariance Matrix (OECM). The relative values of the BECM and of the OECM in the observation space allow weighting the confidence between observations and models. In addition, the BECM matrix will serve to propagate the innovations (observed value minus modeled one) where no observations are available. The principal origin of uncertainties in rCTM 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 species are rapidly transported out of the limited model domain reduces the importance of initial conditions in the overall error budget (Blond and Vautard 2004, Sandu and Chai 2011). Two major strategies are employed to obtain an accurate 4D-analysis of ozone concentrations; the first one is to produce a 4D-Var analysis which allows the correction of unobserved quantities such

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



as emissions rates or even wind fields (Elbern et al., 2007; Semane et al., 2009), and the second one is the use of an EnKF (Ensemble Kalman filter). In the latter method, an ensemble of simulations following a Monte-Carlo approach is used to determine a flow-dependant BECM matrix. By improving the model error representation, the work on ensemble design greatly helps to increase the analysis accuracy and especially the forecast performance of the EnKF (Constantinescu et al., 2007c, d; Wu et al., 2008; Agudelo et al., 2011; Tang et al., 2011; Curier et al., 2012). One approach employed to construct a representative ensemble is to perturb the main model parameters affecting the ozone error variance and correlation (Hanea et al., 2004). Another strategy used in data assimilation to estimate the model errors is to iteratively adjust the BECM using diagnostics such as Desroziers diagnostics which derives model 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 et al., 2004) in order to assimilate tropospheric ozone columns derived from the IASI instrument (Coman et al., 2012). In this paper, we assess the set-up of the Chimere-ENKF in the context of surface ozone data assimilation. We have built a consistent ensemble of assimilation and validation stations for a summertime episode (Fig. 1). A first focus of this paper is to assess the impact of the observation representativeness on the assimilation performance. Following Flemming et al. (2005) (hereafter denoted FLEM05), we have performed a classification of the ozone stations that takes into account the spatial representativeness for model grid cells. A second focus of the paper is to investigate different formulations and diagnostics of the model and observation error and their impact on the assimilation system skills. We investigate sensitivities to the BECM properties by perturbing model parameters together with the ozone state and by using 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. In addition, we have tested two alternative ways in the prescription of the OEMM by using the Desroziers diagnostic or the parameterization described in

Elbern et al. (2007). We compare the performance of the assimilation system using these different error formulations for a 10 days simulation over Europe including a photochemical episode. The paper is organized as follows: The EnKF assimilation scheme is described in Sect. 2; we present the classification of the surface ozone observations in Sect. 3 and the rCTM CHIMERE and its evaluation in Sect. 4. Section 5 shows the different assimilation experiment set-ups and results; the conclusion and future directions are given in Sect. 6.

2 Data assimilation: the EnKF Analysis scheme

First introduced by Evensen (1994, 2003), the EnKF is a sequential filter that allows for a relatively simple implementation of a sophisticated data assimilation scheme appropriate for a large 3-dimensional model. An ensemble of perturbed model states is created using Monte-Carlo methods and evolves forward in time in order to obtain a forecast \bar{x}^f and the associated ensemble mean value over the N ensemble members Eq. (1):

$$\bar{x}^f = \frac{1}{N} \sum_{i=1}^N x_i^f \quad (1)$$

Here x_i^f represents the vector of forecasted ozone concentrations and the subscript i indicates the ensemble number. Also, the BECM, noted as P^f is approximated by the ensemble spread over the N realizations of the model at a given time:

$$P^f = \frac{1}{N} \sum_{i=1}^N (x_i^f - \bar{x}^f)(x_i^f - \bar{x}^f) \quad (2)$$

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Then, when measurements y are available along with the OECM noted R , each ensemble member is updated following Eq. (3):

$$x_i^a = x_i^f + P^f H^T \left(H P^f H^T + R \right)^{-1} (y - H x_i^f) \quad (3)$$

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 (3) yields an analysis model state x_i^a which is used for the initialization of the next forecast. The model error covariance matrix is also updated at the analysis step:

$$P^a = P^f - P^f H^T \left(H P^f H^T + R \right)^{-1} H P^f \quad (4)$$

One of the drawbacks of this method is the introduction of sampling errors due to the limited ensemble size. This leads to an under-estimation of the model errors caused by an artificial decrease of the ensemble variance in Eq. (4) and to spurious correlation in the analyzed fields. This makes it necessary to localize the analysis spatially, 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 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). There are two types of algorithms to solve Eqs. (3) and (4), the stochastic EnKF, where observations are treated as random variables (Burgers et al., 1998; Houtekamer and Mitchell 1998; Evensen, 2003) and several form of deterministic EnKF that used a square root decomposition of the model error covariance matrix (Whitaker and Hamill, 2002; Tippett et al., 2003;

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Evensen et al., 2004; Hunt et al., 2007 and reference therein). These square root filters avoid the measurement's perturbations and thus have a lower analysis error by reducing this additional source of sampling error.

In an EnKF, the model error is sampled by an ensemble of model realizations which has the advantage of evolving in time (in contrast to a static BECM associated to OI methods). As mentioned above, with finite and generally small ensemble sizes, and due to significant model errors, the EnKF estimate of the BECM generally underestimates the true error covariance matrix. Thus, a particular strategy must be employed for inflating the BECM, by treating model and analysis errors in the same framework. In addition, the ensemble design must reflect the model 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 et al., 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. These are derived from a two-dimensional Gaussian distribution which does not introduce bias, with a given variance, and a given spatial correlation (Evensen et al., 2003). Concerning ozone assimilation with an EnKF, Constantinescu et al. (2007d) have shown that the most efficient way to maintain sufficiently dispersive ensemble was to apply an additive perturbation for the covariance inflation. More than just evaluating the consistency of the model error weight in the analysis, we propose here also to use the Desroziers et al. (2005) diagnostic to derive the model error variance. In practice, diagnostics are computed in the observation space over p observations following Eq. (5) for the model error and Eq. (6) for the observation error:

$$(\sigma_B)^2 = \frac{1}{p} \sum_{i=1}^p (y_i^a - y_i^f)(y_i^o - y_i^f) \quad (5)$$

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



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

Errors are thus estimated a posteriori using the observations y^o , the ensemble mean of the forecast y^f and of the analysis y^a . The background error variance has been diagnosed in a variational framework for the assimilation of total ozone columns in global CTM's (Massart et al., 2012; Schwinger and Elbern, 2010) and surface ozone observations (Jaumouillé et al., 2012). The tuning strategy of the covariance inflation factor in EnKF is to adaptively adjust the model error standard deviation prescription (Li et al., 2009b). In this study, the ability of the tuned ensemble to represent a more accurate error statistic has been demonstrated in assimilation exercises that take into account different ranges of true model errors. However, this error estimation requires that the forecast model mean bias is small. Besides, an accurate prior estimate of these observations and model error variances allows their simultaneous estimation although the solution is not unique (Li et al., 2009b; Schwinger and Elbern, 2010). The relation between the diagnostic of the model error in the observation space and in the model space is reliable in the case of a spatially and temporally dense observation network. In this case, the error in the model space can be sampled in the observation space. The background error correlation can also be adjusted (Schwinger and Elbern, 2010; Massart et al., 2012).

Another way to build the ensemble is to perturb physical parameters and uncertain inputs of the model to set up a more physically sound model ensemble, as shown by the work of Hanea et al. (2004). This approach is also tested in our study. For the CHIMERE model, following the work of Boynard et al. (2011), we add stochastic perturbations to the main error sources which are anthropogenic and biogenic emissions, the boundary conditions, the land-use for dry deposition, and meteorological variables (see table in supplementary material).

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



3 Characterization of the surface observations

Ozone observations are operated by national and regional networks across Europe and collected through the European Environment Agency 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, 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 criteria related to ozone concentrations. The area of representativeness associated to 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 meta-data such as population density, land-cover, emission's inventory, topography and transport models (Tarasova et al., 2007; Spangl et al., 2007; Henne et al., 2010). Another approach is based on the identification of the pollution regime following statistics on the ozone concentrations themselves (FLEM05, Joly and Peuch, 2012). This approach provides the advantage of being directly targeted on the pollutant species of interest (ozone) in this paper and 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 on the entire set of available background airbase stations with an altitude below 800 m above sea level (a.s.l), we derive four station types using the P50DV (Table 1 in FLEM05). These are background/mountain stations (MOU, $P50DV < 0.68$), rural stations (RUR, $0.68 < P50DV < 1.07$), suburban stations (SUB, namely 'U1' in FLEM05, $1.07 < P50DV < 1.45$) and urban stations (URB corresponding to "U2", "U3" and street ("S") classes in FLEM05, $P50DV > 1.45$). According

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Regional scale
ozone data
assimilation**

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



to this method, we find that the ozone hourly mean over each station type decreases when variability increases (Table 1). 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 countryside 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 here as not representative to the model grid spacing used for this study (0.5°×0.5°); 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 1). We plotted on the Fig. 2 the daily average ozone profile obtained for summer 2009. In this season, the suburban profile shows a higher ozone concentration in the daytime while keeping the lowest nighttime ozone level, which is in accordance to the variability criterion. Thus, differences in average ozone concentration between station's types are more important in the early morning and are minimum at 15:00 UTC where all station's types reach similar daily maxima. Thus, differences in average ozone concentration between station's types are more important in the early morning and are the lowest in the mid-afternoon when all station's types reach similar daily maxima. Observations associated to the lowest variability are mostly representative of a remote environment instead of a mountain's one because we reject stations with altitude above 800 m. However, background stations located between 300 m a.s.l and 800 m a.s.l exhibit a higher baseline in ozone concentrations probably associated to the tropospheric ozone vertical gradient (Chevalier et al., 2007). These stations are therefore discriminated from the remaining ones that are shared between remote continental and coastal stations under the influence of little polluted marine air masses such like the Mace-Head station in Ireland. Finally, the criterion can be extended to the whole Europe with some exceptions; for instance, we can notice that background stations with low ozone variability which are located in Scandinavia do not have the highest daily mean as observed usually for background 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 emissions regions close to urban areas such as Paris, Berlin or in the Pô Valley. Also, as seen from the Table 1, results obtained with the FLEM05 approach are often consistent with the Airbase stations where around 90 % of the MOU stations are “rural” and also more than 80 % of the SUB stations are “suburban” or “urban”. However, also 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 for a larger environment: these sites constitute around 40 % of our total RUR sites.

4 The regional Chemistry Transport Model CHIMERE

4.1 Model description

CHIMERE is a state-of-the-art rCTM (<http://www.lmd.polytechnique.fr/CHIMERE/>; Menut et al., 2013). The general formulation and the first evaluation at regional scale were presented by Schmidt et al. (2001). Our simulations cover the European continental domain GEMS (-15° , 35° E; 35° , 70° N) for 8 hybrid (σ , p) vertical levels from 995 hPa to 500 hPa. To reduce computational time, we keep a $0.5^{\circ} \times 0.5^{\circ}$ horizontal grid spacing (Fig. 1); also, aerosols (Bessagnet et al., 2004) are not included in the simulation. The mandatory meteorological variables are obtained from the Integrated Forecasting System (IFS) of the European Centre for Medium-Range Weather Forecasts (ECMWF). Analysis at 00:00 UT and 12:00 UT with three-hourly output forecast are linearly interpolated on an hourly basis and on our spatial domain. 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 global chemistry transport model MOZART-4 (Emmons et al., 2010) 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 O_3 , CO ,

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



HCHO, NO_x and SO₂. In this way, we obtain three-hourly boundary conditions that are temporally and spatially interpolated on the CHIMERE grid.

4.2 Evaluation of the reference simulation

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 is performed (i.e. from the 14th August to the 23rd of 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 the 19th to the 21st August (Fig. 6a). 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's accuracy in different environments, the evaluation is conducted separately for each station type using the station 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 (Fig. 3a) and also for the assimilation period (Fig. 3b).

Statistics over the whole summer demonstrate that daily maxima are on average well reproduced by the model. There is a minimum of the RMSE of about 8 ppb for all stations types in the afternoon (Fig. 3a) as well as for the daily maximum during the episode (Table 4). The decreasing amplitude of the diurnal cycle of ozone concentrations from the suburban to the background stations is also captured by the model. At SUB and RUR stations, comparisons between observations and simulations show a good correlation of around 0.7–0.8 (see “REF” statistics in Table 3); on average, observations are overestimated all day long, except mid-afternoon. These errors can be typically attributed to the model resolution that 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 a wrong representation of

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



night-time and early morning chemistry, especially ozone titration, and maybe also dry deposition. The background ozone level (i.e. the MOU stations) is also well captured; average errors range between 8 and 10 ppb, but the correlation is lower (0.53). Simulation of background stations above 300 m a.s.l. exhibits a much stronger positive bias during nighttime. This is probably an artifact in our evaluation process since for these stations we compare observations to the model ground level, while these stations are generally more representative of higher model layers which are less affected by dry deposition.

We also plot the average diurnal profile of ozone concentrations and for the RMSE during the period between the 14th August and the 23rd of August 2009 (Fig. 3b). We notice an increase in the ozone baseline by 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 SUB and RUR stations are underestimated by the model simulation. Background stations do not show the same behavior, the ozone level is even reduced with respect to summer average values; this is because some stations are located in north-eastern Europe or in the United Kingdom, where weather conditions were not favorable to ozone formation.

5 Set-up and results of assimilation experiments

We shall present in Sect. 5.1 the set-up and the results of the assimilation experiments applied to the 10 days summertime period in August 2009 including an ozone pollution episode. An evaluation of the overall performance of the data assimilation system is proposed (cf. Sect. 5.2). Further, we discuss the sensitivity induced by the design and the formulation of the BECM (Sect. 5.3) and the OECM (Sect. 5.4). A review of the different assimilation experiments is shown in the Table 2.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



5.1 Evaluation of the reference simulation

In order to increase the available data and the observation's spatial coverage, we include data from all the three stations types both for assimilation and evaluation. Then, for each observation type (i.e. MOU, RUR, SUB), we randomly divide the station set into two subsets (Fig. 1). In this way, we get one set for the assimilation with 350 stations (MOU: $N = 44$; RUR: $N = 117$; SUB: $N = 189$) and another one for evaluation with 344 stations (MOU: $N = 45$; RUR: $N = 112$; SUB: $N = 187$). The assimilation period extends from the 14th August at midnight to the 23rd August at 23:00 UTC, with an hourly assimilation step.

At each analysis step, we perform a local analysis with a horizontal cut-off radius of 250 km around the analyzed cell and within the boundary layer for vertical levels by using the diagnosed boundary layer height (BLH) for each time step and grid cell. These choices are made to avoid spurious correlation as the ensemble size is quite low. Only tri-dimensional ozone fields are included in the state vector at the analysis step. Ozone perturbations are added to the analysed state, the same noise is applied for all vertical layers inside the boundary layer. Pseudo-random fields are generated with a fixed horizontal decorrelation length (Evensen, 1994) of 200 km (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 with both values.

The inspection of relative errors of the CHIMERE reference run shows that these errors exhibit a clear diurnal cycle (left panel of Fig. 4). This is due to the fact that on the one hand the photochemical build up of ozone and the associated mid day maximum is quite well reproduced by the model (relative RMSE < 20 %); on the other hand, the night time minimum of ozone associated to the lowest daily BLH is partly missed (relative RMSE up to 40 %). We have chosen to prescribe perturbations which represent this error diurnal cycle (light blue curve on the left panel of Fig. 4). The resulting standard deviation of the forecast ensemble (green curve on the left panel of

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Fig. 4) is similar to the relative RMSE of this forecast (purple curve on the right panel of Fig. 4).

The prescription of the observation covariance matrix R is also a key point of the assimilation procedure. In general, assuming no observational error correlation, the OECM matrix R 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 et al. (2003) have determined an observation error of 5 ppb at 15 pm for a 6 km model resolution. Using observations from the AIRNOW database in the USA, Chai et al. (2006) got on average a daily range between 5 ppb and 13 ppb (at night) and assumed an observation error of 8ppb at 60 km horizontal resolution. Using the observational methods (Hollingsworth and Lönnberg, 1986), Flemming et al. (2004) got on average an absolute standard deviation of 5 ppb, independently from the station type of the FLEM05 classification. Following this last study, we first use for all types of stations an observation error standard deviation of 5 ppb which is typically used in data assimilation systems (Hanea et al., 2004; Wu et al., 2008). This 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, these set-ups will be referred to as the REF_ASSIM experiment; it corresponds 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 first assimilation experiment, we compare results from the reference run and the analysis (i.e. the mean of the analyzed ensemble) against observations from validation stations (not assimilated). Figure 6a shows the simulated surface ozone on the 20th August 2009 at 15:00 UTC along with the observations. The reference run correctly displays spatial patterns associated to the

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



**Regional scale
ozone data
assimilation**

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



episode in which a cold front separates Europe into two parts: a western part with marine air masses associated to low ozone values, and an eastern part associated to 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 which seem to be more realistic as confirmed by observations; as an example, 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 in Denmark (10.4° E, 55.4° N) along with the analysis and the reference run is plotted on Fig. 7. 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 the 20th of 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 illustrate the ability of the ensemble to extend innovations along with the ozone flow (directed to north-west), which is a clear advantage of the Kalman filter with a variable BECM as compared to optimal interpolation methods with a fixed BECM.

In addition to these particular elements, we present the statistics for the whole set of sites and for each station type of validation stations on a daily basis for the hourly profile, 8-h mean maxima and daily peaks (REF_ASSIM statistics in Tables 3 and 4). 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 3). Both the analysis and the reference run only show 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 for 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 4) while reducing the negative bias for rural and suburban stations.

The error profile shows globally reduced values and still shows a daily cycle with a minimum of 6 ppb during daytime for all station types (Fig. 5). Contrasting results are obtained for night time regarding the station types; 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 be the result of reduced spatial representativeness of the observations early in the morning. In the next sections, we will investigate how different formulations in the model and observation error formulation can change the assimilation performance.

5.3 Sensitivity to model error

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 adjusting the covariance inflation factor based on diagnostics of previous assimilations, or by perturbing physical parameters that are controlling the model error to get a more physical ensemble and by 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 Table 5), then we present the combined perturbation of ozone and models parameters (NEWPAR, Table 5), and finally the combination of the model parameter perturbation with the online tuning (NEWPAR_MOD_DESR Table 5). We propose to use here an appropriate diagnostic of the ensemble design, useful to compare the different experiments, which is the Reduced Centered Random Variable diagnostic (Candille et al., 2006).

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



near Athens), the ozone field is controlled by only 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.

5 A further sensitivity study is performed by creating different ensemble members using a perturbation of the input data of the model, instead of perturbing the ozone fields directly. Choices of the perturbed parameters, their standard deviation and spatio-temporal correlation (Table in supplementary material) are inspired from several previous studies (Hanna et al., 2001, Beekmann and Derognat, 2003, Hanea et al., 2004, 10 Wu et al., 2008 and Boynard et al., 2010). The scheme of the perturbation is similar to the one applied to the ozone fields with fixed horizontal spatial decorrelation (Evensen et al., 2003), however, the distribution is log-normal and we include a temporal correlation. The standard deviation of the 20 ensemble members of the free run (without assimilation) with perturbed parameter reaches a maximum of 8 ppb at some loca- 15 tions after four days of simulation, but is only about 1 ppb on average, so it generally underestimates the model error. In our case, the model parameters are not included in the state vector, the ensemble size is low and the ensemble spread is reduced at the analysis step, therefore, this ensemble cannot create enough variability (ensemble spread) between two analyses, i.e. during a one hour time step. Thus we choose to 20 combine this perturbation of the model parameters to the classical perturbation of the ozone field itself. We then reproduce the two previous experiments with the addition of the parameter perturbations, namely NEWPAR associated to the REF_ASSIM one and NEWPAR_MOD_DESR for the online tuning experiment (MOD_DESR) applying this new way of making hybrid perturbations.

25 In order to check and compare the ensemble dispersion from these simulations, we calculate the (RCRV) (Candille et al., 2006). For each observation i of the system, the

GMDD

6, 3033–3083, 2013

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



RCRV is defined by the ratio between the innovation and the associated error Eq. (7):

$$\text{RCRV} = \frac{y_i^o - \bar{H}y_i^f}{\sqrt{\sigma_o^2 + \sigma_b^2}} \quad (7)$$

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 variances (σ_o^2) and model error variances (represented by the ensemble variance, σ_b^2). Here the observation error variance is set to 25 ppb² for all experiments; ensemble mean and variance (σ_b^2) are calculated using the Desroziers diagnostic Eq. (5) in the observation space. For a representative ensemble (i.e. when the diagnosed errors are comparable to the innovations), the RCRV should be normally distributed with a zero mean and a standard deviation of 1.

The mean of RCRV (Fig. 9) indicates the relative 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. Figure 9 indicates that ensemble predictions have a weak positive bias in the afternoon and in the evening and a negative bias at 7 a.m. and around 7 p.m. In the morning and in the early evening, the bias is important and is formed rapidly. At that time, the model simulation can be very uncertain, due to increasing or still large emissions and a developing (morning) or fading (evening) PBL. 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 Pô Valley and in the south-eastern Europe (not shown). Globally, a slightly lower bias is found for REF_ASSIM and NEWPAR, for which analyses are closer to the observations, but only small differences are found between the different experiments. In fact, the analysis field is not biased

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



because the bias is removed by the assimilation procedure. The use of the diagnostic leads to an increase of the additive noise, thus, it allows a prediction of the increase of the errors due to the higher bias (Figs. 8 and 9). A step further would be the use of a bias correction procedure or the correction of the model parameters in order to improve the ensemble forecast and subsequently the assimilation performance.

The standard deviation of the RCRV provides a framework for the verification of the ensemble spread, namely, if it is greater (lower) than one, this means that the ensemble is under dispersive (over dispersive) (Fig. 10). The tuned simulations (MOD_DESR and NEWPAR_MOD_DESR) display a correct error budget during day time with a standard deviation of the RCRV close to 1. By contrast, the ensemble is under dispersive for all simulations at night time. The addition of the perturbation of model parameters slightly improves the error representation during nighttime. As the observational error is taken as constant in the calculation of the RCRV statistics, Desroziers diagnostic of the night time error could also suggest that observation errors are underestimated in the night (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, as the ensemble spread is controlled by the added noise, it is preferable to reduce the amount of perturbation. Particularly, the ensemble standard deviation can become important in large areas where observation are absent or scarce, this leads 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 observation are assimilated (Fig. 6b, c). Furthermore, we can notice that the model parameter perturbations (slightly) improve the results in all cases.

GMDD

6, 3033–3083, 2013

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



5.4 Sensitivity to the observation error formulation

Similarly to the previous section (i.e. 5.3) for the BECM, we propose to examine the sensitivity of the system to the formulation of the OECM. In the case of surface ozone observations, measurement errors are generally weak, but representativeness errors can be important. This latter error is model dependent and particularly on its horizontal resolution, in our case 0.5° . 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); first, it shows that the diagnosed errors are rather similar with the prescribed error for remote stations (MOU) but larger for RUR and SUB station types. For SUB sites, this is probably linked to the coarse horizontal resolution of our model (0.5°) that is not representative enough for urban environments; this result is close to the one of Chai et al. (2006) for similar grid spacing. Moreover, the diagnosed error shows a diurnal cycle with lowest values at noon (about ~ 6 ppb for SUB and RUR sites), close to the prescribed value, but are higher at nighttime and in the morning. More specifically, SUB 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 model error, we apply the diagnostic Eq. (6) of the observational error diurnal profile from the previous day to the OECM, as a function of the stations type (TUNING_OBS_TYPE experiment). Figure 11 (middle 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 skill scores do not show a

GMDD

6, 3033–3083, 2013

Regional scale
ozone data
assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



significant improvement or modification with respect to the reference experiment (Table 4), only a slight reduction of the RMSE is found for the mountain stations due to their reduced prescribed error.

As an alternative estimation of the observational error we apply the parameterization following the work of Elbern et al. (2007). Errors are calculated explicitly as the sum of the measurement error variance (set here to 3 % (sd_{meas}) of the instantaneous observation value with a minimum of 3 ppb (Parrish and Fehsenfeld, 2000) and of the representativeness error. The latter is defined by the ratio of the model resolution and the observation representativeness length scale. Here, we used the distance (x) between the stations and the centers of the analyzed cell to the observation location divided by a length scale (L_{repr}) prescribed in Elbern et al. (2007) to 2 km for suburban, 10 km for rural and 20 km for background sites. In particular, for the case when only one suburban station is assimilated within the localized radius of 250 km, we wanted to reduce its influence independently from the decorrelation of model errors by increasing the observational errors variances along with the distance. It means that the distance to the station is explicitly taken into account in the observational error. The formulation of the parameterization is written here as:

$$R = r_{meas}^2 + r_{repr}^2 = (sd_{meas} * y_i^o)^2 + \left(\sqrt{\frac{(x)}{L_{repr}}} * 1.2 \right)^2 \quad (8)$$

In this experiment, called R_EXPLICIT, the observation error variance is therefore changing for each observation site and each analyzed model grid cell. Figure 11 (right panel) shows both the average observational errors for each class of station type determined with Eq. (8) and the diagnostic daily error profile. In the R_EXPLICIT case, as the representativeness term increases with the distance, it is dependent of the model grid and not unique; here we average the observation error variance over the closest grid cell to each site. This method gives a realistic spread of values of average observational errors for different station types between 3 and 6 ppb. It gives slightly better scores for suburban stations. But it actually relies on the observation classification and

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



the specific choice of a representativeness length scale associated to the station types. Following Eqs. (5) and (6) of the diagnostics for each time step, one can see that the sum of the diagnosed observational and the model 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. Also, it should be noted that the diagnostic efficiency can be reduced by the model bias; it can lead to an overestimation of observation error variances (Li et al., 2009b). Although observational errors strongly depend on site representativeness and type, the sensitivity to the observational error is globally small. Then, a change of the observational errors of assimilated stations does not improve the scores for validation stations.

6 Conclusions

In this paper we introduced a new data assimilation chain based on the rCTM CHIMERE in an EnKF framework using ozone surface observations provided by the European database Airbase. The system is based on a local analysis computed by a deterministic square root filter and is applied to an ensemble of 20 perturbed CHIMERE rCTM simulations. For a 10 days period in summer 2009, the reference run shows good performance with root mean square errors (RMSE) around 10 ppb and an average correlation coefficient of 0.7. However, it generally overestimates night time and morning concentrations, while it underestimates midday ozone peaks in particular during the regional ozone pollution episode that occurs during this period. Similar results are found for a longer 3 months summer period. The statistical evaluation of the analysis field using a substantial set of unassimilated observations indicates a 30 % reduction of the RMSE which reaches on average a minimum around 6 ppb in the afternoon, regardless the stations type. The analysis increment can be propagated at synoptic scale, thus allowing corrections downwind to the continental observations, as for example shown for northern Europe. More than 95 % of evaluation stations show a decrease in their

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



RMSE. The average of the ensemble forecast, corrected each hour by the filter, shows also a better performance than the reference run but has a more limited success at transitions between day and night and vice versa. This is due to the formation of a positive bias of the 1-hour ensemble forecast, which is more pronounced for suburban stations. Therefore, the model error is in this case not sensitive to the improvement of ozone initial conditions. As suggested by other studies, a step further would be the simultaneous adjustment of precursor's initial conditions and emissions rate (Tang et al., 2011) or other models parameters (Hanea et al., 2004; Agudelo et al., 2011). Also, the application of a bias correction procedure combine with the data assimilation scheme (Li et al., 2009a) could improve the EnKF performance.

A series of different assimilation experiments are performed with different formulations of the OECM and BECM matrices. Considering values of model and observational errors, the results from different formulations are rather different. We focused on the estimation of the hourly variation of the model and observational errors using Desroziers diagnostics. We used the RCRV standard deviation as a tool to evaluate the result of the tuning of the model error obtained by the diagnostic. The ensemble dispersion is more consistent during daytime while an underestimation of the model error is found during night time; it is attributed to the larger and unaccounted observational error at that time. Generally, the diagnostic indicates a large contribution of the observational error in the evaluated quadratic error of the ensemble forecast. As pointed 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's sites allowed diagnosing observational errors for specific types of sites. As expected, larger representativeness errors were estimated for suburban than for rural and remote sites. The spatial representativeness shows as expected a daily variation with a maximum in the daytime, and a nighttime reduction according to the increase of the ozone precursor's level close to the sources. In addition, skill scores (at evaluation stations) do not show a substantial sensitivity to the modification of observational error variance. This underlies the needs to determine a priori the observation representative-

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



ness such as the FLEM05 classification. The diagnostic can provide a determination of observation data suitable for data assimilation (observation thinning) since the observation error and analysis sensitivity can be estimated individually for each stations. Stations used for assimilation or evaluation are spatially close, where the ozone observations network is dense. Then, in terms of evaluation scores, small changes are found among experiments even for substantial changes (up to a factor of two) in the model and observation errors. Given the low ensemble size, it suggests the robustness of this data assimilation system.

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

Supplementary material related to this article is available online at:

<http://www.geosci-model-dev-discuss.net/6/3033/2013/gmdd-6-3033-2013-supplement.pdf>

Acknowledgements. This work has been supported by the projects MACC and MACC-II (Monitoring Atmosphere and Climate Change) from the European Community's Seventh Framework Programme (FP7 THEME [SPA.2011.1.5-02]) under grant agreement no. 283576 and by the French project ADOMOCA (Assimilation de Données pour les Modèles de Chimie Atmosphérique). Most of the charts have been made using the Python plotting library Matplotlib (Hunter, 2007). B. Gaubert PhD was supported by a grant from the French Ministry of Higher



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

5 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), 50th IEEE Conference, 4430–4435, doi:10.1109/CDC.2011.6160444, 12–15 December 2011.
- 10 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.
- 15 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., 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.
- 20 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.
- 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.
- 25

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Burgers, G., Van Leeuwen, P. J., and Evensen, G.: Analysis scheme in the ensemble Kalman filter, *Mon. Wea. Rev.*, 126, 1719–1724, 1998.
- 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, 112, D12S15, doi:10.1029/2006JD007763, 2007
- 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., 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, 2007c.
- 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, 2007d.
- Courtier, P., Andersson, E., Heckley, W., Vasiljevic, D., Hamrud, M., Hollingsworth, A., Rabier, F., Fisher, M., and Pailleux, J.: The ECMWF implementation of three-dimensional variational assimilation (3D-Var). I: Formulation, *Q. J. Roy. Meteorol. Soc.*, 124, 1783–1807, 1998.
- 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-

**Regional scale
ozone data
assimilation**

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



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., Schmidt, H., and Ebel, A.: Variational data assimilation for tropospheric chemistry modeling, *J. Geophys. Res.*, 102, 15967–15985, 1997.

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

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.

Emmons, L. K., Walters, S., Hess, P. G., Lamarque, J.-F., Pfister, G. G., Fillmore, D., Granier, C., Guenther, A., Kinnison, D., Laepple, T., Orlando, J., Tie, X., Tyndall, G., Wiedinmyer, C., Baughcum, S. L., and Kloster, S.: Description and evaluation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4), *Geosci. Model Dev.*, 3, 43–67, doi:10.5194/gmd-3-43-2010, 2010.

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, 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., 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.

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.

**Regional scale
ozone data
assimilation**

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Flemming, J., Stern, R., and Yamartino R. J.: A new air quality regime classification scheme for O_3 , NO_2 , SO_2 and PM_{10} observations sites, *Atmos. Environ.*, 39, 6121–6129, 2005.
- 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.
- 5 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.
- 10 Hamill, T. M., Whitaker, J. S., and Snyder, C.: Distance-dependent filtering of background error covariance estimates in an ensemble Kalman filter, *Mon. Wea. 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.
- 15 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.
- Honoré, C. C., Rouil, L., Vautard, R., Beekmann, M., Bessagnet, B., Dufour, A., Elichegaray, C., Flaud, J.-M., Malherbe, L., Meleux, F. D. R., 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.
- 20 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.
- 25 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, *Bull. Am. Meteorol. Soc.*, 89, 1147–1164, 2008.
- 30 Houtekamer, P. L. and Mitchell, H. L.: Data assimilation using an Ensemble Kalman Filter technique, *Mon. Weather Rev.*, 126, 796–811, 1998.

**Regional scale
ozone data
assimilation**

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Houtekamer, P. L. and Mitchell, H. L.: A Sequential Ensemble Kalman Filter for Atmospheric Data Assimilation, *Mon. Wea. 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*, 230, 112–126, 2007.
- Hunter, J. D.: Matplotlib: A 2D Graphics Environment, *Comput. Sci. Eng.*, 9, 3, 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.
- 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. Wea. 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, 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. Meteorol. Soc.*, 138, 889–905, 2012.

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- Menut, L., Bessagnet, B., Khvorostyanov, D., Beekmann, M., Colette, A., Coll, I., Curci, G., Foret, G., Hodzic, A., Mailler, S., Meleux, F., Monge, J.-L., Pison, I., Turquety, S., Valari, M., Vautard, R., and Vivanco, M. G.: Regional atmospheric composition modeling with CHIMERE, *Geosci. Model Dev. Discuss.*, 6, 203–329, doi:10.5194/gmdd-6-203-2013, 2013.
- 5 Rouil, L., Honore, C., Vautard, R., Beekmann, M., Bessagnet, B., Malherbe, L., Meleux, F. D. R., 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, *Bull. 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.
- 10 Sandu, A. and Chai, T.: Chemical Data Assimilation – an Overview, *Atmosphere*, 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.*, 15, 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.
- 20 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., Jericevic, A., Kraljevic, L., Miranda, A. I., Nopmongkol, 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.
- 25 Spangl, W., Schneider, J., Moosmann, L., and Nagl, C.: Representativeness and classification of air quality monitoring stations, Umweltbundesamt report, 2007.
- 30

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



- 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.
- 5 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. Wea. Rev.*, 131, 1485–1490, 2003.
- Valari, M. and Menut, L.: Does an Increase in Air Quality Models Resolution Bring Surface
10 Ozone Concentrations Closer to Reality?, *J. Atmos. Ocean. Technol.*, 25, 1955–1968, 2008.
- Van Loon, M., Bultjes, 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., Bultjes, P., Thunis, P., Cuvelier, K., Bedogni, M., Bessagnet, B., Honoré, C., Mousiopoulos, N., Schaap, M., Stern, R., Tarrason, L., and van Loon, M.: Evaluation and inter-
15 comparison 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., Bultjes, 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 uncertainty of a regional air
20 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,
25 *Mon. Wea. 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
30 algorithms for ozone forecasts, *J. Geophys. Res.*, 113, 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.*, 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.

5

GMDD

6, 3033–3083, 2013

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Regional scale
ozone data
assimilation

B. Gaubert et al.

Table 1. Number of stations, daily mean and median of the daily variability (P50DV) average over the different stations types (lines) of the FLEM05 classification (MOU, RUR, SUB). 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 stations type.

	N	mean (ppb) (period: JJA 2009)	P50DV (period: JJA 2009)	Rural (<i>N</i> = 405)	Suburban (<i>N</i> = 256)	Urban (<i>N</i> = 391)
MOU(>300 m)	47	38.8	0.56	42 (89 %)	4 (9 %)	1 (2 %)
MOU(<300 m)	42	36.5	0.58	32 (76 %)	3 (7 %)	7 (17 %)
RUR	228	33.2	0.9	98 (43 %)	40 (18 %)	90 (39 %)
SUB	376	31.6	1.17	63 (17 %)	70 (19 %)	243 (65 %)
URB	359	30,1	1.46	170 (47 %)	139 (39 %)	50 (14 %)

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Table 2. List of assimilations experiments with different formulations of OECM and BECM variances.

	OECM (variance)	BECM (variance)
REF_ASSIM	25 ppb ²	fixed hourly profile
MOD_DESR	25 ppb ²	diagnosed hourly profile
NEWPAR	25 ppb ²	parameter perturbations
NEWPAR_MOD_DESR	25 ppb ²	diagnosed profile and parameter perturbations
TUNNING_OBS_TYPE	f(stations type)	fixed hourly profile
R_EXPLICIT	f(stations type, distance)	fixed hourly profile

Regional scale
ozone data
assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Table 3. Hourly skill scores (ppb) for the CHIMERE run (REF) and the REF_ASSIM experiment (ANA) calculated for the evaluation set for suburban stations ($N = 187$), rural stations ($N = 112$), and for background stations located below an altitude of 300 m above the sea level ($N = 18$) and above 300 m ($N = 27$).

	TOT	MOU < 300	RUR	SUB	MOU > 300
RMSE (REF)	10.79	8.5	11.04	10.86	15.77
RMSE (Ana)	7.76	7.23	8.16	7.58	13.06
Bias (REF)	-0.42	-1.91	0.78	-0.99	7.95
Bias (ana)	0.18	0.67	1.43	-0.62	7.61
Correlation (REF)	0.75	0.53	0.71	0.8	0.43
Correlation (ana)	0.87	0.69	0.85	0.9	0.67

Regional scale
ozone data
assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

⏪

⏩

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Table 4. daily maximum and maximum of the daily eight-hour ozone mean RMSE average over the evaluation data set.

RMSE (ppb)	CHIMERE	REF_ASSIM	TUNNING_OBS_TYPE	R_EXPLICIT
8 h avg max	7.83	4.89	4.89	4.79
Daily Max	10.74	7.21	7.27	7.04
	MOD_DESR	NEWPAR	NEWPAR_MOD_DESR	
8 h avg max	4.91	4.81	4.88	
Daily Max	7.31	7.15	7.25	

Regional scale
ozone data
assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Table 5. Comparison of the statistical score for each assimilation experiment over the validation 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 ($N = 27$).

	TOT	MOU < 300	RUR	SUB	MOU > 300
RMSE (ppb)					
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
REF_ASSIM	7.76	7.23	8.16	7.58	13.06
TUNNING_OBS_TYPE	7.8	7.29	8.14	7.65	12.96
R_EXPLICIT_REF_ASSIM	7.7	7.17	8.11	7.51	13.16
Bias (ppb)					
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
REF_ASSIM	0.18	0.67	1.43	-0.62	7.95
TUNNING_OBS_TYPE	0.08	0.58	1.28	-0.68	7.77
R_EXPLICIT_REF_ASSIM	0.02	0.27	1.23	-0.73	7.77
Correlation coefficient					
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
REF_ASSIM	0.87	0.69	0.85	0.9	0.67
TUNNING_OBS_TYPE	0.87	0.68	0.85	0.9	0.67
R_EXPLICIT_REF_ASSIM	0.87	0.69	0.85	0.91	0.67

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

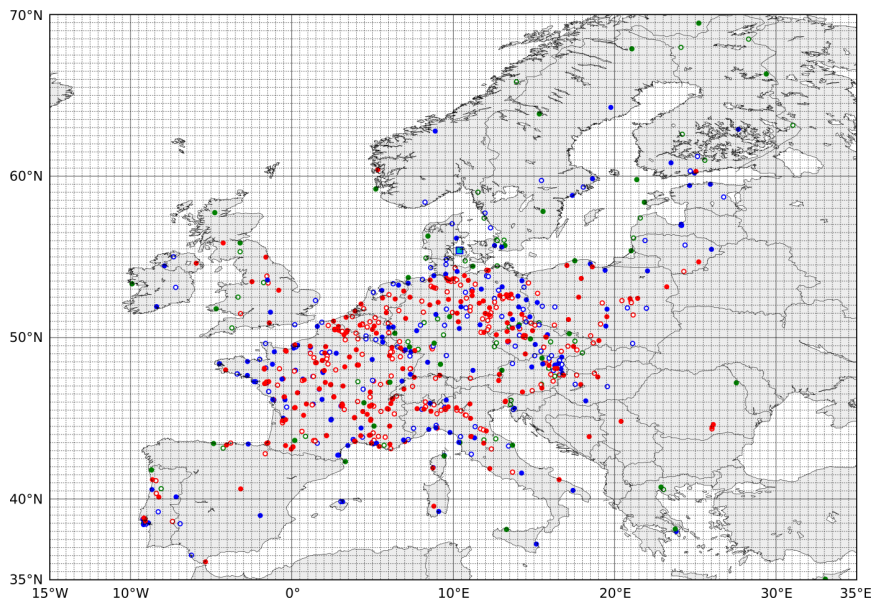


Fig. 1. Modeled domain and horizontal grid spacing (0.5°) with sites retained for assimilation (filled circles) and validation (filled in white circles). Colors indicate the station type with red for suburban, blue for rural and green for background stations. The rural evaluation station for which the ozone concentration time series is plotted in Fig. 7 (located in Odense in Denmark) is shown as a square filled in cyan.

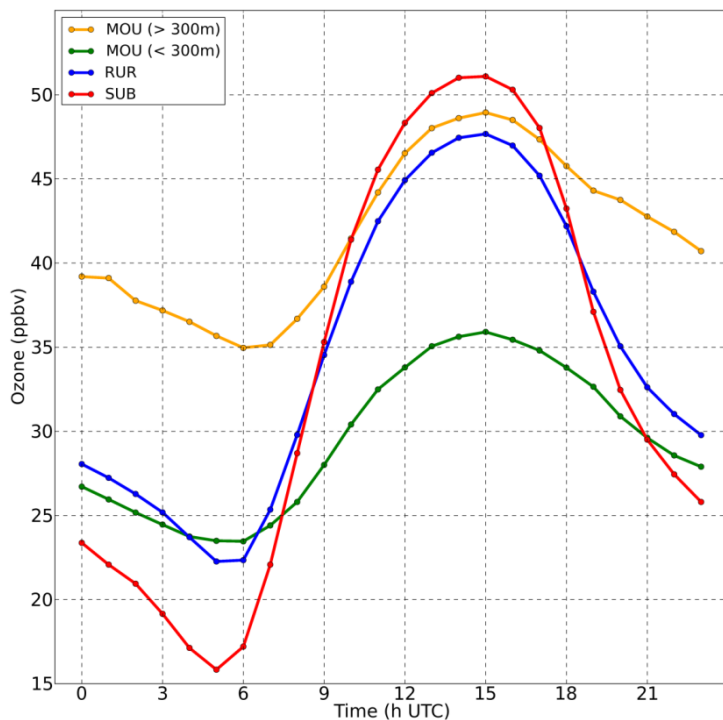


Fig. 2. Average daily ozone profile during JJA 2009 for each type of observation chosen for assimilation and validation. Colors indicate the station type with red for suburban, blue for rural, green for background stations under 300 m a.s.l and orange for background stations between 300 m and 800 m a.s.l.

Regional scale ozone data assimilation

B. Gaubert et al.

Title Page

Abstract Introduction

Conclusions References

Tables Figures

◀ ▶

◀ ▶

Back Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Regional scale
ozone data
assimilation

B. Gaubert et al.

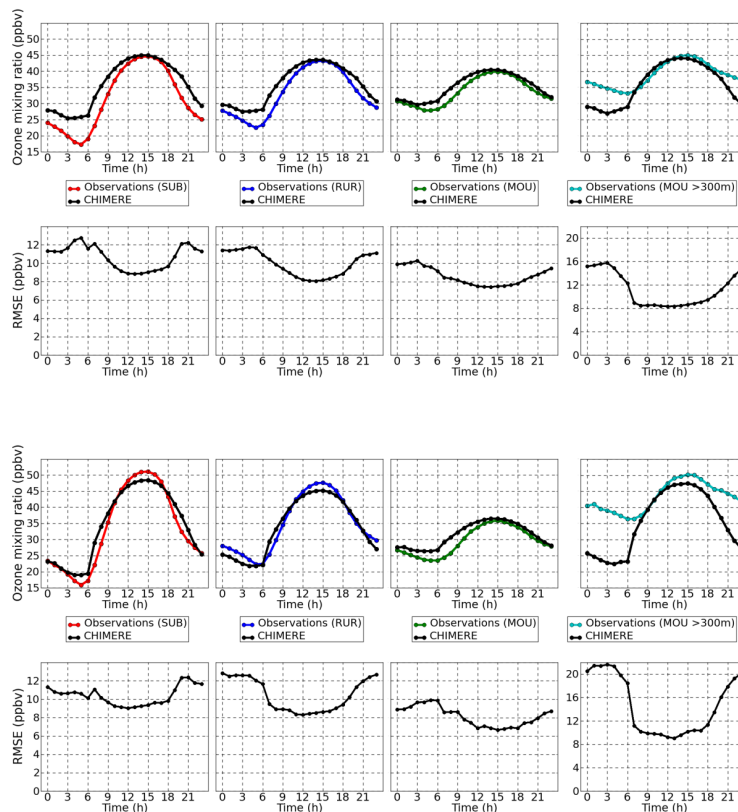


Fig. 3. Simulated (CHIMERE) and observed average diurnal ozone cycle and their associated RMSE during (a) the 2009 summer (top) and (b) the assimilation period (bottom). Results are shown for the suburban stations (left), rural stations (middle left), and for background stations located below (middle right) and above 300 m a.s.l (right). The corresponding number of stations is indicated in Table 1.

Regional scale
ozone data
assimilation

B. Gaubert et al.

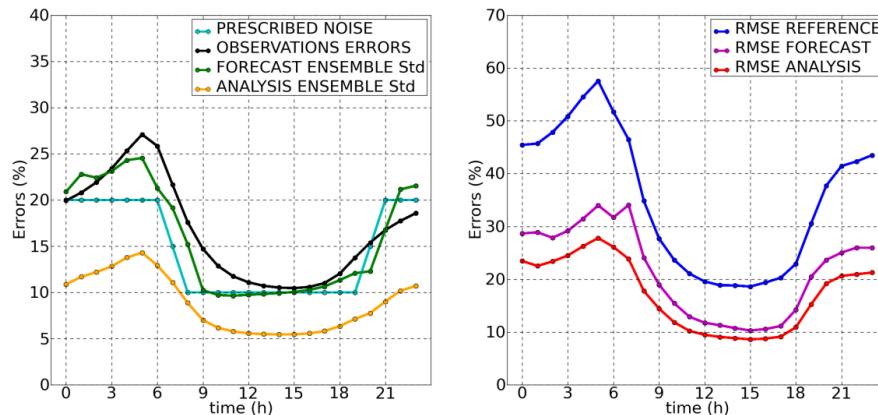


Fig. 4. Relative Errors of the assimilated stations ($N = 295$) for the REF_ASSIM experiment. On the left panel, we plot the diurnal profile of the white noise prescribed to the model ensemble after each assimilation step (light blue curve) and the observation error standard deviation (black curve). 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).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Regional scale
ozone data
assimilation

B. Gaubert et al.

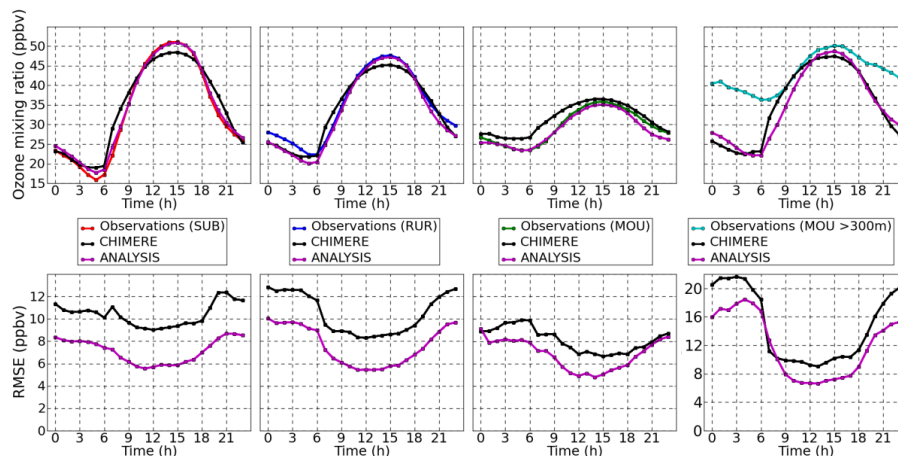


Fig. 5. Average diurnal cycle of ozone and the associated RMSE (for the assimilation period at validation stations) for the CHIMERE reference run and the analysis. Separated panels are shown for the suburban stations (left, $N = 172$), rural stations (middle left, $N = 102$), and for background stations with an altitude below 300 m a.s.l (middle right, $N = 17$) and above (right $N = 27$).

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Regional scale
ozone data
assimilation

B. Gaubert et al.

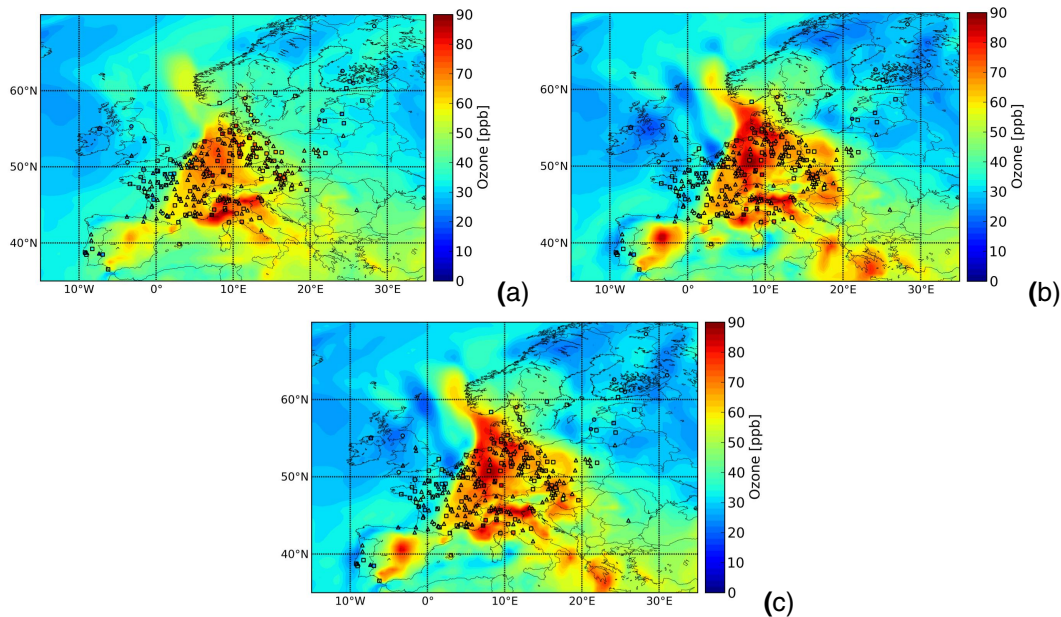


Fig. 6. Simulated ozone fields for **(a)** the CHIMERE reference run (top), **(b)** the REF_ASSIM analysis (middle), and **(c)** the MOD_DESR analysis (bottom) on 15th August at 03:00 UTC. Ozone measurements at validation stations are plotted by circles, squares and triangles for respectively MOU, RUR and SUB stations type.

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Regional scale
ozone data
assimilation

B. Gaubert et al.

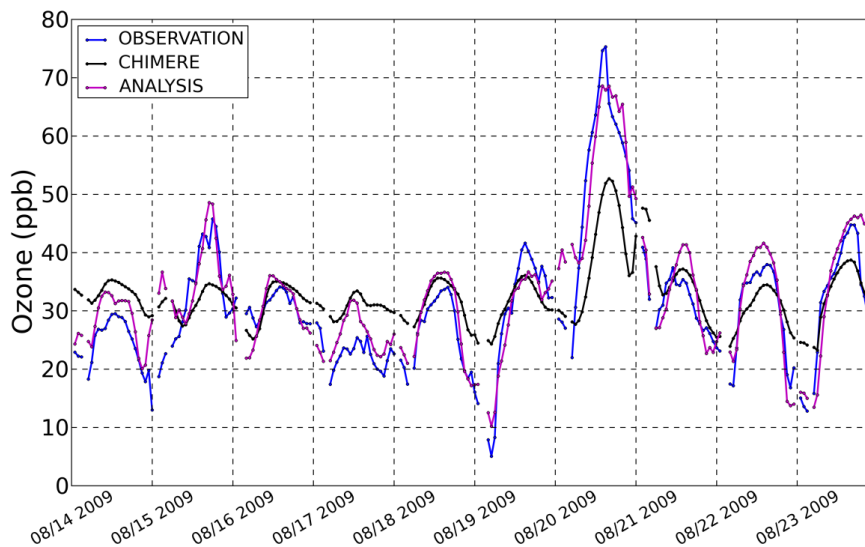


Fig. 7. Ozone measured (blue line) at the Odense site (not assimilated) in Denmark (10.4° E, 55.4° N) compared with CHIMERE reference run (black line) and the analysis (purple line).

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Regional scale
ozone data
assimilation

B. Gaubert et al.

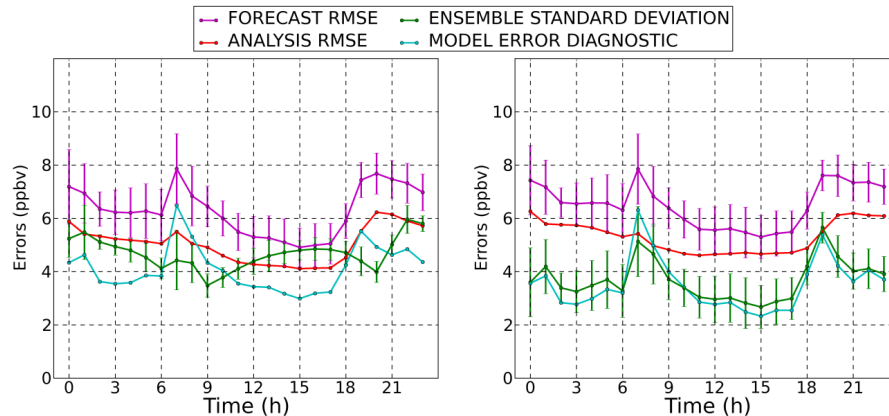


Fig. 8. Mean hourly value (over all assimilated observations) of the RMSE for the ensemble mean analysis (red curve), the one hour ensemble forecast (purple curve), the ensemble standard deviation (green) and the model error Desroziers diagnostic (light blue). The left panel displays these values for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

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures

◀

▶

◀

▶

Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion



Regional scale
ozone data
assimilation

B. Gaubert et al.

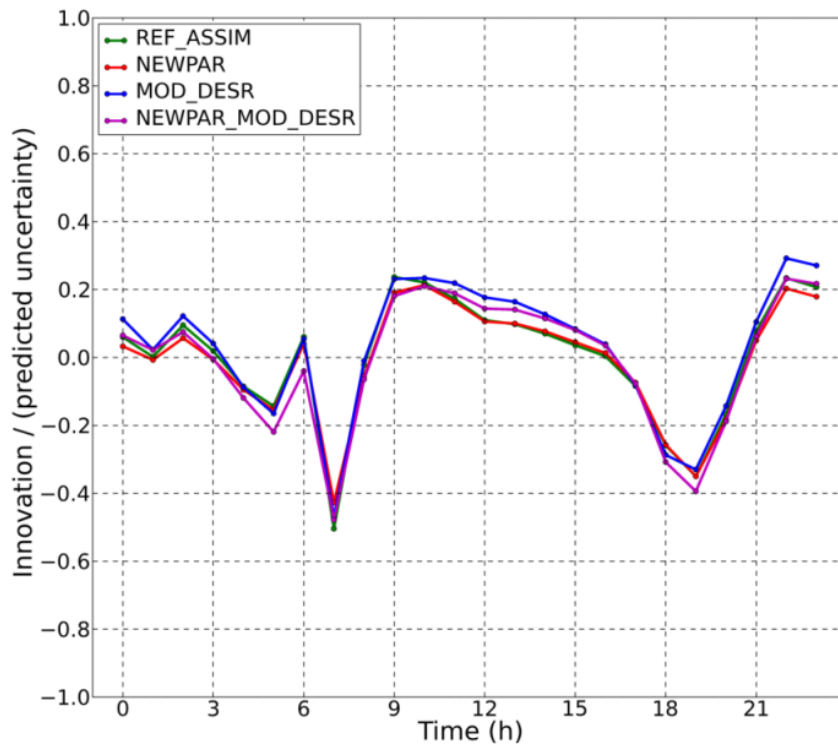


Fig. 9. Mean of the RCRV averaged over the whole assimilation steps for the REF_ASSIM, NEWPAR, MOD_DESR, and the NEWPAR_MOD_DESR assimilation experiment.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[◀](#)[▶](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Regional scale
ozone data
assimilation

B. Gaubert et al.

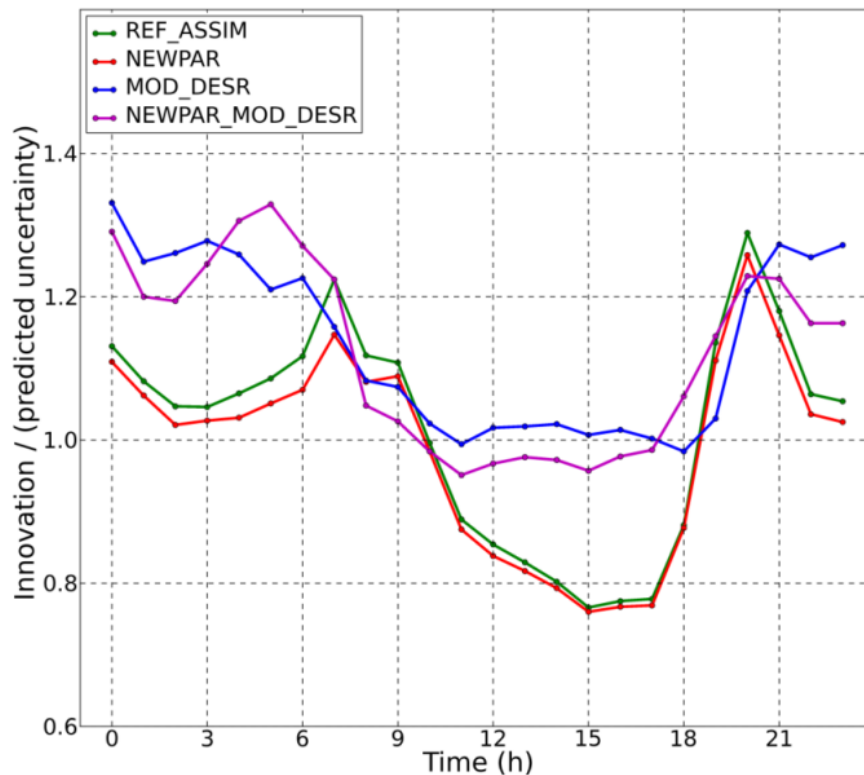


Fig. 10. Standard deviation of the RCRV average over all assimilation steps for the REF_ASSIM, NEWPAR, MOD_DESR, and the NEWPAR_MOD_DESR assimilation experiment.

[Title Page](#)[Abstract](#)[Introduction](#)[Conclusions](#)[References](#)[Tables](#)[Figures](#)[⏪](#)[⏩](#)[◀](#)[▶](#)[Back](#)[Close](#)[Full Screen / Esc](#)[Printer-friendly Version](#)[Interactive Discussion](#)

Regional scale ozone data assimilation

B. Gaubert et al.

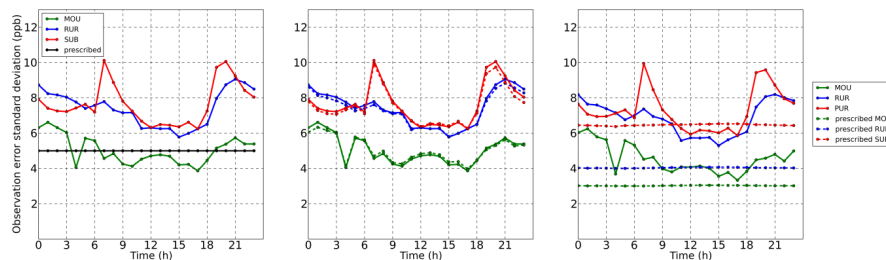


Fig. 11. Prescribed observational error variance and corresponding diagnostics averaged over all assimilation steps for background sites below 300 m a.s.l (23 stations), rural (106 stations) and suburban (166 stations) stations for the REF_ASSIM (left panel), TUNING_OBS_TYPE (middle panel), and R_EXPLICIT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

Title Page

Abstract

Introduction

Conclusions

References

Tables

Figures



Back

Close

Full Screen / Esc

Printer-friendly Version

Interactive Discussion

