Multiconstituent Data Assimilation With WRF-Chem/DART: Potential for Adjusting Anthropogenic Emissions and Improving Air Quality Forecasts Over Eastern China

Chaoqun Ma[1], Tijian Wang[1] , Arthur P. Mizzi[2], Jeffrey L. Anderson[3], Bingliang Zhuang[1] , Min Xie[1], and Rongsheng Wu[1]

[1]School of Atmospheric Sciences, Nanjing University, Nanjing, China, [2]Air Pollution Control Division, Colorado Department of Public Health and Environment, Denver, CO, USA, [3]Computational and Information Systems Laboratory, National Center for Atmospheric Research, Boulder, CO, USA

Abstract We use the Weather Research and Forecasting Model with the chemistry/Data Assimilation Research Testbed (WRF-Chem/DART) chemical weather forecasting/data assimilation system with multiconstituent data assimilation to investigate the improvement of air quality forecasts over eastern China. We assimilate surface in situ observations of sulfur dioxide (SO_2), nitrogen dioxide (NO_2), ozone (O_3), carbon monoxide (CO), particulate matter with diameters less than 2.5 μm ($PM_{2.5}$) and 10 μm (PM_{10}), and satellite aerosol optical depth to adjust the related anthropogenic emissions as well as the chemical initial conditions. We validate our forecast results out to 72 hr by comparison with the in situ observations. Results show that updated emissions improve the model performance between 10% and 65% root mean square error reduction for the assimilated species except particulate matter with a diameter between 2.5 and 10 μm ($PM_{2.5-10}$), which is slightly improved due to the limited anthropogenic contribution to it. In a sensitivity experiment with a different update interval, the CO improvement is found to be sensitive to the cycling time used to update the CO emissions. In another sensitivity experiment when NO_2 observations are not assimilated and nitrogen oxides (NOx) emission are adjusted by only O_3, NO_2 forecasts show similar root mean square error improvement but have lower spatial correlation, indicating the value and limitation of the O_3-NOx cross-variable relationship.

1. Introduction

Eastern China is home to the Yangtze River Delta, Beijing, Shandong Peninsula, and the Chinese Central Plain. It also accommodates most of the Chinese population. Due to that combination of geography, large population, and the associated urban, agricultural, and industrial activity, eastern China's lower troposphere has become one of the most polluted areas in the world. Accurate atmospheric chemistry forecasting is an important tool for developing efficient control policies and reducing the associated human health impacts.

During the past decades, many atmospheric chemistry numerical models have been developed for atmospheric composition forecasting research. However, few of those models have produced sufficiently accurate air quality forecasts for eastern China due to uncertainties in the model parameterizations, meteorological initial/boundary conditions, chemical boundary conditions, chemical initial conditions (CICs), and especially the emissions (Elbern et al., 2007; Hanna et al., 1998; Hanna et al., 2001; Schmidt & Martin, 2003; van Loon et al., 2007). Fortunately, the latter two sources of uncertainties can be reduced by advanced chemical data assimilation methods that integrate observational information into model forecasts. Forecast improvements have been reported in the literature for different constituents by constraining the CICs and/or emissions with observations by the variational approach (Elbern et al., 2007; Elbern & Schmidt, 2001; Vira & Sofiev, 2012; Yumimoto et al., 2012) or the ensemble Kalman filter (EnKF) approach (Barbu et al., 2009; Jiang et al., 2017; Miyazaki et al., 2012; Miyazaki et al., 2017; Mizzi et al., 2016; Tang et al., 2011).

Constraining the emission with model and observations is usually called "top-down" approach to distinguish with building emission inventory from detailed industrial statistical data (which is usually referred

as "bottom-up" approach). Two basic methods exist to do the top-down emissions constraining: (i) adjoint and (ii) EnKF. The adjoint method minimizes a cost function by variational approaches using the adjoint of a model. The drawback of adjoint method is the huge amount of work concerning developing and maintaining the adjoint model. The EnKF uses the flow-dependent covariance generated by an ensemble of model output instead of adjoint model to convert observation information into emission. Therefore, EnKF is more flexible and easy to apply than adjoint method.

However, when performing "top-down" emission adjustments, previous studies generally assimilate observations of only one or two species and adjust only the corresponding emission species. For example, Arellano et al. (2006) and Yumimoto and Uno (2006) used carbon monoxide CO concentration observations to constrain the CO emissions, and Sekiyama et al. (2010), Mao et al. (2014), and Peng et al. (2017) used aerosol observations to constrain the associated aerosol emissions. Similar strategies for constraining emissions can be found for nitrogen oxide (NO_x) emissions (Kurokawa et al., 2009) and sulfur dioxide (SO_2) emissions (Vira & Sofiev, 2012).

Those works for constraining emissions fail to take full advantage of the potential of the numerical algorithm because the emissions of one species and the concentrations of another species are often correlated. In addition, multiconstituent observations are routinely available over China and likely contain information about the cross-species emission/concentration relationship that can be used to constrain the emissions. Some previous studies have shown the advantages of assimilating multiconstituent observations. For example, Barbu et al. (2009) updated sulfur oxide (SO_x) emissions with SO_2 gas and sulfate (SO_4) aerosol observations; forecasts were improved when both kind of observations were assimilated but degraded if assimilating only SO_2 or SO_4. However, care must be taken when using cross-species concentration/emission information to constrain emissions because it can yield negative results. For example, in Tang et al. (2016), NO_x forecasts were found to get worse when constraining the NO_x and volatile organic component (VOC) emissions with only ozone (O_3) observations. The result is attributed to the nonlinear relation between O_3 concentration and NO_x emission who are positively correlated at NO_x limited region (NO_x is relatively low compared to VOC) and negatively correlated at VOC limited region (NO_x is relatively high compared to VOC). That means the filter could adjust the NOx emission in the wrong direction when limited region is wrongly detected. Such degradation was expected to be avoided by assimilating extra NOx observations to directly anchor NOx emissions.

The most recent work by Peng et al. (2018) assimilated near-surface in situ observations of O_3, NO_2, SO_2, and CO and aerosols observations to simultaneously adjust the CICs and emission inputs based on EnKF. Their resultant emission succeeded in improving SO_2 forecast while having little influence on aerosol forecast and even degrading the forecast of NO_2 and O_3. The failure is partly attributed to the lack of VOC emission adjustment and error of chemical mechanisms (Regional Atmospheric Chemistry Mechanism from Stockwell et al., 1997). Considering the potential advantages of top-down emission and the importance of O3 forecast, it is still worth a try to advance the work by solving its deficiencies.

In this paper, we use WRF-Chem/DART (Mizzi et al., 2016; Mizzi et al., 2018) to assimilate near-surface in situ observations of O_3, NO_2, SO_2, CO, and aerosols as well as MODIS aerosol optical depth (AOD) to constrain the related chemical initial states and emissions. Compared to Peng et al. (2018), VOC emission will be adjustable here for better O_3 forecast. Also, model is configured with a different chemical mechanisms and finer horizontal resolution. WRF-Chem/DART is a comprehensive ensemble-based chemical weather forecasting/data assimilation system that uses the ensemble adjustment Kalman filter of J. L. Anderson (2001) and the state augmentation method to constrain the chemical initial conditions and emissions. We studied the impact of independent and cross-species concentration/emission adjustment as well as the dependence on the frequency for updating the emissions. Generally, we found that constraining the emissions improved the forecast skill out to 72 hr.

The paper is organized as follows: The next section discusses the forecast model/data assimilation system, observations, and experimental design; section 3 discusses results of the base, and various chemical data assimilation sensitivity experiments are discussed in section 4; and section 5 presents the summary and conclusions.

2. Data and Methods

2.1. WRF-Chem/DART

Mizzi et al. (2016) developed WRF-Chem/DART by integrating the Weather Research and Forecasting Model with chemistry (WRF-Chem) with the Data Assimilation Research Testbed (DART). WRF-Chem/DART can assimilate atmospheric meteorological and chemical observations from a variety of in situ and remote observation platforms. It updates the corresponding meteorological and chemical initial/boundary conditions and uses the state augmentation method to constrain the related emissions (Liu et al., 2017).

WRF-Chem is used to propagate the initial meteorological and chemical ensemble forward in time. WRF-Chem is an online three-dimensional, Eulerian chemical transport model that considers the complex physical and chemical processes in the troposphere (Grell et al., 2005). It has been applied in various research settings including other chemical data assimilation studies (Li et al., 2013; Schwartz et al., 2012).

DART is an open source community facility for ensemble-based data assimilation research. DART is developed and maintained by the National Center for Atmospheric Research (NCAR; J. Anderson et al., 2009). DART uses a well-documented and flexible coding structure that has facilitated its application to a variety of research projects. See, for example, Kwon et al. (2016) and Rubin et al. (2017). The DART assimilation engine used here is the ensemble adjustment Kalman filter (EAKF) of J. L. Anderson (2001, 2003).

The EAKF is an alternative to the EnKF algorithm advanced by G. Evensen (1994, 2003) with the following formulas:

$$\overline{z^u} = \Sigma^u \left[(\Sigma^p)^{-1} \overline{z^p} + H^T R^{-1} y^o \right], \tag{1}$$

$$\Sigma^u = \left[(\Sigma^p)^{-1} + H^T R^{-1} H \right]^{-1}, \tag{2}$$

$$z_i^u = A^T (z_i^p - \overline{z^p}) + \overline{z^u}, \; i = 1, \ldots, N. \tag{3}$$

In equations (1), (2), and (3), $\overline{z^u}$ is the analysis (posterior) ensemble mean while $\overline{z^p}$ is the background (prior) ensemble mean. z_i^p represents the ith member of the background field ensemble, and z_i^u means the ith member of the analysis field ensemble. y^o is the observations vector with R being its error covariance matrix and H being the corresponding observational operator. Σ^p and Σ^u are the background and analysis error covariance matrix, respectively. A is an arbitrary matrix that satisfies $\Sigma^u = A\Sigma^p A^T$ and could be easily obtained by doing local least squares fit. With A applied to update individual ensemble member like equation (3), EAKF does not need to perturb observation y^o like traditional EnKF in Geir Evensen (2003) and thus eliminates sampling errors associated (Whitaker & Hamill, 2002).

2.2. Forecast Model and Prior Emissions

For our experiments, we use version 3.7.1 of WRF-Chem and the Manhattan version of DART. WRF-Chem uses the Goddard Chemistry Aerosol Radiation and Transport model for aerosols (Colarco et al., 2010). For gas chemistry, the Model for Ozone and Related Chemical Tracers (MOZART-4; Emmons et al., 2010) for gas chemistry is chosen to be different from Peng et al. (2018). As shown in Figure 1, the domain covers eastern China with a horizontal resolution of 15 km and 36 vertical layers. The resolution of 15 km is finer than that (40.5 km) used in Peng et al. (2018) because the result of emission inversion is sensitive to model resolution (Mao et al., 2014).

Meteorological initial and lateral boundary conditions come from the 6-hr 0.25° National Centers for Environmental Prediction Final operational global analysis (https://rda.ucar.edu/datasets/ds083.3). The chemical initial and lateral boundary conditions come from global simulations from the NCAR MOZART-4 model (https://www.acom.ucar.edu/wrf-chem/mozart.shtml).

Within China, the prior anthropogenic emissions were provided by the 0.25° Multi-resolution Emissions Inventory for China in 2015 (MEIC; http://www.meicmodel.org). Outside China, the associated emissions come from the 0.1° Emissions Database for Global Atmospheric Research for 2010. Following Boersma et al. (2008) and Schaap et al. (2008), we scaled the emissions with a prescribed hourly diurnal variation profile, which have two peaks, which are reached during local early morning and evening. No interdiurnal variation was imposed, which means prior emission kept the same for each single day. The Model of Emissions
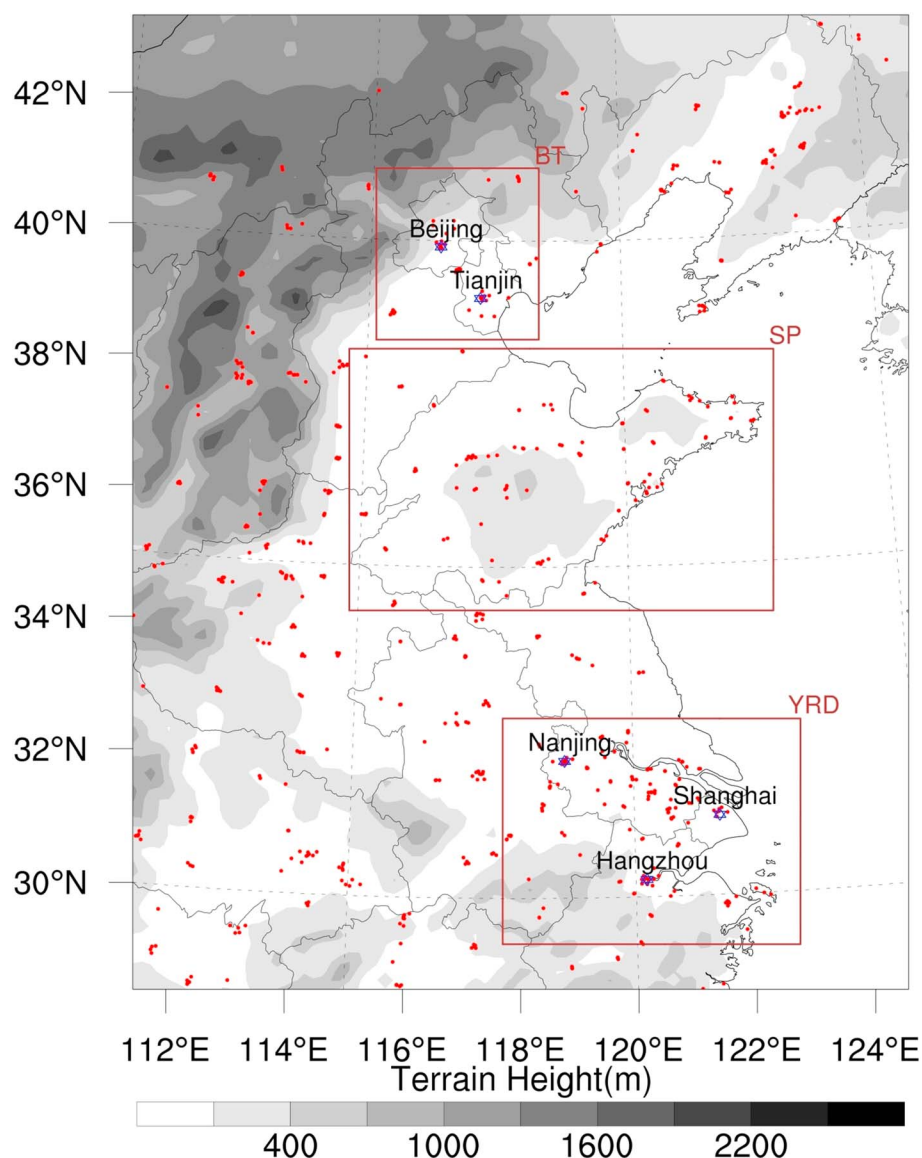
Figure 1. The model domain and observation network (red dots) with model topography (contour plot). The three embedded rectangles identify Beijing-Tianjin (BT), Shandong Province (SP), and the Yangtze River Delta (YRD) regions. Those areas contain most of China's anthropogenic emissions and pollution events.

of Gases and Aerosols from Nature (Guenther et al., 2012) was used to calculate the time-dependent, prior biogenic emissions during the forecast cycle. We did not include biomass burning emissions because they had little to no impact within our domain during the study period.

We use the WRF data assimilation system (WRFDA) to generate the initial meteorology ensemble and NCAR/ACOM WRF-Chem utilities (https://www2.acom.ucar.edu/wrf-chem/wrf-chem-tools-community) to generate the initial chemistry ensemble. The chemistry ensemble is generated in the same way as Mizzi et al. (2016). We impose a Gaussian distribution around a specified mean (the deterministic chemistry initial conditions for MOZART-4) with a 30% standard deviation (STD) about the mean. We use a different approach to generate the emissions ensemble. Following Boynard et al. (2011), we generally impose a Gaussian distribution with a zero mean and a 30% STD on the emission correction factors (introduced in the next paragraph). The exceptions are that for SO_2 we used a 20% STD to account for greater data certainty and for VOC emissions we use a 50% STD to account for greater uncertainty. We impose a horizontal correlation length of 600 km because the MEIC inventory was constructed with data from each province and the

Table 1
The Observation Impact Localization

	Description	Observations that have impact
BC1	Hydrophobic black carbon	$PM_{2.5}$, MODIS AOD
BC2	Hydrophilic black carbon	$PM_{2.5}$, MODIS AOD
OC1	Hydrophobic organic carbon	$PM_{2.5}$, MODIS AOD
OC2	Hydrophilic organic carbon	$PM_{2.5}$, MODIS AOD
SO4	Sulfate	$PM_{2.5}$, MODIS AOD
S1	Sea salt of the 1st size bin	$PM_{2.5}$, MODIS AOD
S2	Sea salt of the 2nd size bin	$PM_{2.5}$, $PM_{2.5\text{-}10}$, MODIS AOD
S3	Sea salt of the 3rd size bin	$PM_{2.5\text{-}10}$, MODIS AOD
D1	Mineral dust of the 1st size bin	$PM_{2.5}$, MODIS AOD
D2	Mineral dust of the 2nd size bin	$PM_{2.5}$, $PM_{2.5\text{-}10}$, MODIS AOD
D3	Mineral dust of the 3rd size bin	$PM_{2.5\text{-}10}$, MODIS AOD
D4	Mineral dust of the 4th size bin	$PM_{2.5\text{-}10}$, MODIS AOD
P25	Fine-mode unspeciated aerosol	$PM_{2.5}$, MODIS AOD
P10	Coarse-mode unspeciated aerosol	$PM_{2.5\text{-}10}$
NO2	Nitrogen dioxide	NO_2
O3	Ozone	O_3
CO	Carbon monoxide	CO
SO2	Sulfur dioxide	SO_2
E_CO	Emission of anthropogenic carbon monoxide	CO (only at 12UTC)
E_NOx	Emission of anthropogenic nitrogen oxides	NO_2, O_3
E_VOC	Emission of anthropogenic volatile organic components	O_3
E_SO2	Emission of anthropogenic sulfur dioxide	SO_2
E_BC	Emission of anthropogenic black carbon	$PM_{2.5}$, MODIS AOD
E_OC	Emission of anthropogenic organic carbon	$PM_{2.5}$, MODIS AOD
E_PM10	Emission of anthropogenic coarse-mode unspeciated aerosol	$PM_{2.5\text{-}10}$
E_PM25	Emission of anthropogenic fine-mode unspeciated aerosol	$PM_{2.5}$, MODIS AOD

Note. The first column lists the model state variable by variable name. The second column provides a description of the model state variable, and the third column lists the observations that were allowed to update the state variable identified in columns one and two.

characteristic length scale for a province in eastern China is 600 km. Cross-species correlations were not considered.

2.3. Data Assimilation Configuration

We conducted continuous 6-hr cycling with WRF-Chem/DART using a 20-member ensemble to update the chemical concentrations and most emission rates. The cycle times were 00:00, 06:00, 12:00, and 18:00 UTC. For updating CO, we used a 24-hr cycle with update at 12:00 UTC. Only observations within $\pm$1-hr window were assimilated. The rationale behind using a different cycle time for CO and why choosing 12:00 UTC are discussed in section 4. In our experiments, we assimilated chemical observations and not meteorological observations. To prevent deviation from the real meteorological condition, meteorological variables ensemble was overwritten with a new one perturbed by WRFDA from Final analysis of the current time after each DA step.

The assimilation state variable vector x can be written as $(m^T, c^T, e^T)^T$. The column vectors m and c represent the three-dimensional fields of meteorological and chemical variables, respectively. The column vector e represents the two-dimensional emissions correction factor fields for each emitted species. The actual emission rate $E_{i,j,s,t}$ is defined as

$$E_{i,j,s,t} = E^p_{i,j,s,t} \cdot \exp\left(e_{i,j,s,t}\right), \quad (4)$$

where $E^p_{i,j,s,t}$ is the prior emission flux at the (i,j) grid location for species s and time t and $e_{i,j,s,t}$ is the corresponding emissions correction factor from e. Because we had no model to propagate e to the next time step, e was assumed to keep constant unless updated by DA (every 6 hr here). So emission for any time after one DA cycle was calculated using equation (4) with the same e. Following Elbern et al. (2007), we use an exponential correction factor to make the error distribution more Gaussian. Additionally, we use the same correction factor for all anthropogenic VOC emissions (15 kinds) to maintain a constant ratio between the different emission species.

For the horizontal localization we conducted a series of sensitivity tests and found that a three-dimensional Gaspari-Cohn type localization function with a horizontal radius half width of 0.05 radians, and 0.075 vertical scale height of the atmosphere gave the best overall results. We used independent data assimilation (also called state-variable localization; see Miyazaki et al., 2012) for updating the chemical state variable and emissions; that is, the observation of a particular chemical species is allowed to update only the same state variable species and emissions rate. We allow one exception. For some experiments, O_3 observations update the NO_x and VOC state variables and emissions. Table 1 summarizes the allowed chemical observation/chemical state variable and emissions interactions. This type of localization is commonly used in chemical data assimilation to reduce spurious correlations and is not the main focus of our research. It does limit the generality of our results because it generally imposes zero correlations on interactions that should have near-zero correlations. To avoid potential ($PM_{2.5}$, PM_{10}) cross-correlations, we assimilated $PM_{2.5}$ and PM_{10} minus $PM_{2.5}$ ($PM_{10\text{-}2.5}$).

For the covariance inflation, we used the DART prior adaptive inflation algorithm of J. L. Anderson (2007). However, that scheme did not work well for the emission. Therefore, we applied a "relaxation to prior spread" method only to emission based on Kotsuki et al. (2017). In our application, we followed Kotsuki et al. (2018) to apply a relaxation parameter of 1.0, meaning that the posterior spread is the same as the prior spread. We imposed time-invariant spread because the observed emissions are thought to have comparable

uncertainty from one day to the next. We experimented with relaxation factors less than 1.0 and found that the emissions spread collapsed after several cycles. That result is expected because there is no forecast step for the emissions. Consequently, the only way to generate/maintain the emissions spread is through inflation. However, the assumption of time-invariant emission uncertainty is only a compromise when knowing nothing about emission uncertainty evolution. Therefore, it may not hold in some cases.

2.4. Observations

We assimilated near-surface observations of CO, SO_2, NO_2, O_3, $PM_{2.5}$, and PM_{10} collected more than 1,000 Chinese Ministry of Environmental Protection (MEP) sites with a 6-hr cycling period and a ± 1-hr observation window. The MEP monitoring site locations are shown in Figure 1. Those sites are evenly distributed around eastern China and become denser near metropolitan areas. We performed quality assurance (QA) reviews to remove suspect observations, including value-range and time-continuity checks. For the remaining observations, we assigned observation errors using the method of Elbern et al. (2007) and Schwartz et al. (2012) to account for measurement and representativeness error with error STDs of 0.3 mg/m^3 for CO, 8 µg/m^3 for SO_2, 18 µg/m^3 for NO_2, 22 µg/m^3 for O3, 8 µg/m^3 for $PM_{2.5}$, and 8 µg/m^3 for $PM_{10-2.5}$.

We also assimilated MODIS AOD observations to provide an extra constraint on the aerosol-related state variables. The AOD observations are total column retrievals. We assimilated the AOD retrievals from the Dark Target data set in Collection 6.1 at 550 nm with a QA flag of 3 (refer to Levy et al. (2013); Sayer et al. (2014) for more details on the MODIS AOD retrievals). Following Hyer et al. (2011), we assigned observations errors based on the following error STDs: (i) 0.1 for $\tau \leq 0.2$, (ii) $0.05 + 0.20\tau$ for $0.2 < \tau \leq 1.4$, and (iii) $0.2 + 0.4\tau$ for $\tau > 1.4$ where τ is the report AOD retrieval. That error specification is slightly larger than the uncertainty recommended by Levy et al. (2013) to account for representativeness differences between AOD retrievals and the model state variables.

Since the horizontal resolution of the AOD retrievals (10 km) and near-surface observations (sometimes more than 1 monitoring site within one model grid) is finer than that of the model, we super-obbed the retrievals and their error variance using the method of Barré et al. (2015). We expect that supper-obbing improved the retrievals' representativeness and reduced the horizontal observation error correlations.

In addition to the QA review discussed above for the MEP in situ and AOD retrieval observations, DART did an outlier check and rejected observations that differed from the model background by more than three times the total prior spread.

2.5. Experimental Design

We use a 10-day study period running from 06:00 UTC 3 September 2016 to 06:00 UTC 13 September 2016. Prior to starting our forecast/assimilation experiments, we spun-up the model by running the model for 2 days from 06:00 UTC 1 September 2016 to 06:00 UTC 3 September 2016 without data assimilation. The model spin-up enabled us to start our experiments with a greater degree of balance between the emissions and the CICs. We used continuous cycling with an assimilation/forecast step every 6 hr with emissions adjustment.

To evaluate the effects of DA on long time forecasts, after 3 days of cycling, starting at 12:00 UTC, 6 September 2016, we ran four 72-hr model integrations after each cycle. The four model integrations, namely, four kinds of forecasts, were initiated with different CICs and emissions input. The first kind of forecast, identified as the "NO DA," was a control forecast that ran the forecast using the prior emissions and CICs from NO DA forecast of the last DA cycle, which means NO DA forecast received no information from DA. The second kind of forecast, identified as the "CICs ONLY," used the posterior mean CICs from DA to initiate while the input emission inventory was the original prior one, which means only CICs were improved by DA. The third kind of forecast, identified as the "EMISS CICs," used the posterior mean CICs and emissions from DA to initiate, which means EMISS CICs forecast used both CICs and emissions updated by DA. The fourth kind, identified as the "EMISS CICs NO-VOC," was a variant of the EMISS CICs where anthropogenic VOC emission was replaced by the original prior one, which can help identify the benefit of updating the VOC emissions through the assimilation of O_3 when it is compared to the EMISS CICs forecast. In the end of the experiment, we are expecting to have 112 (28 cycle × 4 kinds) 72-hr forecasts. Collectively, the assimilation experiment described in section 2 together with its 72-hr forecasts is referred to as the "Base Exp" to distinguish from sensitivity experiments that will be described in

Table 2
Configuration of the Different Kinds of 72-hr Forecast

Forecast name	Whether initiated from updated CICs	Whether initiated from updated emissions	Initiated from which DA experiment	DA experiment configuration
NO DA	No	No		
CICs ONLY (Base)	Yes	No	Base Exp	Described in section 2 and table 1
EMISS CICs (Base)	Yes	Yes	Base Exp	The same to above
EMISS CICs NO-VOC (Base)	Yes	Yes, but VOC emission is the same with NO DA	Base Exp	The same to above
CICs ONLY CO 6-hr	Yes	No	CO 6-hr Exp	The same to Base Exp except CO emission was updated every 6 hours.
EMISS CICs CO 6-hr	Yes	Yes	CO 6-hr Exp	The same to above
CICs ONLY NO NO_2	Yes	No	NO NO_2 Exp	The same to Base Exp except NO_2 observations were not assimilated.
EMISS CICs NO NO_2	Yes	Yes	NO NO_2 Exp	The same to above

Note. "Base" is bracketed in the first column, which means it will be omitted when not comparing forecasts from different DA experiments.

section 4 used to identify the impact of changing some DA configurations. Table 2 gives a summary of DA experiments and forecasts from them.

3. Results From the Base Exp

In this section, we study the changes in the emissions and forecast skill due to the chemical data assimilation and emissions updates in the Base Exp.

3.1. Observation Space Diagnostics

Figure 2 shows three assimilation metrics: (i) root mean square error (RMSE); (ii) total spread, the square root of the sum of the ensemble variance and the observation error variance; and (iii) the number of observations assimilated for the Base Exp.

The left column of Figure 2 shows that for the Base Exp, the posterior ensemble mean is always closer to the observations than the prior ensemble mean, and there is a decreasing trend for both. That result is especially significant for AOD, $PM_{2.5}$, and $PM_{2.5-10}$, which are assimilated jointly. It suggests the AOD, PM2.5 and $PM_{2.5-10}$ observations, which all contain information about aerosol concentration, can be jointly assimilated without degrading model performance on predicting any of them. The check is necessary because sometimes AOD observations could suffer from inappropriate preprocessing and therefore get incorrectly higher than the model background. Then the assimilation of those AODs would result posterior aerosol concentration too high that the posterior $PM_{2.5}$ is even further from observed $PM_{2.5}$.

The middle column of Figure 2 is essentially a diagnostic of chi-square statistics (Menard & Changs, 2000). The prior RMSE, defined as $\sqrt{\frac{1}{P}\sum_{i=1}^{P}\left(y_i^o - \overline{y_i^f}\right)^2}$ where y_i^o is an assimilated observation, $\overline{y_i^f}$ is the prior ensemble mean of the expected observation, and p is the number of assimilated observations, should be comparable to the prior total spread, which is defined as $\sqrt{\frac{1}{P}\sum_{i=1}^{P}\left(\sigma_i^o2 + \sigma_i^f2\right)}$ where $\sigma_i^{o^2}$ is the observation error variance and $\sigma_i^{f^2}$ is the prior ensemble variance in observation space. If the observation error, model ensemble spread, and covariance inflation are properly specified, the total spread and prior RMSE should be balanced. The middle column of Figure 2 shows that our assimilation system is well balanced because the total spread and the prior RMSE have comparable magnitude and trends. The right column of Figure 2 shows the number of assimilated and rejected observations during the study period. Overall, the rejected observations are less than 10% of the total number of observations. This shows that the assimilation system did not reject too many observations and that the ensemble members closely tracked the temporal evolution of the real atmosphere throughout the study period.
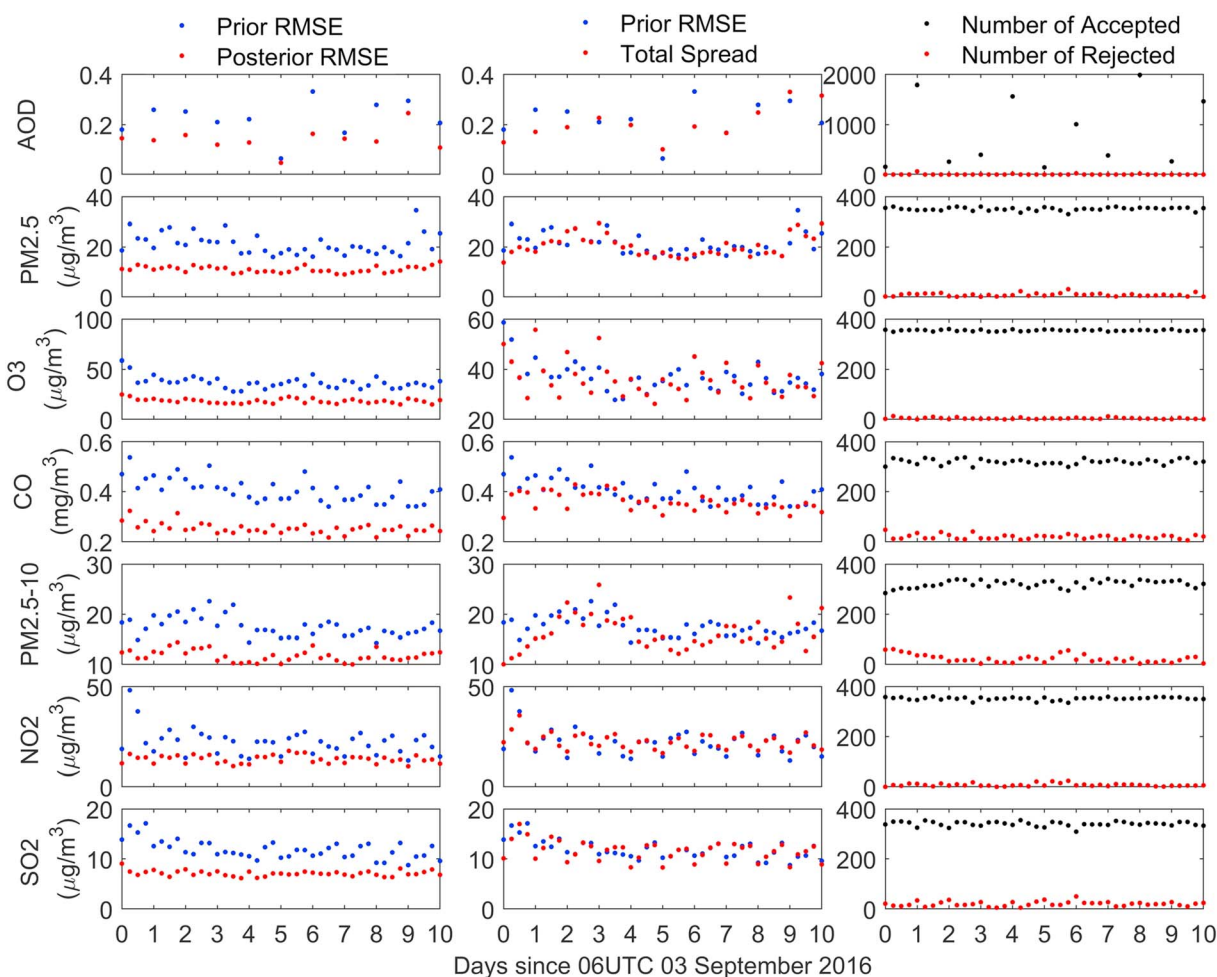
Figure 2. Time series of observation space assimilation diagnostics for Prior and Posterior root mean square error (RMSE) in the left column, Prior RMSE and Total Spread in the middle column, and the Number of Accepted and Rejected Observations in the right column. The different rows show the species-dependent diagnostics.

3.2. Emission Rate Adjustments

3.2.1. Total Regional Emissions

Time evolution plots of the total regional emissions are shown in Figure 3. That figure shows that with time the adjusted emissions converged and that the percentage adjustment was different for different species and regions.

Generally, the total regional emissions converge after 3 days of cycling (~12 assimilation cycles). The amount of time needed for convergence is a function of species and region. It also appears to depend on the observation density. For example, the convergence time for the YRD region is smaller than that for the SP region because the monitoring site density is greater. Similarly, the NO_x emissions (E_NOx) converge more quickly than the VOC emissions (E_VOC) because the former was constrained by O_3 and NO_2 observations while the latter was constrained by O_3 observations only. The CO emissions (E_CO) had the longest convergence time because they were updated once a day rather than four times a day like other species.

To validate the emission update, the bottom-up MEIC inventory for the year 2016 is plotted as dashed lines in Figure 3. Figure 3 also shows that although large uncertainty was assigned to the prior emissions, some of the increments are relatively small so that the posterior emissions are comparable to the prior emissions. In addition, the posterior emissions get closer to the 2016 inventory for SO_2, black carbon, and organic carbon, which means the top-down update approaches the same value as the bottom-up inventory for those species. But for other species, for example, $PM_{2.5}$, NO_x, and VOC in the SP and YRD regions, the posterior emissions
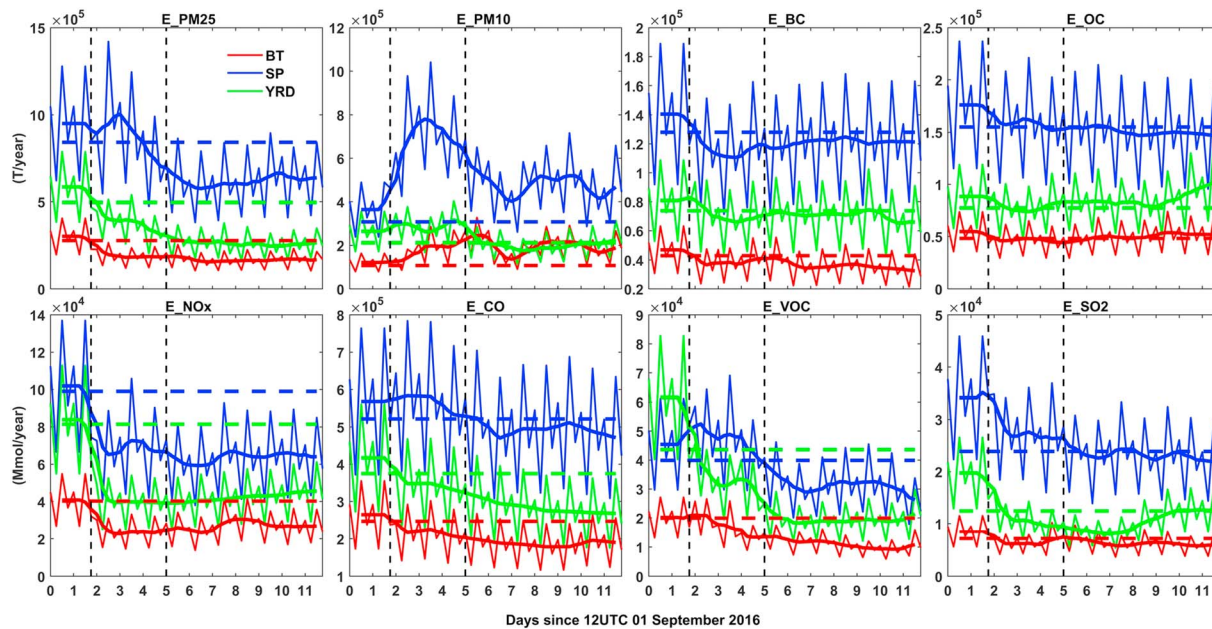
Figure 3. Time series of total regional emissions for the BT (red), SP (blue), and YRD green) regional areas as defined in Figure 1. The posterior emissions are plotted as thin solid lines, and the 1-day moving averages are plotted as thick solid lines. The initial emissions at the left end of the thick solid lines show the total prior MEIC 2015 emissions. The horizontal dashed lines show the moving average of the MEIC 2016 emissions, and the vertical dashed lines identify the first assimilation cycle and the cycle where we started issuing 72-hr forecasts.

were adjusted in the direction resembling the bottom-up MEIC inventories, but the increments are large (e.g., close to 50% of decrease E_VOC for YRD region). Those large increment should not be viewed as overadjustments or errors because (i) the MEIC 2015 and 2016 Inventories, as bottom-up inventories, may contain large uncertainties (E_VOC was assigned ±68% uncertainty in MEIC; Zhang et al., 2009), the large increments are generally within the range of those uncertainties, and (ii) during our study period, the emissions uncertainty is enhanced because China severely restricted mobile and industrial emission sources to improve air quality during the G20 Summit in Hangzhou from 1–6 September 2016 (the actual period of curtailed emissions may have been longer). Those emission reductions would have impacted the YRD and SP regions and may explain the increasing emissions trend that began on 8 September 2016 for SO$_2$ and NO$_x$ in the YRD region. However, the emission adjustments are based on minimizing the error variance between the actual and modeled observations. As such, the emission increments may compensate for model errors by unrealistically adjusting the emissions to reduce that error variance. For example, compared to the prior, the posterior E_PM10 in the SP and BT regions were increased even though they got lower in MEIC 2016. That may be a result of compensating the underestimating of dust whose emissions were simulated according to the land use type and wind speed and were not updated here. The YRD region does not show the same problem because it is in the south where dust has less influence.

3.2.2. Correlation Length Scale

As discussed in Khade et al. (2013), if the data contain long horizontal correlation length scales and long localization length scales are used during assimilation cycling, the correlation length scales will decrease and converge to appropriate length scales. Therefore, in our experiments the emissions correlation length scales should converge to their appropriate values. To understand the change in the emissions correlation length scales during our experiments, we calculate a normalized correlation length scale $\widetilde{L}_s^t$ for species s and cycle t with the following:

At each cycle t, a series of coordinate pairs (x,y) is prepared as $\left(dist\left(i,j,i',j'\right), corr\left(e_{i,j,s}^t, e_{i',j',s}^t\right)\right)$ for each combination of grid points (i,j) and (i',j') where $dist$ is the distance between those points and $corr$ is the ensemble correlation of posterior emissions at those points for species s at time t. At each cycle t, the (x,y) series is then fit to the assumed function $y = \exp\left(-\frac{x^2}{D^2}\right)$ to get the parameter D. The parameter D is then assigned to the
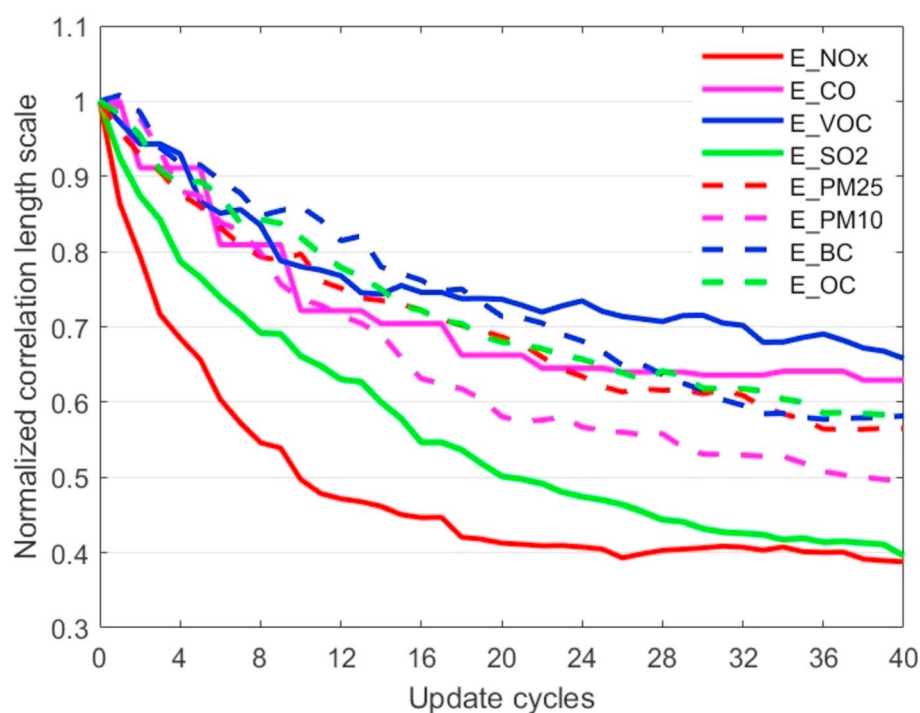
Figure 4. Time series of the normalized horizontal correlation length scale as defined in section 3.2.2.

exponential length scale L_s^t. To address species-dependent initial correlation length-scale differences, we normalize L_s^t by its initial correlation length scale L_s^0 to get $\widetilde{L}_s^t$.

Figure 4 shows time series plots of $\widetilde{L}_s^t$. Like Figure 6 in Khade et al. (2013), the correlation length scale converges with time for each species. The asymptotes are the appropriate correlation lengths. For example, in our results VOC has the largest asymptotic length scale, which is ~0.7 $L_{VOC} \approx$ 420 km and then comes the CO emission (~0.63 $L_{VOC} \approx$ 378 km). The PM25, black carbon, and organic carbon emissions have similar asymptotic length scales ~0.57 $L_s \approx$ 342 km because they are related to fine particulates while PM_{10}, a coarser particulate, has a smaller asymptotic length scale of ~300 km. Similarly, the SO_2 and NO_x emissions have smaller asymptotic length scales of ~240 km. NO_x, VOC, and CO converge the fastest taking ~20 cycles, while the other species take ~30 cycles, which is longer than that shown in Figure 3 for the regional total emissions adjustments. The extra time for convergence of the correlation lengths should be seen as the time needed for WRF-Chem/DART to bring the subregional emissions and concentrations into balance.

3.2.3. Spatial Distribution

Figure 5 shows the spatial distribution of temporal average emissions prior, posterior, and increment for the trace gases, and Figure 6 shows the same for PM. We used only the last eight cycles when calculating those averages because Figures 3 and 4 suggest that the emission adjustments had converged during those cycles.

Comparison of the posterior and the MEIC 2015 inventory in Figure 5 show that (i) the increment is concentrated near urban areas where most of the anthropogenic activities and monitoring sites are located; (ii) emissions are reduced over most of the domain especially around the larger cities like Shanghai and Beijing in eastern China; (iii) emissions are increased near the northern and western boundaries of the domain. The emission reductions near the coastline are reasonable given the emission control policies implemented by most cities in eastern China. The emission increases near the northern and western boundaries may be due to inaccurate trans-boundary pollutant transport. Another explanation is that the effect of the industrial growth exceeds that from emission control policies in those areas, which ends up to increase the emission. This explanation is supported by NOx and CO emissions increase in similar areas reported in the global chemical assimilation/emissions adjustment experiments of Jiang et al. (2017) and Miyazaki et al. (2017). Also, the SO_2 emissions were increased to the east and west of the BT region and the northwest of the YRD region.
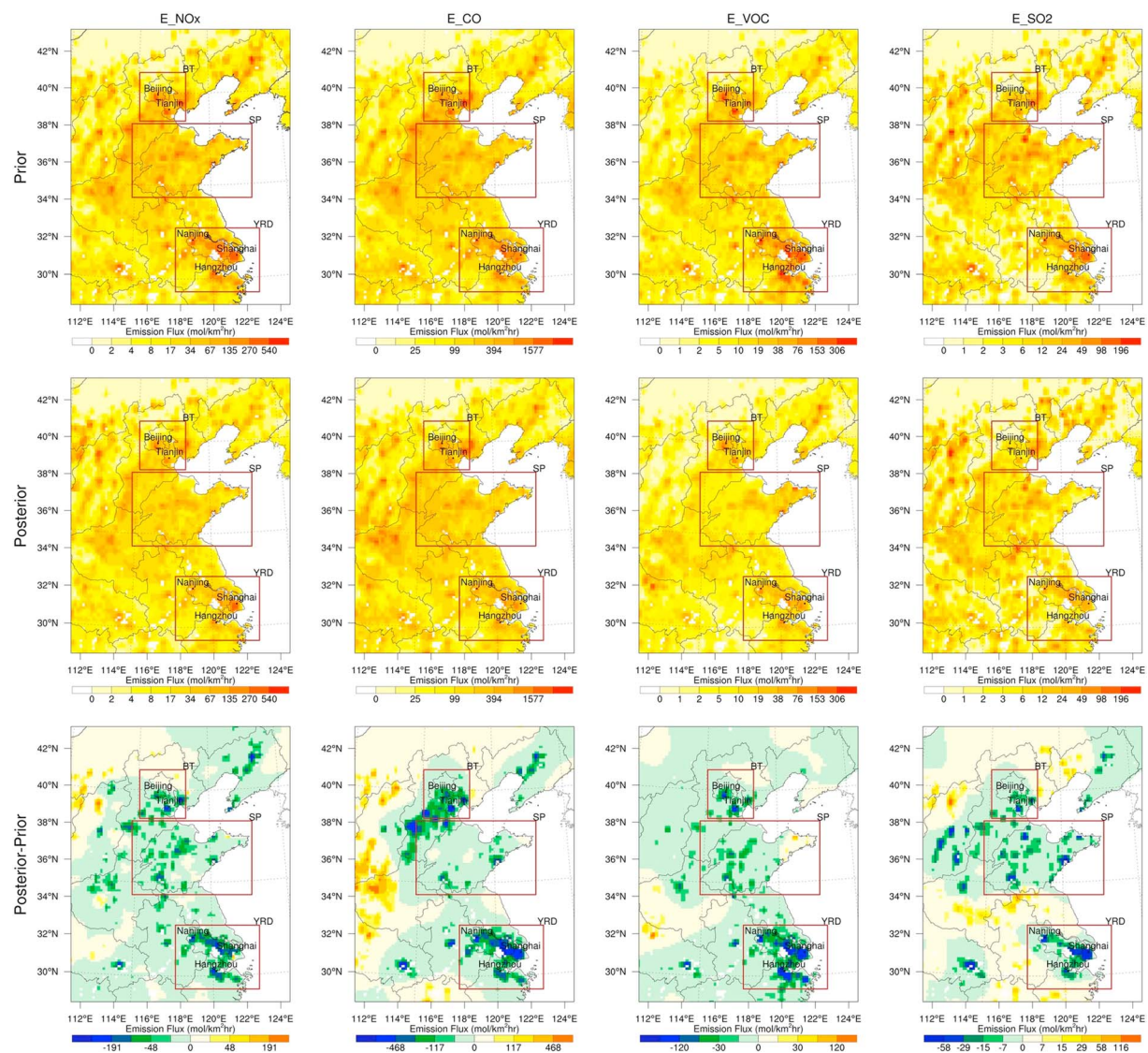
Figure 5. Spatial distribution of the temporal average emissions prior (upper row, representing the MEIC 2015 level), posterior (middle row), and increment (lower row) for the last eight cycles. The different columns are for the different trace gases.

For PM emissions, Figure 6 shows reductions over the entire domain for $PM_{2.5}$, the main component of the fine PM. Those reductions can be attributed to the PM emissions control policy implemented in China since 2013 (Cai et al., 2017). Those reductions are consistent with the reduction of atmospheric PM concentrations found since 2013 (Lin et al., 2018; Luo et al., 2018). In contrast, the PM_{10} emissions, which is a minor component of the coarse PM concentrations, increase for all of China except for areas around Shanghai and Hangzhou. As stated in the end of section 3.2.1, those increased are likely due to the compensation of dust underestimation.

3.3. Impact of Data Assimilation on the 72-hr Forecasts

Next we look at the impact of data assimilation and emission adjustment on 72-hr forecasts. As discussed earlier in section 2.5, we began the 72-hr forecasts at 12 UTC 6 September 2016 after 3 days of assimilation/forecast cycling. The hourly forecasts were verified against the MEP surface observations. We verified only the SO_2, NO_2, $PM_{2.5}$, $PM_{2.5-10}$, O_3, and CO forecasts because those are the six criteria pollutants regulated by the Chinese government. The hourly RMSE is plotted in Figure 7, and the hourly spatial
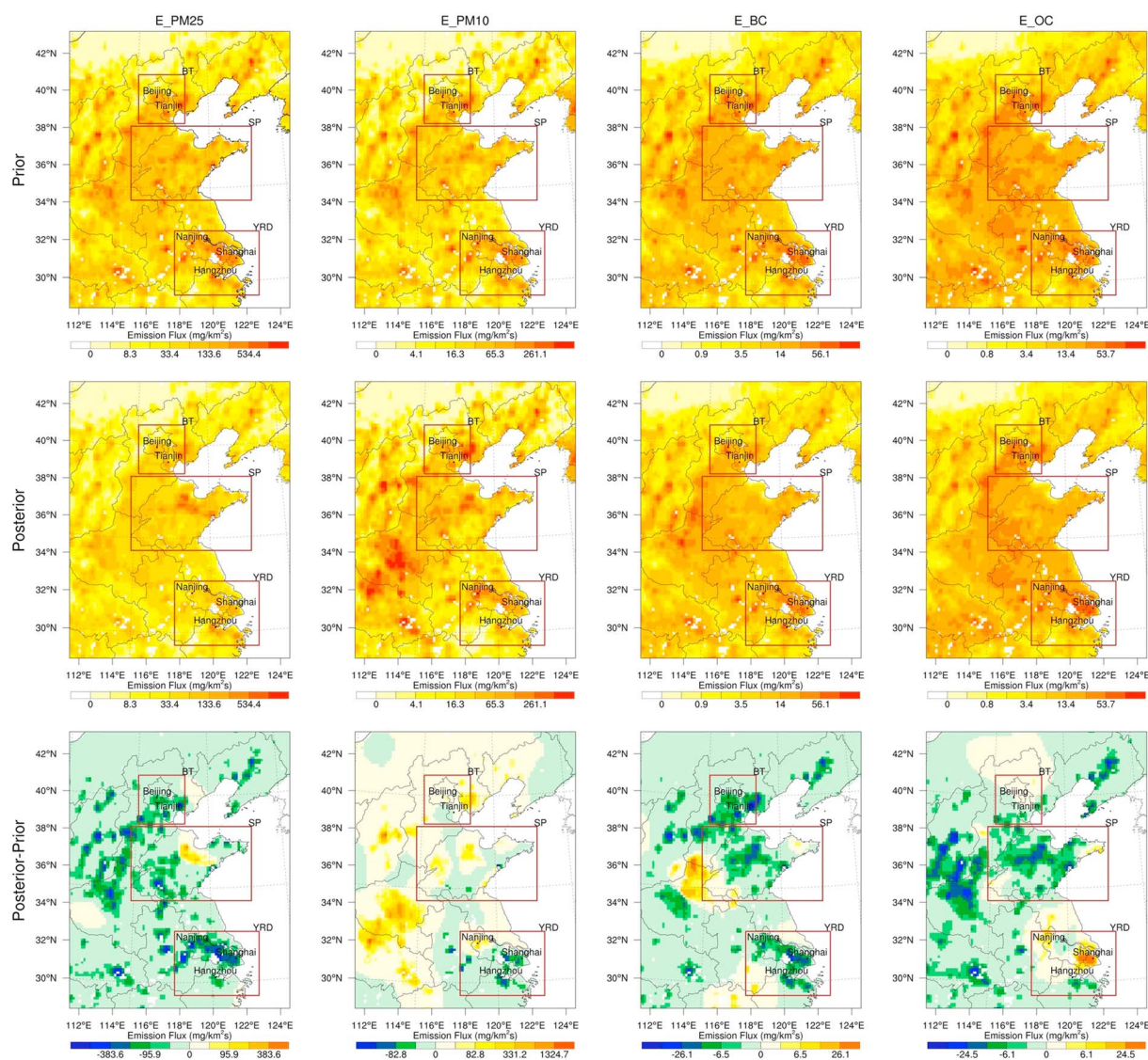
Figure 6. Similar to Figure 5 except for PM.

correlation is plotted in Figure 9. Those figures include error bars that were calculated using the method of Benjamin et al. (2010).

Figure 7 shows that for all species, the RMSE of forecasts with data assimilation is smaller than the NO DA that had no DA influence. Figure 9 shows similar results in that the correlation coefficients from forecasts with data assimilation are higher than those from the NO DA. Due to the improved initial conditions, the forecasts with the help of DA generally perform better than the NO DA forecast for the first 24 hr. After the first 24 to 36 hr, the forecast skill metrics from CICs ONLY converge with those from NO DA. That result is expected because chemical forecasts are less sensitive to correct initial conditions than to correct emissions, and those forecasts share the same emissions. A slight downward (upward) trend of RMSE (correlation) can be observed for NO DA and CICs ONLY, which may come from the rebalancing between CICs and newly introduced initial meteorological conditions after the initiation of each forecast.

The impact of the updated emissions can be seen from the difference between the CICs ONLY and EMISS CICs. Figure 7 shows that the forecasts for all species from the EMISS CICs have improved RMSE for the duration of the forecast period. Figure 9 shows that for the correlation metric, some species show improvement (CO, O_3, SO_2, and $PM_{2.5-10}$) while the remaining species (NO_2 and $PM_{2.5}$) do not show much of an
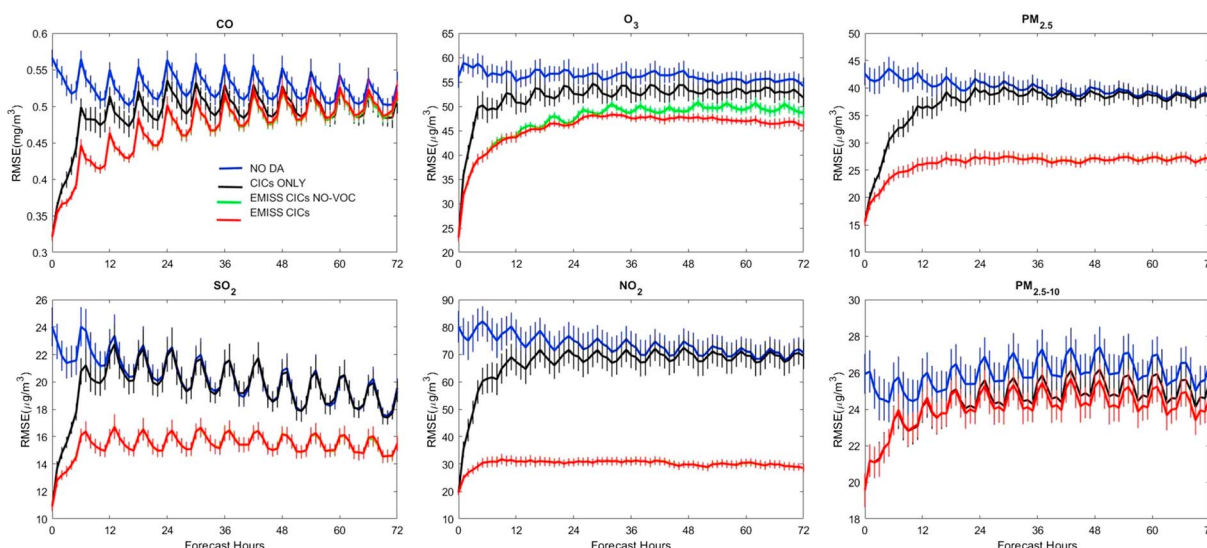
Figure 7. Composite of the 28 72-hr forecast RMSE (mg/m^3 for CO and μg/m^3 for other species) time series for the Chinese criteria pollutants from the four forecasts NO DA, CICs ONLY, EMISS CICs NO-VOC, and EMISS CICs in Base Exp. The forecasts were verified against the MEP surface monitoring data. The error bars represent the uncertainty range.

impact from adjusting the emissions. SO_2 shows the most significant improvement (~25% RMSE reduction and ~50% correlation increase), which is also obtained in Peng et al. (2018). However, the next most significant improvements are for NO_2 and $PM_{2.5}$ (~65% and ~35% RMSE reduction, respectively, but no correlation change), which looks much better than results of Peng et al. (2018). Finally, $PM_{2.5-10}$ shows little improvement because updated emissions contribute only a small part of coarse particles (usually dust contributes the most). The impacts of updating the O_3 and CO are more complicated and discussed below.

Although the emissions adjustments appear to have no impact on the O_3 forecasts in Figure 7, the diurnal variation of RMSE suggests it had a positive impact. Figure 8 shows the diurnal variation of the O_3 RMSE composited from Figure 7. Figure 8 shows that the RMSE reductions from the updated emissions occur during the day when the O_3 concentrations are higher. With lower NOx emission, EMISS CICs could yield more accurate NO_2 forecasts than NO DA and CICs only but may fail to reproduce the low O_3 concentrations during the night because of the low model resolution used here (Wild & Prather, 2006). However, because the overestimated NO_x emissions in the NO DA and CICs Only could help to consume more O_3 during the night, those forecasts are more likely to reproduce the low O_3 concentrations during the night than the EMIS CICs. This explanation is somehow consistent with the different RMSE pattern showed in Figure 8 between urban
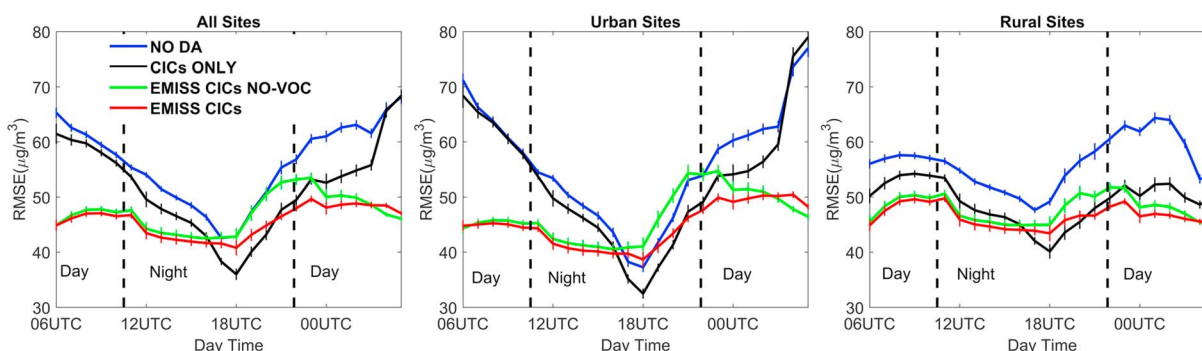

Figure 8. Composite of the 84 24-hr forecasts in Figure 7 showing the diurnal variation of the O_3 root mean square error averaged over all monitor sites (left), urban sites (middle) and rural sites (right).
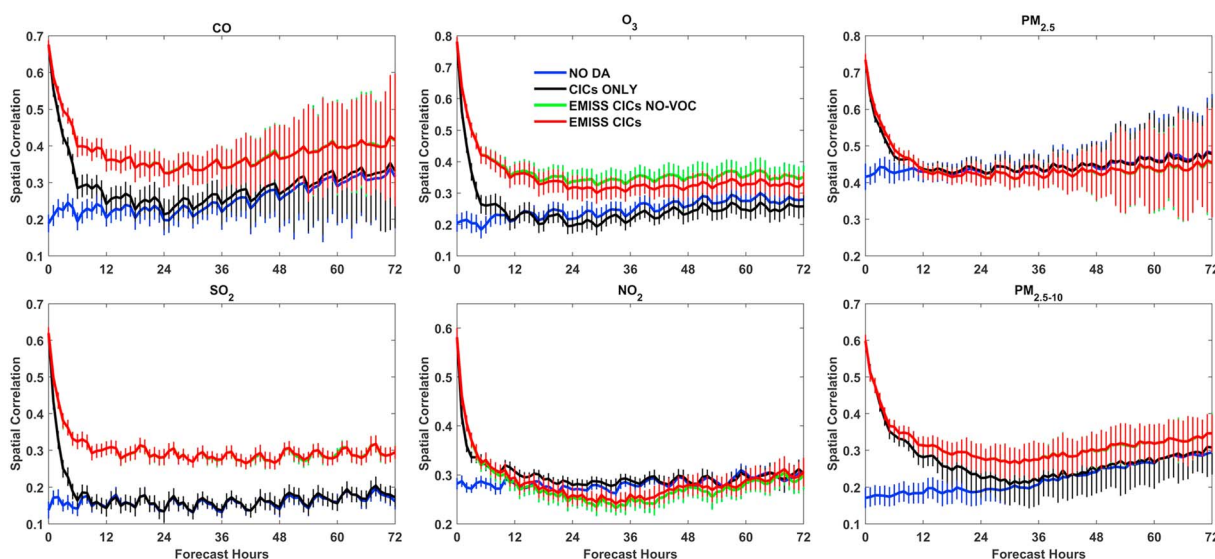
Figure 9. Similar to Figure 7 except for the spatial correlation.

sites and rural sites. At urban sites where NOx emission are high, the RMSE diurnal variation is more significant than that of rural sites. Also, the forecast degradation after emission update during night is less significant in rural sites for less O3 depletion by NOx. In fact, some of our old experiments with 60-km resolution (not shown here) perform even worse during the night, and its overall performance is as bad as Peng et al. (2018). Overall, the emission adjustment is still valuable for O_3 because during the day it significantly reduces the RMSE (~30%) and increases the correlation (~50%).

For CO, the EMISS CICs has improved performance compared to CICs ONLY. That improvement is significant during the first 36 hr where RMSE has ~10% reduction and the correlation has ~40% increase. During the remainder of the 72-hr forecast period, the variability of correlation increases and the RMSE improvement is no longer significant. We speculate that the vanished RMSE improvement is due to transportation of CO. CO is long-lived and has a half-life that is much longer than the other criteria pollutants. Therefore, its concentration emitted at one grid could transport and diffuse to a larger area, causing a more evenly distributed concentration field. This makes the long-term forecast of CO depends more on the meteorological and chemical boundary conditions rather than emission spatial differences introduced by emissions adjustment. Since the CICs Only and EMISS CICs use the same meteorological and chemical boundary conditions, their forecast will gradually become similar after hours of transport and diffusion. Similar results were reported in section S3 of Peters et al. (2007) for CO_2. Their simulated CO_2 pattern is dominated by transport error rather than regional CO_2 flux signal after weeks of model run.

The effect of updating the VOC emissions can be seen by examining differences between the EMISS CICs and EMISS CICs NO-VOC. If the updated VOC emissions are improved relative to their prior, then the EMISS CICs O_3 forecast should be better than that from the EMISS CICs NO-VOC. Figure 7 shows a reduction in O_3 forecast RMSE from the EMISS CICs compared to the EMISS CICs NO-VOC, but the reduction is not significant for some times and the correlation is degraded slightly in Figure 9. Figure 8 shows the RMSE improvement is most prominent near sunrise. That suggests the VOC emission adjustment is positively impacting the O_3 forecasts. The other species show little or no sensitivity to adjusting the VOC emissions.

4. Sensitivity Experiments

4.1. Results From the CO 6-hr Exp

Figure 10 presents a comparison of forecast skill from forecasts issued by the Base Exp with that by the CO 6-hr Exp discussed in the previous paragraph. Generally, the results in Figure 10 are similar to those in Figures 7 and 9. Although performance of the CO experiment is similar to that of the base experiment in CICs ONLY forecasts, performance of EMISS CICs forecasts from the CO 6-hr experiment is consistently
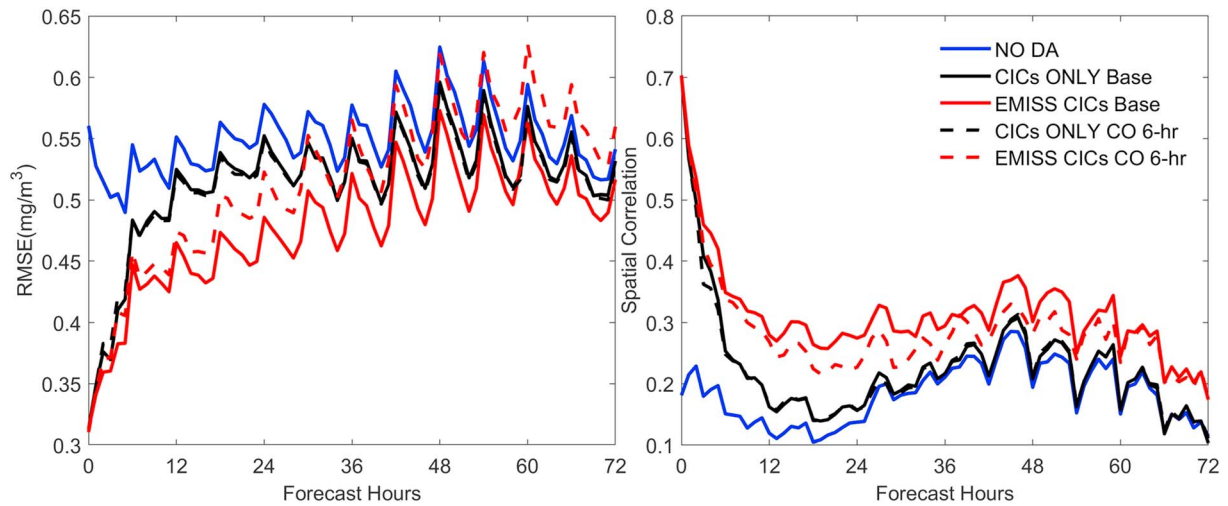
Figure 10. Composite 72-hr CO forecast root mean square error (RMSE; mg/m^3; left) and spatial correlation (right) similar to that in Figures 7 and 9 expect for the Base and CO 6-hr Exp and an abbreviated study period.

worse than its counterpart from the base experiment (about 10% increase of RMSE and 0.05 decrease of correlation). More importantly, forecast RMSE from EMISS CICs CO 6-hr sometimes gets even worse than that from CICs ONLY CO 6 hr, which suggests CO emissions updated in the CO 6-hr Exp may even degrade the CO forecast.

In the -EnKF, observed variables update the state variables through their correlations in the prior ensemble covariance (Jeffrey L Anderson, 2003). The average local ensemble correlation is defined as $\frac{1}{N}\sum\limits_{i,j} corr$

$\left(c_{i,j,s}^t, e_{i,j,s}^t \right)$ where $c_{i,j,s}^t$ and $e_{i,j,s}^t$ are the prior (posterior) concentrations and prior (posterior) emissions at grid
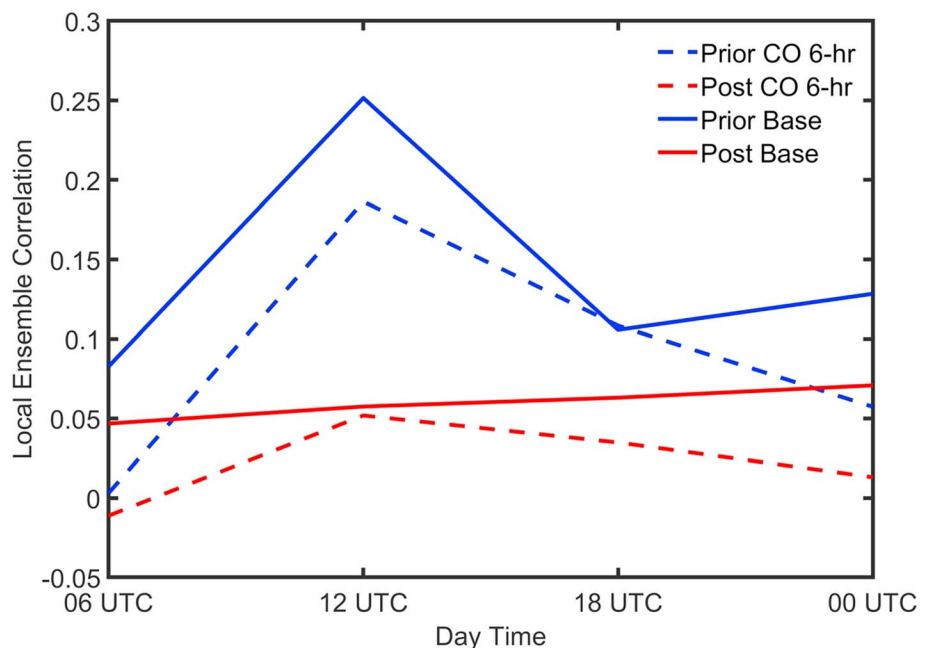

Figure 11. Composite of the local ensemble correlation between surface CO concentration and CO emission at different day time for prior ensemble (blue) and posterior ensemble (red) from Base Exp (solid line) and CO 6-hr Exp (dashed line). Each value is an average from DA cycles of the same day time.
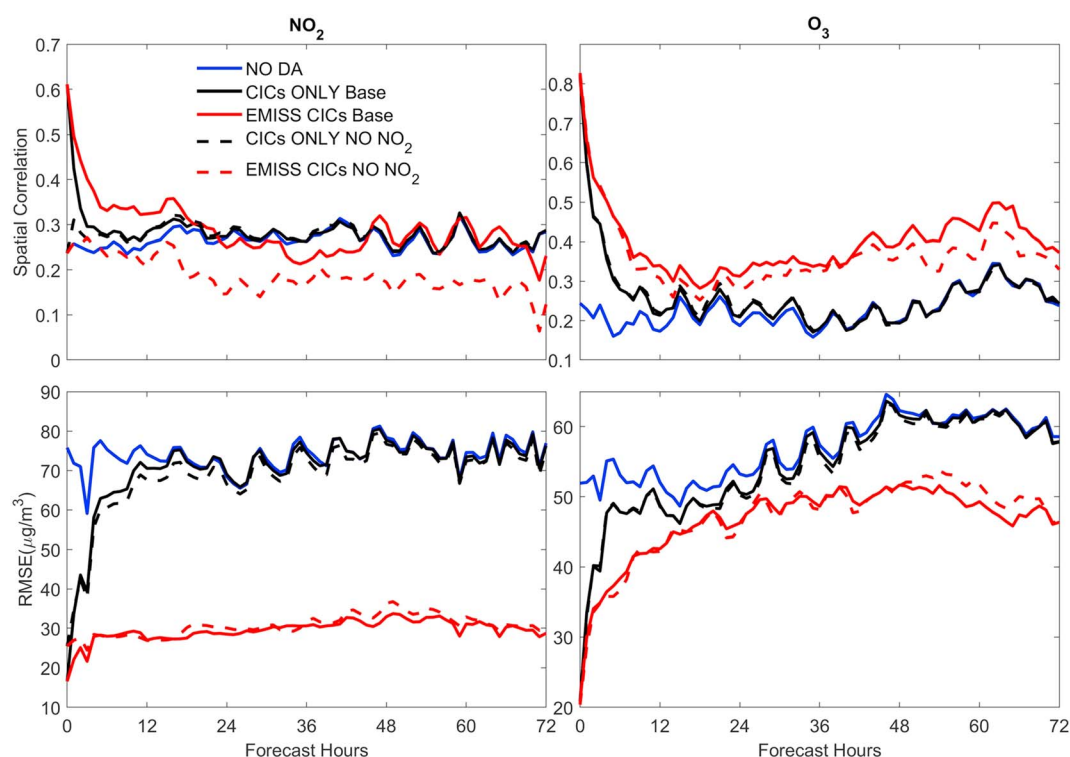
Figure 12. Composite 72-hr NO$_2$ and O$_3$ forecast RMSE (μg/m^3) and spatial correlations for the Base and NO NO$_2$ Exps and an abbreviated study period similar to Figure 10.

point (*i,j*) for update cycle *t* and species *s* and *N* is the number of surface grid points. Figure 11 shows the diurnal variation of the average prior and posterior local ensemble correlations from the base and modified experiments. Figure 11 shows three interesting results: (i) The prior correlations show a diurnal variation with the greatest correlation at 12 UTC (the time used for updating the emissions in the base experiment); (ii) the posterior correlation is less than the prior; and (iii) the prior correlation of the base experiment is higher than that of the modified experiment.

The last two results provide a possible explanation for why the modified experiment had degraded performance for the CO forecasts. According to result 2, signals between CO emission and concentration will inevitably get weaker after one DA cycle. To restore the signal of CO—a long lifetime atmospheric constituent whose emission rate is relatively small compared with its concentration—the model may need longer than 6 hr to integrate enough emission information into the concentration ensemble. Therefore, an update interval of 6 hr, which works for other shorter lifetime species, does not work well for CO. However, as discussed in section 3.3, too long an update interval may also cause trouble because signals about emission differences could diffuse. The emission update interval may be an important factor for EnKF emission inversion studies, at least for CO.

4.2. Results From NO NO$_2$ Exp

The composite 72-hr NO$_2$ and O$_3$ forecast skill metrics for the Base and NO NO2 Exps are plotted in Figure 12. Generally, assimilation of the O$_3$ observations may be able to remove most of the bias in the NO$_x$ emissions without the help of NO$_2$ observations. That could be seen by similar forecast RMSE scores of the EMISS CICs Base and EMISS CICs NO NO$_2$. However, the EMISS CICs NO NO$_2$ degraded correlations for the NO$_2$ forecasts (a decrease of more than 0.1). Tang et al. (2016) found a similar degradation in skill for the NO$_x$ forecast and attributed it to spurious NO$_x$ emission updates caused by the nonlinearity of the NO$_x$-VOC-O$_3$ system.

Figure 13 shows maps of the NO$_x$ emission differences for the Base and NO NO$_2$ Exps. The emission increments are similar for Base and NO NO$_2$ Exps. Elbern et al. (2007) reported similar results that DA
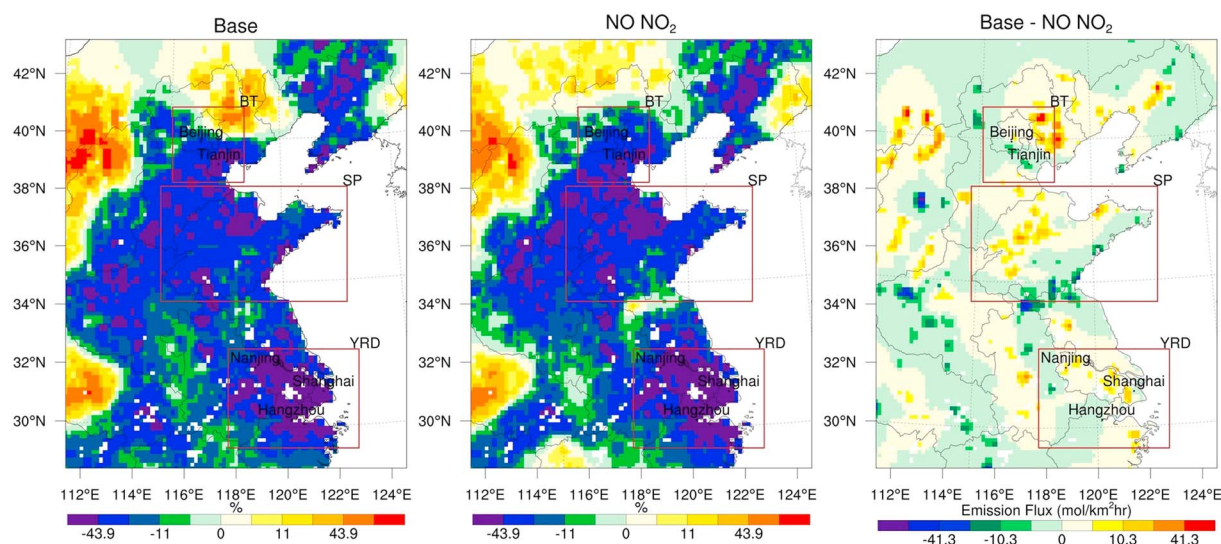
Figure 13. Average NO$_x$ emission increments from the last four cycles for the Base and NO NO$_2$ Exps. The right panel shows the absolute difference between posterior NOx emissions from Base and NO NO$_2$.

improvements can rest on assimilating ozone observations only and the inclusion of NOx observations yields little change with their coarse model resolution. The similarity between our resultant emissions might be expected due to the similarity between our NO$_2$ forecast RMSE. However, when reviewed in greater detail, our increments have different signs and magnitudes (see the northeastern part of the BT region and the southern edge of the SP region). Normalized correlation length scales were calculated according to section 3.2.2 for the NOx emissions in the modified experiment, the value changed from 0.40 to 0.62, which means NOx emissions were more spatially correlated when the NO2 observations were not assimilated.

The nonlinearity of the chemical system may explain the sign differences, but it cannot explain the larger normalized correlation length scale in the NO$_x$ emissions for the NO NO$_2$ Exp. The larger correlations suggest that some of the NOx emissions fine structure is not resolved by assimilating only O$_3$ observations. That failure is possible because the spatial correlation length scale of O$_3$ concentration is usually larger than that of NO$_2$ (in an ensemble sense). In other words, O$_3$ is usually more uniform in space than NO$_2$ due to its longer lifetime and therefore may not contain the information on the NOx emissions fine structure. The difference between 0.40 and 0.62 also indicates that the "appropriate" length scale as described in sec. 3.2.2 is determined not only by the real emission error distribution but also by the correlation length scales of related observations.

5. Conclusions

With the help of WRF-Chem/DART, an EAKF was used to assimilate surface monitor observations of SO$_2$, NO$_2$, O$_3$, CO, PM$_{2.5}$, and PM$_{10}$ as well as MODIS AOD. Related background concentration fields were updated along with the anthropogenic emission inventory of NOx, SO$_2$, CO, VOC, and four components of PM. Independent variable localization was adopted to avoid potential spurious correlation across different species. Posterior CICs and emissions were examined by initiating 72-hr forecasts after each DA cycle and validating the forecasts against in situ observations. Results suggest that WRF-Chem/DART is a powerful tool to adjust anthropogenic emissions and improve air quality forecasts over the eastern part of China.

After a series of DA cycles, the posterior emissions could converge to a relatively stable state from the view of either regional total emissions or horizontal ensemble correlation length scale. Compared to the bottom-up original MEIC2015 and MEIC2016 inventory, the posterior inventory obtained from the base configuration looks reasonable for every emitted species from perspectives of regional total and spatial distribution except anthropogenic coarse-mode aerosol.

With the help of updated anthropogenic emissions, model forecasts are improved for each species verified. Significant decreases of RMSE (about 25%, 57%, and 35%, respectively) are seen for SO_2, NO_2, and $PM_{2.5}$ for the whole forecast period. The forecast RMSE of O_3 is significantly decreased (about 30%) during the day-time and CO enjoys RMSE reduction (about 10%) during the first 36 hr of forecasts. The RMSE of $PM_{2.5-10}$ is only slightly reduced due to the limited anthropogenic contribution to it. Also, the spatial correlation of forecasts has an increase of around 0.1 for all species except NO_2 for the whole forecast period. The results are totally different and much better than those obtained in a previous work by Peng et al. (2018).

CO emission was updated every 24 hr at 12 UTC in the Base Exp because the experiment with a 6-hr update interval shows increased forecast error (about 10% increase of RMSE and 0.05 decrease of correlation) and decreased emission-concentration signal. The results also indicate the update interval is an important factor to take into account when estimating emissions of species with long lifetime. Other update intervals besides 6 and 24 hr may yield even better results and will be tested in future work.

An experiment without assimilating NO_2 observations yields similar forecast RMSE and spatial correlation improvement for O_3 as well as NO_2 forecast RMSE decrease. However, emissions adjusted without NO_2 observations degrade the forecast spatial correlation of NO_2 (a decrease of more than 0.1). The result indicates (1) O_3 is valuable during NOx emission inversion for its cross-species relationship with NO_2; (2) O_3 alone is not enough because of nonlinear effects within the O_3-NOx-VOC system and a large decorrelation length scale of O_3 concentration. The experiment also highlights the importance of multi-constituent data assimilation. More experiments are needed in the future to figure out the potential benefits of other cross-variable influences during DA.

Also, by comparing our experiments with Peng et al. (2018), further increasing model resolution is worth a try and observations of VOC should be assimilated if available. Considering DA experiments were carried out for only 10 days with only 20 members, applying more DA cycles and increasing the ensemble members may yield results with better representativeness.

In the end, it must be noted that the posterior emission inventory obtained from DA is only an inventory that makes the model yield forecasts closer to the assimilated observations. Therefore, it contains the error of the model and does not necessarily represent the real emission scenario. Readers should be careful when interpreting the emissions presented here.

References

Anderson, J., Hoar, T., Raeder, K., Liu, H., Collins, N., Torn, R., & Avellano, A. (2009). The data assimilation research testbed: A community facility. *Bulletin of the American Meteorological Society*, *90*(9), 1283–1296. https://doi.org/10.1175/2009BAMS2618.1

Anderson, J. L. (2001). An ensemble adjustment Kalman filter for data assimilation. *Monthly Weather Review*, *129*(12), 2884–2903. https://doi.org/10.1175/1520-0493(2001)129<2884:aeakff>2.0.co;2

Anderson, J. L. (2003). A local least squares framework for ensemble filtering. *Monthly Weather Review*, *131*(4), 634–642. https://doi.org/10.1175/1520-0493(2003)131<0634:ALLSFF>2.0.CO;2

Anderson, J. L. (2007). An adaptive covariance inflation error correction algorithm for ensemble filters. *Tellus Series A-Dynamic Meteorology and Oceanography*, *59*(2), 210–224. https://doi.org/10.1111/j.1600-0870.2006.00216.x

Arellano, A. F., Kasibhatla, P. S., Giglio, L., van der Werf, G. R., Randerson, J. T., & Collatz, G. J. (2006). Time-dependent inversion estimates of global biomass-burning CO emissions using Measurement of Pollution in the Troposphere (MOPITT) measurements. *Journal of Geophysical Research*, *111*, D09303. https://doi.org/10.1029/2005JD006613

Barbu, A. L., Segers, A. J., Schaap, M., Heemink, A. W., & Builtjes, P. J. H. (2009). A multi-component data assimilation experiment directed to sulphur dioxide and sulphate over Europe. *Atmospheric Environment*, *43*(9), 1622–1631. https://doi.org/10.1016/j.atmosenv.2008.12.005

Barré, J., Gaubert, B., Arellano, A. F., Worden, H. M., Edwards, D. P., Deeter, M. N., et al. (2015). Assessing the impacts of assimilating IASI and MOPITT CO retrievals using CESM-CAM-chem and DART. *Journal of Geophysical Research: Atmospheres*, *120*, 10,501–10,529. https://doi.org/10.1002/2015JD023467

Benjamin, S. G., Jamison, B. D., Moninger, W. R., Sahm, S. R., Schwartz, B. E., & Schlatter, T. W. (2010). Relative short-range forecast impact from aircraft, profiler, radiosonde, VAD, GPS-PW, METAR, and mesonet observations via the RUC hourly assimilation cycle. *Monthly Weather Review*, *138*(4), 1319–1343. https://doi.org/10.1175/2009mwr3097.1

Boersma, K. F., Jacob, D. J., Eskes, H. J., Pinder, R. W., Wang, J., & van der A, R. J. (2008). Intercomparison of SCIAMACHY and OMI tropospheric NO2 columns: Observing the diurnal evolution of chemistry and emissions from space. *Journal of Geophysical Research*, *113*, D16S26. https://doi.org/10.1029/2007JD008816

Boynard, A., Beekmann, M., Foret, G., Ung, A., Szopa, S., Schmechtig, C., & Coman, A. (2011). An ensemble assessment of regional ozone model uncertainty with an explicit error representation. *Atmospheric Environment*, *45*(3), 784–793. https://doi.org/10.1016/j.atmosenv.2010.08.006

Cai, S., Wang, Y., Zhao, B., Wang, S., Xing, C., & Hao, J. (2017). The impact of the "Air Pollution Prevention and Control Action Plan" on PM 2.5 concentrations in Jing-Jin-Ji region during 2012–2020. *Science of the Total Environment*, *580*, 197–209. https://doi.org/10.1016/j.scitotenv.2016.11.188

Colarco, P., Silva, A. D., Chin, M., & Diehl, T. (2010). Online simulations of global aerosol distributions in the NASA GEOS-4 model and comparisons to satellite and ground-based aerosol optical depth. *Journal of Geophysical Research*, *115*, D14207. https://doi.org/10.1029/2009JD012820

Elbern, H., & Schmidt, H. (2001). Ozone episode analysis by four-dimensional variational chemistry data assimilation. *Journal of Geophysical Research*, *106*(D4), 3569–3590. https://doi.org/10.1029/2000JD900448

Elbern, H., Strunk, A., Schmidt, H., & Talagrand, O. (2007). Emission rate and chemical state estimation by 4-dimensional variational inversion. *Atmospheric Chemistry and Physics*, *7*(14), 3749–3769. https://doi.org/10.5194/acp-7-3749-2007

Emmons, L. K., Walters, S., Hess, P. G., Lamarque, J. F., Pfister, G. G., Fillmore, D., et al. (2010). Description and evaluation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4). *Geoscientific Model Development*, *3*(1), 43–67. https://doi.org/10.5194/gmd-3-43-2010

Evensen, G. (1994). Sequential data assimilation with a nonlinear quasi-geostrophic model using Monte-Carlo methods to forecast error statistics. *Journal of Geophysical Research*, *99*(C5), 10143–10162. https://doi.org/10.1029/94JC00572

Evensen, G. (2003). The Ensemble Kalman Filter: theoretical formulation and practical implementation. *Ocean Dynamics*, *53*(4), 343–367. https://doi.org/10.1007/s10236-003-0036-9

Grell, G. A., Peckham, S. E., Schmitz, R., McKeen, S. A., Frost, G., Skamarock, W. C., & Eder, B. (2005). Fully coupled "online" chemistry within the WRF model. *Atmospheric Environment*, *39*(37), 6957–6975. https://doi.org/10.1016/j.atmosenv.2005.04.027

Guenther, A., Jiang, X., Heald, C., Sakulyanontvittaya, T., Duhl, T., Emmons, L., & Wang, X. (2012). The Model of Emissions of Gases and Aerosols from Nature version 2.1 (MEGAN2. 1): An extended and updated framework for modeling biogenic emissions. *Geoscientific Model Development*, *5*, 1471–1492. https://doi.org/10.5194/gmd-5-1471-2012

Hanna, S. R., Chang, J. C., & Fernau, M. E. (1998). Monte Carlo estimates of uncertainties in predictions by a photochemical grid model (UAM-IV) due to uncertainties in input variables. *Atmospheric Environment*, *32*(21), 3619–3628. https://doi.org/10.1016/S1352-2310(97)00419-6

Hanna, S. R., Lu, Z., Frey, H. C., Wheeler, N., Vukovich, J., Arunachalam, S., et al. (2001). Uncertainties in predicted ozone concentrations due to input uncertainties for the UAM-V photochemical grid model applied to the July 1995 OTAG domain. *Atmospheric Environment*, *35*(5), 891–903. https://doi.org/10.1016/S1352-2310(00)00367-8

Hyer, E. J., Reid, J. S., & Zhang, J. (2011). An over-land aerosol optical depth data set for data assimilation by filtering, correction, and aggregation of MODIS Collection 5 optical depth retrievals. *Atmospheric Measurement Techniques*, *4*(3), 379–408. https://doi.org/10.5194/amt-4-379-2011

Jiang, Z., Worden, J. R., Worden, H., Deeter, M., Jones, D. B. A., Arellano, A. F., & Henze, D. K. (2017). A 15-year record of CO emissions constrained by MOPITT CO observations. *Atmospheric Chemistry and Physics*, *17*(7), 4565–4583. https://doi.org/10.5194/acp-17-4565-2017

Khade, V. M., Hansen, J. A., Reid, J. S., & Westphal, D. L. (2013). Ensemble filter based estimation of spatially distributed parameters in a mesoscale dust model: experiments with simulated and real data. *Atmospheric Chemistry and Physics*, *13*(6), 3481–3500. https://doi.org/10.5194/acp-13-3481-2013

Kotsuki, S., Ota, Y., & Miyoshi, T. (2017). Adaptive covariance relaxation methods for ensemble data assimilation: experiments in the real atmosphere. *Quarterly Journal of the Royal Meteorological Society*, *143*(705), 2001–2015. https://doi.org/10.1002/qj.3060

Kotsuki, S., Terasaki, K., Yashiro, H., Tomita, H., Satoh, M., & Miyoshi, T. (2018). Online model parameter estimation with ensemble data assimilation in the real global atmosphere: A Case With the Nonhydrostatic Icosahedral Atmospheric Model (NICAM) and the Global Satellite Mapping of Precipitation data. *Journal of Geophysical Research: -Atmospheres*, *123*, 7375–7392. https://doi.org/10.1029/2017JD028092

Kurokawa, J., Yumimoto, K., Uno, I., & Ohara, T. (2009). Adjoint inverse modeling of NOx emissions over eastern China using satellite observations of NO2 vertical column densities. *Atmospheric Environment*, *43*(11), 1878–1887. https://doi.org/10.1016/j.atmosenv.2008.12.030

Kwon, Y., Yang, Z. L., Zhao, L., Hoar, T. J., Toure, A. M., & Rodell, M. (2016). Estimating snow water storage in North America using CLM4, DART, and snow radiance data assimilation. *Journal of Hydrometeorology*, *17*(11), 2853–2874. https://doi.org/10.1175/JHM-D-16-0028.1

Levy, R. C., Mattoo, S., Munchak, L. A., Remer, L. A., Sayer, A. M., Patadia, F., & Hsu, N. C. (2013). The Collection 6 MODIS aerosol products over land and ocean. *Atmospheric Measurement Techniques*, *6*(11), 2989–3034. https://doi.org/10.5194/amt-6-2989-2013

Li, Z., Zang, Z., Li, Q. B., Chao, Y., Chen, D., Ye, Z., et al. (2013). A three-dimensional variational data assimilation system for multiple aerosol species with WRF/Chem and an application to PM2.5 prediction. *Atmospheric Chemistry and Physics*, *13*(8), 4265–4278. https://doi.org/10.5194/acp-13-4265-2013

Lin, C. Q., Li, Y., Lau, A. K. H., Li, C. C., & Fung, J. C. H. (2018). 15-year PM2.5 trends in the Pearl River Delta Region and Hong Kong from satellite observation. *Aerosol and Air Quality Research*, *18*(9), 2355–2362. https://doi.org/10.4209/aaqr.2017.11.0437

Liu, X., Mizzi, A. P., Anderson, J. L., Fung, I. Y., & Cohen, R. C. (2017). Assimilation of satellite NO2 observations at high spatial resolution using OSSEs. *Atmospheric Chemistry and Physics*, *17*(11), 7067–7081. https://doi.org/10.5194/acp-17-7067-2017

Luo, Y., Deng, Q.-F., Yang, K., Yang, Y., Shang, C.-X., & Yu, Z.-Y. (2018). Spatial-temporal change evolution of PM2.5 in typical regions of China in recent 20 years. *Huan jing ke xue= Huanjing kexue*, *39*(7), 3003–3013. https://doi.org/10.13227/j.hjkx.201709174

Mao, Y., Li, Q., Chen, D., Zhang, L., Hao, W.-M., & Liou, K.-N. (2014). Top-down estimates of biomass burning emissions of black carbon in the Western United States. *Atmospheric Chemistry and Physics*, *14*(14), 7195–7211. https://doi.org/10.5194/acp-14-7195-2014

Menard, R., & Changs, L. P. (2000). Assimilation of stratospheric chemical tracer observations using a Kalman filter. Part II: Chi(2)-validated results and analysis of variance and correlation dynamics. *Monthly Weather Review*, *128*(8), 2672–2686. https://doi.org/10.1175/1520-0493(2000)128<2672:aoscto>2.0.co;2

Miyazaki, K., Eskes, H., Sudo, K., Boersma, K. F., Bowman, K., & Kanaya, Y. (2017). Decadal changes in global surface NOx emissions from multi-constituent satellite data assimilation. *Atmospheric Chemistry and Physics*, *17*(2), 807–837. https://doi.org/10.5194/acp-17-807-2017

Miyazaki, K., Eskes, H. J., Sudo, K., Takigawa, M., van Weele, M., & Boersma, K. F. (2012). Simultaneous assimilation of satellite NO2, O-3, CO, and HNO3 data for the analysis of tropospheric chemical composition and emissions. *Atmospheric Chemistry and Physics*, *12*(20), 9545–9579. https://doi.org/10.5194/acp-12-9545-2012

Mizzi, A. P., Arellano, A. F., Edwards, D. P., Anderson, J. L., & Pfister, G. G. (2016). Assimilating compact phase space retrievals of atmospheric composition with WRF-Chem/DART: A regional chemical transport/ensemble Kalman filter data assimilation system. *Geoscientific Model Development*, *9*(3), 965–978. https://doi.org/10.5194/gmd-9-965-2016

Mizzi, A. P., Edwards, D. P., & Anderson, J. L. (2018). Assimilating compact phase space retrievals (CPSRs): comparison with independent observations (MOZAIC in situ and IASI retrievals) and extension to assimilation of truncated retrieval profiles. *Geoscientific Model Development*, *11*(9), 3727–3745. https://doi.org/10.5194/gmd-11-3727-2018

Peng, Z., Lei, L., Liu, Z., Su, J., Ding, A., Ban, J., et al. (2018). The impact of multi-species surface chemical observation assimilation on air quality forecasts in China. *Atmospheric Chemistry and Physics*, *18*(23), 17387–17404. https://doi.org/10.5194/acp-18-17387-2018

Peng, Z., Liu, Z., Chen, D., & Ban, J. (2017). Improving PM2.5 forecast over China by the joint adjustment of initial conditions and source emissions with an ensemble Kalman filter. *Atmospheric Chemistry and Physics*, *17*(7), 4837–4855. https://doi.org/10.5194/acp-17-4837-2017

Peters, W., Jacobson, A. R., Sweeney, C., Andrews, A. E., Conway, T. J., Masarie, K., et al. (2007). An atmospheric perspective on North American carbon dioxide exchange: CarbonTracker. *Proceedings of the National Academy of Sciences of the United States of America*, *104*(48), 18,925–18,930. https://doi.org/10.1073/pnas.0708986104

Rubin, J. I., Reid, J. S., Hansen, J. A., Anderson, J. L., Holben, B. N., Xian, P., et al. (2017). Assimilation of AERONET and MODIS AOT observations using variational and ensemble data assimilation methods and its impact on aerosol forecasting skill. *Journal of Geophysical Research: Atmospheres*, *122*, 4967–4992. https://doi.org/10.1002/2016JD026067

Sayer, A. M., Munchak, L. A., Hsu, N. C., Levy, R. C., Bettenhausen, C., & Jeong, M. J. (2014). MODIS Collection 6 aerosol products: Comparison between Aqua's e-Deep Blue, Dark Target, and "merged" data sets, and usage recommendations. *Journal of Geophysical Research: Atmospheres*, *119*, 13,965–13,989. https://doi.org/10.1002/2014JD022453

Schaap, M., Timmermans, R. M. A., Roemer, M., Boersen, G. A. C., Builtjes, P. J. H., Sauter, F. J., et al. (2008). The LOTOS-EUROS model: Description, validation and latest developments. *International Journal of Environment and Pollution*, *32*(2), 270–290. https://doi.org/10.1504/ijep.2008.017106

Schmidt, H., & Martin, D. (2003). Adjoint sensitivity of episodic ozone in the Paris area to emissions on the continental scale. *Journal of Geophysical Research*, *108*(D17), 8561. https://doi.org/10.1029/2001JD001583

Schwartz, C. S., Liu, Z. Q., Lin, H. C., & McKeen, S. A. (2012). Simultaneous three-dimensional variational assimilation of surface fine particulate matter and MODIS aerosol optical depth. *Journal of Geophysical Research*, *117*, D13202. https://doi.org/10.1029/2011JD017383

Sekiyama, T. T., Tanaka, T. Y., Shimizu, A., & Miyoshi, T. (2010). Data assimilation of CALIPSO aerosol observations. *Atmospheric Chemistry and Physics*, *10*(1), 39–49. https://doi.org/10.5194/acp-10-39-2010

Stockwell, W. R., Kirchner, F., Kuhn, M., & Seefeld, S. (1997). A new mechanism for regional atmospheric chemistry modeling. *Journal of Geophysical Research*, *102*(D22), 25,847–25,879. https://doi.org/10.1029/97JD00849

Tang, X., Zhu, J., Wang, Z. F., & Gbaguidi, A. (2011). Improvement of ozone forecast over Beijing based on ensemble Kalman filter with simultaneous adjustment of initial conditions and emissions. *Atmospheric Chemistry and Physics*, *11*(24), 12,901–12,916. https://doi.org/10.5194/acp-11-12901-2011

Tang, X., Zhu, J., Wang, Z. F., Gbaguidi, A., Lin, C. Y., Xin, J. Y., et al. (2016). Limitations of ozone data assimilation with adjustment of NOx emissions: Mixed effects on NO2 forecasts over Beijing and surrounding areas. *Atmospheric Chemistry and Physics*, *16*(10), 6395–6405. https://doi.org/10.5194/acp-16-6395-2016

van Loon, M., Vautard, R., Schaap, M., Bergstrom, R., Bessagnet, B., Brandt, J., et al. (2007). Evaluation of long-term ozone simulations from seven regional air quality models and their ensemble. *Atmospheric Environment*, *41*(10), 2083–2097. https://doi.org/10.1016/j.atmosenv.2006.10.073

Vira, J., & Sofiev, M. (2012). On variational data assimilation for estimating the model initial conditions and emission fluxes for short-term forecasting of SOx concentrations. *Atmospheric Environment*, *46*, 318–328. https://doi.org/10.1016/j.atmosenv.2011.09.066

Whitaker, J. S., & Hamill, T. M. (2002). Ensemble data assimilation without perturbed observations. *Monthly Weather Review*, *130*(7), 1913–1924. https://doi.org/10.1175/1520-0493(2002)130<1913:EDAWPO>2.0.CO;2

Wild, O., & Prather, M. J. (2006). Global tropospheric ozone modeling: Quantifying errors due to grid resolution. *Journal of Geophysical Research*, *111*, D11305. https://doi.org/10.1029/2005JD006605

Yumimoto, K., & Uno, I. (2006). Adjoint inverse modeling of CO emissions over Eastern Asia using four-dimensional variational data assimilation. *Atmospheric Environment*, *40*(35), 6836–6845. https://doi.org/10.1016/j.atmosenv.2006.05.042

Yumimoto, K., Uno, I., Sugimoto, N., Shimizu, A., Hara, Y., & Takemura, T. (2012). Size-resolved adjoint inversion of Asian dust. *Geophysical Research Letters*, *39*, L24807. https://doi.org/10.1029/2012GL053890

Zhang, Q., Streets, D. G., Carmichael, G. R., He, K. B., Huo, H., Kannari, A., et al. (2009). Asian emissions in 2006 for the NASA INTEX-B mission. *Atmospheric Chemistry and Physics*, *9*(14), 5131–5153. https://doi.org/10.5194/acp-9-5131-2009