Manuscript prepared for Atmos. Chem. Phys.
with version 2014/07/29 7.12 Copernicus papers of the LaTeX class copernicus.cls.
Date: 29 August 2015

Ensemble data assimilation of total column ozone using a coupled meteorology-chemistry model and its impact on the structure of Typhoon Nabi (2005)

S. Lim[1,3,4], S. K. Park[1,2,3,4], and M. Zupanski[5]

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

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

Abstract. Ozone (O_3) plays an important role in chemical reactions and is usually incorporated in chemical data assimilation (DA). In tropical cyclones (TCs), O_3 usually shows a lower concentration inside the eyewall and an elevated concentration around the eye, impacting meteorological as well as chemical variables. To identify the impact of O_3 observations on TC structure, including meteorological and chemical information, we developed a coupled meteorology-chemistry DA system by employing the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem) and an ensemble-based DA algorithm – the maximum likelihood ensemble filter (MLEF). For a TC case that occurred over East Asia, Typhoon Nabi (2005), our results indicate that the ensemble forecast is reasonable, accompanied with larger background state uncertainty over the TC, and also over eastern China. Similarly, the assimilation of O_3 observations impacts meteorological and chemical variables near the TC and over eastern China. The strongest impact on air quality in the lower troposphere was over China, likely due to the pollution advection. In the vicinity of the TC, however, the strongest impact on chemical variables adjustment was at higher levels. The impact on meteorological variables was similar in both over China and near the TC. The analysis results are verified using several measures that include the cost function, root-mean-squared (RMS) error with respect to observations, and degrees of freedom for signal (DFS). All measures indicate a positive impact of DA on the analysis – the cost function and RMS error have decreased by 16.9 and 8.87 %, respectively. In particular, the DFS indicates a strong positive impact of observations in the TC area, with a weaker maximum over northeast China.

1 Introduction

The air quality forecast is related to emissions, transport, transformation and removal processes, and to the prevailing meteorological conditions. Therefore, the coupled meteorology–chemistry model is essential for the air quality and weather forecasting (e.g., Carmichael et al., 2008). The coupled system forecast is improved through coupled meteorology–chemistry data assimilation (DA), which estimates the best initial conditions by combining the information from the model and observations in a mathematically consistent manner (e.g., Houtekamer and Mitchell, 1998; Elbern and Schmidt, 1999; Wang et al., 2001; Evensen, 2003; Park and Zupanski, 2003; Navon, 2009; Zupanski, 2009; Park et al., 2015).

Ozone (O_3) has a relatively long photochemical lifetime and high concentrations at high latitude and in the stratosphere, except during ozone hole conditions. It is a passive tracer at synoptic scale or smaller; thus variations of total column O_3 in space and time are a result of the atmospheric flow, and is highly correlated to many meteorological variables in the upper troposphere (Wu and Zou, 2008). Assimilation of O_3 has several motivations such as (Lahoz et al., 2007): 1) taking better account of stratospheric O_3 when assimilating satellite radiance data; 2) leading to better radiative forcing when used by the model radiation scheme; 3) providing useful dynamical information via the motion of O_3 in the atmosphere; and 4) improving the accuracy of UV index forecasting. Moreover, the improved stratospheric O_3 distribution by DA can affect meteorological variables such as stratospheric winds and temperature as well as other chemical variables (e.g., Lahoz et al., 2007; Park et al., 2015).

O_3 is also relevant to the structure of tropical cyclones (TCs), showing a lower concentration just inside the eyewall and elevated concentration around the eye (e.g., Carsey and Willoughby, 2005; Zou and Wu, 2005; Wu and Zou, 2008), which is caused by the updraft in the eyewall and subsidence in the eye (Zou and Wu, 2005). Using these relations, the daily total column O_3 from Total Ozone Mapping Spectrometer (TOMS) showed that mutual adjustment occurred between the TC and its upper tropospheric environment on a synoptical timescale (Rodgers et al., 1990; Stout and Rodgers, 1992). The linear relationship between total column O_3 from TOMS and mean vertically-integrated potential vorticity (MPV) was used to improve hurricane or winter storm prediction (e.g., Jang et al., 2003; Zou and Wu, 2005; Wu and Zou, 2008). However, these studies employed a meteorological model, not the coupled meteorology–chemistry model. They used the standard dynamical variables as control variables and empirical regressions to develop a cross-correlation between O_3 and dynamical model variables.

In this study, we directly assimilate the total column O_3 from the Ozone Monitoring Instrument (OMI) to identify the impact of O_3 observations on TC structure including meteorological and chemical information in a coupled meteorology–chemistry model (e.g., WRF-Chem) with an ensemble-based DA system (e.g., Maximum Likelihood Ensemble Filter; MLEF). We define an augmented control variable that contains both meteorological and chemical variables. Here meteorological variables consist of dynamical variables (e.g., wind components) and physical variables (e.g., water

vapor, cloud water, etc.). Therefore, the cross-correlations between meteorological and chemical variables are obtained directly from ensemble forecasts (e.g., Park et al., 2015). Section 2 describes the methodology, and Sect. 3 presents results. Conclusions are provided in Sect. 4.

2 Methodology

2.1 Model

In this research, we use the Weather Research and Forecasting (WRF) model coupled with Chemistry (WRF-Chem) version 3.4.1 as a prediction model on a regional scale. It simulates the emission, transport, mixing and chemical transformation of trace gases and aerosols simultaneously with meteorology (Grell et al., 2005). The WRF-Chem uses configuration options for various meteorological processes such as the WRF Single-Moment 6-class (WSM6) scheme for the microphysics, the Community Atmospheric Model (CAM) scheme for the radiation physics, the Monin–Obukhov scheme for the surface layer, the Noah land surface model for the land surface, the Yonsei University (YSU) scheme for the planetary boundary layer, and the Kain–Fritsch scheme for the cumulus parameterization. These are the recommended physics options for the regional climate case at 10–30 km grid size. As an advection option, the monotonic transport is applied to turbulent kinetic energy and scalars such as mixing ratios of water vapor, cloud water, rain, snow and ice and chemical species~~, which~~. The monotonic transport is commonly used for real-time and research applications (e.g., Chapman et al., 2009; Yang et al., 2011). Regarding the chemical mechanism, the Carbon Bond Mechanism version Z (CBM-Z) without Dimethylsulfide scheme is used for gas-phase chemistry. The CBM-Z includes the prediction of $\mathrm{O_3}$ and several other chemical constituents (Fast et al., 2006).

In terms of the DA system, we use an ensemble-based DA method called the Maximum Likelihood Ensemble Filter (MLEF; Zupanski, 2005; Zupanski et al., 2008). The MLEF generates the analysis solution which maximizes the likelihood of the posterior probability distribution, obtained by minimization of a cost function. The MLEF belongs to the family of deterministic ensemble filters, hence it is a hybrid between variational and ensemble DA methods. The MLEF employs a cost function derived using a Gaussian probability density function and produces both the analysis and the background error covariance (Zupanski, 2005). It is well suited for use with highly nonlinear observation operators, for a small additional computational cost of minimization using the Hessian preconditioning (Zupanski, 2005; Zupanski et al., 2007b, 2008), and has been employed in many studies including uncertainty analysis, parameter estimation and data assimilation (e.g., Zupanski and Zupanski, 2006; Zupanski et al., 2007a; Lokupitiya et al., 2008; Kim et al., 2010; Apodaca et al., 2014; Tran et al., 2014; Park et al., 2015).

The coupling between the MLEF and WRF-Chem is made through an interface module that transforms the MLEF control variables into the `netcdf` file of WRF-Chem, and vice versa. This interface module is a component of MLEF, and hence the WRF-Chem is not altered.

2.2 Observations

Satellite retrievals often provide estimates of chemical concentration as a total vertical column, and they cover a wide geographical range compared to other measurements (e.g., Silver et al., 2013). In our study, the total column O_3 obtained by OMI is used as an observation. The OMI is a nadir-viewing near-UV/Visible ~~CCD~~ charge-coupled device (CCD) spectrometer aboard NASA's Aura satellite (OMI Team, 2012). The total column O_3 is Level 2 data (OMTO3) based on the Total Ozone Mapping Spectrometer (TOMS) v8.5 algorithm, which is obtained from an orbital swath with a resolution of $13\,\mathrm{km} \times 24\,\mathrm{km}$ at nadir (OMI Team, 2012). It achieves global coverage in one day. In this experiment, we did not apply the quality flags because the first appearance of the row anomaly that affects particular viewing directions, corresponding to the rows on the CCD detectors (OMI Team, 2012) did not occur in 2005, when the TC case considered occurred (i.e., Typhoon Nabi, 2005). Therefore, we employ the OMI data without quality flags.

Figure 1 shows the total column O_3 from OMI at 04:05 UTC 3 September 2005. It shows a lower concentration just inside the eyewall and elevated concentration around the eye. This distinct distribution is well described when the TC has the strongest intensity in the intensifying stages (e.g., Carsey and Willoughby, 2005). Note that OMI switches from its normal global mode to zoom-in mode, to perform spatial zoom (higher resolution) measurements, for a 24 h period about once a month. It occurs when OMI finishes its last orbital pass over Europe, and returns to global mode after 14–15 orbits or about 24 h later. During this period of zoom-in mode, OMI has no global coverage of data (OMI Team, 2012). Typhoon Nabi (2005) reached the maximum intensity on 2 September when OMI entered ~~in~~ into the zoom-in mode. Due to the lack of O_3 data in our domain on 2 September, we have alternatively chosen 3 September for the analysis of O_3 properties during the maximum development of the TC case.

2.3 Experimental design

For the TC case, we choose Typhoon Nabi (2005), which lasted several days from 29 August 2005 until 8 September 2005. Nabi moved westward after its formation and passed near Saipan on 31 August as an intensifying TC, transformed to a super typhoon on 1 September, and reached its peak with winds of $175\,\mathrm{km\,h^{-1}}$ (10-min average) on 2 September. It became weak while turning to the north and striking Kyushu on 6 September. Nabi turned to the northeast after passing by South Korea, and transformed to an extratropical cyclone passing over Hokkaido on 8 September.

In general the DA is composed of two components – *prediction* and *analysis*. A meaningful cycling of DA is inherently related to the prediction component, as every new cycle begins from the forecast guess from the previous cycle. The analysis component of DA is also important, as it provides the impact of observations on the analysis produced by DA. In the current research, we focus

on the analysis component of DA, as the first step towards the eventual DA system for OMI observations.

Conducting the DA cycling with several cycles can make DA more powerful. Although one can potentially have 4 cycles with a 6-hour assimilation window in a day, the ~~limited~~ infrequent availability of OMI observations over the model domain allows only one DA cycle per day. Therefore, we only perform the first DA cycle, which has the strongest impact among the cycles. ~~We believe~~ It is our view that this single cycle DA experiment is sufficient to illustrate the effect of coupled meteorology-chemistry DA.

We focused on a single DA cycle from 00:00 to 06:00 UTC 3 September 2005, which is one of the strongest periods of ~~its~~ the typhoon lifetime. We conduct the experiment with 32 ensembles and 6 h assimilation window. Note that the OMI observations have an approximate frequency of once per day over the typhoon and the surrounding geographical area. Therefore, adding more DA cycles would not be beneficial since no additional data are available. In the future we plan to add a capability to assimilate other observations, such as meteorological observations and all-sky infrared radiances from a geostationary satellite.

The initial and lateral boundary conditions for meteorological states are provided by the National Centers for Environmental Prediction (NCEP) Global Forecasting System (GFS), while those for chemical variables are obtained from the Model for Ozone and Related chemical Tracers (MOZART) chemistry global model of the National Center for Atmospheric Research (NCAR)/Atmospheric Chemistry Division (ACD). The WRF-Chem is set up with a horizontal resolution of $30\,\mathrm{km}$ and 51 vertical levels with the bottom at the ground and the top at $10\,\mathrm{hPa}$ using a terrain-following hydrostatic pressure coordinate (Skamarock et al., 2008).

The model domain is centered over the Korean Peninsula, covering an area of approximately $3900\,\mathrm{km} \times 4400\,\mathrm{km}$ with 132×147 horizontal grid points. The control variables defined in the coupled meteorology–chemistry DA are the WRF-Chem prognostic variables that contain meteorological variables such as winds, perturbation potential temperature, perturbation geopotential, water vapor mixing ratio and perturbation dry air mass in a column, and the chemical variables such as ozone (O_3), nitrates (NO, NO_2, NO_3), and sulfur dioxide (SO_2). The experiments consist of (i) the forecast (without DA) which is useful to understand the synoptic situation and background error covariance, and (ii) the analysis (with DA) which is useful to understand the assimilation ~~impacts~~impact.

2.4 Bias correction of total column O_3

We define the observation operator transforming the WRF-Chem O_3 forecast to the total column O_3 observation. It contains the calculation of total column O_3 with unit conversion and bi-linear interpolation, that is; 1) to transform the physical units of O_3 from the model-produced concentrations in parts-per-million-volume (ppmv) units to the OMI data in Dobson Units (DU), and 2) to transform the O_3 amount from the model grid levels to vertically integrated value at the observation location.

Mathematically, the operator can be written as

$$h(x) = h_i(h_c(h_u(x))) \tag{1}$$

where x denotes an input model variable (e.g., concentration), and h_i represents the horizontal interpolation operator, h_c the vertical column integration and h_u the unit transformation from ppmv to DU. The unit transformation for ozone, h_u, is given by

$$h_u(x) = \frac{A \cdot 10^{-8} \Delta p}{g \cdot m_d} \cdot x \tag{2}$$

where $A = 6.02252 \times 10^{23}$ is the Avogadro number, Δp is the vertical increment of pressure in the layer (hPa), g is the gravity constant, and m_d is the molecular weight of dry air (kg ~~/~~mol^{-1}). The vertical column integration, h_c, is ~~simply~~

$$h_c(s) = \sum_{k=1}^{K} s_k \tag{3}$$

where s_k is the ozone in DU at layer k, and K denotes the number of vertical layers. Finally the bi-linear horizontal interpolation, h_i, is

$$h_i(r) = \sum_{i=1}^{I} w_i r_i \tag{4}$$

where r_i is the vertically-integrated ozone at grid point i, w_i is the bi-linear observation ~~weights~~ weight at grid point i, and I denotes the number of grid points used in the interpolation ($I = 4$ in our case). After combining (2)-(4) into (1), the observation operator for OMI observations becomes

$$h(x) = \sum_{i=1}^{I} w_i \left(\sum_{k=1}^{K} \frac{A \cdot 10^{-8} (\Delta p)_k}{g \cdot m_d} \cdot x_k \right)_i . \tag{5}$$

In these processes, the most demanding part of the observation operator is bias correction of total column O_3 ~~observation~~observations. Although we use the reference pressure at the model top as 10 hPa, which is the highest value we could use in the current model version, there are still considerable amounts of O_3 in the stratosphere that could not be included in the calculation of the model guess (e.g., background). Since this creates a negative bias in the mean observation error, we introduce a multiplicative bias correction ε to preserve positive-definiteness of the bias-corrected guess (Apodaca et al., 2014) as

$$h_B(\boldsymbol{x}) = \varepsilon \cdot h(\boldsymbol{x}) \tag{6}$$

where $\boldsymbol{x}$ is the model state vector. With the multiplicative bias correction in Eq. (6), we can make a new cost function in unbiased form as

$$J(\boldsymbol{x}) = \frac{1}{2}(\boldsymbol{x} - \boldsymbol{x}_\mathrm{b})^\mathrm{T} \mathbf{P}_\mathrm{f}^{-1} (\boldsymbol{x} - \boldsymbol{x}_\mathrm{b}) + \frac{1}{2}[\boldsymbol{y} - h_B(\boldsymbol{x})]^\mathrm{T} \mathbf{R}^{-1} [\boldsymbol{y} - h_B(\boldsymbol{x})] \tag{7}$$

where $\boldsymbol{x}_\mathrm{b}$ is the prior (background) state, $\boldsymbol{y}$ is the observation vector, and the superscript T means a transpose. Here, h is the nonlinear observation operator, $\mathbf{P}_\mathrm{f}$ is the background (forecast) error covariance matrix in the ensemble subspace, and $\mathbf{R}$ is the observation error covariance matrix. Equation (7) is the cost function used in DA, provided ε can be estimated. The optimal value of parameter ε is obtained by implicitly assuming lognormal probability density function errors for a multiplicative bias correction in Eq. (6) (e.g., Apodaca et al., 2014) as

$$\varepsilon = \varepsilon_0 \exp\left[\frac{\frac{1}{N}\sum_{i=1}^{N}\log\left(\frac{\boldsymbol{y}_i}{\varepsilon_0 h(\boldsymbol{x})_i}\right)}{1+\frac{r_0}{w_0}}\right] \tag{8}$$

where ε_0 is a guess parameter value and N is the number of observations. The empirical weighting values are set to $r_0 = w_0 = 0.5$ which implies having the same confidence in observations and the guess. We assume the starting value of the bias to be

$$\varepsilon_0 = \frac{\overline{\boldsymbol{y}}}{\overline{h(\boldsymbol{x})}} \quad \text{where} \quad \overline{\boldsymbol{y}} = \frac{1}{N}\sum_{i=1}^{N}\boldsymbol{y}_i, \quad \overline{h(\boldsymbol{x})} = \frac{1}{N}\sum_{i=1}^{N}h(\boldsymbol{x})_i \tag{9}$$

Equation (8) is calculated once in every DA cycle.

3 Results

A specific characteristic of our experiments is that both meteorological and chemical variables are used as control variables in DA. Regarding the meteorological variables, we focus on what is related to the TC formation and development, such as the temperature, wind, and water vapor. Regarding the chemical variables, we select the chemical constituents such as O_3, NO_2 and SO_2. These are used to identify the impact of O_3 observations on the TC structure in a WRF-Chem-MLEF system.

3.1 Synoptic situation with ensemble WRF-Chem forecast

In general, observations show that SO_2 has larger concentrations in the troposphere while O_3 and NO_2 have larger concentrations in the stratosphere (e.g., Meena et al., 2006). However, in East Asia, especially in eastern China, there is a significant tropospheric NO_2 concentration because of the industrialized and urbanized part of China (Richter et al., 2005; Ohara et al., 2007). Regarding the meteorological variables, temperature and water vapor have higher values in the troposphere, while wind has larger speed near the tropopause. To consider these characteristics, we focused on two pressure levels: (i) 850 hPa (lower troposphere) and (ii) 200 hPa (upper troposphere/lower stratosphere; UTLS). ~~Our ensemble~~ The WRF-Chem forecast ~~also supports these general distributions of control variables~~is in general agreement with the observed synoptic situation, which are not shown in this paper.

3.2 Background error covariance

The background error covariance represents the background state uncertainty (e.g., Zupanski and Zupanski, 2007; Kim et al., 2010). These are estimated by taking the difference between the ensemble perturbation forecasts (total of 32) and the control forecast in the ensemble system (Zupanski, 2005; Zhang et al., 2013). In our study, the ensemble WRF-Chem-MLEF estimates the background error covariance defined in Zupanski (2005) as

$$\mathbf{P}_{\mathrm{f}} = \mathbf{P}_{\mathrm{f}}^{1/2}\left(\mathbf{P}_{\mathrm{f}}^{1/2}\right)^{\mathrm{T}}, \quad \mathbf{P}_{\mathrm{f}}^{1/2} = \left(p_1^{\mathrm{f}} \cdots p_N^{\mathrm{f}}\right), \quad p_n^{\mathrm{f}} = m(\boldsymbol{x}_n^0) - m(\boldsymbol{x}^0) \tag{10}$$

where the index n is an ensemble member, N is the total number of ensemble forecasts, m is the WRF-Chem model, and the subscript 0 denotes the initial time of the forecast with corresponding initial conditions $\boldsymbol{x}^0$ (i.e., control forecast) and ensemble initial conditions $\boldsymbol{x}_n^0$ (i.e., ensemble forecasts). In this experiment, the initial ensemble perturbations are generated by using the lagged forecast outputs (Zhang et al., 2013).

Being calculated from the WRF-Chem ensemble forecast, the flow-dependent background error covariance is defined for meteorological and chemical variables, which allows chemistry observations to impact meteorological variables in DA. In Zhang et al. (2013), a larger background state uncertainty was found in the storm region. Our results also identify the larger background state uncertainty near the TC, similar to Kim et al. (2010). Figure 2 shows the standard deviation (SD) of background error covariance for chemical variables. $\mathrm{O_3}$ in particular (Fig. 2a and d, respectively) shows a large background state uncertainty near the TC, with the maximum of $0.024\,\mathrm{ppmv}$ at $200\,\mathrm{hPa}$ (Fig. 2d). The background state uncertainties of $\mathrm{NO_2}$ and $\mathrm{SO_2}$ at $200\,\mathrm{hPa}$ (Fig. 2e and f, respectively) are located near the TC, characterized by small magnitude and weak influence on tropospheric pollution. On the other hand, the background state uncertainties of $\mathrm{NO_2}$ and $\mathrm{SO_2}$ at $850\,\mathrm{hPa}$ (Fig. 2b and c, respectively) have more impact on central eastern China, implying no visible (or obvious) impact of the low-level $\mathrm{NO_2}$ and $\mathrm{SO_2}$ on the TC.

The SD of background error covariance for meteorological variables appear to be more related to the TC structure (see Fig. 3). In particular, wind (Fig. 3a and d, respectively) shows a larger background state uncertainty near the TC at both pressure level, especially in the eye region at $850\,\mathrm{hPa}$ (Fig. 3a). Temperature (Fig. 3b and e, respectively) also shows a larger background state uncertainty near the TC, especially at $200\,\mathrm{hPa}$ (Fig. 3d). Regarding the water vapor mixing ratio (Fig. 3c and f, respectively), there is a larger background state uncertainty in the eye region at both pressure levels. Larger background state uncertainty potentially implies a stronger analysis correction, provided that total column $\mathrm{O_3}$ observations are available.

3.3 Analysis increment through the $\mathrm{O_3}$ data assimilation

We assess the impact of the assimilated $\mathrm{O_3}$ observations using analysis increments $(\boldsymbol{x}_{\mathrm{a}} - \boldsymbol{x}_{\mathrm{b}})$, which show the correction of the background state using the observations. It is calculated by the following

variable transformation (Zhang et al., 2013; Zupanski, 2005)

$$\boldsymbol{x}_{\mathrm{a}} - \boldsymbol{x}_{\mathrm{b}} = \mathbf{P}_{\mathrm{f}}^{1/2} \left\{ \mathbf{I} + [\mathbf{Z}(\boldsymbol{x}_{\mathrm{b}})]^{\mathrm{T}} \mathbf{Z}(\boldsymbol{x}_{\mathrm{b}}) \right\}^{-1/2} \zeta \tag{11}$$

where ζ is the control variable in the ensemble space; the matrix in Eq. (11) is equal to the inverse of the square root Hessian of the cost function in Eq. (7); $\mathbf{Z}$ is the observation information matrix with column vectors $\mathbf{z}_i = \mathbf{R}^{-1/2}[h(\boldsymbol{x}_i) - h(\boldsymbol{x}_{\mathrm{b}})]$, where the index i denotes the ensemble member.

Figure 4 shows the analysis increments $(\boldsymbol{x}_{\mathrm{a}} - \boldsymbol{x}_{\mathrm{b}})$ of chemical variables obtained by assimilating $\mathrm{O_3}$ observations. By comparing Figs. 2 and 4 one can notice that the $\mathrm{O_3}$ analysis increments are in agreement with background state uncertainties, as expected from Eq. (11). At 850 hPa, the $\mathrm{O_3}$ analysis increment has an increase near the TC, but a decrease over China (Fig. 4a). At 200 hPa, however, there is an increase of $\mathrm{O_3}$ near the TC, and marginal change over China (Fig. 4d). The strong positive response has the largest value of approximately 0.024 ppmv. At 200 hPa, positive $\mathrm{O_3}$ analysis increments are correlated with positive $\mathrm{NO_2}$ (Fig. 4e) and $\mathrm{SO_2}$ (Fig. 4f) increments in the TC region, while no clear correlation is found in other regions. Note that $\mathrm{NO_2}$ and $\mathrm{SO_2}$ are not related to the TC at 850 hPa while the $\mathrm{O_3}$ analysis increments are correlated with $\mathrm{NO_2}$ (Fig. 4b) and $\mathrm{SO_2}$ (Fig. 4c), increasing in central eastern China and Korea and decreasing in northeastern China.

Figure 5 shows the analysis increments $(\boldsymbol{x}_{\mathrm{a}} - \boldsymbol{x}_{\mathrm{b}})$ of meteorological variables by $\mathrm{O_3}$ assimilation. Corresponding to background state uncertainties, the analysis increments of wind show ~~notable~~ significant impact on both lower and upper pressure levels. Positive $\mathrm{O_3}$ increments correspond to positive wind increments at 850 hPa (Fig. 5a), especially in the eye region, and to positive wind increments at 200 hPa (Fig. 5d) in the TC and in the northeastern China and Korea. Regarding the temperature impact (Fig. 5b and e, respectively), the positive $\mathrm{O_3}$ increments generate temperature cooling near the TC and warming over northeastern China. Regarding the water vapor mixing ratio, positive $\mathrm{O_3}$ increments generate a reduction of water vapor mixing ratio (Fig. 5c and f, respectively) near the TC as well as in the eye region at both pressure levels. At 850 hPa, the water vapor mixing ratio is increasing with positive $\mathrm{O_3}$ increments over the northeastern China (Fig. 5c). These results illustrate that chemical observations can impact not only the chemical variables but also the meteorological variables, due to using the ensemble-based coupled meteorology–chemistry background error covariance, as indicated by Park et al. (2015).

3.4 Verification of $\mathrm{O_3}$ data assimilation

As a verification measure, we examine the $\mathrm{O_3}$ assimilation impact on the cost function and on the root mean square (RMS) error with respect to $\mathrm{O_3}$ observations, the same data used in the analysis. The cost function of $\mathrm{O_3}$ driven by Eq. (7) has decreased from 0.36924×10^4 (background) to 0.30689×10^4 (analysis), i.e., it is reduced by approximately 16.9 %. The RMS error, calculated as

$$\mathrm{RMS_a} = \sqrt{\frac{1}{N} \sum [\boldsymbol{y} - h(\boldsymbol{x}_{\mathrm{a}})]^2}, \quad \mathrm{RMS_b} = \sqrt{\frac{1}{N} \sum [\boldsymbol{y} - h(\boldsymbol{x}_{\mathrm{b}})]^2} \tag{12}$$

where subscripts a and b ~~denotes~~ denote analysis and background, respectively, has also decreased from 0.16684×10^2 DU (background) to 0.15204×10^2 DU (analysis), i.e., by about 8.87 %. These results suggest that O_3 assimilation has produced a significant improvement in the initial conditions.

In addition, the impact of total column O_3 observations is also quantified in terms of the uncertainty reduction. With the Gaussian probability assumption, the information content of observations can be represented as the degrees of freedom for signal (DFS; Rodgers, 2000), d_s, as

$$d_s = \mathrm{tr}\left[\mathbf{I} - \mathbf{P}_a \mathbf{P}_f^{-1}\right] \tag{13}$$

where tr is ~~trace functions~~the trace function, $\mathbf{I}$ is the identity matrix, and $\mathbf{P}_a$ and $\mathbf{P}_f$ are the analysis and background error covariances, respectively. Here d_s can also be expressed as

$$d_s = \sum_i \frac{\lambda_i^2}{1+\lambda_i^2} \tag{14}$$

where λ_i are the eigenvalues of the observation information matrix (e.g., Zupanski et al., 2007b). Note from Eq. (14) that the d_s are strictly a non-negative measure: zero values indicate no impact of observations, while positive values indicate a reduction of uncertainty due to assimilation. As shown in Zupanski et al. (2007b), the estimation of Eq. (14) is also useful in a reduced-rank setting of ensemble DA.

Figure 6 shows the DFS of assimilated total column O_3 observations. One can note that ~~it generally coincides~~ the largest values of DFS coincide with the satellite path, and thus ~~with~~ the observations, as expected. The area with the maximum impact is near the TC location, indicating that it is the area where the total column O_3 observation had the strongest impact. In agreement with the analysis increments, there exists a secondary maximum over northeast China, and a smaller one over the Yellow Sea. Given that the DA system includes meteorological and chemical control variables, this result also indicates that O_3 total column observations have a positive impact on both the meteorological and chemical components of the WRF-Chem system, especially in the TC area.

4 Conclusions

In this study, we investigated the impact of ozone (O_3) assimilation on the structure of a tropical cyclone (TC). We directly assimilated the total column O_3 from the Ozone Monitoring Instrument (OMI) in a coupled meteorology–chemistry modelling system – the Weather Research and Forecasting (WRF) model coupled with Chemistry (WRF-Chem). An ensemble-based data assimilation (DA) method, the maximum likelihood ensemble filter (MLEF), is employed and interfaced with the WRF-Chem. We include only a single DA cycle since the OMI observations are covering the model domain only once per day (i.e., 06:00 UTC), and no other observations are available at that time.

Our results show that the O_3 assimilation has a ~~notable~~ significant impact on the analyses of other chemical variables (e.g., NO_2 and SO_2) as well as O_3 itself, and meteorological variables (e.g.,

wind, temperature, water vapor, etc.), especially near the TC case considered. These meteorological variables are closely related to the TC structure and other properties. The O_3 observations can affect other chemical and meteorological variables, and thus the TC itself. For example, temperature is related to development, wind to intensity, and water vapor to precipitation of the TC. Therefore, the implied corrections of these variables in TC regions have a potential to improve the ~~understanding, and eventually the~~ forecast of TCs.

In our DA experiments, the ensemble forecast error, given by the background error standard deviation, appears reasonable with larger uncertainty over the TC area and also over eastern China. The root mean square error reduction indicates an improvement of the optimal analysis state, while the degrees of freedom for signal indicate a reduction of the uncertainty of the optimal analysis.

The use of a single DA cycle limits the conclusions that can be drawn regarding the robustness of the DA system, but it does not impact the performance and implications of using a coupled meteorology–chemistry DA system. It is desired to perform a DA cycling with multiple cycles (i.e., the prediction component of DA); however, it has several difficult aspects that are not possible to resolve in the current setup. It is known that the realistic DA is not perfect in providing dynamically balanced initial conditions, typically resulting in a forecast spin-up period where some of the analysis adjustments are filtered out (Kalnay, 2002). A practical remedy is to produce an improved fit to observations, bringing about the related stronger impact on dynamical model variables (e.g., wind, temperature and pressure), which would eventually result in a longer, sustained influence into the forecast. However, given that the assimilation of OMI observations exerts a stronger impact on chemical variables than dynamical initial conditions, the 24-hour forecast that we need for the next cycle would not be strongly influenced by the OMI observations. Thus we need to assimilate additional observations.

As a future study, we plan to explore the longer DA periods (e.g., several days) to assess the impact of O_3 observation on the track, intensity and precipitation of TCs. Although we have only one available observation product per day for O_3, we anticipate a positive impact of assimilation. In order to obtain more improved DA effects, in addition to O_3, we plan to assimilate NO_2 and SO_2 observations, as well as meteorological observations and all-sky infrared satellite radiances from a geostationary satellite that will be launched in the near future. Noting that NO_2 and SO_2 show high concentrations in East Asia, especially over eastern China, we expect to improve our understanding of the TC structure and the transboundary air pollution as well through assimilation of such chemical compositions from satellite 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 would also like to ac-

knowledge a partial support from the National Science Foundation Collaboration in Mathematical Geosciences Grant 0930265 and the NASA Modeling, Analysis and Prediction (MAP) Program Grant NNX13AO10G.

References

Apodaca, K., Zupanski, M., DeMaria, M., Knaff, J. A., and Grasso, L. D.: Development of a hybrid variational-ensemble data assimilation technique for observed lightning tested in a mesoscale model, Nonlin. Processes Geophys., 21, 1027–1041, doi:10.5194/npg-21-1027-2014, 2014.

Buehner, M.: Ensemble-derived stationary and flow-dependent background-error covariances, Q. J. Roy. Meteorol. Soc., 131, 1013–1043, 2005.

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

Carsey, T. P. and Willoughby, H. E.: Ozone measurements from eyewall transects of two Atlantic tropical cyclones, Mon. Weather Rev., 133, 166–174, 2005.

Chapman, E. G., Gustafson Jr., W. I., Easter, R. C., Barnard, J. C., Ghan, S. J., Pekour, M. S., and Fast, J. D.: Coupling aerosol-cloud-radiative processes in the WRF-Chem model: Investigating the radiative impact of elevated point sources, Atmos. Chem. Phys., 9, 945–964, doi:10.5194/acp-9-945-2009, 2009.

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

Evensen, G.: The ensemble Kalman filter: theoretical formulation and practical implementation, Ocean Dynam., 53, 343–367, 2003.

Fast, J. D., Gustafson Jr., W. I., Easter, R. C., Zaveri, R. A., Barnard, J. C., Chapman, E. G., Grell, G. A., and Peckham, S. E.: Evolution of ozone, particulates, and aerosol direct radiative forcing in the vicinity of Houston using a fully coupled meteorology-chemistry-aerosol model, J. Geophys. Res., 111, D21305, doi:10.1029/2005JD006721, 2006.

Fletcher, S. J. and Zupanski, M.: A data assimilation method for log-normally distributed observational errors, Q. J. Roy. Meteorol. Soc., 132, 2505–2519, 2006.

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., and Eder, B.: Fully coupled "online" chemistry within the WRF model, Atmos. Environ., 39, 6957–6975, 2005.

Houtekamer, P. L. and Mitchell, H. L.: Data assimilation using an ensemble Kalman filter technique, Mon. Weather Rev., 126, 796–811, 1998.

Jang, K. I., Zou, X., De Pondeca, M. S. F. V., Shapiro, M., Davis, C., and Krueger, A.: Incorporating TOMS ozone measurements into the prediction of the Washington, DC, winter storm during 24–25 January 2000, J. Appl. Meteorol., 42, 797–812, 2003.

Kalnay, E.: Atmospheric Modeling, Data Assimilation and Predictability, Cambridge University Press, 364 pp., 2002.

Kim, H. H., Park, S. K., Zupanski, D., and Zupanski, M.: 2010: Uncertainty analysis using the maximum likelihood ensemble filter and WRF and comparison with dropwindsonde observations in Typhoon Sinlaku (2008), Asia-Pac. J. Atmos. Sci., 46, 317–325, 2010.

Lahoz, W. A., Errera, Q., Swinbank, R., and Fonteyn, D.: Data assimilation of stratospheric constituents: a review, Atmos. Chem. Phys., 7, 5745–5773, doi:10.5194/acp-7-5745-2007, 2007.

Lokupitiya, R. S., Zupanski, D., Denning, A. S., Kawa, S. R., Gurney, K. R., and Zupanski, M.: Estimation of global CO_2 fluxes at regional scale using the maximum likelihood ensemble filter, J. Geophys. Res., 113, D20110, doi:10.1029/2007JD009679, 2008.

Meena, G. S., Bhosale, C. S., and Jadhav, D. B.: Retrieval of stratospheric O_3 and NO_2 vertical profiles using zenith scattered light observations, J. Earth Syst. Sci., 115, 333–347, 2006.

Navon, I. M.: Data assimilation for numerical weather prediction: a review, in: Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications, edited by: Park, S. K. and Xu, L., Springer, Berlin, Heidelberg, 21–65, 2009.

Ohara, T., Akimoto, H., Kurokawa, J., Horii, N., Yamaji, K., Yan, X., and Hayasaka, T.: An Asian emission inventory of anthropogenic emission sources for the period 1980–2020, Atmos. Chem. Phys., 7, 4419–4444, doi:10.5194/acp-7-4419-2007, 2007.

OMI Team: Ozone Monitoring Instrument (OMI) Data User's Guide, NASA, Greenbelt, 62, 2012.

Park, S. K. and Zupanski, D.: Four-dimensional variational data assimilation for mesoscale and storm-scale applications, Meteorol. Atmos. Phys., 82, 173–208, 2003.

Park, S. K., Lim, S., and Zupanski, M.: Structure of forecast error covariance in coupled atmosphere–chemistry data assimilation, Geosci. Model Dev., 8, 1315–1320, 2015.

Richter, A., Burrows, J. P., Nüß, H., Granier, C., and Niemeier, U.: Increase in tropospheric nitrogen dioxide over China observed from space, Nature, 437, 129–132, 2005.

Rodgers, C. D.: Inverse Methods for Atmospheric Sounding: theory and Practice, World Scientific, Singapore, 256 pp., 2000.

Rodgers, E. B., Stout, J., Steranka, J., and Chang, S.: Tropical cyclone-upper atmospheric interaction as inferred from satellite total ozone observations, J. Appl. Meteorol., 29, 934–954, 1990.

Silver, J. D., Brandt, J., Hvidberg, M., Frydendall, J., and Christensen, J. H.: Assimilation of OMI NO_2 retrievals into the limited-area chemistry-transport model DEHM (V2009.0) with a 3-D OI algorithm, Geosci. Model Dev., 6, 1–16, doi:10.5194/gmd-6-1-2013, 2013.

Skamarock, W. C., Klemp, J. B., Dudhia, J., Gill, D. O., Barker, D. M., Duda, M. G., Huang, X.-Y., Wang, W., and Powers, J. G.: A description of the Advanced Research WRF version 3. NCAR/TN-475+ STR, National Center For Atmospheric Research, Boulder, CO, 113 pp., 2008.

Stout, J. and Rodgers, E. B.: Nimbus-7 total ozone observations of western North Pacific tropical cyclones, J. Appl. Meteorol., 31, 758– 783, 1992.

Tran, A. P., Vanclooster, M., Zupanski, M., and Lambot, S.: Joint estimation of soil moisture profile and hydraulic parameters by ground-penetrating radar data assimilation with maximum likelihood ensemble filter, Water Resour. Res., 50, 3131–3146, doi:10.1002/2013WR014583, 2014.

Wang, K.-Y., Lary, D. J., Shallcross, D. E., Hall, S. M., and Pyle, J. A.: A review on the use of the adjoint method in four-dimensional atmospheric–chemistry data assimilation, Q. J. Roy. Meteor. Soc., 127, 2181–2204, 2001.

Wu, Y. and Zou, X.: Numerical test of a simple approach for using TOMS total ozone data in hurricane environment, Q. J. Roy. Meteor. Soc., 134, 1397–1408, 2008.

Yang, Q., W. I. Gustafson Jr., Fast, J. D., Wang, H., Easter, R. C., Morrison, H., Lee, Y.-N., Chapman, E. G., Spak, S. N., and Mena-Carrasco, M. A.: Assessing regional scale predictions of aerosols, marine stratocumu-

lus, and their interactions during VOCALS-REx using WRF-Chem, Atmos. Chem. Phys., 11, 11951–11975, doi:10.5194/acp-11-11951-2011, 2011.

Zhang, R., Sanger, N. T., Orville, R. E., Tie, X., Randel, W., and Williams, E. R.: Enhanced NOx by lightning in the upper troposphere and lower stratosphere inferred from the UARS global NO2 measurements, Geophys. Res. Lett., 27, 685–688, 2000.

Zhang, S. Q., Zupanski, M., Hou, A. Y., Lin, X., and Cheung, S. H.: Assimilation of precipitation-affected radiances in a cloud-resolving WRF ensemble data assimilation system, Mon. Weather Rev., 141, 754–772, 2013.

Zou, X. and Y. Wu.: On the relationship between Total Ozone Mapping Spectrometer (TOMS) ozone and hurricanes, J. Geophys. Res.–Atmos., 110, D06109, doi:10.1029/2004JD005019, 2005.

Zupanski, D. and Zupanski, M.: lel error estimation employing an ensemble data assimilation approach, Mon. Wea. Rev., 134, 1337–1354, 2006.

Zupanski, D., Denning, A. S., Uliasz, M., Zupanski, M., Schuh, A. E., Rayner, P. J., Peters, W., and Corbin, K. D.: Carbon flux bias estimation employing Maximum Likelihood Ensemble Filter (MLEF), J. Geophys. Res., 112, D17107, doi:10.1029/2006JD008371, 2007a.

Zupanski, D., Hou, A. Y., Zhang, S. Q., Zupanski, M., Kummerow, C. D., and Cheung, S. H.: Applications of information theory in ensemble data assimilation, Q. J. Roy. Meteor. Soc., 133, 1533–1545, 2007b.

Zupanski, M.: Maximum likelihood ensemble filter: theoretical aspects, Mon. Weather Rev., 133, 1710–1726, 2005.

Zupanski, M.: Theoretical and practical issues of ensemble data assimilation in weather and climate, in: Data Assimilation for Atmospheric, Oceanic and Hydrologic Applications, edited by: Park, S. K. and Xu, L., Springer, Berlin, Heidelberg, 67–84, 2009.

Zupanski, M., Navon, I. M., and Zupanski, D.: The Maximum Likelihood Ensemble Filter as a non-differentiable minimization algorithm, Q. J. Roy. Meteor. Soc., 134, 1039–1050, 2008.

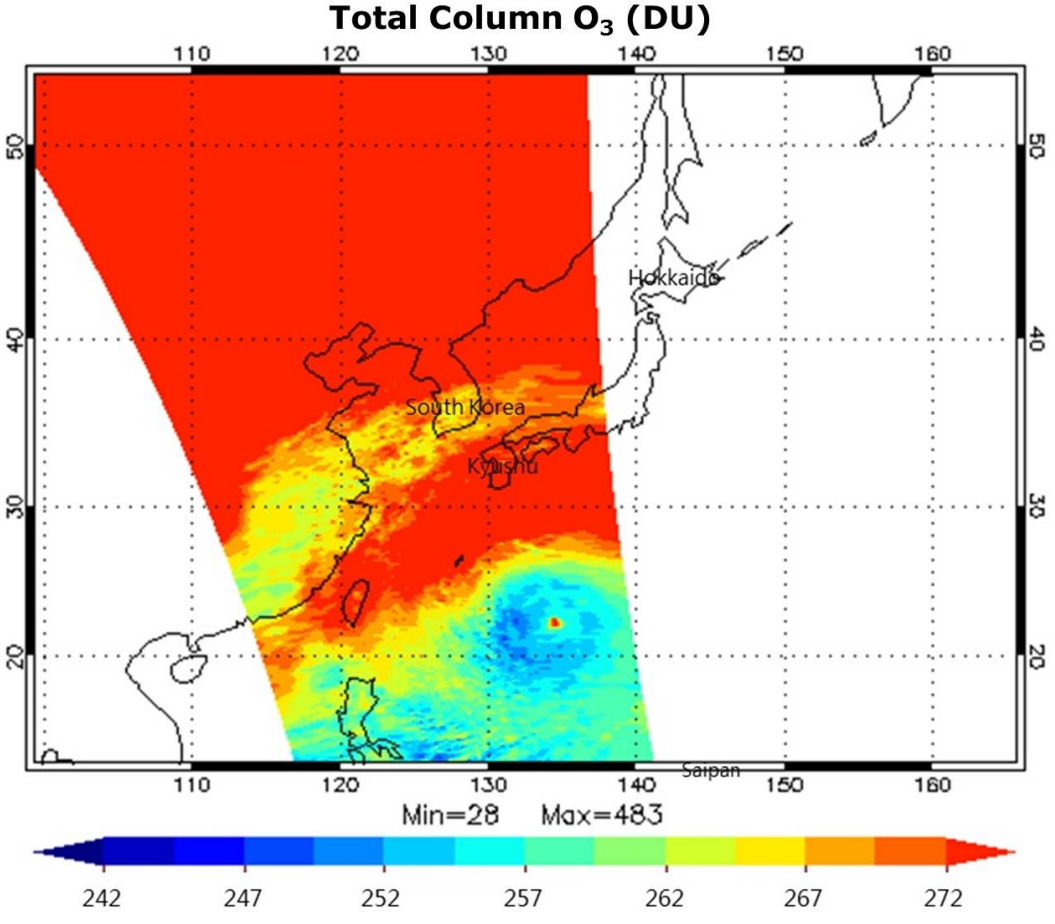

Figure 1. Total column O_3 (in DU) from OMI at 04:05 UTC, 3 September 2005.

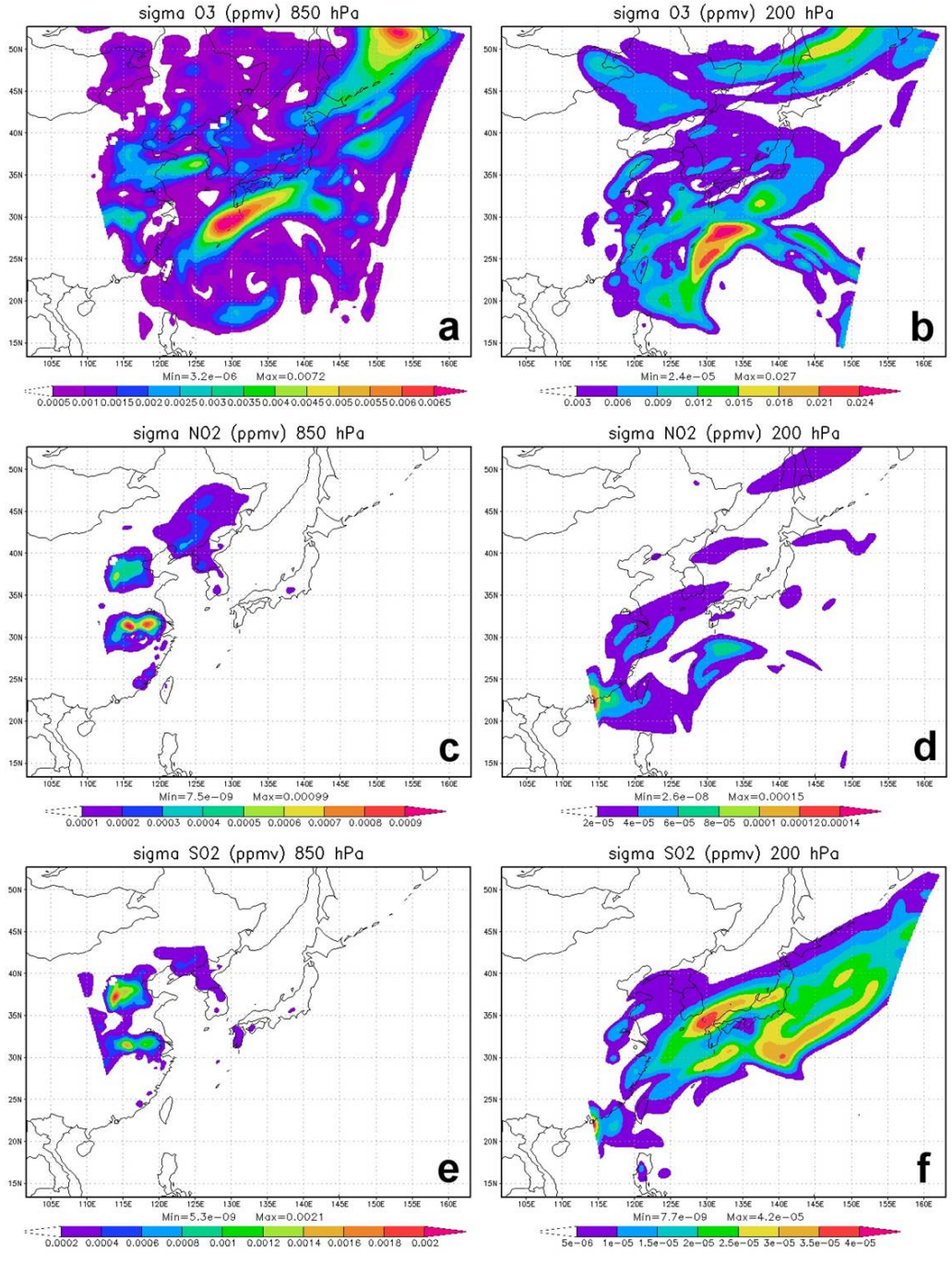

Figure 2. Standard deviation of background error covariance for chemical variables valid on 06:00 UTC, 3 September 2005 at 850 hPa (left panel) for **(a)** O_3, **(b)** NO_2 and **(c)** SO_2, and at 200 hPa (right panel) for **(d)** O_3, **(e)** NO_2 and **(f)** SO_2. Units are ppmv.

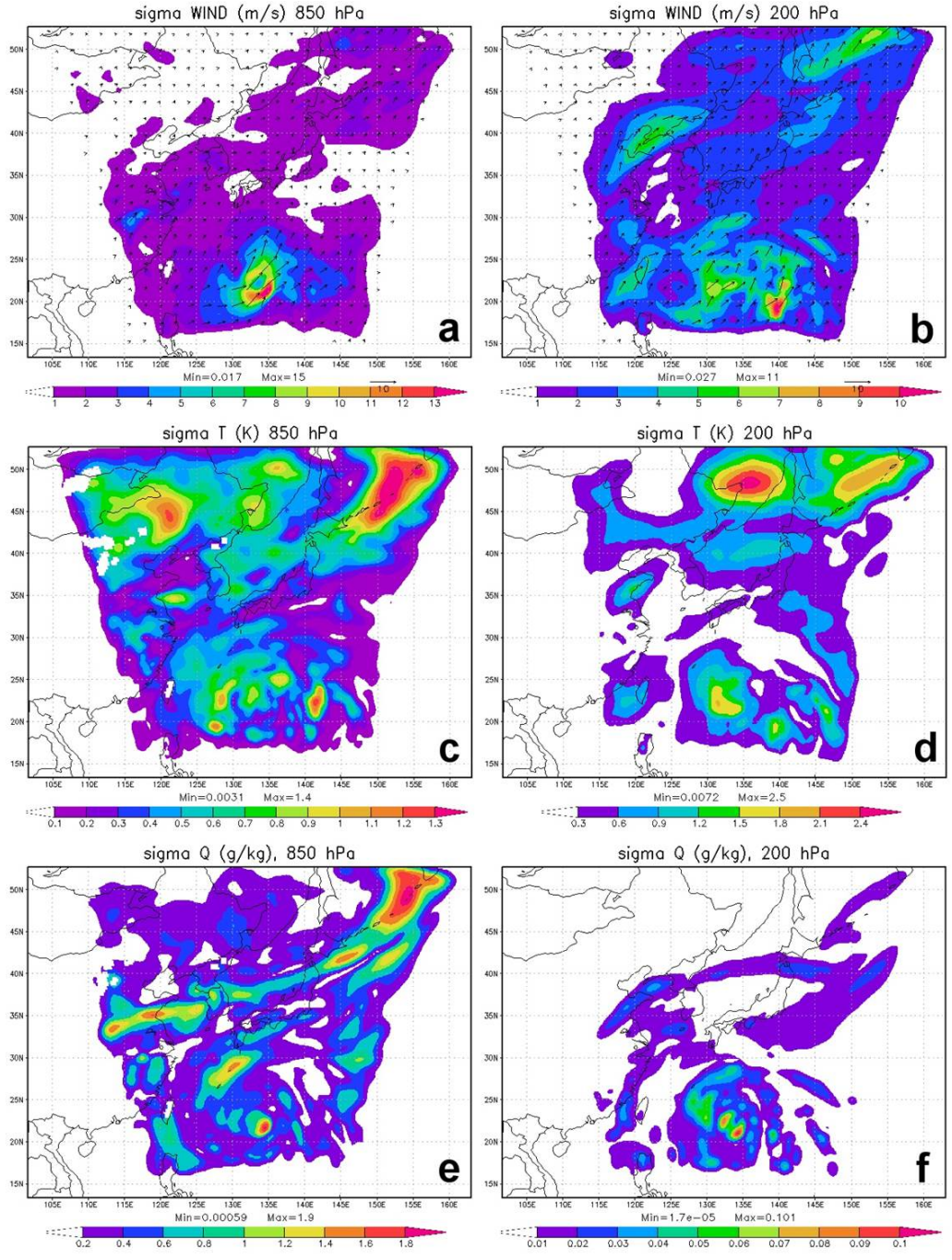

Figure 3. Standard deviation of background error covariance for atmospheric variables valid on 06:00 UTC, 3 September 2005 at 850 hPa (left panel) for **(a)** wind, **(b)** temperature and **(c)** water vapor mixing ratio, and at 200 hPa (right panel) for **(d)** wind, **(e)** temperature and **(f)** water vapor mixing ratio. Units are m s^{-1} for wind, K for temperature and g kg^{-1} for water vapor mixing ratio.

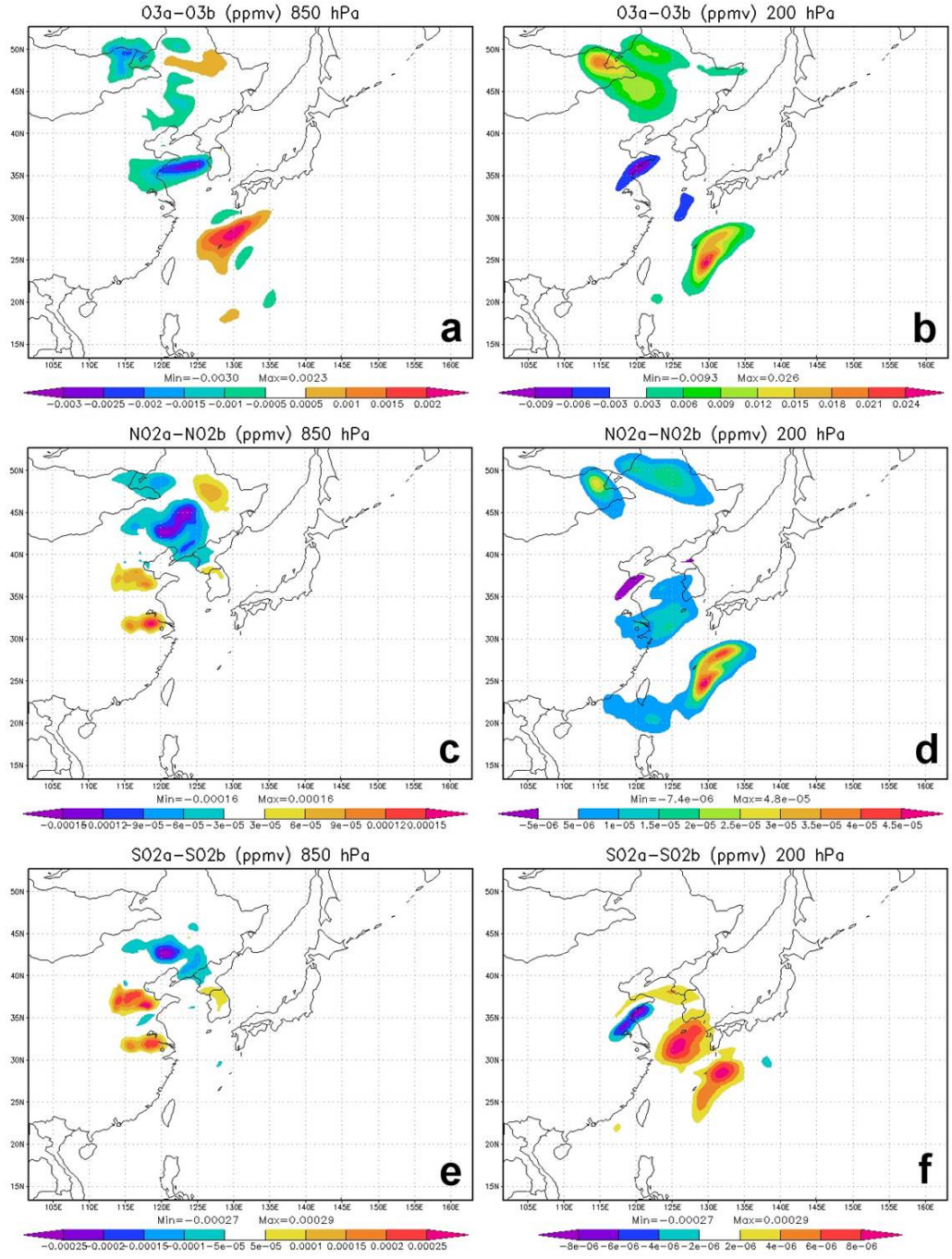

Figure 4. Same as in Fig. 2 except for analysis increment ($\boldsymbol{x}_\mathrm{a} - \boldsymbol{x}_\mathrm{b}$) of chemical variables in response to total column O_3. Units are ppmv.

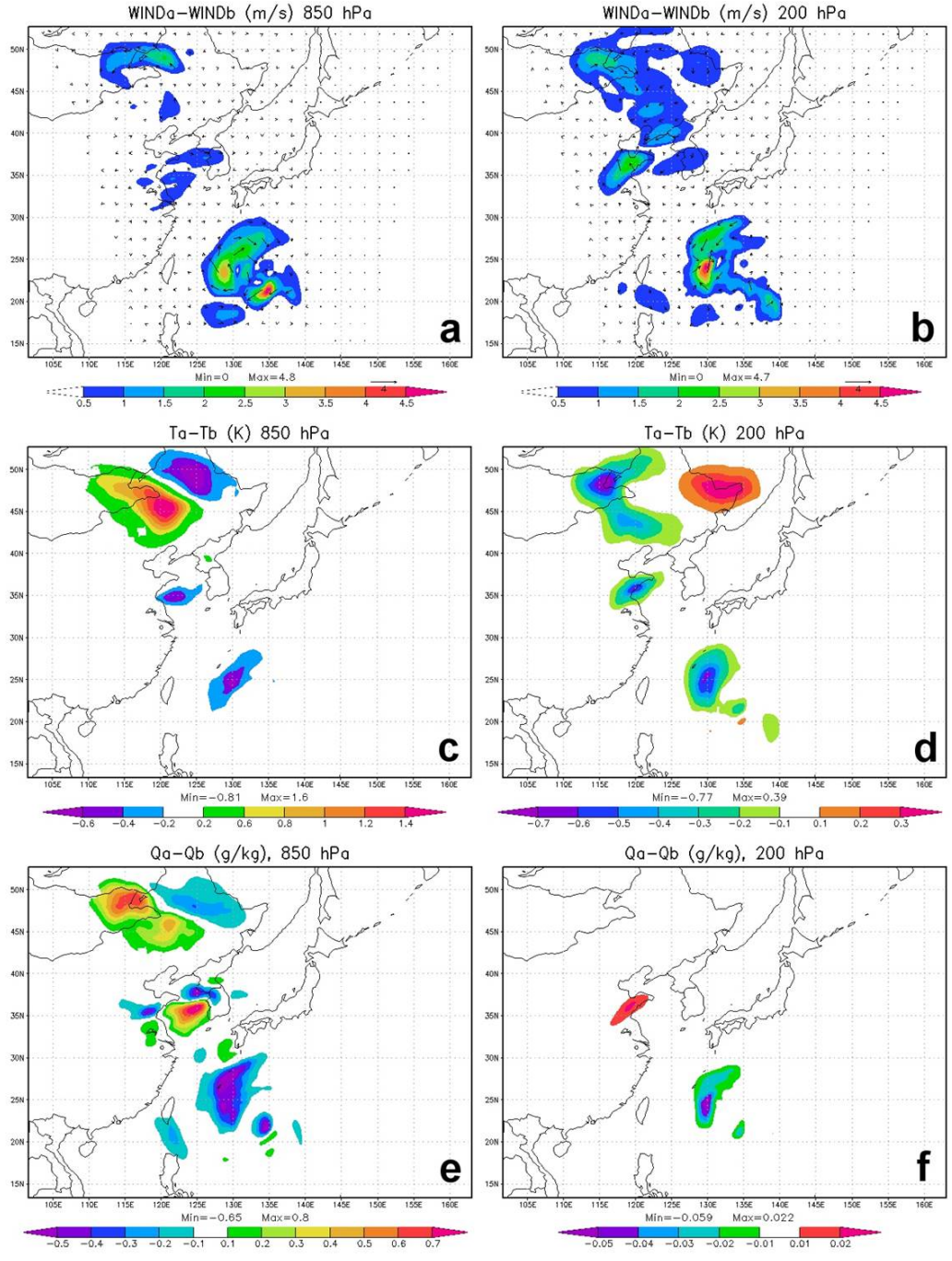

Figure 5. Same as in Fig. 3 except for analysis increment ($\boldsymbol{x}_\mathrm{a} - \boldsymbol{x}_\mathrm{b}$) of atmospheric variables in response to total column O_3. Units are m s^{-1} for wind, K for temperature and g kg^{-1} for water vapor mixing ratio.

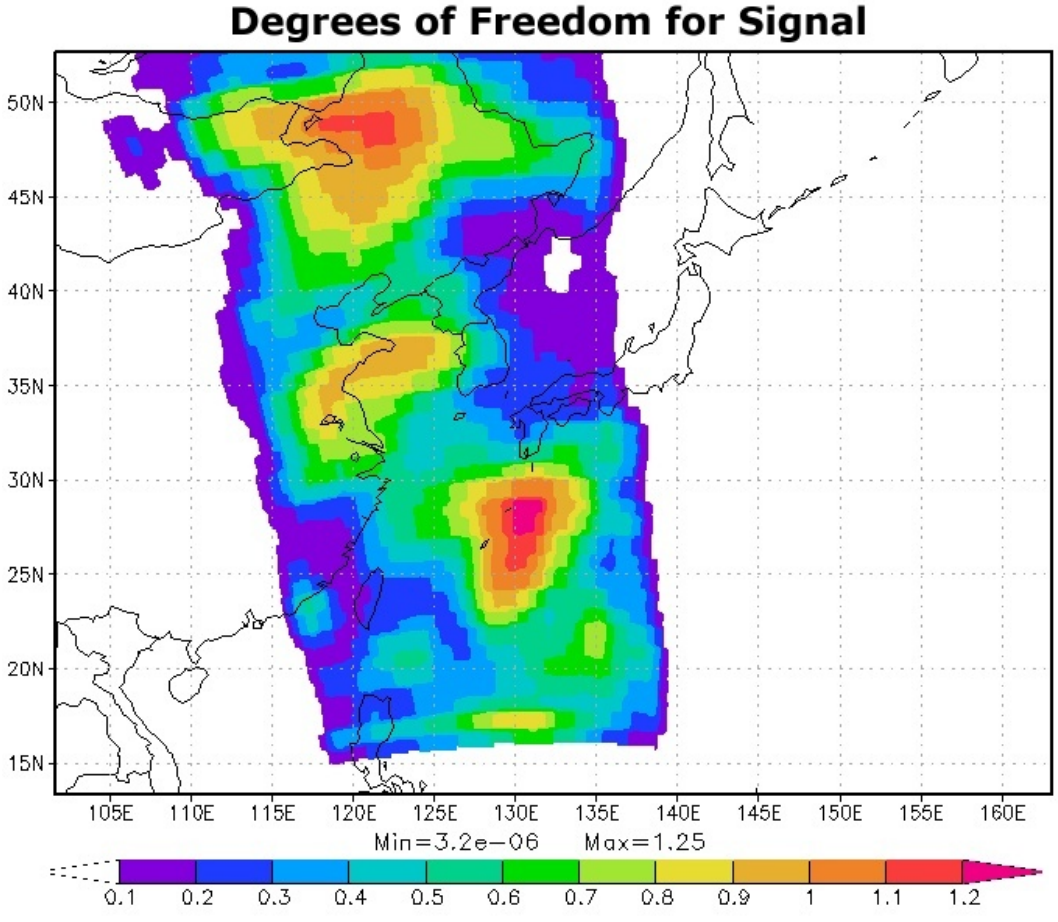

Figure 6. Degrees of freedom for signal (DFS) of assimilated total column O_3 observation valid at 06:00 UTC, 3 September 2005. The units are non-dimensional.