An Inversion of NO\textsubscript{x} and NMVOC Emissions using Satellite Observations during the KORUS-AQ Campaign and Implications for Surface Ozone over East Asia

Amir H. Souri\textsuperscript{1,*}, Caroline R. Nowlan\textsuperscript{1}, Gonzalo González Abad\textsuperscript{1}, Lei Zhu\textsuperscript{1,2}, Donald R. Blake\textsuperscript{3}, Alan Fried\textsuperscript{4}, Andrew J. Weinheimer\textsuperscript{5}, Jung-Hun Woo\textsuperscript{6}, Qiang Zhang\textsuperscript{7}, Christopher E. Chan Miller\textsuperscript{1}, Xiong Liu\textsuperscript{1}, and Kelly Chance\textsuperscript{1}

\textsuperscript{1}Harvard-Smithsonian Center for Astrophysics, Cambridge, MA, USA
\textsuperscript{2}School of Environmental Science and Engineering, Southern University of Science and Technology, Shenzhen, China
\textsuperscript{3}Department of Chemistry, University of California, Irvine, Irvine, CA, USA
\textsuperscript{4}Institute of Arctic & Alpine Research, University of Colorado, Boulder, CO, USA
\textsuperscript{5}National Center for Atmospheric Research, Boulder, CO, USA
\textsuperscript{6}Department of Advanced Technology Fusion, Konkuk University, Seoul, South Korea
\textsuperscript{7}Department of Earth System Science, Tsinghua University, Beijing, China

* corresponding author: ahsouri@cfa.harvard.edu

Abstract. The absence of up-to-date emissions has been a major impediment to accurately simulate aspects of atmospheric chemistry, and to precisely quantify the impact of changes of emissions on air pollution. Hence, a non-linear joint analytical inversion (Gauss-Newton method) of both volatile organic compounds (VOC) and nitrogen oxides (NO\textsubscript{x}) emissions is made by exploiting the Smithsonian Astrophysical Observatory (SAO) Ozone Mapping and Profile Suite Nadir Mapper (OMPS-NM) formaldehyde (HCHO) and the National Aeronautics and Space Administration (NASA) Ozone Monitoring Instrument (OMI) tropospheric nitrogen dioxide (NO\textsubscript{2}) retrievals during the Korea-United States Air Quality (KORUS-AQ) campaign over East Asia in May-June 2016. Effects of the chemical feedback of NO\textsubscript{x} and VOCs on both NO\textsubscript{2} and HCHO are implicitly included through iteratively optimizing the inversion. Emissions estimates are greatly improved (averaging kernels>0.8) over medium- to high-emitting areas such as cities and dense vegetation. The amount of total NO\textsubscript{x} emissions is mainly dictated by values reported in the MIX-Asia 2010 inventory. After the inversion we conclude a decline in the emissions (before, after, change) for China (87.94±44.09 Gg/day, 68.00±15.94 Gg/day, -23%), North China Plain (NCP) (27.96±13.49 Gg/day, 19.05±2.50 Gg/day, -32%), Pearl River Delta (PRD) (4.23±1.78 Gg/day, 2.70±0.32 Gg/day, -36%), Yangtze River Delta (YRD) (9.84±4.68 Gg/day, 5.77±0.51
Gg/day, -41%), Taiwan (1.26±0.57 Gg/day, 0.97±0.33 Gg/day, -23%), and Malaysia (2.89±2.77 Gg/day, 2.25±1.34 Gg/day, -22%), all of which have effectively implemented various stringent regulations. In contrast, South Korea (2.71±1.34 Gg/day, 2.95±0.58 Gg/day, +9%) and Japan (3.53±1.71 Gg/day, 3.96±1.04 Gg/day, +12%) experience an increase in NOx emissions potentially due to risen number of diesel vehicles and new thermal power plants. We revisit the well-documented positive bias of the model in terms of biogenic VOC emissions in the tropics. The inversion, however, suggests a larger growth of VOC (mainly anthropogenic) over NCP (25%) than previously reported (6%) relative to 2010. The spatial variation in both magnitude and sign of NOx and VOC emissions results in non-linear responses of ozone production/loss. Due to simultaneous decrease/increase of NOx/VOC over NCP and YRD, we observe an ~53% reduction in the ratio of the chemical loss of NOx (LNOx) to the chemical loss of ROx (ROx+HO2) transitioning toward NOx-sensitive regimes, which in turn, reduces/increases the afternoon chemical loss/production of ozone through NOx+OH (-0.42 ppbv hr⁻¹)/HO2 (and ROx)+NO (+0.31 ppbv hr⁻¹). Conversely, a combined decrease in NOx and VOC emissions in Taiwan, Malaysia, and the southern China suppresses the formation of ozone. Ultimately, model simulations indicate enhancements of maximum daily 8-hour average (MDA8) surface ozone over China (0.62 ppbv), NCP (4.56 ppbv), and YRD (5.25 ppbv) due to the non-linear ozone chemistry, suggesting that emissions standards should be extended to regulate VOCs to be able to curb ozone production rates. Taiwan, Malaysia, and PRD stand out as the regions undergoing lower MDA8 ozone levels resulting from the NOx reductions occurring predominantly in NOx-sensitive regimes.
Introduction

The study of ozone (O$_3$) formation within the troposphere in East Asia is of global importance. This significant pollutant is not confined to the source, as it spreads hemispherically through the air, affecting background concentrations as far away as U.S. A study by Lin et al. [2017] provided modeling evidence of enhancements of springtime surface ozone levels (+0.5 ppbv yr$^{-1}$) in the western U.S. in 1980-2014 solely due to the tripling of Asian anthropogenic emissions over the period. As more studies have informed the impact of ozone pollution on both human health and crop yields, Chinese governmental regulatory agencies have begun to take action on cutting the amount of NO$_x$ (NO+NO$_2$) emissions since 2011-2012 [Gu et al., 2013; Reuter et al., 2014; Krotkov et al., 2016; de Foy et al., 2016; Souri et al., 2017]; however no effective policy on volatile organic compound (VOC) emissions, emitted from various sources such as solvent use, mobile, and chemical industries [Liu et al., 2008a,b], had been put into the effect prior to 2016 [Stavrakou et al., 2017; Souri et al., 2017; Shen et al., 2019; Li et al., 2019], with an exception to Pearl River Delta (PRD) [Zhong et al. 2013]. In addition to China, a number of governments including those of Malaysia and Taiwan have put a great deal of effort into shifting their energy pattern from consuming fossil fuels to renewable sources [Trappey et al., 2012; Chua and Oh, 2011]. On the other hand, using satellite observations, Irie et al. [2016] and Souri et al. [2017] revealed a systematic hiatus in the reduction of NO$_x$ over South Korea and Japan potentially due to increases in the number of diesel vehicles and new thermal power plants built to compensate for the collapse of the Fukushima nuclear power plant in 2011. Therefore, it is interesting to quantify to what extent these policies have impacted ozone pollution.

Unraveling the origin of ozone is complicated by a number of factors encompassing the nonlinearity of ozone formation to its sources, primarily from NO$_x$ and VOCs. Therefore, to be able to quantify the impact of recent emission changes, we have developed a top-down estimate of emission inventories using well-characterized observations. There are a myriad of studies focusing on optimizing the bottom-up anthropogenic and biogenic emissions using satellites observations, which provide high spatial coverage, in conjunction with chemical transport models for VOCs [e.g., Palmer et al., 2003; Shim et al., 2005; Curci et al., 2010; Stavrakou et al., 2009, 2011], and NO$_x$ [Martin et al., 2003; Chai et al., 2009; Miyazaki et al., 2017; Souri et al., 2016a, 2017, 2018]. Most inverse modeling studies do not consider both NO$_2$ and formaldehyde (HCHO) satellite-based observations to perform a joint-inversion. It has been shown that VOC and NO$_x$ emissions
can affect the production/loss of each other [Marais et al., 2012; Wolfe et al. 2016; Valin et al., 2016; Souri et al., 2020]. Consequently, a joint method that incorporates both species while minimizing the uncertainties in their emissions is better suited to address this problem. Dealing with this tangled relationship between VOC-NO₂ and NOₓ-HCHO requires an iteratively non-linear inversion framework able to incrementally consider the relationships derived from a chemical transport model. Here we will provide an optimal estimate of NOₓ and VOC emissions during the KORUS-AQ campaign using the Smithsonian Astrophysical Observatory (SAO) Ozone Mapping and Profile Suite Nadir Mapper (OMPS-NM) HCHO and the National Aeronautics and Space Administration (NASA) Ozone Monitoring Instrument (OMI) NO₂ retrievals whose accuracy and precisions are characterized against rich observations collected during the campaign. Having a top-down constraint on both emissions permits a more precise quantification of the impact of the recent emission changes on different chemical pathways pertaining to ozone formation and loss.

**Measurements, Modeling and Method**

**Remote sensing measurements**

**OMPS HCHO**

OMPS-NM onboard the Suomi National Polar-orbiting Partnership (Suomi NPP) is a UV-backscattered radiation spectrometer launched in October 2011 [Flynn et al., 2014]. Its revisit time is the same as other NASA A-Train satellites, including Aura at approximately 13:30 local time at the equator in ascending mode. OMPS-NM covers 300-380 nm with a resolution of 1 nm full-width half maximum (FWHM). The sensor has a 340×740 pixel charge-coupled device (CCD) array measuring the UV spectra at a spatial resolution of 50×50 km² at nadir. The HCHO retrieval has been fully described in González Abad et al. [2015; 2016]. Briefly, OMPS HCHO slant columns are fit using direct radiance fitting [Chance, 1998] in the spectral range 327.7-356.5 nm. The spectral fit requires a reference spectrum as function of the cross-track position as it attempts to determine the number of molecules with respect to a reference (i.e., a differential spectrum fitting). To account for this, we use earthshine radiances over a relatively clear area in the remote Pacific Ocean within -30° to +30° latitudes. An upgrade to this reference correction is the use of daily HCHO profiles over the mean climatological ones from simulations done by the GEOS-Chem chemical transport model. The scattering weights describing the sensitivity of the light path through a simulated atmosphere are calculated using VLIDORT. The shape factors used for
calculating air mass factors (AMFs) are derived from a regional chemical transport model (discussed later) that is used for carrying out the inversion in the present study. We remove unqualified pixels based on cloud fraction < 40%, solar zenith angle < 65°, and a main quality flag provided in the data. We oversample the HCHO columns for the period of May-June 2016 using a Cressman spatial interpolator with a 1° radius of influence.

**OMI Tropospheric NO$_2$**

We use NASA OMI tropospheric NO$_2$ (version 3.1) level 2 data whose retrieval is made in the violet/blue (402-465 nm) due to strong absorption of the molecule in this wavelength range [Levelt et al., 2018]. The sensor has a nadir spatial resolution of 13×24 km$^2$ which can extend to 40×160 km$^2$ at the edge of scanlines. A more comprehensive description of the retrieval and the uncertainty associated with the data can be found in Krotkov et al. [2017] and Choi et al. [2019]. We remove bad pixels based on cloud fraction < 20%, solar zenith angle < 65°, without the row anomaly, vertical column density (VCD) quality flag = 0, and Terrain Reflectivity < 30%. Similar to the OMPS HCHO, we recalculate AMFs by using shape factors from the chemical transport model used in this study. We oversample the OMI granules using the Cressman interpolator with a 0.25° radius of influence.

**Model simulation**

To be able to simulate the atmospheric composition, and to perform an analytical inverse modeling, we set up a 27-km grid resolution regional chemical transport model using the Community Multiscale Air Quality Modeling System (CMAQ) model [Byun and Schere, 2006] that consists of 328×323 grids covering China, Japan, South Korea, Taiwan and some portions of Russia, India and South Asia (Figure 1). The time period covered by the simulation is from April to June 2016. We use the month of April for spin-up. The anthropogenic emissions are based on the monthly MIX-Asia 2010 inventory [Li et al., 2015] in the CB05 mechanism. The anthropogenic emissions are mainly grouped into three different sectors, namely mobile, point, and residential (area) sources. We apply a diurnal scale to the mobile sectors used in the national emission inventory (NEI)-2011 emission platform to represent the first-order approximation of traffic patterns. We include biomass burning emissions from the Fire Inventory from NCAR (FINN) v1.6 inventory [Wiedinmyer et al., 2011], and consider the plume rise parametrization used in the GEOS-Chem model (i.e., 60% of emissions are distributed uniformly in the planetary boundary layer (PBL)). We use the offline Model of Emissions of Gases and Aerosols from Nature
(MEGAN) v2.1 model [Guenther et al., 2006] following the high resolution inputs described in Souri et al. [2017]. The diurnally lateral chemical conditions are simulated by GEOS-Chem v10 [Bey et al., 2001] using the full chemistry mechanism (NO$_x$-O$_x$-HC-Aer-Br) spun up for a year. With regard to weather modeling, we use the Weather Research and Forecasting model (WRF) v3.9.1 [Skamarock et al., 2008] at the same resolution to that of the CMAQ (~27km), but with a wider grid (342×337), and 28 vertical pressure sigma levels. The lateral boundary conditions and the grid nudging inputs are from the global Final (FNL) 0.25° resolution model. The major configurations for the WRF-CMAQ model are summarized in Table 1 and Table 2.

**Inverse modeling**

We attempt to improve our high-dimensional imperfect numerical representation of atmospheric compounds using the well-characterized NO$_2$ and HCHO columns from satellites. We use an analytical inversion using the WRF-CMAQ model to constrain the relevant bottom-up emission estimation [Souri et al., 2016; Souri et al., 2017; Souri et al., 2018]. The inversion seeks to solve the following cost function under the assumptions that i) both observation and emission error covariances follow Gaussian probability density functions with a zero bias, ii) the observation and emission error covariances are independent and iii) the relationship between observations and emissions is not grossly non-linear:

$$J(x) = \frac{1}{2}(y - F(x))^T S_o^{-1} (y - F(x)) + \frac{1}{2}(x - x_a)^T S_e^{-1} (x - x_a)$$

where $x$ is the inversion estimate (a posteriori) given two sources of data: a priori ($x_a$) and observation ($y$). $S_o$ and $S_e$ are the error covariance matrices of observation (instrument) and emission. $F$ is the forward model (here WRF-CMAQ) to project the emissions onto columns. The first term of Eq.1 attempts to reduce the distance between observations and the simulated columns. The second term incorporates some prior understanding and expectation of the true state of the emissions, that is, it does not allow the a posteriori to deviate largely from the a priori, even though the observations could be far from our estimation. The weight of each term is dictated by its covariance matrix. If $S_e$ is large compared to $S_o$, the a posteriori will be independent of the a priori knowledge and, conversely, if $S_o$ dominates, the final solution will consist mostly of the a priori.

Following the Gauss-Newton method described in Rodger [2000], we derive iteratively (i.e., $i$ is the index of iteration) the posterior emissions by:

$$x_{i+1} = x_a + G[y - F(x_i) - K_i(x_i - x_a)]$$
where $G$ is the Kalman gain,

$$G = S_e K_i^T \left( K_i S_e K_i^T + S_o \right)^{-1}$$  \hspace{1cm} (3)

and $K_i (= K(x_i))$ is the Jacobian matrix calculated explicitly from the model (discussed later). The covariance matrix of the a posteriori is calculated by:

$$S_e = (I - GR^T) S_e$$  \hspace{1cm} (4)

where $R$ is the Jacobian from the $i$th iteration. Here we iterate Eq.2 three times. The averaging kernels ($A$) are given by:

$$A = I - S_e S_e^{-1}$$  \hspace{1cm} (5)

The inversion system is complicated by the commonly overlooked fact that observations are biased. For instance, Souri et al. [2018] found that airborne remote sensing observations were high relative to surface Pandora measurements. The overestimation of the VCDs was problematic, since it could have been propagated in the inversion, inducing a bias in the top-down estimation. The authors partly mitigated it by constraining the MODIS albedo which was assumed to be responsible for the bias. Attempts to reduce the bias resulting from coarse profiles from a global model in calculating gas shape profiles were made by recalculating the shape factors using those from higher spatial resolution regional models in other studies [e.g., Souri et al., 2017; Laughner et al., 2018]. For this study, we use abundant observations from the KORUS-AQ campaign and follow the intercomparison platform proposed by Zhu et al. [2016; 2020] using aircraft observations collected during the campaign to be able to mitigate the biases in HCHO columns.

Based on the corrected global model as a benchmark, we scale up all OMPS HCHO columns by 20%. To mitigate the potential biases in OMI NO$_2$, we followed exclusively the values reported over the KORUS-AQ period in Choi et al. [2019]. We increase the NO$_2$ concentration uniformly by 33.9% (see table A3 in the paper).

We calculate the covariance matrix of observations using the column uncertainty variable provided in the satellite datasets and consider them as random errors. Therefore, these values are significantly lowered down by oversampling the data over the course of two months. In addition to that, we take into account a fixed error for all pixels due to variability that exists in the applied bias correction. This error is based on the RMSE obtained from the mentioned studies used for removing biases.
To increase the degree of freedom for the optimization, we combine all sector emissions including anthropogenic, biomass burning and biogenic emissions for NO\textsubscript{x} and VOCs. Therefore, we use the following formula to estimate the covariance of the a priori:

\[
\sigma_{\text{Total}}^2 = f_{\text{Anthro}}^2 \times \sigma_{\text{Anthro}}^2 + f_{\text{BB}}^2 \times \sigma_{\text{BB}}^2 + f_{\text{Bio}}^2 \times \sigma_{\text{Bio}}^2
\]

where \(f\) denotes the fraction of the emission sector with respect to the total emissions, and \(\sigma\) is the standard deviation of each sector category which is calculated from the average of each sector to a relative error listed in Table 3.

For the same purpose (enhancing the amount of information gained from satellite observation) and to increase computational speed, we reduce the dimension of the state vectors (emissions) by aggregating them. However, grouping emissions into certain zones could also introduce another type of uncertainty, known as the aggregation error. We choose optimally aggregated zones by running the inversion multiple times, each with a certain selection of state vectors [Turner and Jacob, 2015]. As in our previous study in Souri et al. [2018], we use the Gaussian Model Mixture (GMM) method to cluster emissions into certain zones that share roughly similar features and investigate which combinations will lead to a minimum of the sum of aggregation and smoothing errors.

In order to create the \(K\) matrix, one must estimate the impact of changes in emissions for each of the aggregated zones to the concentrations of a target compound which is calculated using CMAQ-Direct Decoupled Method (DDM) [Dunker et al., 1989; Cohan et al., 2005]. For instance, the first row and column of \(K\) denoting the response of the first grid cell to a zonal emission can be obtained by:

\[
K_{(1,1)} = \frac{S_{NO_2}^{(1,1)}}{ENO_2^T_{\text{Total Zone}}}
\]

where \(S_{NO_2}^{(1,1)}\) is the sensitivity result in units of molecule cm\(^{-2}\) for the first grid indicating how concentration of NO\textsubscript{2} column will change at the first row and column of the domain by changing one unit of emission of total NO\textsubscript{x} emissions. We do not consider the interconnection between the zonal emissions and concentrations due to computational burdens. The same concept will be applied to HCHO and VOC emissions. The advantage of using CMAQ-DDM to estimate the sensitivity lies in the fact that it calculates the local gradient which better represents the non-linear relationship existing between the emissions and the columns [Souri et al., 2017; Souri et al., 2018], which in turn, reduces the number of iterations.
Validation of the model in terms of meteorology

It is essential to first evaluate some key meteorological variables, because large errors in the weather can complicate the inversion [e.g., Liu et al. 2017]. In order to validate the performance of the WRF model in terms of a number of meteorological variables including surface temperature, relative humidity, and winds, we use more than 1100 surface measurements from integrated surface database (ISD) stations (https://www.ncdc.noaa.gov/isd) over the domain in May-June 2016. Table 4 lists the comparison of the model and the observations for the mentioned variables. Our model demonstrates a very low bias (0.6°C) with regard to surface temperature. We find a reasonable correspondence in terms of relative humidity indicating a fair water vapor budget in the model. The largest discrepancy between the model and observations in terms of temperature and humidity occurs in those grid cells that are in the proximity of the boundary conditions (not shown). Concerning the wind components, the deviation of the model from the observations is smaller than results obtained in a relatively flat area like Houston in Souri et al. [2016].

Comparison of the model and the satellite observations

Prior to updating the emissions, we find it necessary to shed light on the spatial distribution of tropospheric NO$_2$ and HCHO total columns from both observations and model, and their potential differences relative to their key precursors emissions.

NO$_2$

The first row in Figure 2 illustrates tropospheric NO$_2$ columns from the regional model, OMI (using adjusted AMF and bias corrected), and the logarithmic ratio of both quantities in May-June 2016 at ~1330 LST over Asia. The second row depicts daily-mean values of dominant sources of NO$_x$, namely as, biogenic, anthropogenic, and biomass burning emissions (that are subject to change after the inversion). A high degree of correlation between the anthropogenic NO$_x$ emissions and NO$_2$ columns implies the predominant production of NO$_2$ from the anthropogenic sources [Logan, 1983]. We find a reasonable two-dimensional Pearson correlation ($r=0.73$) between the modeled and the observed columns. Generally, the WRF-CMAQ largely underestimated (56%) tropospheric NO$_2$ columns with respect to those of OMI over the entire domain. Segregating intuitively the domain into high emission areas (NO$_x$ > 10 ton/day) and low ones (NO$_x$ < 10 ton/day) allows for a better understanding of the discrepancy between the model and the observations. In the high NO$_x$ areas, the model tends to overestimate tropospheric NO$_2$ columns by 73%, whereas for the low NO$_x$ regions, the model shows a substantial underestimation by 68%.
Such a conflicting bias is confirmed by the contour map of the logarithm ratio of the model to OMI in Figure 2. The large overestimation of the model in terms of NO$_2$ over the polluted areas is explained by stringent regulations enacted in various countries in Asia; for instance, Chinese regulatory agencies have taken aggressive actions recently to cut anthropogenic NO$_x$ emissions by implementing selective catalytic reduction in power plants, closing a number of coal power plants, and policies on transportation [Zhang et al., 2012, Liu et al., 2016]. The highest positive bias in the model is observed over Shanxi Province in China, home to coal production, underscoring the effectiveness of the emission standards at controlling air pollution. Likewise, we observe a positive bias in the model over major cities in Japan and South Korea; but the magnitude of the reduction over these cities is substantially smaller than what we observe in China. In particular, Irie et al. [2016] and Souri et al. [2017] found a hiatus in the NO$_x$ reduction over Japan and South Korea during the 2010-2014 period mainly due to rapid increases in the number of diesel cars in South Korea, and thermal power plants built as a substitution for the Fukushima nuclear plant in Japan.

The underestimation of the model in the low NO$_x$ regions is related to a number of factors such as i) the widely-reported underestimation of soil (biogenic) NO$_x$ emissions due to the lack of precise knowledge of fertilizers use, soil biota, or canopy interactions [Jaeglé, et al., 2005; Hudman et al., 2010; Souri et al., 2016], ii) the underestimation of the upper-troposphere NO$_2$ due to non-surface emissions (aviation/lightning) or errors in the vertical mixing or moist convection [e.g., Souri et al., 2018], and iii) a possible overprediction of the lifetime of organic nitrates diminishing background NO$_2$ levels [Canty et al., 2015]. Addressing the second issue requires a very high resolution model with explicit resolving microphysics and large eddy simulations, and the last problem requires more experimental studies to improve organic nitrates chemistry [Romer Present et al., 2020]. In this study, we attempt to mitigate the discrepancy between the model and the satellite observations solely by adjusting the relevant emissions. Accordingly, future approaches improving models the physical/chemical processes can offset top-down emissions estimates inevitably.

**HCHO**

A comparison between HCHO columns from the model and OMPS along with the major sources of VOCs in May-June 2016 is depicted in Figure 3. A reasonable correlation ($r=0.78$) between the model and OMPS suggests a good confidence in the location of emissions. However, the magnitude of HCHO columns between the two datasets strongly disagrees, especially over the
tropics where biogenic emissions are large. A myriad of studies have reported a largely positive bias (by a factor of 2-3) associated with isoprene emissions estimated by MEGAN using satellite measurements [e.g., Millet et al., 2008; Stavrakou et al., 2009; Marais et al., 2012; Bauwens et al., 2016]. To compound, Stavrakou et al. [2011] found a large overestimation in methanol emissions from the same model that can further preclude the accurate estimation of the yield of HCHO. This is especially the case for the tropics.

As a response to the overestimation of the biogenic VOCs by MEGAN, we observe a largely positive bias in the simulated HCHO columns ranging from 50% over the south of China to ~400% over Malaysia and Indonesia. As we move away from the hotspot of the biogenic emissions in lower latitudes, the positive bias of the model declines, ultimately turning into a negative bias at higher latitudes. OMPS HCHO columns suggest that the yield of HCHO over North China Plain (NCP) and Yangtze River Delta (YRD) is comparable to those over the tropics suggesting that the anthropogenic emissions over NCP are the dominant source of HCHO [Souri et al., 2017; Jin and Holloway, 2015]. We do not see a significant deviation in the model from the observations over this region indicating that no noticeable efforts on controlling VOC emissions in NCP and YRD have been made which is very likely due to the fact that the recent regulations over China have overlooked cutting emissions from several industrial sectors [Liu et al., 2016] prior to 2016 [Li et al. 2019]. This finding lines up with results reported by Souri et al. [2017] and Shen et al. [2019]. We observe both underestimated and overestimated values in the simulated HCHO columns over areas in South Korea and Japan. The underestimation of HCHO in the model over regions with low VOCs (such as Mongolia and Pacific Ocean) can be either due to missing sources or the incapability of CMAQ to account for moist convective transport. As shown here, it is necessary to adjust the emissions to better match the simulated columns with the satellites observations given their errors, and by doing so, there is a chance for a better simulation of the formation of tropospheric ozone.

Updated Emissions

In this section we report the results from the inverse modeling and the associated uncertainty associated with the top-down estimation; moreover, we wish to assess how much information is gained from utilizing satellite observations via the calculation of averaging kernels. Finally, observations are used to verify, to some extent, the accuracy of our top-down emission estimations.
The first row in Figure 4 shows the a priori, the a posteriori, and their ratios in terms of the total NO\textsubscript{x} emissions in May-June 2016. We observe that the ratios are highly correlated with those of CMAQ/OMI shown in Figure 2, suggesting that the inversion attempts to reduce the distance between the model and the observations. Major reductions occur over China. We attribute them to strict emissions policies [Liu et al., 2016; Reuter et al., 2015; de Foy et al., 2016; Krotkov et al., 2016; Souri et al., 2017]. The enhancements in NO\textsubscript{x} emissions are commonly found in rural areas, especially over grasslands located in the western/central China and Mongolia. The changes in NO\textsubscript{x} emissions over South Korea and Japan are positive [Irie et al., 2016; Souri et al., 2017]. This is especially the case for Japan for which we observe a larger enhancement in total NO\textsubscript{x} emissions (12%) essentially due to new thermal power-plants. The second row in Figure 4 depicts the relative errors in the a priori, the a posteriori, and AKs. Relative errors in the a priori are mostly confined to values close to 50% in polluted areas. They increase further, up to 100%, in areas experiencing relatively large contributions from biomass burning or biogenic (soil) emissions. Encouragingly, OMI tropospheric NO\textsubscript{2} columns in conjunction with the solid mathematical inversion method [Rodger, 2000] greatly reduce the uncertainties associated with the emissions in polluted areas; we observe AKs close to 1 over major cities or industrial areas. We see the lowest values in AKs over rural areas due to weaker signal/noise ratios from the sensor. Therefore, it is desirable but very difficult to improve the model using the sensor in terms of NO\textsubscript{x} chemistry/emissions in remote areas, evident in the low values of AKs. Table 5 lists the magnitude of the total NO\textsubscript{x} emissions in several regions (refer to Figure 1) before and after carrying out the inversion. If we assume that the dominant source of NO\textsubscript{x} emissions is anthropogenic, the most successful countries at cutting emissions (before, after) are China (87.94±44.09 Gg/day, 68.00±15.94 Gg/day), Taiwan (1.26±0.57 Gg/day, 0.97±0.33 Gg/day), and Malaysia (2.89±2.77 Gg/day, 2.25±1.34 Gg/day). All three countries have successfully implemented plans to reduce anthropogenic emissions since 2010-2011 [Zhang et al., 2012; Trappey el al., 2012; Chua and Oh, 2011]. The uncertainty associated with the top-down estimate improves considerably. The largest reduction in the uncertainty of the emissions is observed over China, a response to a strong signal from OMI.

An interesting observation lies in the discrepancy between the ratio of OMI/CMAQ (Figure 2) to that of the a posteriori to the a priori over the North China Plain, suggesting that using a bulk ratio [Martin et al., 2003] cannot fully account for possible chemical feedback. The ratio of OMI/CMAQ is consistently lower than changes in the emission. Two reasons contribute to this
effect: i) as NOx emissions decrease in NOx-saturated areas (i.e., the dominant sink of radicals is through NO2+OH), OH levels essentially increase resulting in a shorter lifetime in NO2; therefore to reduce NO2 concentrations, a substantial reduction in NOx (suggested by OMI/CMAQ) is unnecessary coinciding with results from the inverse modeling, ii) the CMAQ-DDM (Figure S1) suggests that NO2 columns decrease due to increasing VOC emissions over the region; accordingly, the cross-relationship between NO2 concentrations and VOC emissions which has been implicitly taken into consideration by iteratively optimizing the cost function partly adds to the discrepancy. It is because of the chemical feedback that recent studies have attempted to enhance the capability of inverse modeling by iteratively adjusting relevant emissions [e.g., Cooper et al., 2017; Li et al., 2019].

To assess the resulting changes in the tropospheric NO2 columns after the inversion, and to validate our results, we compare the simulated values using the a priori and the a posteriori with OMI in Figure 5. We observe 64% reduction in the tropospheric NO2 columns on average over NCP despite only 32% reduction in the total NOx emissions over the region, a result of the chemical feedback. The two-dimensional Pearson correlation between the simulation using the a posteriori and OMI increases from 73% (using the a priori) to 83%. Both datasets now are in a better agreement as far as the magnitude goes. However, we do not see a significant change in the background values in the new simulation compared to those of OMI. This is primarily because of the consideration of higher covariances over low-emitting areas that weighs up the inversion towards the prior values.

To further validate the results, we compare the NO2 data from the NCAR’s four-channel chemiluminescence instrument onboard the DC-8 aircraft during the campaign (not shown). These data are not interfered by NOx family. The aircraft collected the data in the Korean Peninsula around 23 days in May-June 2016 covering various altitudes and hours (https://www-air.larc.nasa.gov/cgi-bin/ArcView/korusaq, access date: December 2019). We observe an underestimation of NO2 at the near surface levels (<900 hPa) by 19% (DC8 = 4.50 ppbv, CMAQ = 3.67 ppbv). The updated emissions increase the near surface levels over the Korean Peninsula, which in turn, reduce the bias to 11% (CMAQ = 4.02 ppbv).

**VOC**

Figure 6 illustrates the total VOC emissions before and after the inversion along with their errors. Immediately apparent is the large reduction of VOC emissions in the tropics and subtropics
due to the overestimation of isoprene from MEGAN v2.1. In contrast, enhancements of the emissions are evident at higher latitudes. We observe that the dominantly anthropogenic VOC emissions over NCP increase (~25%) after the adjustment highlighting the minimal efforts made to reduce this particular source of emissions [Souri et al., 2017; Shen et al., 2019]. For instance, Stavrakou et al. [2017] reported ~6% increases in anthropogenic VOC emissions over China from 2010 to 2014. Despite the presence of vegetation over Japan and South Korea, we do not see largely overestimated values in the emissions. Hence, the overestimation of isoprene emissions is more pronounced in the tropics possibly because of an overestimation in the emission factors used for specific plants. Nevertheless, a non-trivial oversight in models could be an insufficient representation of both HOx chemistry and dry deposition in forest canopies [Millet et al., 2008]; as a result, the net amount of HCHO in the atmosphere over forest areas is higher than what should be if removal through either a chemical loss or a faster dry deposition is considered.

Owing to the fact that we assume anthropogenic VOC emissions to be less uncertain relative to other sectors, the errors in the a priori are smaller in populated areas. We observe that OMPS HCHO columns are able to significantly reduce the uncertainty associated with the total VOC emissions over areas showing a strong HCHO signal (>1016 molec.cm⁻²). Over clean areas, it is the other way around; we see less confidence in our top-down estimate (AK<0.4) in areas such as Tibet and Mongolia.

We then compare the simulated HCHO column using two different emission inventories with those of OMPS in Figure 7. We observe a substantial improvement both in the spatial structure and the magnitude of simulated HCHO columns using the a posteriori with respect to OMPS. The two-dimensional Pearson correlation increases from 0.78 to 0.91 after applying the adjustments to the emissions. In response to the increases in the total VOC emissions over the North China Plain, we observe ~11% enhancements in the simulated HCHO total columns. The updated emissions lead to a reduction in HCHO total columns as large as 70% in the tropics.

Validation of the model in terms of VOCs is not a straightforward task because the chemical mechanism used for our model has lumped several VOC species such as terminal/internal olefin or paraffin, only a handful of which were measured during the campaign. Besides, the MIX-Asia inventory estimates the anthropogenic emissions for a selected number of VOCs in the CB05 mechanism. Here, we focus only on six compounds including isoprene, HCHO, ethene, ethane, acetaldehyde, and methanol whose emissions are adjusted (with the same factor) based on satellite
measurements. The comparison of the simulated values with the DC-8 measurements showed a noticeable mitigation in the discrepancy between two datasets at lower boundaries (<900 hPa) in terms of isoprene (Figure S2), ethane (Figure S3), ethene (Figure S4), and acetaldehyde (Figure S5). Surprisingly, we observe a large underestimation of methanol over the Korean Peninsula by a factor of ten (Figure S6). Same tendency was observed in other regions in Wells et al. [2014] (see Figure 8 in the paper). Our inversion obviously fails at mitigating the bias as there is not much direct information from the satellite observations on this compound. Wells et al. [2014] and Hu et al. [2011] demonstrated that methanol can be a secondary source of HCHO up to 10-20% in midlatitudes in warm seasons. We tend to underestimate HCHO concentrations (by 15%) in the lower atmosphere (<900 hPa) after using the a posteriori over the Korean Peninsula (Figure S7).

Implications for surface ozone

The results we have generated can be further exploited to elucidate changes in the ozone production rates P(O₃) owing to have constrained both NOₓ and VOC emissions. We calculate P(O₃) by subtracting the ozone loss driven by HOₓ (HO+HO₂), reaction with several VOCs (i.e., alkenes and isoprene), the formation of HNO₃, and O₃ photolysis followed by the reaction of O¹(¹D) with water vapor, from the ozone formation via removal of NO through HO₂ or RO₂:

\[ P(O_3) = k_{HO_2+NO}[HO_2][NO] + \sum k_{RO_2+NO}[RO_2][NO] - k_{OH+NO_2+M}[OH][NO_2][M] - k_{HO_2+O_3}[HO_2][O_3] - k_{OH+O_3}[OH][O_3] - k_{D(1D)+HO_2}[O(1D)][H_2O] - L(O_3 + VOCs) \]  

Since P(O₃) is a non-linear function of NOₓ and VOC emissions, it is advantageous to look at the ratio of chemical loss of NOₓ to that of ROₓ (RO₂+HO₂), a robust indicator to pinpointing underlying drivers for ROₓ cycle. LROₓ is defined through the sum of primarily radical-radical reactions:

\[ LRO_x = k_{HO_2+HO_2}[HO_2]^2 + \sum k_{RO_2+HO_2}[RO_2][HO_2] + \sum k_{RO_2+RO_2}[RO_2]^2 \]  

LNOₓ mainly occurs via the NO₂+OH reaction:

\[ LNO_x = k_{OH+NO_2+M}[OH][NO_2][M] \]  

Typically, a value of LNOₓ/LROₓ~2.7 defines the transition line between VOC-sensitive and NOₓ-sensitive regimes [Schroeder et al., 2017; Souri et al., 2020].

Figure 8 depicts a contour map of LNOₓ/ROₓ ratios before and after the inversion. As expected, the larger ratios are confined within major cities or industrial areas due to abundant NOₓ.
emissions. The hotspot of VOC-sensitive regimes is located in NCP and YRD. Also of interest in Figure 8 is that advection renders a major fraction of the Yellow Sea (the sea connecting China to Korea) VOC-sensitive. Using the a posteriori leads to precipitous changes in the chemical condition regimes. As a result of a large reduction in the isoprene emissions in both the tropics and subtropics, we observe a shift toward VOC-limited, though the values of LNOₓ/ROₓ are yet too far from the transition line (i.e., <<2.7). The substantial reduction in NOₓ emissions and an increase in VOC emissions over NCP and YRD go hand-in-hand transitioning towards NOₓ-sensitive regime. The ratios over South Korea and Japan are found to be variable and somehow in sync with the changes in NOₓ emissions.

The resultant changes in the LNOₓ/LROₓ ratios shed some light on ozone sensitivity with respect to its major precursors, but P(O₃) is also dependent on the absolute values of emissions, the degree of reactivity of VOCs, and the abundance of radicals. Here we use the integrated reaction rates (IRR) to determine the most influential reactions pertaining to ozone loss and production at the surface. We focus on 1200 to 1800 China standard time (CST) hours. Figure 9 shows the differences in the major pathways for the loss and the formation of ozone at the surface within the time window. The differences are computed based on the subtraction of the simulation with the a posteriori from that with the a priori. In Figure 9 we see a strong degree of correlation between the changes in magnitude of P(O₃) through HO₂+NO reaction with those of NOₓ emissions (Figure 4), whereas the changes in magnitude of P(O₃) via RO₂+NO reaction primarily are par with those of VOC emissions (Figure 6). We observe P(O₃) increases through HO₂+NO and RO₂+NO reactions in Japan, South Korea, Myanmar, and Philippines because of increases in NOₓ emissions in NOₓ-sensitive regions. The simultaneous decrease in NOₓ and VOC in PRD and Taiwan causes the production of ozone via the same pathways to reduce.

Normally, in VOC-rich environments, reduction in VOC emissions boosts OH concentrations (Figure S8). Consequently, we observe an enhancement in NO₂+OH reaction in the tropics and subtropics. A substantial reduction in the chemical loss of ozone through NO₂+OH over NCP and YRD arises from a considerable decrease of NOₓ emissions and an increase in OH (due to chemical feedback of NOₓ). Because of increases in HOₓ concentrations over NCP (Figure S8-S9), we also observe an enhancement in ozone loss through reacting with HOₓ. Changes in the ozone photolysis (O¹D+H₂O) are majorly dictated by photolysis and water vapor mixing ratios, both of which are roughly constant in both simulations; accordingly the difference of the reaction
rate is mainly reflecting those in ozone (shown later). Interestingly, we observe a large reduction in the loss of ozone through reaction with VOCs at lower latitudes. This is essentially because of the reduction in ISOP+O$_3$, a VOC that prevails in those latitudes. Despite a much slower reaction rate for ISOP+O$_3$ compared to ISOP+OH and ISOP+$hv$ [Karl et al. 2004], this specific chemical pathway can be important as a way to oxidize isoprene and forming HO$_x$ in forests [Paulson and Orlando, 1996].

Figure 10 sums the differences of all mentioned chemical pathways involved in formation/loss of surface ozone at 1200-1600 CST. Because of a complex non-linear relationship between P(O$_3$) and its precursors, we observe a variability in both the sign and amplitude of P(O$_3$). On average, changes in O$_3$ production dominate over changes in O$_3$ sinks except in Malaysia which underwent a significant reduction in isoprene emissions, thus slowing down the ISOP+O$_3$ reaction. Following the patterns of NO$_x$-limited and VOC-limited in Figure 8, it is possible to conclude that the P(O$_3$) differences are mainly driven by those of NO$_x$ depending at which chemical condition the changes in emissions have occurred.

Much of the above analysis is based on ozone production rates, however, various parameters encompassing dry deposition, vertical diffusion, and advection can also affect ozone concentrations. Therefore we further compute the difference between the simulated maximum daily 8-h average (MDA8) surface ozone levels before and after the inversion depicted in Figure 11. We see a striking correlation between P(O$_3$) (right panel in Figure 10) and MDA8 surface ozone indicating that the selected chemical pathways in this study can explain ozone changes.

Nonetheless, the transport obviously plays a vital role in the spatial variability associated with the differences of surface ozone [e.g., Souri et al., 2016b]. Figure 11 suggests a significant enhancement of ozone over NCP (~4.56 ppbv, +5.6%) and YRD (5.2 ppbv, +6.8%) due to simultaneous decreases/increases in NO$_x$/VOCs which is in agreement with Li et al. [2019]. On the other hand, reductions in NO$_x$ mitigate ozone pollution in PRD (-5.4%), Malaysia (-5.6%) and Taiwan (-11.6%). Table 6 lists the simulated MDA8 surface ozone levels for several regions before and after updating the emissions. Increases in MDA8 ozone over NCP and YRD overshadow decreases in southern China resulting in 1.1% enhancement for China. This provides strong evidence that regulations on cutting VOC emissions should not be ignored. The largest reduction/increase of MDA8 ozone is found over Taiwan/YRD.
Summary

In this paper we have focused on providing a top-down constraint on both volatile organic compound (VOC) and nitrogen oxides (NOx) emissions using a combination of error-characterized Smithsonian Astrophysical Observatory (SAO) Ozone Mapping and Profile Suite Nadir Mapper (OMPS-NM) formaldehyde (HCHO) and National Aeronautics and Space Administration (NASA) Ozone Monitoring Instrument (OMI) nitrogen dioxide (NO2) retrievals during the Korean and United States (KORUS) campaign over East Asia in May-June 2016. Here, we include biogenic, biomass burning and anthropogenic emissions from MEGAN, FINN, and MIX-Asia 2010 inventory, respectively. A key point is that by considering together the satellite observations, we have been able to not only implicitly take the chemical feedback existing between HCHO-NOx and NO2-VOC into account through iteratively optimizing an analytical non-linear inversion, but also to quantify the impact of recent changes in emissions (since 2010) on surface ozone pollution.

Concerning total NOx emissions, the inversion estimate suggests a substantial reduction over China (-23%), North China Plain (NCP) (-32%), Pearl River Delta (PRD) (-36%), Yangtze River Delta (YRD) (-41%), Taiwan (-23%), and Malaysia (-22%) with respect to the values reported in the prior emissions mostly dictated by the MIX-Asia 2010 inventory. In essence these values reflect recent actions to lower emissions in those countries [Zhang et al., 2012; Trappey et al., 2012; Chua and Oh, 2011]. The analytical inversion also paves the way for estimating the averaging kernels (AKs), thereby informing the amount of information acquired from satellites on the emissions estimation. We observe AKs>0.8 over major polluted areas indicating that OMI is able to improve the emission estimates over medium to high-emitting regions. Conversely, AKs are found to be small over pristine areas suggesting that little information can be gained from the satellite over rural areas given retrieval errors. In line with the studies of Irie et al. [2016] and Souri et al. [2017], we observe a growth in the total NOx emissions in Japan (12%) and South Korea (+9%) which are partially explained by construction of new thermal power plants in Japan, and an upward trend in the number of diesel vehicles in South Korea.

MEGAN v2.1 estimates too much isoprene emissions in the tropics and subtropics, a picture that emerges from the latitudinal dependence of the posterior VOC emissions to the prior ones. It is readily apparent from the top-down constrained VOC emissions that the prevailing anthropogenic VOC emissions in NCP is underestimated by 25%, a direction that is in agreement with studies by Souri et al. [2017] and Shen et al. [2019]. We find out that OMPS HCHO columns
can greatly reduce the uncertainty associated with the total VOC emissions (AKs>0.8) over regions having a moderate-strong signal (>10^{16} molec.cm^{-2}).

A large spatial variability associated with both NOx and VOC results in great oscillation in chemical conditions regimes (i.e., NOx-sensitive or VOC-sensitive). Due to considerable reduction/increase in NOx/VOC emissions in NCP and YRD, we observe a large increase (53%) in the ratio of the chemical loss of NOx (LNOx) to the chemical loss of ROx (RO2+HO2) shifting the regions towards NOx-sensitive. As a result, a substantial reduction in afternoon NO2+OH reaction rate (a major loss of O3), and an increase in afternoon NO+HO2 and RO2+NO (a major production pathway for O3) are observed, leading to enhancements of the simulated maximum daily 8-hr average (MDA8) surface ozone concentrations by ~5 ppbv. Therefore, additional regulations on VOC emissions should be implemented to battle ozone pollution in those areas. On the other hand, being predominantly in NOx-sensitive regimes favors regions including Taiwan, Malaysia and PRD to benefit from reductions in NOx, resulting in noticeable decreases in simulated MDA8 surface ozone levels.

It has taken many years to develop satellite-based gas retrievals, and weather and chemical transport models accurate enough to enable observationally-based estimates of emissions with reasonable confidence and quantified uncertainty, and produce credible top-down emission inventories over certain areas. However it is essential to improve certain aspects to be able to narrow the range of uncertainty associated with the estimation: i) getting the bias of the satellite gas retrievals about right for some areas (which requires a rigorous construction of representivity factor when it comes to comparing two datasets) would be insufficient because the bias can vary substantially over time and space depending on underlying surface properties and the atmospheric state, ii) there is a need for a proper quantification of the errors in the prior emissions based on the very raw information used in the emission inventory to discourage some arbitrariness, iii) the model parameter errors including those from PBL, radiation, and winds should be propagated to the final output [e.g., Rodger 2000], iv) due to intertwined chemical feedback between various chemical compounds, inverse modeling needs to properly incorporate all available information (beyond HCHO and NO2) considering the cross-relationship either explicitly or implicitly.

Acknowledgment
We are thankful for the funding from NASA Aura Science Team (#NNX17AH47G), and NOAA AC4 program (#NA18OAR4310108). We acknowledge the publicly available WRF, CMAQ,
GEOS-Chem models, and KORUS-AQ data that make this study possible. The simulations were run on the Smithsonian Institution High Performance Cluster (SI/HPC).

Authors’ contributions

A.H.S designed the research, analyzed the data, conducted the inverse modeling, CMAQ, GEOS-Chem, WRF, and MEGAN, made all figures and wrote the manuscript. C.R.N, G.G, C.E.C.M, X.L. and K.C retrieved OMPS HCHO columns and conceived the study. L.Z. validated OMPS HCHO. D.R.B, A.F, and A.J.W measured different compounds during the campaign. J.W and Q.Z provided MIX-Asia inventory. All authors contributed to discussions and edited the manuscript.
References


de Foy, B., Lu, Z., Streets, D.G., 2016. Satellite NO 2 retrievals suggest China has exceeded its NO x reduction goals from the twelfth Five-Year Plan. Scientific Reports 6, 1–9. https://doi.org/10.1038/srep35912


of the global atmospheric methanol emission fluxes. Atmospheric Chemistry and Physics 11, 4873–4898. https://doi.org/10.5194/acp-11-4873-2011


retrievals using observations from 12 aircraft campaigns. Atmospheric Chemistry and Physics Discussions 1–25. https://doi.org/10.5194/acp-2019-1117

Table 1. CMAQ major configurations

<table>
<thead>
<tr>
<th>CMAQ version</th>
<th>V5.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Chemical Mechanism</td>
<td>CB05 with chlorine chemistry</td>
</tr>
<tr>
<td>Lightning NOx emission</td>
<td>Included using inline code</td>
</tr>
<tr>
<td>Photolysis</td>
<td>Inline including aerosol impacts</td>
</tr>
<tr>
<td>Horizontal advection</td>
<td>YAMO (hyamo)</td>
</tr>
<tr>
<td>Vertical advection</td>
<td>WRF omega formula (vwrf)</td>
</tr>
<tr>
<td>Horizontal mixing/diffusion</td>
<td>Multiscale (multiscale)</td>
</tr>
<tr>
<td>Vertical mixing/diffusion</td>
<td>Asymmetric Convective Model version 2 (acm2)</td>
</tr>
<tr>
<td>Aerosol</td>
<td>AERO 6 for sea salt and thermodynamics (aero6)</td>
</tr>
<tr>
<td>IC/BC source</td>
<td>GEOS-Chem v10</td>
</tr>
</tbody>
</table>

Table 2. WRF physics options

<table>
<thead>
<tr>
<th>WRF Version</th>
<th>V3.9.1</th>
</tr>
</thead>
<tbody>
<tr>
<td>Microphysics</td>
<td>WSM-6</td>
</tr>
<tr>
<td>Long-wave Radiation</td>
<td>RRTMG</td>
</tr>
<tr>
<td>Short-wave Radiation</td>
<td>RRTMG</td>
</tr>
<tr>
<td>Surface Layer Option</td>
<td>Monin-Obukhov</td>
</tr>
<tr>
<td>Land-Surface Option</td>
<td>Noah LSM</td>
</tr>
<tr>
<td>Boundary Layer</td>
<td>ACM2</td>
</tr>
<tr>
<td>Cumulus Cloud Option</td>
<td>Kain-Fritsch</td>
</tr>
<tr>
<td>IC/BC</td>
<td>FNL 0.25°</td>
</tr>
</tbody>
</table>

Table 3. The uncertainty assumptions used for estimating the covariance matrix of the a priori.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Anthropogenic</th>
<th>Biogenic</th>
<th>Biomass Burning</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOx</td>
<td>50%</td>
<td>200%</td>
<td>100%</td>
</tr>
<tr>
<td>VOC</td>
<td>150%</td>
<td>200%</td>
<td>300%</td>
</tr>
</tbody>
</table>

Table 4. Statistics of surface temperature, relative humidity, and wind. Corr – Correlation; RMSE – Root Mean Square Error; MAE – Mean Absolute Error; MB – Mean Bias; O – Observation; M - Model; O_M – Observed Mean; M_M – Model Mean; SD – Standard Deviation; Units for RMSE/MAE/MB/O_M/M_M/O_SD/M_SD: °C for temperature, percentage for relative humidity, and m s⁻¹ for wind.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Corr</th>
<th>RMSE</th>
<th>MAE</th>
<th>MB</th>
<th>O_M</th>
<th>M_M</th>
<th>O_SD</th>
<th>M_SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Temperature</td>
<td>0.74</td>
<td>7.0</td>
<td>2.8</td>
<td>0.6</td>
<td>22.2</td>
<td>22.8</td>
<td>9.5</td>
<td>8.7</td>
</tr>
<tr>
<td>Relative Humidity</td>
<td>0.76</td>
<td>12.1</td>
<td>9.5</td>
<td>-1.1</td>
<td>67.8</td>
<td>66.6</td>
<td>14.3</td>
<td>18.6</td>
</tr>
<tr>
<td>U Wind</td>
<td>0.58</td>
<td>1.3</td>
<td>0.7</td>
<td>0.1</td>
<td>0.1</td>
<td>0.2</td>
<td>1.2</td>
<td>1.4</td>
</tr>
<tr>
<td>V Wind</td>
<td>0.49</td>
<td>1.6</td>
<td>0.7</td>
<td>0.3</td>
<td>0.2</td>
<td>0.5</td>
<td>1.6</td>
<td>1.2</td>
</tr>
</tbody>
</table>
Table 5. NO\textsubscript{x} emissions before and after carrying out the inversion using OMI/OMPS for different countries in May-June 2016.

<table>
<thead>
<tr>
<th>Countries</th>
<th>The a priori (Gg/day)</th>
<th>The a posteriori (Gg/day)</th>
<th>Changes in magnitudes</th>
<th>Changes in errors</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>87.94±44.09\textsuperscript{1}</td>
<td>68.00±15.94\textsuperscript{2}</td>
<td>-23%</td>
<td>-63%</td>
</tr>
<tr>
<td>North China Plain</td>
<td>27.96±13.49</td>
<td>19.05±2.50</td>
<td>-32%</td>
<td>-81%</td>
</tr>
<tr>
<td>Pearl River Delta</td>
<td>4.23±1.78</td>
<td>2.70±0.32</td>
<td>-36%</td>
<td>-84%</td>
</tr>
<tr>
<td>Yangtze River Delta</td>
<td>9.84±4.68</td>
<td>5.77±0.51</td>
<td>-41%</td>
<td>-89%</td>
</tr>
<tr>
<td>Thailand</td>
<td>4.38±3.24</td>
<td>4.20±2.28</td>
<td>-4%</td>
<td>-29%</td>
</tr>
<tr>
<td>Japan</td>
<td>3.53±1.71</td>
<td>3.96±1.04</td>
<td>+12%</td>
<td>-39%</td>
</tr>
<tr>
<td>Malaysia</td>
<td>2.89±2.77</td>
<td>2.25±1.34</td>
<td>-22%</td>
<td>-49%</td>
</tr>
<tr>
<td>Vietnam</td>
<td>2.87±2.04</td>
<td>2.79±1.57</td>
<td>-3%</td>
<td>-23%</td>
</tr>
<tr>
<td>South Korea</td>
<td>2.71±1.34</td>
<td>2.95±0.58</td>
<td>+9%</td>
<td>-56%</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>1.72±1.06</td>
<td>2.10±0.87</td>
<td>+22%</td>
<td>-18%</td>
</tr>
<tr>
<td>Philippines</td>
<td>1.30±1.10</td>
<td>1.54±0.98</td>
<td>+18%</td>
<td>-11%</td>
</tr>
<tr>
<td>Taiwan</td>
<td>1.26±0.57</td>
<td>0.97±0.33</td>
<td>-23%</td>
<td>-42%</td>
</tr>
<tr>
<td>Cambodia</td>
<td>0.54±0.50</td>
<td>0.57±0.45</td>
<td>+5%</td>
<td>-11%</td>
</tr>
<tr>
<td>Mongolia</td>
<td>0.19±0.13</td>
<td>0.28±0.12</td>
<td>+44%</td>
<td>-8%</td>
</tr>
</tbody>
</table>

1- The errors in the a priori are estimated from equation 6.
2- The errors in the a posteriori are calculated by equation 4.
Table 6. MDA8 surface ozone levels before and after carrying out the inversion for different regions in May–June 2016.

<table>
<thead>
<tr>
<th>Regions</th>
<th>The a priori (ppbv)</th>
<th>The a posteriori (ppbv)</th>
<th>Changes in magnitudes</th>
</tr>
</thead>
<tbody>
<tr>
<td>China</td>
<td>56.10±16.34</td>
<td>56.72±16.71</td>
<td>+1.1%</td>
</tr>
<tr>
<td>North China Plain</td>
<td>81.15±9.57</td>
<td>85.71±10.39</td>
<td>+5.6%</td>
</tr>
<tr>
<td>Pearl River Delta</td>
<td>65.94±9.39</td>
<td>62.37±8.93</td>
<td>-5.4%</td>
</tr>
<tr>
<td>Yangtze River Delta</td>
<td>76.79±5.90</td>
<td>82.04±5.21</td>
<td>+6.8%</td>
</tr>
<tr>
<td>Thailand</td>
<td>50.86±8.84</td>
<td>48.85±7.94</td>
<td>-3.9%</td>
</tr>
<tr>
<td>Japan</td>
<td>64.29±7.98</td>
<td>65.52±7.78</td>
<td>+1.9%</td>
</tr>
<tr>
<td>Malaysia</td>
<td>46.87±21.87</td>
<td>44.22±12.90</td>
<td>-5.6%</td>
</tr>
<tr>
<td>Vietnam</td>
<td>49.90±9.20</td>
<td>48.88±8.65</td>
<td>-2.0%</td>
</tr>
<tr>
<td>South Korea</td>
<td>84.23±3.57</td>
<td>84.90±3.69</td>
<td>+0.8%</td>
</tr>
<tr>
<td>Bangladesh</td>
<td>65.79±12.08</td>
<td>65.21±12.20</td>
<td>-0.9%</td>
</tr>
<tr>
<td>Philippines</td>
<td>27.92±9.11</td>
<td>28.69±7.92</td>
<td>+2.8%</td>
</tr>
<tr>
<td>Taiwan</td>
<td>61.55±10.88</td>
<td>54.38±8.00</td>
<td>-11.6%</td>
</tr>
<tr>
<td>Cambodia</td>
<td>39.87±3.62</td>
<td>40.20±3.46</td>
<td>+0.8%</td>
</tr>
<tr>
<td>Mongolia</td>
<td>40.11±2.52</td>
<td>40.16±2.40</td>
<td>+0.1%</td>
</tr>
</tbody>
</table>
Figures:

**Figure 1.** The CMAQ 27-km domain covering the major proportion of Asia. The background picture is retrieved from publicly available NASA’s blue marble (© NASA).
Figure 2. (first row), tropospheric NO$_2$ columns from the WRF-CMAQ model, OMI (using adjusted AMFs based on the shape factors derived from the model and bias corrected following Choi et al. [2019]), and the logarithmic ratio of CMAQ/OMI during May-June 2016 at ~1330 LST. (second row) The major sources of NO$_x$ emissions in the region including biogenic (soil) emissions simulated by MEGAN, anthropogenic emissions estimated by MIX Asia (2010), and biomass burning emissions made by FINN. The emissions are the daily-mean values based on the emissions in May-June.
Figure 3. (first row), HCHO total columns from the WRF-CMAQ model, OMPS (using adjusted AMFs based on the shape factors derived from the model and bias corrected following the method proposed in Zhu et al. [2020]), and the logarithmic ratio of CMAQ/OMPS during May-June 2016 at ~1330 LST. (second row) The major sources of VOC emissions in the area including biogenic emissions simulated by MEGAN, anthropogenic emissions estimated by MIX Asia (2010), and biomass burning emissions made by FINN. The emissions are the daily-mean values based on the emissions in May-June. The VOC emissions only add up those compounds that are included in the CB05 mechanism.
Figure 4. (first row), total NO$_x$ emissions (i.e., the a priori), constrained by the satellite observations (i.e., the a posteriori) in May-June 2016, and the ratio of the a posteriori to the a priori. (second row) the errors in the a priori based on Table 3, the errors in the top-down estimation, and the averaging kernels (AKs) obtained from the estimation.
Figure 5. (from left to right), tropospheric NO$_2$ columns from OMI, WRF-CMAQ simulated with the prior emissions, and the same model but with the top-down emissions constrained by OMI/OMPS in May-June 2016.
Figure 6. (first row), total VOC emissions (i.e., the a priori), constrained by the satellite observations (i.e., the a posteriori) in May-June 2016, and the ratio of the a posteriori to the a priori. (second row) the errors in the a priori based on Table 3, the errors in the top-down estimation, and the averaging kernels (AKs) obtained from the estimation.
Figure 7. (from left to right), HCHO total columns from OMPS, the WRF-CMAQ simulated with the prior emissions, and the same model but with the top-down emissions constrained by the satellite in May-June 2016.

Figure 8. (from left to right), ratio of LNO\textsubscript{x}/LRO\textsubscript{x} simulated by the prior and the posterior emissions, and their differences at 1200-1800 CST, averaged over May-June 2016.
Figure 9. Differences between the simulations with the updated emissions and the default ones of six major pathways of ozone production/loss. The time period is May-June 2016, 1200-1800 CST.

Figure 10. Changes in the major chemical pathways of ozone production/loss, and the net of ozone production $P(O_3)$ after updating the emissions. The time period is May-June 2016, 1200-1800 CST.
Figure 11. Simulated MDA8 surface ozone using the updated emissions constrained by OMI/OMPS observations (left), the default ones (middle), and their difference (right) in May-June 2016.