Forecast error covariance structure in coupled atmosphere–chemistry data assimilation

S. K. Park¹,²,³,⁴, S. Lim²,³, and M. Zupanski⁵

¹Department of Environmental Science and Engineering, Ewha Womans University, Seoul, Republic of Korea
²Department of Atmospheric Science and Engineering, Ewha Womans University, Seoul, Republic of Korea
³Center for Climate/Environment Change Prediction Research, Ewha Womans University, Seoul, Republic of Korea
⁴Severe Storm Research Center, Ewha Womans University, Seoul, Republic of Korea
⁵Cooperative Institute for Research in the Atmosphere, Colorado State University, Fort Collins, Colorado, USA

Received: 26 October 2014 – Accepted: 20 November 2014 – Published: 10 December 2014

Correspondence to: S. K. Park (spark@ewha.ac.kr)

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

In this study, we examined the structure of an ensemble-based coupled atmosphere–chemistry forecast error covariance. The Weather Research and Forecasting (WRF) model coupled with Chemistry (WRF-Chem), a coupled atmosphere–chemistry model, was used to create an ensemble error covariance. The control variable includes both the dynamical and chemistry model variables. A synthetic single observation experiment was designed in order to evaluate the cross-variable components of a coupled error covariance. The results indicate that the coupled error covariance has important cross-variable components that allow a physically meaningful adjustment of all control variables. The additional benefit of the coupled error covariance is that a cross-component impact is allowed, e.g., atmospheric observations can exert impact on chemistry analysis, and vice versa. Given the realistic structure of ensemble forecast error covariance produced by the WRF-Chem, we anticipate the ensemble-based coupled atmosphere–chemistry data assimilation will respond similarly to assimilation of real observations.

1 Introduction

The regional air quality is affected by synoptic weather situations or air masses with special chemical properties (Grell et al., 2000). In prediction of air quality, the coupled physical and chemical processes are essential, which include transport, deposition, emission, chemical transformation, aerosol interactions, photolysis, and radiation (Grell et al., 2005). Optimized initial conditions for a numerical model, including such coupled processes, can be obtained by data assimilation (DA; e.g., Houtekamer and Mitchell, 1998; Eibert and Schmidt, 1999; Wang et al., 2001; Evensen, 2003; Park and Zupanski, 2003; Navon, 2009; Zupanski, 2009). Therefore, DA for an air quality prediction system could be approached as a coupled atmosphere–chemistry DA, with interaction between atmospheric and chemistry components. In typical data assimila-
tion methodologies, such as variational and ensemble, the interaction between different variables is achieved by forecast error covariance, in particular its cross-variable components. Therefore, it is of fundamental interest for the development of atmosphere–chemistry DA to investigate the coupled forecast error covariance. Here, we investigate the structure of the atmosphere–chemistry forecast error covariance using ensemble forecasting, which corresponds to the prediction step of an ensemble data assimilation algorithm (e.g., Zupanski, 2005, 2009).

2 Methodology and synoptic case

In this research, we use the Weather Research and Forecasting (WRF) model coupled with Chemistry (WRF-Chem) as a prediction model (Grell et al., 2005). It simulates the emission, transport, mixing and chemical transformation of trace gases and aerosols simultaneously with meteorology and investigates the regional scale air quality using the Carbon Bond Mechanism version Z (CBMZ) chemistry option.

We chose a synoptic case on 3 September 2005 related to Typhoon Nabi (2005), characterized by an increased impact on the Korean Peninsula. The experiment begins at 00:00 UTC and ends at 06:00 UTC on 3 September 2005. The WRF-Chem is set up with a horizontal resolution of 30 km and 28 vertical levels. Model domain is centered over the Korean Peninsula, covering an area of approximately 3900 km × 4400 km with 132 × 147 horizontal grid points.

The ensemble forecast includes 32 ensemble members with a 6 h assimilation window. The lateral boundary conditions are provided by the National Center for Environmental Prediction (NCEP) Global Forecasting System (GFS). The control variables defined in DA (i.e., variables adjusted during DA) are the WRF-Chem prognostic variables that include dynamical variables such as winds, perturbation potential temperature, perturbation geopotential, water vapor mixing ratio and perturbation dry air mass in column, and the chemical variables such as ozone (O₃), nitrates (NO, NO₂, NO₃) and sulfur dioxide (SO₂) as well.
3 Experimental design

A common approach to investigating forecast error covariance in data assimilation is to conduct a single observation experiment (Thepaut et al., 1996; Whitaker et al., 2009; Buehner et al., 2010), in which only one observation is assimilated using the full DA system. The analysis increments (i.e., analysis minus guess) from such an experiment show how the observation information is distributed spatially and among different analysis variables (e.g., Buehner, 2005). However, in order to investigate the structure of a coupled forecast error covariance before real observations are available and even before the full DA algorithm is developed, one can consider the assimilation of a single synthetic observation located at a chosen model grid point. In particular, we define the synthetic observation as

\[ y_{\text{synth}} = x^f + \sigma_o \]  

(1)

where \( x^f \) is the forecast and \( \sigma_o \) is the observation error standard deviation (SD). Following Thepaut et al. (1996, Eq. 3), with some modifications and using Eq. (1), the analysis increment in a single synthetic observation experiment is

\[ x^a - x^f = P_f \left( \frac{\sigma_o}{\sigma_f + \sigma_o} \right)_{i,j,k} \]  

(2)

where \( x^a \) is the analysis, \( \sigma_f \) is the forecast error SD, and the subscript \( i,j,k \) defines the grid location of the pseudo-observation point. Equation (2) indicates that analysis increment represents the \( i,j,k \)th column of the forecast error covariance scaled by SDs of observation error and forecast error. In our experiments the forecast error covariance is ensemble-based, as defined in Zupanski (2005) as:

\[ P_f = P_{f}^{1/2} \left( P_{f}^{1/2} \right)^T, \quad P_{f}^{1/2} = \left( p_1^f \cdots p_N^f \right), \quad p_n^f = m(x_0^n) - m(x_0) \]  

(3)
where the superscript T denotes the transpose, the index \( n \) refers to ensemble member, \( N \) is the total number of ensemble forecasts, \( m \) represents the nonlinear WRF-Chem model, and the subscript 0 denotes the initial time of the forecast with corresponding initial conditions \( x_0 \) and ensemble initial conditions \( x_0^n \). In this experiment, the control initial conditions are obtained by interpolation from the NCEP GFS model, while the initial ensemble perturbations are created using the lagged forecast outputs.

Since we are interested in the coupled atmosphere–chemistry forecast error covariance, we design two experiments with: (i) synthetic temperature observation at 250 hPa located at a grid point near \((132^\circ E, 23^\circ N)\), on the northwest side of the typhoon, and (ii) synthetic ozone observation at 250 hPa located at a grid point near the eye of the typhoon \((134^\circ E, 21^\circ N)\).

4 Results

We show the impact of single synthetic temperature \((T)\) and ozone \(\left(O_3\right)\) observations in terms of the analysis increments \(x^a - x^f\) impacting all control variables. As mentioned earlier, our main interest is to examine the cross-variable covariance structure between atmospheric and chemistry variables, since the cross-variable analysis impact is possible only because of the multivariate structure of the coupled ensemble forecast error covariance.

In Fig. 1 we show the impact of synthetic \(T\) observation at 250 hPa on the analysis increments of \(T, O_3,\) nitrogen-dioxide \((NO_2)\), and sulfur dioxide \((SO_2)\). The analysis increment of \(T\) at 250 hPa (e.g., at the same level of synthetic \(T\) observation) shows a typical response with nearly circular isolines with the maximum of 0.4 K at the observation location (Fig. 1a). The analysis increments of \(O_3, NO_2,\) and \(SO_2\) are also shown in vertical cross-sections. One can see that \(O_3\) (Fig. 1b) and \(NO_2\) (Fig. 1c) analyses have the largest change at the level of single \(T\) observation, while the \(SO_2\) analysis (Fig. 1d) is mostly impacted near 700 hPa (approximately \(\sigma\)-level 13). This is likely a consequence of the vertical structure of \(O_3\) and \(NO_2\) with the largest values in
the upper troposphere and the stratosphere, while $SO_2$ has typically the largest values in the lower troposphere (e.g., Meena et al., 2006). The strongest impact of $T$ observation is on $O_3$, with the magnitude up to 0.001 ppmv, while the magnitude is somewhat smaller for $NO_2$ and $SO_2$. One can also infer that an increase of $T$ will imply a decrease of $O_3$, $NO_2$, and $SO_2$. Probably the most important implication of these results is that observations of an atmospheric variable (e.g., temperature) can change the analysis of chemical variables in a physically meaningful way. This means that even with no chemistry observations in the local area, the analysis of chemical variables can still be adjusted in agreement with standard dynamical variables of the model. On the other hand, if there are chemistry observations in the area, the chemistry analysis change introduced by atmospheric observations will act as an additional dynamical constraint to the final analysis.

In Fig. 2 the impact of $O_3$ single observation at 250 hPa on itself and the other variables is shown. As before, we focus on the vertical cross-section of the analysis response. The impact of $O_3$ observation on its own analysis shows the anticipated response with the largest magnitude at observation location, approximately 0.02 ppmv (Fig. 2a). Although smaller in magnitude, the analysis increments of $NO_2$ (Fig. 2b) and $SO_2$ (Fig. 2c) show the vertical structure with maxima in the upper and lower troposphere, respectively. It is also notable that an increase of $O_3$ brings about an increase of $NO_2$ and $SO_2$, confirming the direct relationship between these variables as noticed in Fig. 1. The $T$ analysis increment indicates that there is a cooling at the level of $O_3$ observation, while there is a warming above and below (Fig. 2d).

The results shown in Fig. 2 indirectly confirm that the improved stratospheric ozone distribution by DA can make a better representation of stratospheric, temperature and other constituents (e.g., Lahoz et al., 2007).
5 Conclusions

The structure of an ensemble-based coupled atmosphere–chemistry forecast error covariance was examined in the context of the WRF-Chem model. A synthetic single observation experiment was designed in order to evaluate the cross-variable components of the coupled error covariance. Our results indicate that the coupled error covariance has important cross-variable components that allow a physically meaningful adjustment of all control variables, and a much wider impact of observations (e.g., atmospheric observation on chemistry analysis, and vice versa). The analysis increments created in response to synthetic temperature and ozone observations illustrate the complexity of atmosphere–chemistry cross-correlations and the forecast error covariance structure. Given the realistic structure of ensemble forecast error covariance produced by the WRF-Chem, we anticipate the ensemble-based coupled atmosphere–chemistry data assimilation will respond similarly to assimilation of real observations. Therefore, our next step is to apply the WRF-Chem with an ensemble-based data assimilation algorithm (e.g., the maximum likelihood ensemble filter (MLEF); Zupanski, 2005) to assimilation of real chemical and atmospheric observations.

Acknowledgements. This work is supported by the Korea Environmental Industry & Technology Institute through the Eco Innovation Program (ARQ201204015), and partly by the National Research Foundation of Korea grant (No. 2009-0083527) funded by the Korean government (MSIP). The third author acknowledges a partial support from the National Science Foundation Collaboration in Mathematical Geosciences Grant 0930265 and from the NASA Modeling, Analysis and Prediction (MAP) Program Grant NNX13AO10G.

References

Buehner, M., Houtekamer, P. L., Charette, C., Mitchell, H. L., and He, B.: Intercomparison of variational data assimilation and the ensemble kalman filter for global deterministic NWP.
Figure 1. The analysis increments \((x^a - x^f)\) in response to a single \(T\) observation at 250 hPa (near \(\sigma\)-level 24): (a) horizontal response of \(T\) at 250 hPa, and vertical responses of (b) \(O_3\), (c) \(NO_2\) and (d) \(SO_2\). In (b–d), the vertical axis represents the vertical \(\sigma\)-levels. Units are ppmv for chemical variables and K for temperature.
Figure 2. Same as in Fig. 1 but for vertical cross-section of the analysis increments ($x^a - x^f$) in response to a single O$_3$ observation at 250 hPa for (a) O$_3$, (b) NO$_2$, (c) SO$_2$ and (d) T.