RMetS
Royal Meteorological Society

Three-dimensional variational data assimilation of ozone and fine particulate matter observations: some results using the Weather Research and Forecasting – Chemistry model and Grid-point Statistical Interpolation†

M. Pagowski,[a,b]* G. A. Grell,[a,c] S. A. McKeen,[a,c] S. E. Peckham[a,c] and D. Devenyi[a,c]

[a]*NOAA Earth System Research Laboratory (ESRL), Boulder, Colorado, USA*
[b]*Cooperative Institute for Research in the Atmosphere (CIRA), Colorado State University, Fort Collins, Colorado, USA*
[c]*Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado at Boulder, Colorado, USA*
*Correspondence to: M. Pagowski, NOAA Earth System Research Laboratory, Global Systems Division, R/GSD1, 325 Broadway, Boulder, CO 80305-3328, USA. E-mail: mariusz.pagowski@noaa.gov

In operational air-quality forecasting, initial concentrations of chemical species are often obtained using previous-day forecasts with limited or no account for the observations. In this article we assess the role that assimilation of surface measurements of ozone and fine aerosols can play in improving the skill of air-quality forecasts. An assimilation experiment is performed using the Weather Research and Forecasting – Chemistry model and Grid-point Statistical Interpolation, a three-dimensional variational assimilation tool. The modelling domain covers the northeastern region of North America. The measurements come from the United States Environmental Protection Agency AIRNow network and are available hourly. Background error covariance statistics are derived from forecasts in July 2004. Comparison of forecasts issued in August and September 2006 and initialized with and without the assimilation follows. Results show that forecasts of ozone and fine aerosol concentrations benefit from the assimilation in terms of standard verification scores for a period of at least 24 hours. However, significant reduction of errors as a consequence of the assimilation is accompanied by fast model error growth in the early forecast hours. Published in 2010 by John Wiley & Sons, Ltd.

Key Words: data assimilation; chemistry; air quality; modelling

Received 13 March 2009; Revised 28 May 2010; Accepted 19 August 2010; Published online in Wiley Online Library 22 October 2010

Citation: Pagowski M, Grell GA, McKeen SA, Peckham SE and Devenyi D. 2010. Three-dimensional variational data assimilation of ozone and fine particulate matter observations: Some results using the Weather Research and Forecasting – Chemistry model and Grid-point Statistical Interpolation. *Q. J. R. Meteorol. Soc.* **136**: 2013–2024. DOI:10.1002/qj.700

1. Introduction

Data assimilation is an essential part of weather forecasting in all major meteorological centres. Until now, however, operational air-quality forecasts have been often initialized using concentrations of species obtained from the previous day's forecast with limited or no account for the observations.

This lack of widespread application of data assimilation in air-quality modelling, compared to weather forecasting, is both a result of the complexity of the problem (the number of chemical species varies in the model from tens to hundreds and can be many times the number of atmospheric state variables) and the scarcity of chemical observations compared to meteorology (especially with respect to vertical profiles). The increasing presence of satellites, unmanned

aircraft systems, and chemical measurements by instruments mounted on commercial aircraft will likely improve availability of data for assimilation into chemical transport models in the near future and potentially contribute to the improvement of air-quality forecasts.

Inaccuracies in predicted winds, atmospheric stability and insolation are recognized as prominent sources of errors in air-quality forecasts. Chemical transport models add new error sources to forecasts that arise from uncertainties in parametrizations of chemical processes, initial concentrations of species, emission sources, and, for regional models, concentrations of species on the lateral boundaries. The role of these factors features prominently in the literature devoted to air-quality modelling (e.g. Guenther *et al.*, 1993; Jonson and Isaksen, 1993; Mathur *et al.*, 1994; Alapaty *et al.*, 1995; Sillman and Samson, 1995; Zhang and Rao, 1999; Biswas and Rao, 2001; Hanna *et al.*, 2001; Barna and Knipping, 2006; Lee *et al.*, 2008).

In meteorology, data assimilation has been traditionally applied to improve initial conditions. In air quality, data assimilation can have a broader application that not only improves initial conditions of species concentrations but also provides tools for better estimation of emission sources known to contain large errors. To account for interactions between physical and chemical processes (e.g. cloud processes and aerosol formation, effect of aerosols on atmospheric radiation), an ideal data assimilation system would allow a simultaneous assimilation of meteorological data and concentration of species into a meteorological–chemical model.

Applications of data assimilation to modelling atmospheric chemistry include the work of Elbern and collaborators (Elbern *et al.*, 1997, 2000, 2007; Elbern and Schmidt, 1999, 2001) who developed a four-dimensional variational (4D-Var) system that is being used with the MM5 meteorological model (www.mmm.ucar.edu/mm5) and the European Air pollution Dispersion chemical transport model (EURAD, www.eurad.uni-koeln.de/index_e.html). In a series of publications they demonstrated the feasibility and usefulness of 4D-Var assimilation in idealized and real data settings. Assimilated species included ozone, NO_x, SO_2 and CO. Through an inversion procedure they also adjusted emission rates of species and explicitly addressed the relative role of initial conditions vs. emission source estimates of ozone precursors, SO_2 and $SO_4{}^{-2}$. In conclusion, the optimized emission factors markedly improved SO_2 forecasts and moderately improved ozone forecasts; in the latter case, initial conditions were a major factor in the improvement while emissions had a smaller effect, possibly due to the inadequate model resolution, chemical mechanism biases and the coarse NO_x observation network.

Daescu and Carmichael (2003), Carmichael *et al.* (2003), Sandu *et al.* (2005), Chai *et al.* (2007) and Constantinescu *et al.* (2007a, 2007b, 2007c, 2007d) describe an Ensemble Kalman Filter (EnKF) and 4D-Var assimilation systems that include the Sulphur Transport Eulerian Model (STEM: Carmichael *et al.*, 1991). Studies with different applications of EnKF and with 4D-Var in idealized and real data settings concentrated on ozone forecasting and showed that the EnKF approach to chemical data assimilation is promising but methods of covariance inflation and localisation, that were essential for filter performance, required further investigation. 4D-Var and EnKF were also used to reduce uncertainties in emission source estimates and to adjust boundary conditions. Even though improvement in the forecasting skill was noted, no conclusion could be drawn on the relative importance of initial conditions or emission and lateral boundary condition adjustments.

Kahnert (2008) applied three-dimensional variational (3D-Var) data assimilation to assimilate lidar observations of an aerosol backscattering coefficient with the High Resolution Limited-Area Model (HIRLAM: Gustafsson *et al.*, 2001). Currently, the European Centre for Medium-range Weather Forecasts (ECMWF) pursues assimilation of aerosol optical depths from Moderate-resolution Imaging Spectroradiometer (MODIS) instruments to predict bulk aerosol concentrations (Benedetti and Fisher, 2007; Benedetti *et al.*, 2009; Morcrette *et al.*, 2009). Despite these research achievements, use of chemical data assimilation in real-time air-quality forecasting remains rare.

Below, we assess the role that assimilation of surface measurements of ozone and fine aerosols can play in improving the skill of air-quality forecasts. We present results that were obtained with a 3D-Var system that includes the Weather Research and Forecasting – Chemistry model (WRF-Chem: Grell *et al.*, 2005) and Grid-point Statistical Interpolation (GSI: Wu *et al.*, 2002; Purser *et al.*, 2003a, 2003b). We show that even a simple approach in chemical data assimilation of ozone and fine particulate matter leads to gains in the skill of chemical model forecasts. This can be accredited to the improvement in initial conditions with only a slight increase of computational time.

We describe observations, our modelling system, experiment design and results that we obtained from the assimilation. Discussion and recommendations for future endeavours are presented in the conclusion.

2. Observations

Ozone and fine particulate matter are the main components of smog, and their presence in the atmosphere poses a serious health hazard for people. Thus, the accurate prediction of these species' concentration levels has significant human and economic cost implications.

We define fine particulate matter as aerosols with a diameter smaller than 2.5 μm, hereafter referred to as $PM_{2.5}$.

Surface ozone and $PM_{2.5}$ concentrations are obtained through the US Environmental Protection Agency (EPA) AIRNow network. Ozone from the AIRNow network is measured by ultraviolet (UV) absorption with instrument requirements specified under the US National Ambient Air Quality Standards (NAAQS) (see Federal Register, Vol. 73, No. 60, 40 CFR Parts 50 and 58, available at www.epa.gov/air/ozonepollution/actions.html). Uncertainty of the ozone measurements, which is a necessary parameter within the assimilation procedure, is taken as 1 ppbv. $PM_{2.5}$ from the AIRNow network is measured by Tapered Element Oscillating Microbalance (TEOM), which has been certified since 1990 for $PM_{2.5}$ measurements under the NAAQS (see Federal Register, Vol. 71, No. 200, 40 CFR Parts 50, available at www.epa.gov/air/particles/actions.html). Uncertainty of $PM_{2.5}$ measurements is taken to be 2 $\mu g\,m^{-3}$. Large measurement errors of $PM_{2.5}$ mass that are due to species volatility and depend on atmospheric conditions are possible (e.g. Hitzenberger *et al.*, 2004). However, formulations to estimate such errors for general applications are not available.

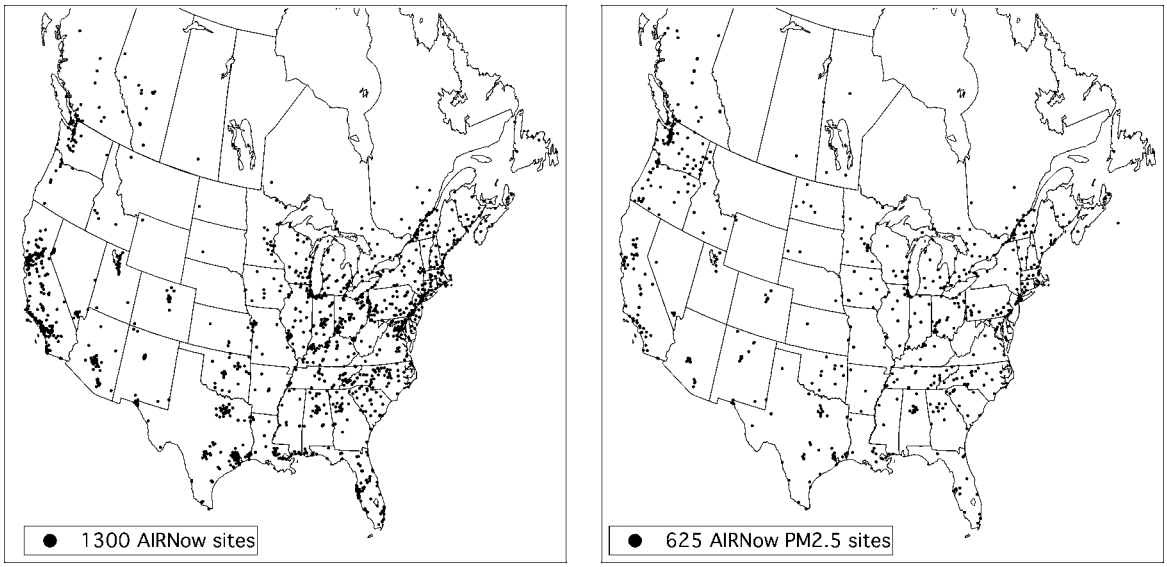

Figure 1. Maps of measurement sites in AIRNow network for real-time ozone (left) and $PM_{2.5}$ (right).

The AIRNow network provides hourly-averaged ozone and $PM_{2.5}$ concentrations over the USA and Canada. These measurements are readily available around the clock with minimal delay and can be used in real time for data assimilation or model evaluation. About 1300 stations measure ozone, with roughly half that number measuring $PM_{2.5}$. Maps of the stations measuring both constituents are shown in Figure 1. The density of stations is highest in the eastern part of the USA, followed by California and eastern Texas, while observations are relatively sparse in the middle of the continent. Monitors for both species are located mostly in urban and suburban settings. In this presentation, only results from the assimilation of surface monitors shown in Figure 1 are presented. Work on assimilation of upper-air ozone and $PM_{2.5}$ measurements within the GSI framework is currently underway.

3. Modelling system

WRF-Chem is an atmospheric model with online chemistry that simultaneously predicts weather and atmospheric composition and, consequently, allows interactions between meteorology and chemistry such as aerosol–radiation or aerosol–microphysics feedbacks. The model includes two dynamical cores and a wide range of parametrizations of boundary layer, radiation, microphysics, convection, ocean and soil as well as multiple choices of formulations for gaseous and aerosol reactions. Specific choices of parametrizations in this study are given in section 4.

GSI is a 3D-Var assimilation tool in which analysis is obtained by minimization of a cost function given by

$$\mathrm{J}(\mathbf{x}) \equiv (\mathbf{x}-\mathbf{x_b})^{\mathrm{T}}\,\mathbf{B}^{-1}\,(\mathbf{x}-\mathbf{x_b}) + \left(\mathbf{y}-\mathrm{H}(\mathbf{x})\right)^{\mathrm{T}}\mathbf{R}^{-1}\left(\mathbf{y}-\mathrm{H}(\mathbf{x})\right). \quad (1)$$

In this equation, $\mathbf{x}$ is a vector of analysis, $\mathbf{x_b}$ is the forecast or background vector, $\mathbf{y}$ is an observation vector, H is an observation operator and $\mathbf{B}$ is the background error covariance matrix. $\mathbf{R}$ is the observation error covariance matrix; each element of the matrix combines measurement and representativeness errors. In the GSI, the matrix $\mathbf{B}$ is separated into vertical and horizontal components (the latter can be anisotropic); it is represented as a product of error variances and correlation matrices which are modelled with recursive filters (Purser *et al.*, 2003a, 2003b). Application of the latter requires specification of background error correlation length scales.

It is noteworthy that the analysis variables that the GSI uses are: stream function, unbalanced part of velocity potential, unbalanced part of temperature, unbalanced part of surface pressure and pseudo-relative humidity. In our case, analysis variables also include ozone and $PM_{2.5}$. The analysis is univariate in ozone and $PM_{2.5}$. Even though one aspect of the GSI is that unbalanced variables are used in the minimization, and this currently has no implication for chemical applications, it is possible that in the future, regression might be employed to establish relationships between the balanced part of state variables and concentration of species. If employed, this might also minimize the adjustment of concentrations of species due to imbalances introduced by the assimilation.

4. Experimental design

In the data assimilation experiments we use an Advanced Research WRF dynamical core (ARW: Skamrock *et al.*, 2005) and the following parametrizations of physical and chemical processes: Yonsei University boundary-layer parametrization (YSU: Hong and Pan, 1996); Noah land-surface model (Chen and Dudhia, 2001); Grell–Dévényi convective parametrization (Grell and Dévényi, 2002); WSM-5 microphysics (Hong *et al.*, 2004); Dudhia short-wave (Dudhia, 1989); Rapid Radiative Transfer Model long-wave radiation (RRTM: Mlawer *et al.*, 1997); Regional Acid Deposition Model gaseous chemistry (RADM-2: Stockwell *et al.*, 1990, 1997); Modal Aerosol Dynamics for Europe/Secondary Organic Aerosol Model mechanism (MADE/SORGAM: Ackermann *et al.*, 1998; Schell *et al.*, 2001).

Lateral boundary conditions for meteorological state variables in the WRF-Chem come from the operational Weather Research and Forecasting – Nonhydrostatic Mesoscale Model (WRF-NMM, www.dtcenter.org/wrf-nmm/users/docs/user_guide/V3/index.pdf) forecasts at

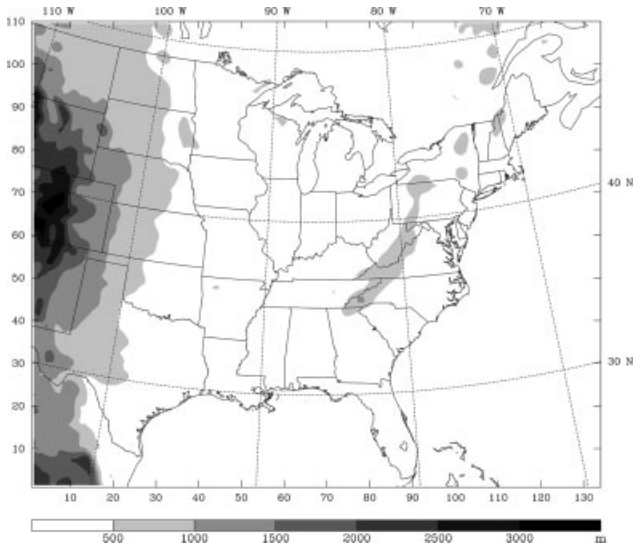

Figure 2. Modelling domain and topography.

12 km grid resolution available from the Environmental Modeling Center (EMC) of the National Oceanic and Atmospheric Administration (NOAA). Data assimilation was not attempted for meteorological state variables in this work; instead, meteorological state variables in WRF-Chem are initialized daily with the operational analyses obtained at EMC using WRF-NMM forecasts and the GSI.

Boundary conditions for chemical species in the simulations are laterally and temporally invariant. They are specified based on measurements onboard National Aeronautics and Space Administration (NASA)-sponsored aircraft missions as derived by McKeen *et al.* (2002). We believe that application of such lateral boundary conditions in the model is defensible given difficulties in matching grid resolutions, emission inventories and chemical speciation of regional and global models. For example, Tang *et al.* (2007) showed that predictions of ozone by global models differed substantially and thus were a major source of uncertainty for lateral boundary conditions in a regional model. They also found by comparing regional model simulations with aircraft observations that mean ozone distributions below 3 km were insensitive to global boundary conditions. We are not aware of an analogue study for $PM_{2.5}$ but expect that similar conclusions would hold, since concentrations of both species in the lower troposphere are dominated by local processes.

Initial conditions for chemical species at the very beginning of simulation periods are also specified based on the measurements derived by McKeen *et al.* (2002). Similar to model evaluation studies by McKeen *et al.* (2007, 2009), a five-day model spin-up period was employed to allow for the adjustment of concentrations of chemical species.

Emissions of ozone and $PM_{2.5}$ precursors are those referenced in Kim *et al.* (2006). These are based on the US EPA 1999 National Emissions Inventory (version 3) with updates of major electrical generating facilities to the observed July 2004 emissions from the Continuous Emissions Monitoring Network. The modelling domain covers the eastern part of North America with a horizontal grid spacing equal to 27 km and 34 vertically stretched levels up to the model top at 50 hPa. The domain is shown in Figure 2.

The observation operator represents linear horizontal interpolation. No extrapolations are performed in the vertical, and surface observations are assumed to coincide with the first model level. Assumptions on fluxes for extrapolations to the first model level would be quite uncertain when using similarity theory and not likely to improve reliability of the assimilation. Since observations are largely confined to flat terrain and altitudes of some measurement sites are not available, differences between model topography and reality were neglected. Unlike in observations, the observation operator does not include temporal averaging. For consistency, the same simplifications (with respect to vertical location and temporal averaging) were applied in the model evaluation. A one-hour assimilation window matches the averaging time of observations.

Observations that exceeded unrealistic values (200 ppbv for ozone and 150 μg m^{-3} for $PM_{2.5}$) and those for which the difference between the measurements and interpolated model concentrations exceeded threshold values were eliminated from the assimilation. Otherwise, no quality control was performed on observations.

Our observation covariance estimate error, represented by diagonal matrix **R** in the expression for the cost function in section 3, is a sum of measurement and representativeness errors. Measurement errors were described in section 2. Representativeness error is parametrized with Elbern *et al.*.'s (2007) formula given by $\varepsilon_{\mathrm{repr}} = \varepsilon_{\mathrm{abs}} \times (\Delta x / L_{\mathrm{repr}})^{1/2}$, where $\varepsilon_{\mathrm{abs}}$ is 'characteristic absolute error', a tunable parameter, Δx is model grid size and L_{repr} represents the radius of influence of an observation. Experimentally, we found $\varepsilon_{\mathrm{abs}} = 1/2\, \varepsilon_{\mathrm{obs}}$ to be adequate and prescribed L_{repr} equal to 10 km, 4 km and 2 km for rural, suburban and urban sites, respectively. We prefer this approach to thinning of observations since we aim to obtain grid concentration in the analyses that remains in relation to all measurements within the model grid. Given the observation operator's simplicity and the relatively small number of observations, the computational time of the assimilation procedure is minimal.

To determine horizontal and vertical length-scales of error correlations and variances for ozone and $PM_{2.5}$, the US National Meteorological Centre (NMC) method (Parrish and Derber, 1992) was used. In the NMC method, background-forecast errors are approximated by differences between forecasts valid at the same time but issued at different times. To eliminate effects of model biases due to the diurnal cycle, 48-hour forecasts were issued at 0000 UTC and 1200 UTC in July and August 2004 and differences were calculated for forecasts at 24 and 48 hours verified at 0000 UTC and 1200 UTC. Error variances were averaged horizontally for each model level. Horizontal and vertical error correlation length-scales were obtained at all model levels by a least-square fitting of Gaussian functions to correlation-distance relations (in the vertical the distance was measured in model levels rather than physical distance).

Vertical variability of horizontal length-scales of error correlation and background error standard variation for ozone and $PM_{2.5}$ at 0000 UTC and 1200 UTC are shown in Figure 3. The increase of horizontal correlation scales in the upper troposphere is comparable to meteorology, where a similar trend is observed (e.g. Daley, 1991), except for somewhat erratic model behaviour at the tropopause where the length-scales decrease; this being a consequence of the prescribed climatological boundary conditions. Horizontal length-scales for ozone are generally shorter than length-scales for $PM_{2.5}$. Diurnal variation of the vertical length-scales is primarily associated with the evolution of the boundary layer and is small in the upper troposphere.

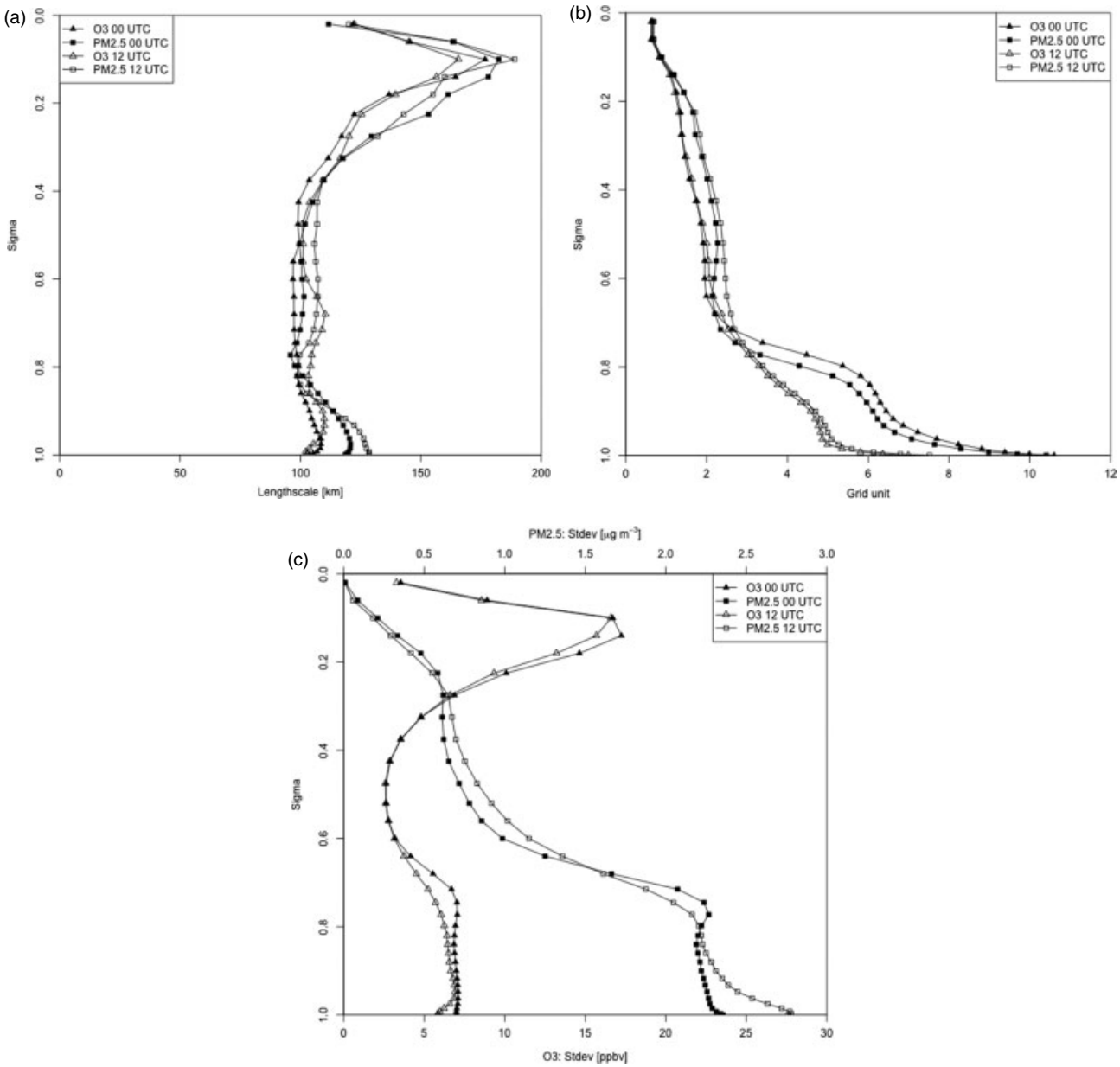

Figure 3. Background error correlation metrics at 0000 UTC and 1200 UTC for ozone and $PM_{2.5}$: top – horizontal length-scales, middle – vertical length scales, bottom – standard deviations. Note that the vertical length-scales are plotted in units of the vertical grid. In this figure values of sigma (1.0, 0.8, 0.6, 0.4, 0.2, 0.0) on the ordinate correspond approximately to values of atmospheric pressure (965, 780, 600, 415, 230, 50) hPa.

Standard deviation of the background error is directly related to the species concentrations. For ozone, the largest values are in the upper troposphere, followed by a high concentration in the lower troposphere, and minimum concentration in the middle troposphere. For $PM_{2.5}$, the largest values are observed at the surface and at the top of the boundary layer in the late afternoon (0000 UTC) and at the surface in the early morning (1200 UTC). Above the boundary layer, the standard deviation of the background error of $PM_{2.5}$ decreases rapidly.

It is important to note that the aerosol parametrizations employed in the model carry multiple species and sizes of aerosol particles. For consistency with $PM_{2.5}$ measurements, the total model concentration of aerosol species with diameters below 2.5 μm is used. An ad hoc approach was employed in which total $PM_{2.5}$ mass following the assimilation was distributed to different species and sizes based on their *a priori* contributions to the total $PM_{2.5}$ mass. Alternatively, the distribution could have been achieved through the background-forecast-derived covariances between aerosol constituents and total $PM_{2.5}$. However, given the lack of concentration measurements of aerosol species and sizes, and uncertain reliability of background error covariances, this approach, which also requires considerably larger effort, was not pursued.

In the experiment, a 12-hour assimilation cycle was employed to issue 24-hour forecasts at 0000 UTC and 1200 UTC in August and September 2006 to be compared with forecasts without data assimilation (control).

5. Results and evaluation

The evaluation of the control and forecasts with the assimilation is valid for 10 August to 30 September 2006. To prevent 'overfitting', experiments were performed

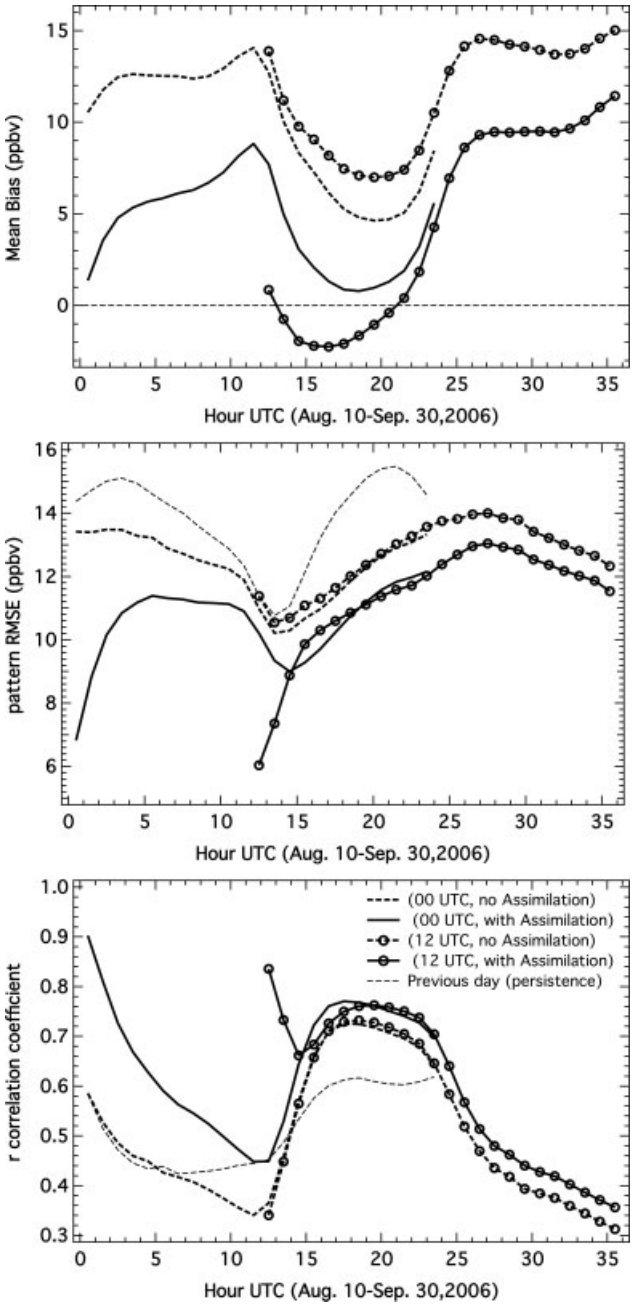

Figure 4. Time series of evaluation statistics for control forecasts (thick dashed line) and forecasts with assimilation (solid line) for ozone: top – bias, middle – pattern RMSE, bottom – correlation coefficient. Previous-day persistence plotted with thin dashed line.

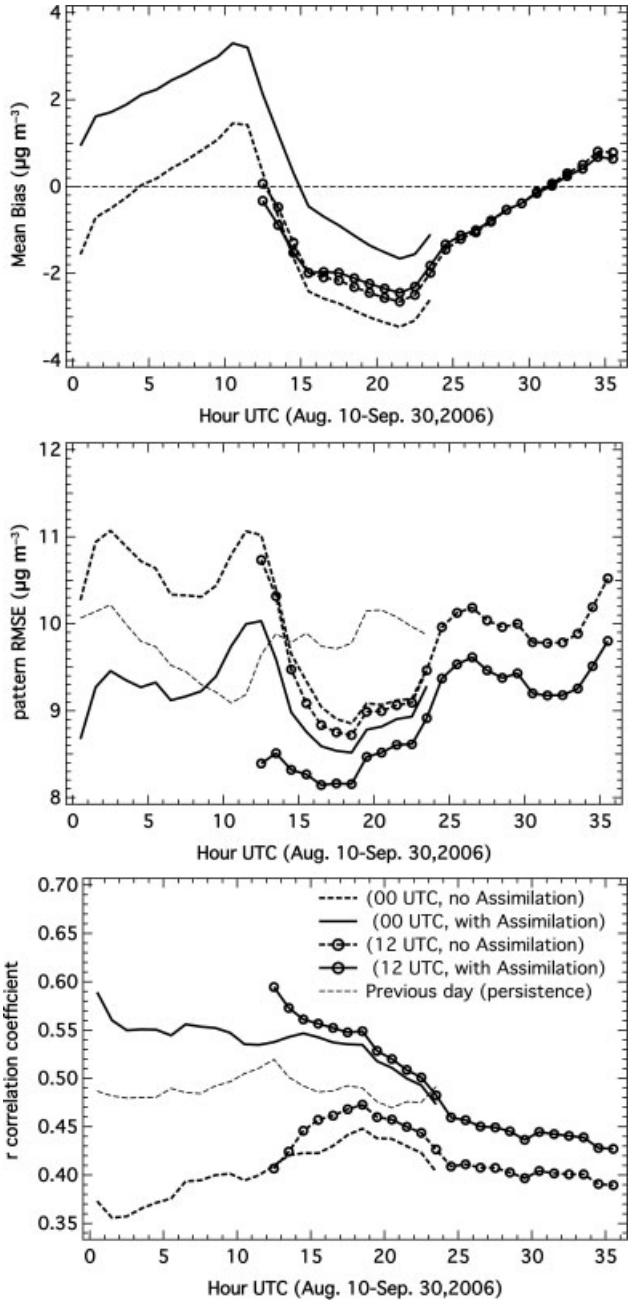

Figure 5. Time series of evaluation statistics for control forecasts (thick dashed line) and forecasts with assimilation (solid line) for $PM_{2.5}$: top – bias, middle – pattern RMSE, bottom – correlation coefficient. Previous-day persistence plotted with thin dashed line.

with assimilation of 50% and 75% of randomly chosen observations for both ozone and $PM_{2.5}$. Only minimal deterioration of bias (less than 1 ppbv for ozone and 0.1 μg m^{-3} for $PM_{2.5}$), pattern root-mean-square error (pattern RMSE, Taylor (2001), about 1 ppbv for ozone and less than 0.1 μg m^{-3} for $PM_{2.5}$), and spatial correlation (about 0.02 for ozone and less than 0.01 for $PM_{2.5}$) between the analyses and the independent set of observations was noted. Description of evaluation statistics is given in the appendix.

The assumption on the Gaussian distribution of model and observation errors is implicit in deriving the cost function. Examination of distributions of background-forecast and observation errors with respect to the analyses in our simulations showed that a first-order autoregressive model would be a more suitable alternative.

Bias correction of forecasts was not applied even though presence of bias violates theoretical assumptions of the data assimilation methodology. However, evaluation showed that statistical verification scores of the analyses obtained from surface 'biased' and 'debiased' forecasts are nearly identical.

In Figures 4 and 5, time series of bias, pattern RMSE, and correlation for 24-hour forecasts of both species with and without data assimilation are shown in comparison to the persistence. (The bias of a domain-averaged persistence remains close to zero and was, therefore, not plotted.) To better match observation times and continuous character of the measurements, the statistics were calculated for forecasts averaged at two consecutive full hours.

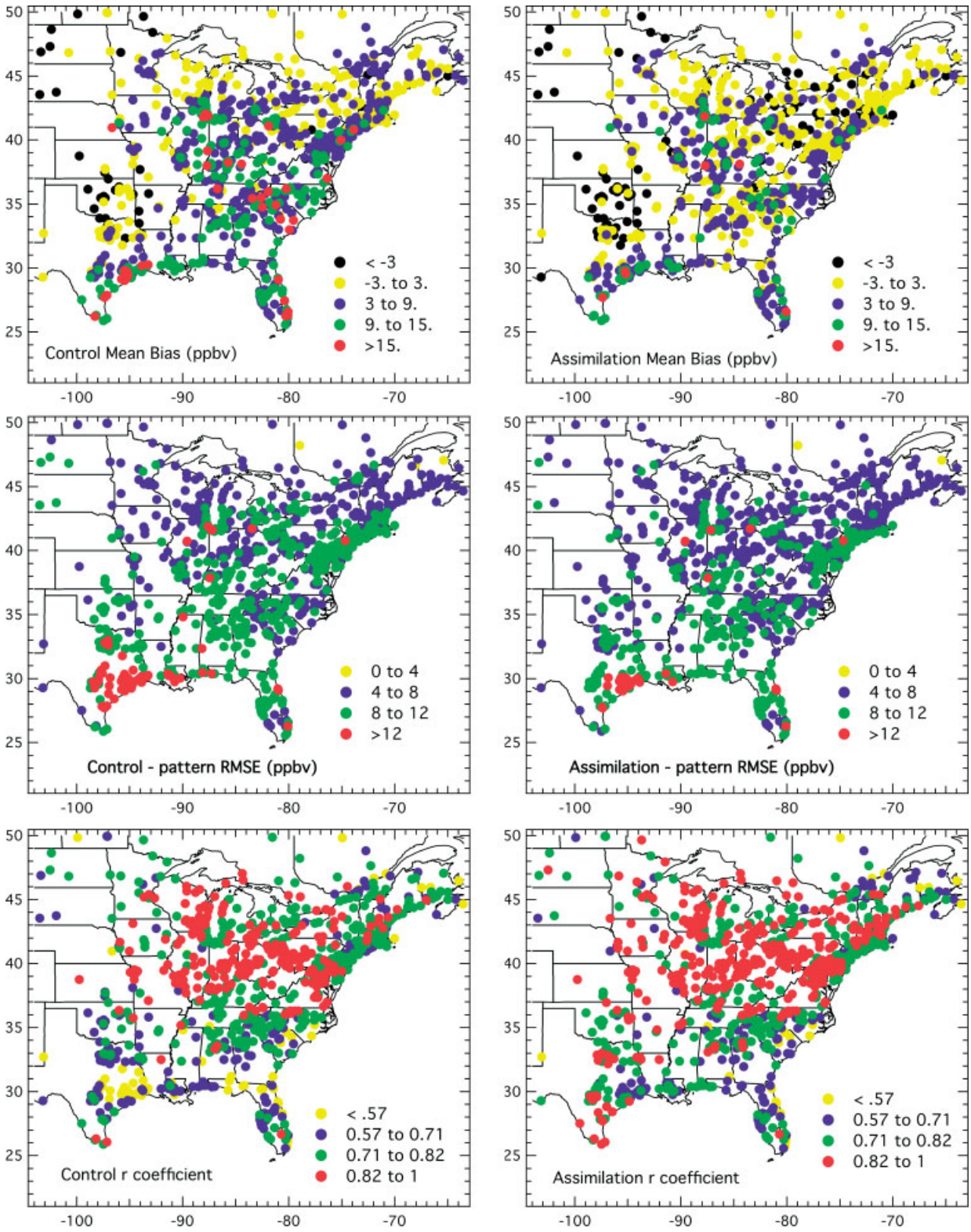

Figure 6. Evaluation statistics for forecasts of the next-day 8-hour average maximum ozone concentration in the control (left) and forecasts with assimilation (right): top – mean bias, middle – pattern RMSE, bottom – correlation coefficient.

For ozone forecasts with assimilation, the improvement in all metrics was significant but large increases in bias, pattern RMSE and a decrease in correlation are seen in the early hours of the forecasts. Generally, poor skill of models in predicting night-time and early morning boundary-layer meteorology and ozone is likely a main reason for this quick deterioration of the forecasts. Another factor for the deterioration might be lack of accounting for the effect that the assimilation should have on the concentration of ozone precursors. In terms of bias and pattern RMSE reduction, assimilations at 0000 UTC and 1200 UTC have a similar effect but the drop in correlation for the early morning assimilation is more pronounced. Assimilation in the early morning provides very modest improvement in forecast skill for the following night. Yet, despite the simplicity of the current assimilation approach, a positive effect of

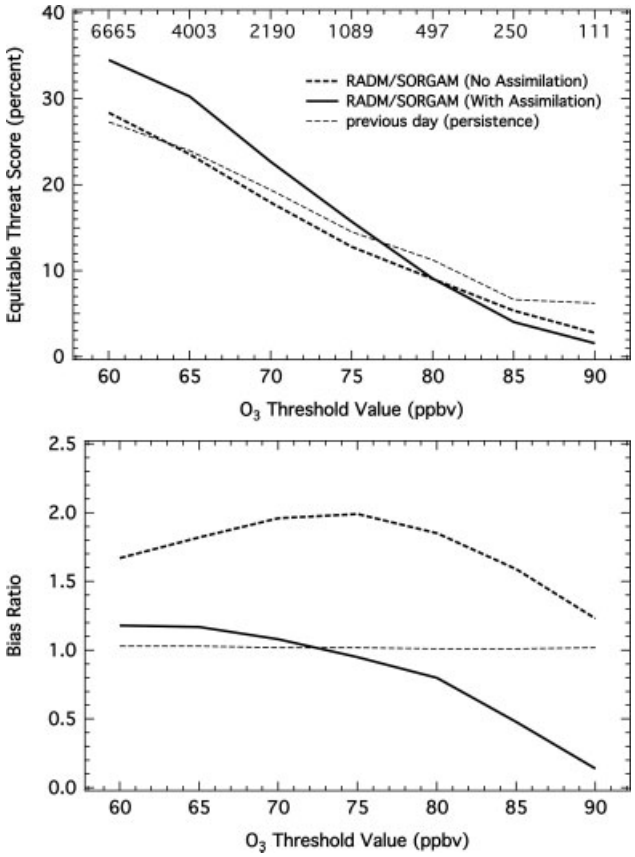

Figure 7. Equitable threat score and bias ratio for different thresholds of the maximum 8-hour average ozone concentration for control forecasts (thick dashed line) and forecasts with assimilation (solid line). Score for persistence plotted with thin dashed line. Number of observations of the maximum 8-hour average ozone concentration given on the upper axis.

assimilations on forecasts in all cases is quite apparent during the whole 24-hour evaluation period.

For $PM_{2.5}$ forecasts, the positive impact of the assimilation is also evident in the evaluation metrics presented in Figure 5. Unlike for ozone, the quality of forecasts does not deteriorate rapidly after the assimilation. Similar to ozone, the assimilation in the early morning only slightly improves forecast skill on the following night.

Similar to previous air-quality forecast evaluations with AIRNow data (Eder *et al.*, 2006; McKeen *et al.*, 2007), below we present statistics for the quantities used in regulatory applications. In May 2008 the regulatory limit for ozone compliance within the USA was changed from 85 ppbv to 75 ppbv for a maximum 8-hour average. Since December 2006 the regulatory limit for $PM_{2.5}$ compliance within the USA has been set at 35 μg/m^3 for a 24-hour average. To allow for sufficient forecast lead time, only forecasts issued at 0000 UTC are considered below.

Bias, pattern RMSE and the correlation coefficient (calculated as for the previous evaluations) for the next-day 8-hour average maximum ozone concentration in the control and forecasts with assimilation are compared in Figure 6. An increase in the number of yellow dots (small bias and pattern RMSE) and a decrease in the number of red points (large bias and pattern RMSE) from left to right in the two upper rows of the figure are apparent. In the third row of this figure, an increase in the number of red dots that corresponds to better correlation in the right-hand plot is also visible. In summary, out of 954 points in Figure 6, mean bias and pattern RMSE improved for over 80% of points, and correlation improved for over 75% of points.

The equitable threat score (ETS, a measure of forecast skill: Schaefer, 1990) and bias ratio (BR, a ratio of the predicted to the observed area) for different thresholds of the maximum 8-hour ozone concentration are plotted in Figure 7. The improvement in quality of forecasts due to assimilation is obvious given a known deficiency of ETS which favours forecasts with larger BR. We note that the maximum 8-hour average ozone concentration is usually centred at about 1500 LDT (1900 UTC on the east coast of the USA and 2000 UTC in the central region of the USA). Despite the fact that data assimilation was performed 20–21 hours earlier, the improvement for thresholds less than 80 ppbv is apparent.

In Figure 8, bias, pattern RMSE and the correlation coefficient for the 24-hour average $PM_{2.5}$ concentration are shown. As for ozone, there is a visible decrease of bias and pattern RMSE, which are reflected in the yellow and blue shifts from left to right in the two upper rows. Also, an increase in the number of red dots from the left to the right column in the third row points to a higher correlation in forecasts with assimilation. In summary, mean bias improved for over 60% of 355 points in Figure 8, pattern RMSE improved for over 80% of the points and correlation improved for over 90% of the points.

As for ozone, ETS and BR for different thresholds of the 24-hour average $PM_{2.5}$ concentration are shown in Figure 9. Here, very significant improvement of both verification scores is achieved when assimilation is applied. Poor performance of the chemical transport model with respect to persistence in predicting $PM_{2.5}$ concentrations can be noted. However, in this respect WRF-Chem is by no means inferior to other models as noted in model comparison studies by McKeen *et al.* (2009) and Djalalova *et al.* (2010).

We separately compared forecasts with and without the assimilation during two stagnation episodes in August 2006 when measurements at numerous sites exceeded mandated regulatory limits for ozone and $PM_{2.5}$ compliance. Qualitatively, ETS and BR for both types of forecasts during the episodes displayed features similar to those shown in Figures 7 and 9.

6. Discussion and conclusions

This experiment demonstrated that improvement of initial conditions of ozone and $PM_{2.5}$ concentrations via assimilation can lead to much improved forecasts of concentrations of these species in terms of standard statistical measures. Despite a quick drop of forecast skill in the early hours following assimilation (particularly for ozone) a positive impact of assimilation was observed in forecasts out to at least 24 hours.

That result for ozone was encouraging, especially considering that it is a very reactive species, dependent on the presence of precursors and sunlight. Measurements of concentrations of precursors were not available for this evaluation period so we could not determine if the assimilation of ozone had any positive impact on their predicted concentrations. Currently, modelled concentrations of the precursors are only affected implicitly by assimilation of ozone. In future we plan to influence concentrations of

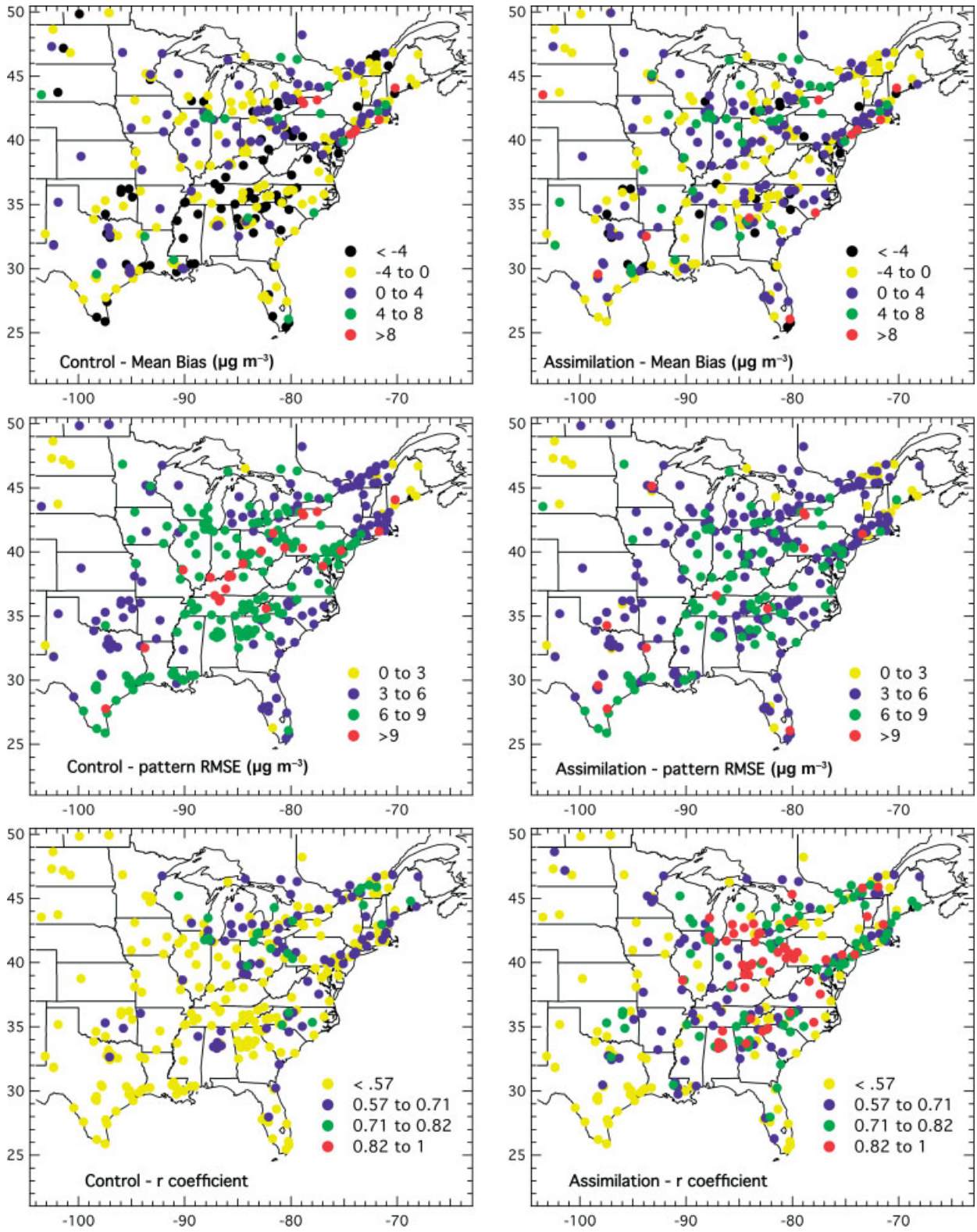

Figure 8. Evaluation statistics for forecasts of the 24-hour average $PM_{2.5}$ concentration in the control (left) and forecasts with assimilation (right): top – mean bias, middle – pattern RMSE, bottom – correlation coefficient.

the precursors and to minimize adjustment of their concentrations due to chemical imbalances introduced by the assimilation.

Improvement in the forecasting skill of $PM_{2.5}$ concentrations with assimilation was expected. First, $PM_{2.5}$ is generally less volatile than ozone (with the possible exception of decrease in concentration of hydrophilic aerosols due to wet scavenging) and the impact of assimilation should have longer-term effects. Second, $PM_{2.5}$ forecasts generally showed large errors in terms of basic statistics, and their correction via assimilation should be substantial. Experiments with a much simpler aerosol parametrization had similar improvements, and forecasts with assimilation were comparable in skill to those using the MADE/SORGAM module. Given generally poor performance of models in predicting aerosol concentrations, it can be expected that

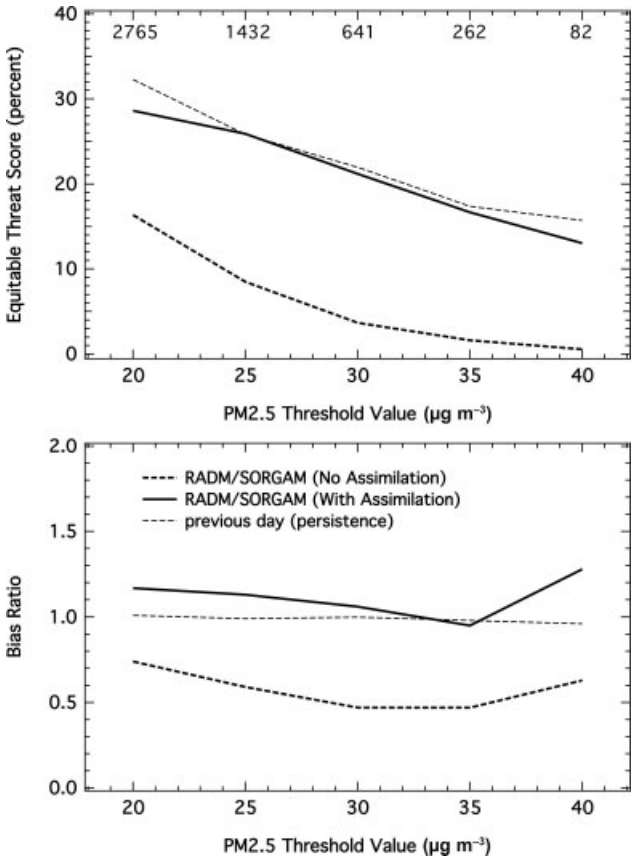

Figure 9. Equitable threat score and bias ratio for different thresholds of the 24-hour average $PM_{2.5}$ concentration for control forecasts (thick dashed line) and forecasts with assimilation (solid line). Score for persistence plotted with thin dashed line. Number of observations of the maximum 8-hour average ozone concentration given on the upper axis.

forecasting with simple aerosol parametrizations and data assimilation will have advantages over the more complex parametrizations not using assimilation.

High-ozone episodes commonly occur in atmospheric conditions that limit horizontal and vertical transport of species, leading to the accumulation of pollutants at the surface (Vukovich *et al.*, 1977). Such conditions often occur in the summer within stationary high-pressure systems characterized by subsidence, weak winds, few clouds and high insolation. Further investigation will determine if correlations between stream function and ozone, and temperature and ozone justify regression dependence that could be used in data assimilation. If sufficient dependence exists, further improvements in ozone predictions would be expected for regions where observations are sparse but where meteorological observations exist.

Extensions to the work presented here will include an investigation of forecast performance at different assimilation times with higher frequency of the assimilation cycle, and the use of GSI to assimilate measurements from aircraft, ozonesondes and satellites.

Horizontal and vertical length-scales and background error variances will receive further tuning since the current data are derived from a monthly series of 48-hour forecasts that are issued continuously without data assimilation. The tuning will be accomplished on a seasonal basis when the assimilation cycle is implemented in real time. With an increased sample size and greater statistical confidence, zonal dependence of background error scales will be introduced.

Our near-term goals also include simultaneous assimilation of meteorological and chemical data.

This study focussed on the importance of initial conditions in air-quality prediction. The assimilation in our experiments had a positive effect on the skill of short-range forecasts of ozone and $PM_{2.5}$ concentrations but errors that develop during the simulations have origins that cannot be ameliorated with 3D-Var. Major sources of forecast errors can be attributed to deficiencies in parametrizations of chemical processes at night and, likely, poor representation of the nocturnal boundary layer and mixing of species. As noted by others, applications of inverse methods with adjoints or ensemble filters to address uncertainties in emission source estimates should have a more long-lasting effect on the skill of forecasts. Finally, but maybe most importantly, significant improvement in air-quality forecasts will not occur until more comprehensive observations of atmospheric composition become available; particularly for smog prediction, concentration measurements of ozone precursors and aerosol constituents are critical.

Appendix

Verification statistics employed in the manuscript are described below.

Taylor (2001) defined pattern RMSE as

$$E' \equiv \left[\frac{1}{N} \sum_{n=1}^{N} \left\{ \left(f_n - \overline{f} \right) - \left(r_n - \overline{r} \right) \right\}^2 \right]^{\frac{1}{2}}, \qquad (2)$$

where f, r and overbar denote a test field, a reference field and averaging, respectively. In our case, f represents model concentrations interpolated temporally (model output available at the (exact) beginning of an hour) and spatially to observation locations, and r represents observations (available 30 minutes after the hour).

Equitable threat score (ETS: Schaefer, 1990) is defined as

$$\mathrm{ETS} \equiv \frac{a - ch}{a + b + c - ch}, \qquad (3)$$

$$ch = \frac{(a + b)(a + c)}{a + b + c + d}. \qquad (4)$$

In the above formulae, a (hits), b (false alarms), c (misses), and d (correct rejections) stand for probabilities of an event exceeding a certain threshold.

Bias ratio (BR) is defined as

$$\mathrm{BR} \equiv \frac{a + b}{a + c}. \qquad (5)$$

To calculate correlation between model and observations, a series of observations valid at a given forecast time is matched with a series of interpolated model values. Therefore, this statistic represents spatial rather than temporal correlation between model and observations. Also, for these matching series model biases and RMSEs were calculated.

Acknowledgements

This research is funded by Early Start Funding from the NOAA Office of Atmospheric Research, the NOAA Air Quality Program, and the NOAA/NWS Office of Science and Technology. Credit for programme support and management is given to Paula Davidson (NOAA/NWS/OST), Steve Fine (NOAA/ARL), and Jim Meagher (NOAA/ESRL/CSD). Ozone and $PM_{2.5}$ observations were obtained through the US EPA AIRNow Data Exchange programme. Comments of anonymous reviewers and Tom Schlatter from NOAA/ESRL are highly appreciated. Hendrik Elbern and members of his research group at the University of Cologne are thanked for discussions on some aspects of this work.

References

Ackermann IJ, Hass H, Memmesheimer M, Ebel A, Binkowski FS, Shankar U. 1998. Modal aerosol dynamics model for Europe: Development and first applications. *Atmos. Environ.* **32**: 2981–2999.

Alapaty K, Olerud Jr DT, Schere KL, Hanna AF. 1995. Sensitivity of Regional Oxidant Model predictions to prognostic and diagnostic meteorological fields. *J. Appl. Meteorol.* **34**: 1787–1801.

Barna MG, Knipping EM. 2006. Insights from the BRAVO study on nesting global models to specify boundary conditions in regional air quality modeling simulations. *Atmos. Environ.* **40**(S2): 574–582.

Benedetti A, Fisher M. 2007. Background error statistics for aerosols. *Q. J. R. Meteorol. Soc.* **133**: 391–405.

Benedetti A, Morcrette J-J, Boucher O, Dethof A, Engelen RJ, Fisher M, Flentje H, Huneeus N, Jones L, Kaiser JW, Kinne S, Mangold A, Razinger M, Simmons AJ, Suttie M. 2009. Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System. 2: Data assimilation. *J. Geophys. Res.* **114**: D13205, DOI:10.1029/2008JD011115.

Biswas J, Rao ST. 2001. Uncertainties in episodic ozone modeling stemming from uncertainties in the meteorological fields. *J. Appl. Meteorol.* **40**: 117–136.

Carmichael GR, Peters LK, Saylor RD. 1991. The STEM-II regional scale acid deposition and photochemical oxidant model. I: An overview of model development and applications. *Atmos. Environ.* **25A**: 2077–2090.

Carmichael GR, Daescu DN, Sandu A, Chai T. 2003. '*Computational aspects of chemical data assimilation into atmospheric models*.' Science Computational ICCS 2003, Melbourne. Pp 269–278 in *Springer Lecture Notes in Computer Science IV*. Springer: Berlin.

Chai T, Carmichael GR, Tang Y, Sandu A, Hardesty M, Pilewskie P, Whitlow S, Browell EV, Avery MA, Nédélec P, Merrill JT, Thompson AM, Williams E. 2007. Four-dimensional data assimilation experiments with International Consortium for Atmospheric Research on Transport and Transformation ozone measurements. *J. Geophys. Res.* **112**: D12S15, DOI:10.1029/2006JD007763.

Chen F, Dudhia J. 2001. Coupling an advanced land surface–hydrology model with the Penn State–NCAR MM5 modeling system. Part I: Model implementation and sensitivity. *Mon. Weather Rev.* **129**: 569–585.

Constantinescu EM, Sandu A, Chai T, Carmichael GR. 2007a. Assessment of ensemble-based chemical data assimilation in an idealized setting. *Atmos. Environ.* **41**: 18–36.

Constantinescu EM, Chai T, Sandu A, Carmichael GR. 2007b. Autoregressive models of background errors for chemical data assimilation. *J. Geophys. Res.* **112**: D12309, DOI: 10.1029/2006JD008103.

Constantinescu EM, Sandu A, Chai T, Carmichael GR. 2007c. Ensemble-based chemical data assimilation. I: General approach. *Q. J. R. Meteorol. Soc.* **133**: 1229–1243.

Constantinescu EM, Sandu A, Chai T, Carmichael GR. 2007d. Ensemble-based chemical data assimilation. II: Covariance localization. *Q. J. R. Meteorol. Soc.* **133**: 1245–1256.

Daescu DN, Carmichael GR. 2003. An adjoint sensitivity method for the adaptive location of the observations in air quality modeling. *J. Atmos. Sci.* **60**: 434–450.

Daley R. 1991. *Atmospheric data analysis*. Cambridge University Press: Cambridge.

Djalalova I, Wilczak J, McKeen S, Grell G, Peckham S, Pagowski M, DelleMonache L, McQueen J, Tang Y, Lee P, McHenry J, Gong W, Bouchet V, Mathur R. 2010. Ensemble and bias-correction techniques for air quality model forecasts of surface O_3 and $PM_{2.5}$ during the TEXAQS-II experiment of 2006. *Atmos. Environ.* **44**: 455–467.

Dudhia J. 1989. Numerical study of convection observed during the Winter Monsoon Experiment using a mesoscale two-dimensional model. *J. Atmos. Sci.* **46**: 3077–3107.

Eder B, Kang D, Mathur R, Yu S, Schere K. 2006. An operational evaluation of the Eta–CMAQ air quality forecast model. *Atmos. Environ.* **40**: 4894–4905.

Elbern H, Schmidt H. 1999. A four-dimensional variational chemistry data assimilation scheme for Eulerian chemistry transport modelling. *J. Geophys. Res.* **104**: 18583–18598.

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

Elbern H, Schmidt H, Talagrand O, Ebel A. 2000. 4D-variational data assimilation with an adjoint air quality model for emission analysis. *Environ. Modelling Software* **15**: 539–548.

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

Elbern H, Strunk A, Schmidt H, Talagrand O. 2007. Emission rate and chemical state estimation by 4-dimensional variational inversion. *Atmos. Chem. Phys.* **7**: 3749–3769.

Grell GA, Dévényi D. 2002. A generalized approach to parameterizing convection combining ensemble and data assimilation techniques. *Geophys. Res. Lett.* **29**: 1693, DOI:10.1029/2002GL015311.

Grell GA, Peckham SE, Schmitz R, McKeen SA, Frost GJ, Skamarock WC, Eder B. 2005. Fully-coupled 'online' chemistry within the WRF model. *Atmos. Environ.* **39**: 6957–6975.

Guenther AB, Zimmerman PR, Harley PC, Monson RK, Fall R. 1993. Isoprene and monoterpene emission rate variability: Model evaluations and sensitivity analyses. *J. Geophys. Res.* **98**: 12609–12617.

Gustafsson N, Berre L, Hörnquist S, Huang X-Y, Lindskog M, Navascués B, Mogensen KS, Thorsteinsson S. 2001. Three-dimensional variational data assimilation for a limited area model. Part I: General formulation and the background error constraint. *Tellus* **53A**: 425–446.

Hanna SR, Lu Z, Frey HC, Wheeler N, Vukovich J, Arunachalam S, Fernau M, Hansen DA. 2001. Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain. *Atmos. Environ.* **35**: 891–903.

Hitzenberger R, Berner A, Galambos Z, Maenhaut W, Cafmeyer J, Schwarz J, Müller K, Spindler G, Wieprecht W, Acker K, Hillamo R, Mäkelä T. 2004. Intercomparison of methods to measure the mass concentration of the atmospheric aerosol during INTERCOMP2000: Influence of instrumentation and size cuts. *Atmos. Environ.* **38**: 6467–6476.

Hong S-Y, Pan H-L. 1996. Nonlocal boundary layer vertical diffusion in a medium-range forecast model. *Mon. Weather Rev.* **124**: 2322–2339.

Hong S-Y, Dudhia J, Chen S-H. 2004. A revised approach to ice microphysical processes for the bulk parameterization of clouds and precipitation. *Mon. Weather Rev.* **132**: 103–120.

Jonson JE, Isaksen ISA. 1993. Tropospheric ozone chemistry: The impact of cloud chemistry. *J. Atmos. Chem.* **16**: 99–122.

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

Kim S-W, Heckel A, McKeen SA, Frost GJ, Hsie E-Y, Trainer MK, Richter A, Burrows JP, Peckham SE, Grell GA. 2006. Satellite-observed U.S. power plant NO_x emission reductions and their impact on air quality. *Geophys. Res. Lett.* **33**: L22812, DOI:10.1029/2006GL027749.

Lee P, Kang D, McQueen J, Tsidulko M, Hart M, DiMego G, Seaman N, Davidson P. 2008. Impact of domain size on modeled ozone forecast for the northeastern United States. *J. Appl. Meteorol. Clim.* **47**: 443–461.

McKeen SA, Wotawa G, Parrish DD, Holloway JS, Buhr MP, Hübler G, Fehsenfeld FC, Meagher JF. 2002. Ozone production from Canadian wildfires during June and July of 1995. *J. Geophys. Res.* **107**: D14, DOI:10.1029/2001JD000697.

McKeen SA, Chung SH, Wilczak J, Grell G, Djalalova I, Peckham S, Gong W, Bouchet V, Moffet R, Tang Y, Carmichael GR, Mathur R, Yu S. 2007. Evaluation of several $PM_{2.5}$ forecast models using data collected during the ICARTT/NEAQS 2004 field study. *J. Geophys. Res.* **112**: D10S20, DOI:10.1029/2006JD007608.

McKeen SA, Grell G, Peckham S, Wilczak J, Djalalova I, Hsie E-Y, Frost G, Peischl J, Schwarz J, Spackman R, Holloway J, de Gouw J, Warneke C, Gong W, Bouchet V, Gaudreault S, Racine J, McHenry J, McQueen J, Lee P, Tang Y, Carmichael GR, Mathur R. 2009. An evaluation of real-time air quality forecasts and their urban emissions over eastern Texas during the summer of 2006 Second Texas Air Quality Study field study. *J. Geophys. Res.* **114**: D00F11, DOI:10.1029/2008JD011697.

Mathur R, Schere KL, Nathan A. 1994. Dependencies and sensitivity of tropospheric oxidants to precursor concentrations over the northeast United States: A model study. *J. Geophys. Res.* **99**: 10535–10552.

Mlawer EJ, Taubman SJ, Brown PD, Iacono MJ, Clough SA. 1997. Radiative transfer for inhomogeneous atmospheres: RRTM, a validated correlated-k model for the longwave. *J. Geophys. Res.* **102**: 16663–16682.

Morcrette J-J, Boucher O, Jones L, Salmond D, Bechtold P, Beljaars A, Benedetti A, Bonet A, Kaiser JW, Razinger M, Schulz M, Serrar S, Simmons AJ, Sofiev M, Suttie M, Tompkins AM, Untch A. 2009. Aerosol analysis and forecast in the European Centre for Medium-Range Weather Forecasts Integrated Forecast System: Forward modeling. *J. Geophys. Res.* **114**: D06206, DOI:10.1029/2008JD011235.

Parrish DF, Derber JC. 1992. The National Meteorological Center's spectral statistical interpolation analysis system. *Mon. Weather Rev.* **120**: 1747–1763.

Purser RJ, Wu W-S, Parrish DF, Roberts NM. 2003a. Numerical aspects of the application of recursive filters to variational statistical analysis. Part I: Spatially homogeneous and isotropic Gaussian covariances. *Mon. Weather Rev.* **131**: 1524–1535.

Purser RJ, Wu W-S, Parrish DF, Roberts NM. 2003b. Numerical aspects of the application of recursive filters to variational statistical analysis. Part II: Spatially inhomogeneous and anisotropic general covariances. *Mon. Weather Rev.* **131**: 1536–1548.

Sandu A, Daescu DN, Carmichael GR, Chai T. 2005. Adjoint sensitivity analysis of regional air quality models. *J. Comput. Phys.* **204**: 222–252.

Schaefer JT. 1990. The critical success index as an indicator of warning skill. *Weather and Forecasting* **5**: 570–575.

Schell B, Ackermann IJ, Hass H, Binkowski FS, Ebel A. 2001. Modeling the formation of secondary organic aerosol within a comprehensive air quality model system. *J. Geophys. Res.* **106**: 28275–28293.

Sillman S, Samson PJ. 1995. Impact of temperature on oxidant photochemistry in urban, polluted rural and remote environments. *J. Geophys. Res.* **100**: 11497–11508.

Skamarock W, Klemp JB, Dudhia J, Gill DO, Barker DM, Wang W, Powers JG. 2005. '*A description of the Advanced Research WRF Version 2*.' Mesoscale and Microscale Meteorology Division, NCAR/TN–468+STR, National Center for Atmospheric Research, Boulder, Colorado, USA.

Stockwell WR, Middleton P, Chang JS, Tang X. 1990. The second-generation regional acid deposition model chemical mechanism for regional air quality modeling. *J. Geophys. Res.* **95**: 16343–16367.

Stockwell WR, Kirchner F, Kuhn M, Seefeld S. 1997. A new mechanism for regional atmospheric chemistry modeling. *J. Geophys. Res.* **102**: 25847–25879.

Tang Y, Carmichael GR, Thongboonchoo N, Chai T, Horowitz LW, Pierce RB, Al-Saadi JA, Pfister G, Vukovich JM, Avery MA, Sachse GW, Ryerson TB, Holloway JS, Atlas EL, Flocke FM, Weber RJ, Huey LG, Dibb JE, Streets DG, Brune WH. 2007. Influence of lateral and top boundary conditions on regional air quality prediction: A multiscale study coupling regional and global chemical transport models. *J. Geophys. Res.* **112**: D10S18, DOI:10.1029/2006JD007515.

Taylor KE. 2001. Summarizing multiple aspects of model performance in a single diagram. *J. Geophys. Res.* **106**: 7183–7192.

Vukovich FM, Bach Jr WD, Crissman BW, King WJ. 1977. On the relationship between high ozone in rural surface layer and high pressure systems. *Atmos. Environ.* **11**: 967–983.

Wu W-S, Purser RJ, Parrish DF. 2002. Three-dimensional variational analysis with spatially inhomogeneous covariances. *Mon. Weather Rev.* **130**: 2905–2916.

Zhang J, Rao ST. 1999. The role of vertical mixing in the temporal evolution of the ground-level ozone concentrations. *J. Appl. Meteorol.* **38**: 1674–1691.