Journal of Geophysical Research: Atmospheres

RESEARCH ARTICLE
10.1029/2018JD028452

Estimator of Surface Ozone Using Formaldehyde and Carbon Monoxide Concentrations Over the Eastern United States in Summer

Ye Cheng[1], Yuhang Wang[1], Yuzhong Zhang[1,2], James H. Crawford[3], Glenn S. Diskin[3], Andrew J. Weinheimer[4], and Alan Fried[5]

[1]School of Earth and Atmospheric Sciences, Georgia Institute of Technology, Atlanta, GA, USA, [2]Now at the School of Engineering and Applied Sciences, Harvard University, Cambridge, MA, USA, [3]NASA Langley Research Center, Hampton, VA, USA, [4]Atmospheric Chemistry Observations and Modeling Laboratory, National Center for Atmospheric Research, Boulder, CO, USA, [5]Institute of Arctic and Alpine Research, University of Colorado Boulder, Boulder, CO, USA

Key Points:
- Observations show robust correlations and consistent linear regression slopes of O_3-CH_2O, O_3-CO and CO-CH_2O
- Biogenic isoprene oxidation makes the largest contribution to the regression slope of O_3-CH_2O across much of the eastern United States
- We build a fast-response O_3 estimator using input near-surface CH_2O and CO concentrations to improve surface O_3 distribution monitoring

Supporting Information:
- Supporting Information S1
- Figure S1

Correspondence to:
Y. Wang,
yuhang.wang@eas.gatech.edu

Citation:
Cheng, Y., Wang, Y., Zhang, Y., Crawford, J. H., Diskin, G. S., Weinheimer, A. J., & Fried, A. (2018). Estimator of surface ozone using formaldehyde and carbon monoxide concentrations over the eastern United States in summer. *Journal of Geophysical Research: Atmospheres*, *123*, 7642–7655. https://doi.org/10.1029/2018JD028452

Received 1 FEB 2018
Accepted 19 JUN 2018
Accepted article online 3 JUL 2018
Published online 20 JUL 2018

Abstract Strong correlations of O_3-CH_2O, O_3-CO and CO-CH_2O were observed during the Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ) aircraft experiment in July 2011 over the Washington-Baltimore area. The linear regression slopes of observed O_3-CH_2O, O_3-CO, and CO-CH_2O do not vary significantly with time (11 a.m. to 4 p.m.) or altitude in the boundary layer. These observed relationships are simulated well by a regional chemical transport model. Using tagged-tracer simulations, we find that biogenic isoprene oxidation makes the largest contribution to the regression slope of O_3-CH_2O across much of the eastern United States, providing a good indicator for O_3 enhanced by biogenic isoprene oxidation. In contrast, the regression slope of O_3-CO is controlled by both anthropogenic and biogenic emissions. Therefore, we use the CO-CH_2O relationship to separate biogenic from anthropogenic contributions to CO. By combining these regressions, we can track the contributions to surface O_3 by anthropogenic and biogenic factors and build a fast-response ozone estimator using near-surface CH_2O and CO concentrations as inputs. We examine the quality of O_3 estimator by increasing or decreasing anthropogenic emissions by up to 50%. The estimated O_3 distribution is in reasonably good agreement with the full-model simulations ($R^2 > 0.77$ in the range of −30% to +50% of anthropogenic emissions). The analysis provides the basis for using high-quality geostationary satellites with UV, thermal infrared, or near-infrared instruments for observing CH_2O and CO to improve surface O_3 distribution monitoring. The estimation model can also be applied to derive observation-derived regional metrics to evaluate and improve full-fledged 3-D air quality models.

1. Introduction

Ozone (O_3) is a major pollutant in the troposphere (e.g., Lelieveld & Dentener, 2000; Logan et al., 1981; Wang et al., 1998; Wang & Jacob, 1998). Thus, monitoring tropospheric O_3 at regional and global scales is important for environmental protection. Spaceborne remote sensing utilizing its absorption features in the ultraviolet (UV) and the thermal infrared (TIR) bands is the most convenient way to provide O_3 spatial distributions around the globe. However, because of the molecular scattering of UV (X. Liu, Bhartia, et al., 2010) and lack of contrast of TIR (Beer, 2006), the satellite observations for O_3 still show a limited sensitivity in the lowermost troposphere, especially near the surface, which is directly relevant to air quality (e.g., Cuesta et al., 2013).

Three-dimensional air quality models can provide information on the distribution of surface O_3, but the accuracy of surface O_3 simulations is limited by uncertainties in precursor emissions, atmospheric processes, and nonlinear photochemistry. Some model uncertainties can be mitigated through probabilistic approaches (e.g., Dabberdt et al., 2004; Delle Monache, Deng, et al., 2006; Delle Monache, Hacker, et al., 2006; Vautard et al., 2009). Statistical methods have also been applied to estimate O_3 distributions, including classification and regression trees, linear regression, and neural networks (e.g., Biancofiore et al., 2015; Burrows et al., 1995; Cobourn, 2007; Perez & Reyes, 2006; Shad et al., 2009; Van der Wal & Janssen, 2000). The advantage of such statistical modeling is that it offers moderate to high accuracy at a moderate cost (Y. Zhang et al., 2012). However, the nature of statistical modeling often requires a suite of input variables and does not enable better understanding of chemical and physical processes (e.g., Guillas et al., 2008).

An alternative to numerical or statistical modeling is to make use of the observations of O_3 surrogates. The correlations between O_3 and other chemical species might be used effectively to diagnose O_3 chemical and physical processes (e.g., Chin et al., 1994; Koo et al., 2012; Parrish et al., 1993; Wang & Zeng, 2004). Carbon monoxide (CO) is often observed and simulated to have a linear relationship with O_3 in the lower atmosphere (e.g., Buhr et al., 1996; Cardenas et al., 1998; Cheng et al., 2017; Chin et al., 1994; Cooper, Moody, Parrish, Trainer, Holloway, et al., 2002; Cooper, Moody, Parrish, Trainer, Ryerson, et al., 2002; Fishman & Seiler, 1983; Honrath et al., 2004; Huntrieser et al., 2005; Q. Li et al., 2002; Mao & Talbot, 2004; Parrish et al., 1993, 1998). Over the eastern United States, the observed slope of O_3 to CO at ~0.3 reflects contributions by CO from primary anthropogenic emissions and that from biogenic isoprene oxidation (Cheng et al., 2017) due in part to the production of both O_3 and CO from biogenic volatile organic compound (VOC) oxidation (Atkinson & Arey, 1998; Choi et al., 2010; Geng et al., 2011; Guenther et al., 1995; Hudman et al., 2008; K.-Y. Lee et al., 2014; Pang et al., 2009; Pierce et al., 1998; Y. Zhang & Wang, 2016). It implies that the observations of CO concentrations can be potentially applied to track the contributions to surface O_3 by anthropogenic and biogenic factors. In addition to Environmental Protection Agency surface monitoring networks, satellite observations of lower tropospheric CO are more promising than O_3 over polluted regions since CO concentrations are usually higher in the boundary layer than the free troposphere and unlike O_3 it does not have high concentrations in the stratosphere, although both CO and O_3 have substantial free tropospheric columns relative to the boundary layer. CO can be detected by satellite TIR, near-infrared (NIR), and joint TIR and NIR instruments, such as the NIR SCanning Imaging Absorption spectroMeter for Atmospheric CHartographY (SCIAMACHY; e.g., De Laat et al., 2012), the Measurement of Pollution in the Troposphere (e.g., Emmons et al., 2004; Straume et al., 2005), and Atmospheric Infrared Sounder Aumann et al., 2003).

Formaldehyde (CH_2O) is a principal intermediate species in the oxidation of atmospheric hydrocarbons (e.g., Duane et al., 2002; Fried et al., 2011; Pang et al., 2009; Wiedinmyer et al., 2001, and references therein). It is also a major radical source leading to ozone production in the presence of nitrogen oxides (NO_x; e.g., Z. Liu, Wang, Vrekoussis, et al., 2012). We will show that regional CH_2O is also correlated to O_3 in section 3. Since the sources of CO and CH_2O are often different, the two correlations can provide separate constraints on O_3 distributions. While regulatory monitoring of surface CH_2O is unavailable, CH_2O is detectable from space with good sensitivities in the boundary layer by measuring backscattered solar UV between 325 and 360 nm (Chance et al., 2000), including SCIAMACHY (Wittrock et al., 2006), Ozone Monitoring Instrument (González Abad et al., 2015; Kurosu et al., 2004), Global Ozone Monitoring Experiment-2 (GOME-2; De Smedt et al., 2012), and Ozone Mapping Profiler Suite (González Abad et al., 2016; C. Li et al., 2015). The total uncertainty of the HCHO vertical column data is typically in the range of 50–105% for each measurement. Through averaging, the uncertainties for monthly means are down to 20–40% for GOME-2A and SCIAMACHY (De Smedt et al., 2008), 38% for Ozone Monitoring Instrument, and 46% for GOME-2B (Zhu et al., 2016). Therefore, the uncertainty of the monthly HCHO vertical column data from polar-orbiting instruments are in the range of 20–50%.

In this study, we apply a 3-D chemical transport model to quantitatively study factors contributing to the observed correlations and regression slopes of O_3 with CH_2O and CO and those of CO with CH_2O using the Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality (DISCOVER-AQ) measurements. The analysis will show that it is feasible to use observed aircraft CO and CH_2O concentrations to improve estimates of the surface O_3 distribution over the eastern United States. The uncertainties of current satellite-derived near-surface CO and CH_2O are still too large (e.g., Buchwitz et al., 2007; Gloudemans et al., 2005; L. Zhang, Jiang, et al., 2016) to be applied in this method. The Environmental Protection Agency surface CO monitoring data are also unusable since CO concentrations are often below the reporting limit (e.g., Zeng & Wang, 2011). We therefore use the model simulated data, which reasonably captures the observed aircraft concentrations and relationships of O_3, CO, and CH_2O during the DISCOVER-AQ campaign, to develop and evaluate a surface O_3 estimation model with surface CO and CH_2O as input parameters. In the future, geostationary satellite instruments such as Tropospheric Emissions: Monitoring of Pollution (TEMPO; Chance et al., 2013) and Geostationary Carbon Observatory (Polonsky et al., 2014) will greatly improve the monitoring of near-surface measurements of O_3 precursors with sufficient accuracy and therefore make it possible to use observed CH_2O and CO in the O_3 estimator we develop here. We describe the 3-D chemical transport model, DISCOVER-AQ data set,

analysis methods, the O_3 estimation model, and validation method in section 2. Section 3 describes the analysis, modeling, and evaluation results. Discussion of implementing the O_3 estimator and conclusions are given in sections 4 and 5, respectively.

2. Data and Analysis Methods

2.1. Observations and Simulations

The observation data used in this study were obtained from the NASA 2011 DISCOVER-AQ airborne campaign (http://www-air.larc.nasa.gov/missions/discover-aq/discover-aq.html). Sampling by the NASA P-3B aircraft was conducted from Washington's Beltway northeast to Baltimore and continuing on to the Delaware state line and occasionally over the Chesapeake Bay. Fourteen flights over six locations selected for aircraft spirals were carried out to measure the vertical structures of pollutants. Two hundred fifty-three daytime vertical profiles from 300 m to 5 km were measured between 27 June and 31 July. CO was measured by a diode laser spectrometer (Sachse et al., 1987). O_3 was measured by the National Center for Atmospheric Research four-channel chemiluminescence instrument (Weibring et al., 2010). The uncertainties of the measurements on these two species are 2% and 5%, respectively. CH_2O was measured by a difference frequency generation absorption spectrometer (Weibring et al., 2010). For CH_2O levels above 1 ppbv the total measurement uncertainty at the 1σ level was estimate to be around 5%, which folds in systematic and limits of detection uncertainties. To evaluate model simulations with the observations, we identify the model profiles corresponding to the locations of aircraft spirals and the time of aircraft sampling. Corresponding model vertical profiles and observations are used in correlation analysis to evaluate model performance.

We use a 3-D Regional chEmical trAnsport Model (REAM) to represent the observations and then conduct further correlation analyses. The REAM model was applied in previous studies to analyze vertical mixing, large-scale transport, emission estimates, and tropospheric chemistry over North America and East Asia (e.g., Cheng et al., 2017; Choi et al., 2005; Choi, Wang, Yang, et al., 2008; Choi, Wang, Zeng, et al., 2008; Gu et al., 2013, 2014, 2016; Jing et al., 2006; Z. Liu, Wang, et al., 2010; Z. Liu, Wang, Gu, et al., 2012; Z. Liu, Wang, Vrekoussis, et al., 2012; Z. Liu et al., 2014; Wang et al., 2006, 2007; Yang et al., 2011; Zeng et al., 2003, 2006; Zhao & Wang, 2009; Zhao, Wang, Choi, & Zeng, 2009; Zhao, Wang, & Zeng, 2009; Zhao et al., 2010; Y. Zhang, Wang, et al., 2016; Y. Zhang & Wang, 2016; R. Zhang et al., 2017). The model domain covers the contiguous United States with a horizontal resolution of 36×36 km^2. The chemistry mechanism in REAM is the GEOS-Chem standard chemical mechanism (V9-02; Bey et al., 2001) with updates of kinetics data (http://jpldataeval.jpl.nasa.gov). The anthropogenic emission inventory used in the model is the 2011 National Emission Inventory (https://www.epa.gov/air-emissions-inventories/2011-national-emissions-inventory-nei-data). The biogenic isoprene emissions are the results of the Model of Emissions of Gases and Aerosols from Nature version 2.1 (Guenther et al., 2012). Initial and boundary conditions for chemical tracers are taken from the GEOS-Chem (V9-02) 2°× 2.5°simulation results (Bey et al., 2001). Meteorology fields are from the Weather Research and Forecasting model, which assimilated products from the Climate Forecast System Reanalysis (http://cfs.ncep.noaa.gov/cfsr).

The previous study by Cheng et al. (2017) shows that REAM simulates well the observed vertical and temporal variations of O_3, CO, NO_x, isoprene, and CH_2O, as well as the correlation between O_3 and CO, and the O_3-CO regression slopes during the 2011 DISCOVER-AQ campaign. In this work, we therefore focus on analyzing the correlation and regression slope of O_3-CH_2O. We trace separately via tagged tracers three different CH_2O sources, primary anthropogenic emissions and the oxidation of anthropogenic VOCs ($CH_2O_{anthroVOCs}$), the oxidation of biogenic isoprene ($CH_2O_{biolSOP}$), and transport from model lateral and upper boundaries (CH_2O_{BC}), to analyze the contribution from each source to the observed O_3-CH_2O relationship. Other biogenic VOCs are not taken into account because isoprene provides the source for the vast majority biogenic CH_2O (e.g., Guenther et al., 2012; Kesselmeier & Staudt, 1999; Lathiere et al., 2006; Sindelarova et al., 2014). In tagged-tracer simulations, relevant species and radicals, such as O_3, NO_x, and HO_x (OH and HO_2), are fixed using results archived from the standard simulation. The sum of the three individual tagged tracers is within 2% of the total CH_2O concentrations in the standard simulation for grid cells over the Washington-Baltimore region. We carry out minor scaling adjustments in postprocessing, assuming that relative CH_2O attributions stay the same, to ensure that the sum of the CH_2O tracers is the same as the total CH_2O for each grid cell in the standard simulation. We evaluate the scaling adjustments by calculating the relative error of the sum of

the three tagged CH_2O to the total simulated CH_2O (Figure S1 in the supporting information). In the study domain (mainly the southeast United States), the error is <0.5%. Since the magnitude of total CH_2O is lower in other areas, the error is lager but is still <1.5% for western United States and 3% for other rural regions. With simulated CH_2O attribution results, we can decompose the O_3-CH_2O regression slope into three subslopes of the corresponding CH_2O tracers (equation (1), derived in Appendix A),

$$\begin{aligned}&\text{Least squares regression slope of } O_3\text{-}CH_2O\\&=\frac{\mathrm{Cov}(CH_2O_{\mathrm{anthroVOCs}},O_3)}{\mathrm{Var}(CH_2O_{\mathrm{total}})}+\frac{\mathrm{Cov}(CH_2O_{\mathrm{BC}},O_3)}{\mathrm{Var}(CH_2O_{\mathrm{total}})}+\frac{\mathrm{Cov}(CH_2O_{\mathrm{biolSOP}},O_3)}{\mathrm{Var}(CH_2O_{\mathrm{total}})},\end{aligned}\tag{1}$$

where Cov and Var denote covariance and variance, respectively. Equation (1) shows that the contribution of each CH_2O tracer, that is, the subslope values, to the O_3-CH_2O regression slope is proportional to its covariance with O_3. It is therefore possible to have both positive and negative slope contributions.

For the DISCOVER-AQ region during the Baltimore-Washington study, where the majority of CH_2O is biogenic, we can use the regression slope of O_3-CH_2O in the evaluation of model results using the observations. When extending the analysis using equation (1) to remote regions, the small variance of background CH_2O leads to abnormally large slopes, making it difficult to show the spatial distribution of the regression slope. We therefore use an *inversed slope* of $\Delta CH_2O/\Delta O_3$ (equation (2), see the Appendix A) to illustrate the spatial distribution over the United States since the variance of O_3 is a more stable denominator that that of CH_2O,

$$\begin{aligned}&\text{Least squares regression slope of } CH_2O\text{-}O_3\\&=\frac{\mathrm{Cov}(CH_2O_{\mathrm{anthroVOCs}},O_3)}{\mathrm{Var}(O_3)}+\frac{\mathrm{Cov}(CH_2O_{\mathrm{BC}},O_3)}{\mathrm{Var}(O_3)}+\frac{\mathrm{Cov}(CH_2O_{\mathrm{biolSOP}},O_3)}{\mathrm{Var}(O_3)}.\end{aligned}\tag{2}$$

2.2. Surface Ozone Distribution Estimation

The total O_3 concentration is contributed by three major sources: anthropogenic O_3 production ($O_{3\mathrm{anthro}}$), biogenic O_3 production ($O_{3\mathrm{biolSOP}}$), and the transport from the lateral and upper model boundaries ($O_{3\mathrm{background}}$; equation (3)). In the ozone estimation model, we estimate $O_{3\mathrm{anthro}}$ and $O_{3\mathrm{biolSOP}}$ using the regressions of O_3 with source-tagged CO and CH_2O. We will show in the next section that the correlations and regression slopes of O_3-CH_2O and CO-CH_2O are almost entirely due to biogenic isoprene over the eastern United States, where CH_2O concentrations are mostly due to oxidation of biogenic isoprene. We make use of this finding and use CH_2O as a proxy for O_3 related to biogenic emissions. We decompose surface O_3 concentrations in equation (3) into three components related to regional anthropogenic emissions, biogenic emissions, and background (not related to the emissions within the estimation domain). We approximate the emission related components using CO (equation (4)). Recognizing that we would like to make use of observation-based CO and CH_2O concentrations, we replace CO from anthropogenic emissions (CO_{anthro}) with CO-CO_{biogenic}-$CO_{\mathrm{background}}$, where CO_{biogenic} is CO from biogenic isoprene oxidation that proceeds through CH_2O, and further compute CO_{biogenic} as a function of CH_2O from biogenic isoprene oxidation (CH_2O_{biolSOP}; equation (5)). Considering that most of CH_2O is biogenic over the region in the summer, we replace CH_2O_{biolSOP} as CH_2O-$CH_2O_{\mathrm{background}}$ (equation (6)).

$$[O_3]=[O_3]_{\mathrm{anthro}}+[O_3]_{\mathrm{bigenic}}+[O_3]_{\mathrm{background}}\tag{3}$$

$$\approx\frac{\Delta[O_3]}{\Delta[CO]_{\mathrm{anthro}}}\times[CO]_{\mathrm{anthro}}+\frac{\Delta[O_3]}{\Delta[CO]_{\mathrm{biolSOP}}}\times[CO]_{\mathrm{biolSOP}}+[O_3]_{\mathrm{background}}\tag{4}$$

$$\begin{aligned}&=\frac{\Delta[O_3]}{\Delta[CO]_{\mathrm{anthro}}}\times\left([CO]_{\mathrm{total}}-\frac{\Delta[CO]_{\mathrm{biolSOP}}}{\Delta[CH_2O]_{\mathrm{biolSOP}}}\times[CH_2O]_{\mathrm{biolSOP}}-[CO]_{\mathrm{background}}\right)\\&+\frac{\Delta[O_3]}{\Delta[CO]_{\mathrm{biolSOP}}}\times\frac{\Delta[CO]_{\mathrm{biolSOP}}}{\Delta[CH_2O]_{\mathrm{biolSOP}}}\times[CH_2O]_{\mathrm{biolSOP}}+[O_3]_{\mathrm{background}}\end{aligned}\tag{5}$$

$$\approx \frac{\Delta[O_3]}{\Delta[CO]_{anthro}} \times \left([CO]_{total} - \frac{\Delta[CO]_{bioISOP}}{\Delta[CH_2O]_{bioISOP}} \times \left([CH_2O]_{total} - [CH_2O]_{background} \right) - [CO]_{background} \right) + \frac{\Delta[O_3]}{\Delta[CH_2O]_{bioISOP}} \times \left([CH_2O]_{total} - [CH_2O]_{background} \right) + [O_3]_{background} \quad (6)$$

$$= A \times [CO]_{total} + B \times [CH_2O]_{total} + C \quad (7)$$

where $A = \frac{\Delta[O_3]}{\Delta[CO]_{anthro}}$, $B = -\frac{\Delta[O_3]}{\Delta[CO]_{anthro}} \times \frac{\Delta[CO]_{bioISOP}}{\Delta[CH_2O]_{bioISOP}} + \frac{\Delta[O_3]}{\Delta[CH_2O]_{bioISOP}}$, $C = -\frac{\Delta[O_3]}{\Delta[CO]_{anthro}} \times [CO]_{background} + \left(\frac{\Delta[O_3]}{\Delta[CO]_{anthro}} \times \frac{\Delta[CO]_{bioISOP}}{\Delta[CH_2O]_{bioISOP}} - \frac{\Delta[O_3]}{\Delta[CH_2O]_{bioISOP}} \right) \times [CH_2O]_{background} + [O_3]_{background}$.

We will show that this function works well for the eastern United States in the summer. The study domain (to be shown in Figure 4) is selected where >90% of monthly mean surface CH_2O is biogenic. Over the domain of surface O_3 estimation, we compute the $\frac{\Delta[O_3]}{\Delta[CO]_{anthro}}$, $\frac{\Delta[O_3]}{\Delta[CO]_{bioISOP}}$, and $\frac{\Delta[CO]_{bioISOP}}{\Delta[CH_2O]_{bioISOP}}$ values using the least squares regression slope formulations of O_3-CO by Cheng et al. (2017) and O_3-CH_2O in equation (1). The tagged-tracer simulations show that background values do not have a significant spatial dependence and we use a value of 60 ppbv for $CO_{background}$ and a value of 200 pptv for $CH_2O_{background}$. Regression of equation (6) yields a value of 10 ppbv for $O_{3background}$. As we will discuss in section 4, the unique feature of this surface O_3 estimation is that it is based on the temporospatial stability of the predicting parameters of the regression slopes and background values. To demonstrate the feasibility of this O_3 estimator, for hourly estimation from 11:00 a.m. to 4:00 p.m., we group the corresponding model data in all grid cells in the same hour and compute the regional regression slope values. Therefore, the estimated O_3 spatial variation is due to those of CO and CH_2O only. In order to test the reliability, the above estimation method is validated through *leave-one-out cross validation* (LOOCV). The method uses a single sample from the original data sets as the validation data, and the remaining samples (excluding the selected validation data point) are used in the estimator. Each sample in the data sets is used once as validation data. We conduct the validation individually for each hour of a day from 11:00 a.m. to 4:00 p.m. Of the 31 days from 1 to 31 July (d_1, d_2... d_{31}), the day i (d_i) is selected as the validation data, and the remaining days (d_1, d_2... $d_{(i-1)}$, $d_{(i+1)}$... d_{31}) are used to compute the regression slopes and estimate the surface O_3 distribution of day i. We exclude the data before 11:00 a.m. or after 4 p.m. when the estimation has large biases because photochemical production of O_3, CO, and CH_2O is still slow and the correlations among the species are not photochemically driven.

3. Results

3.1. Correlations and Regression Slopes of O_3-CH_2O and CO-CH_2O

We compare simulated O_3-CH_2O correlations to the DISCOVER-AQ observations as a function of altitude or as a function of time in Figure 1. Simulated and observed correlation coefficient (R) values are in good agreement. Strong correlation between O_3 and CH_2O is found from 300 m to the top of the boundary layer (~2.5 km) with a narrow range of R values (~0.75). Near the surface, simulated R values show a slight decrease due to the increase of the contribution of surface primary emissions of CH_2O, which are not directly related to photochemical O_3 production. From the upper boundary layer to the free troposphere, R values show a drastic drop and changes sign from positive to negative. The sign change reflects the rapid decrease of CH_2O with altitude in the lower free troposphere where O_3 increases with altitude (Figure S2 in the supporting information). In the middle and upper boundary layers, the relative contribution by photochemical CH_2O production increases with altitude. Therefore, the concurrent photochemical production of O_3 and CH_2O is a major factor contributing to the observed positive correlation between O_3 and CH_2O in the boundary layer. We also compare the simulated and observed O_3-CH_2O correlation coefficient as a function of time of the day with data from 300 m to 2.5 km (Figure 1b). These R values represent the spatial correlation in a given hour and are somewhat lower than the spatial correlations in the vertical just discussed. The model is in good agreement with the observations except the underestimation at 9 a.m. and overestimation at 5 p.m., which again is related to the fact that photochemistry becomes less important here.

The observed regression slope of O_3 to CH_2O is also captured by model simulation with satisfactory agreement in both vertical distribution and diurnal variation (Figures 1c and 1d). The O_3-CH_2O regression slope at a given altitude is underestimated by the model at 0.3–2.5 km by ~15%. However, the observed slight

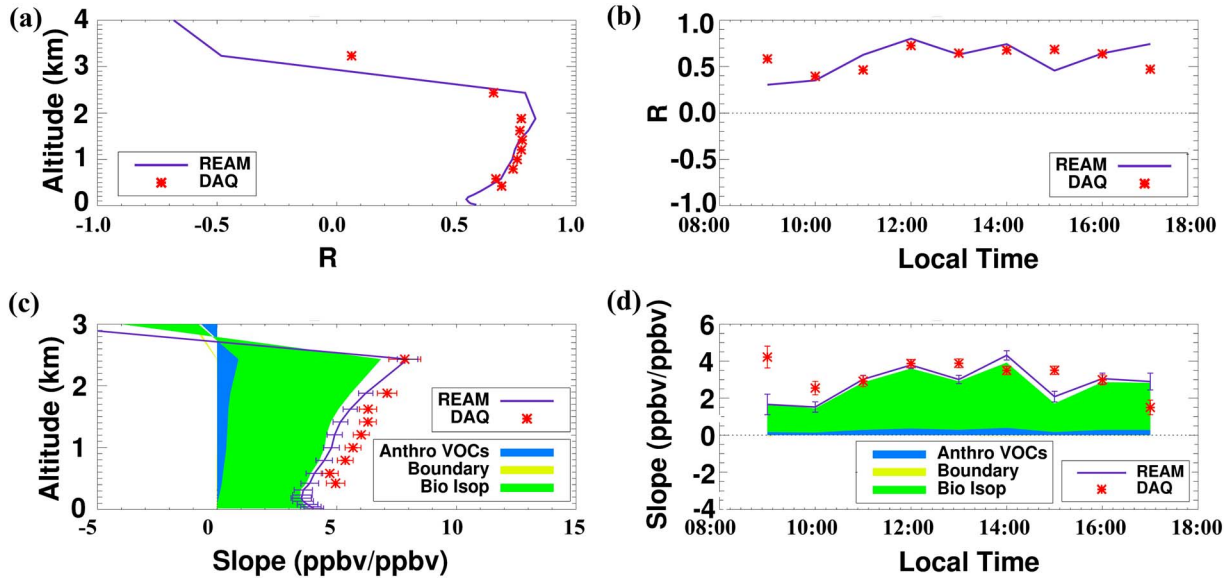

Figure 1. Observed and simulated O_3-CH_2O correlation coefficients (*R*) and regression slope and subslopes (equation (1)) as a function of altitude (of data for 11 a.m. to 4 p.m.; a, c) and as a function of local time (of data for altitude of 0.3–2.5 km; b, d). The subslopes due to varied CH_2O sources are shown using areas filled with different colors (equation (1)). The legends for different CH_2O sources are the same as in equation (1). The horizontal bars in (c) and vertical bars in (d) show the observed or simulated standard deviations of the regression slopes. The *R*, slope, and subslope values are computed using the DISCOVER-AQ (DAQ) observations or corresponding model data at a given altitude bin or for a given period. DISCOVER-AQ = Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality; REAM = Regional chEmical trAnsport Model.

increase (from ~5 to ~8 ppbv/ppbv) in the boundary is captured by the model. The observed O_3-CH_2O regression slope at a given hour is better simulated than that at a given altitude. We further quantify source contributions by decomposing the O_3-CH_2O regression slope into the different CH_2O sources using equation (1). During DISCOVER-AQ, the regression slope of O_3-CH_2O is contributed almost exclusively by that of biogenic CH_2O.

Over the United States, we use the model results in Figure 2 to understand the relative contributions of CH_2O sources to the CH_2O-O_3 regression slope near the surface (equation (2)). The slope decomposition results show that the contribution from biogenic isoprene to the slope of O_3 to CH_2O is overwhelming over most regions of the eastern United States. Over the regions where it dominates, biogenic CH_2O has positive correlations with O_3 due to the concurrent production of these two species from the oxidation of biogenic VOCs. The exception is in central United States, where biogenic CH_2O concentrations are high but O_3 concentrations are low due to low NO_x concentrations. The chemical loss of O_3 leads to a negative regression slope, which appears high because of the low variance of O_3 in the region (equation (2)). Without significant photochemical production, the variation of O_3 in this region is low and is therefore relatively easier to estimate than the other regions. Overall, the short lifetimes of biogenic isoprene and CH_2O make it possible to use its concentrations to estimate the spatial variation of O_3 using relationships like equation (7).

We also compare observed and simulated correlation and regression slopes of CO and CH_2O using an equation similar to equation (1) for the Baltimore-Washington area. This information is also used in equation (7). The simulation results are in good agreement with the observations as a function of altitude or time of a day (Figure 3). Strong correlation between CO and CH_2O (R = ~0.85) is observed and simulated from 300 m to 2.5 km (Figure 3a). This correlation is most likely due to the coemissions from anthropogenic sources and the coproduction of mostly biogenic CO and CH_2O. In this altitude range, the regression slope of CO to CH_2O is about ~20 ppbv/ppbv without little variation between 9 a.m. and 5 p.m. (Figure 3c). Below 300 m, simulated *R* value decreases toward the surface because the contribution of CO primary emissions increases significantly near the surface while the secondary formation is still the major source of CH_2O (Figure S3 in the supporting information). These two processes are not correlated. Correspondingly, the slope of CO to CH_2O increases from 200 m to surface because the gradient of CO is larger than that of CH_2O (Figure S2) due to the

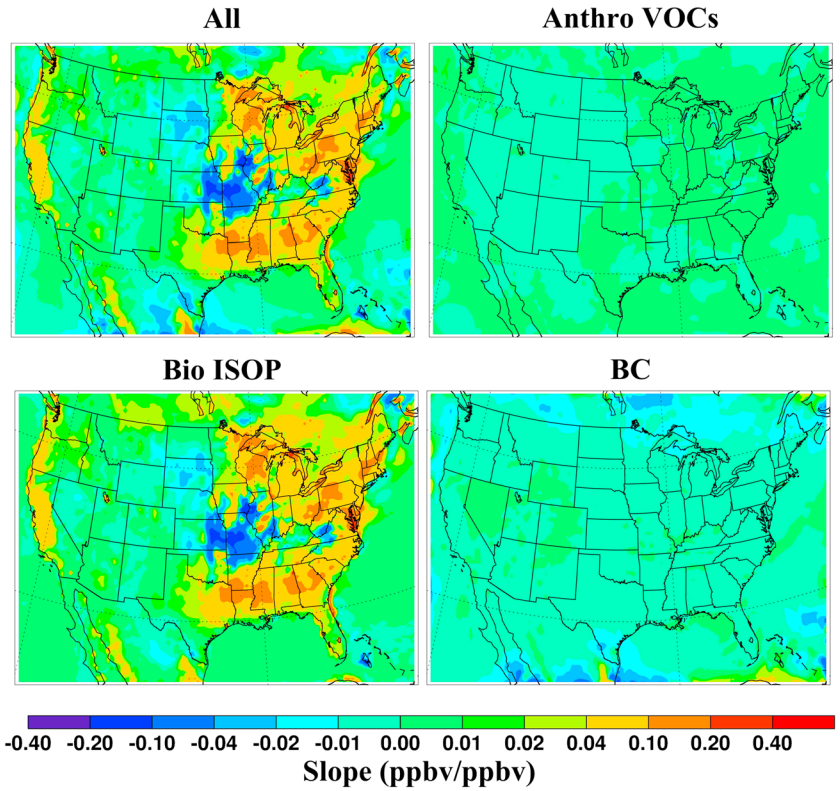

Figure 2. Distribution of the regression slopes and subslopes of surface CH_2O-O_3 (equation (2)) over the United States. The title *All* stands for the slope computed by total O_3 and total CH_2O concentration; the titles of *Bio ISOP*, *Anthro VOCs*, and *BC* stand for the subslopes computed by total O_3 and CH_2O from primary anthropogenic emissions and oxidation of anthropogenic VOCs, oxidation of biogenic isoprene, and transport from the lateral and upper model boundaries, respectively. The slope and subslope values on each grid are computed using selected hourly data on daytime of 11: 00 a.m. to 04: 00 p.m. from 1 to 30 July. VOCs = volatile organic compounds.

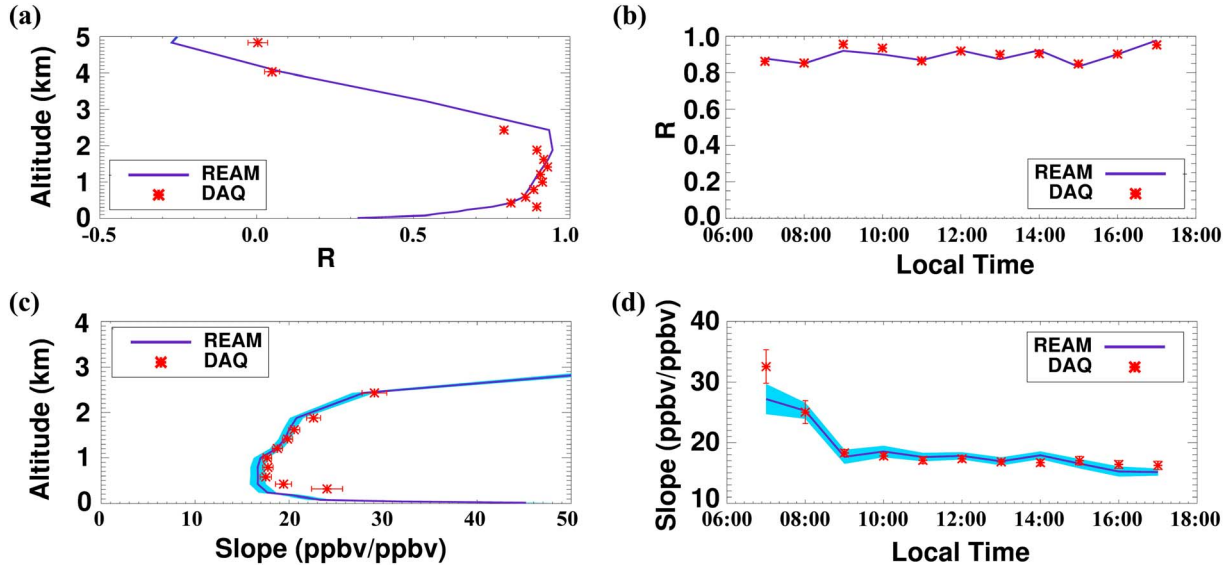

Figure 3. Observed and simulated CO-CH_2O correlation coefficients (R) and regression slope for the Baltimore-Washington area as a function of altitude for daytime of 11:00 a.m. to 04:00 p.m. (a, c) and local time for altitude of 0.3–2.5 km (b, d). The horizontal bars in (c) and vertical bars in (d) show the observed standard deviations of the regression slopes. Shaded blue areas in (c) and (d) show simulated standard deviations of the regression slopes. The R and slope values are computed using the DISCOVER-AQ (DAQ) observations or corresponding model data at a given altitude bin or for a given time period. DISCOVER-AQ = Deriving Information on Surface Conditions from Column and Vertically Resolved Observations Relevant to Air Quality; REAM = Regional chEmical trAnsport Model..

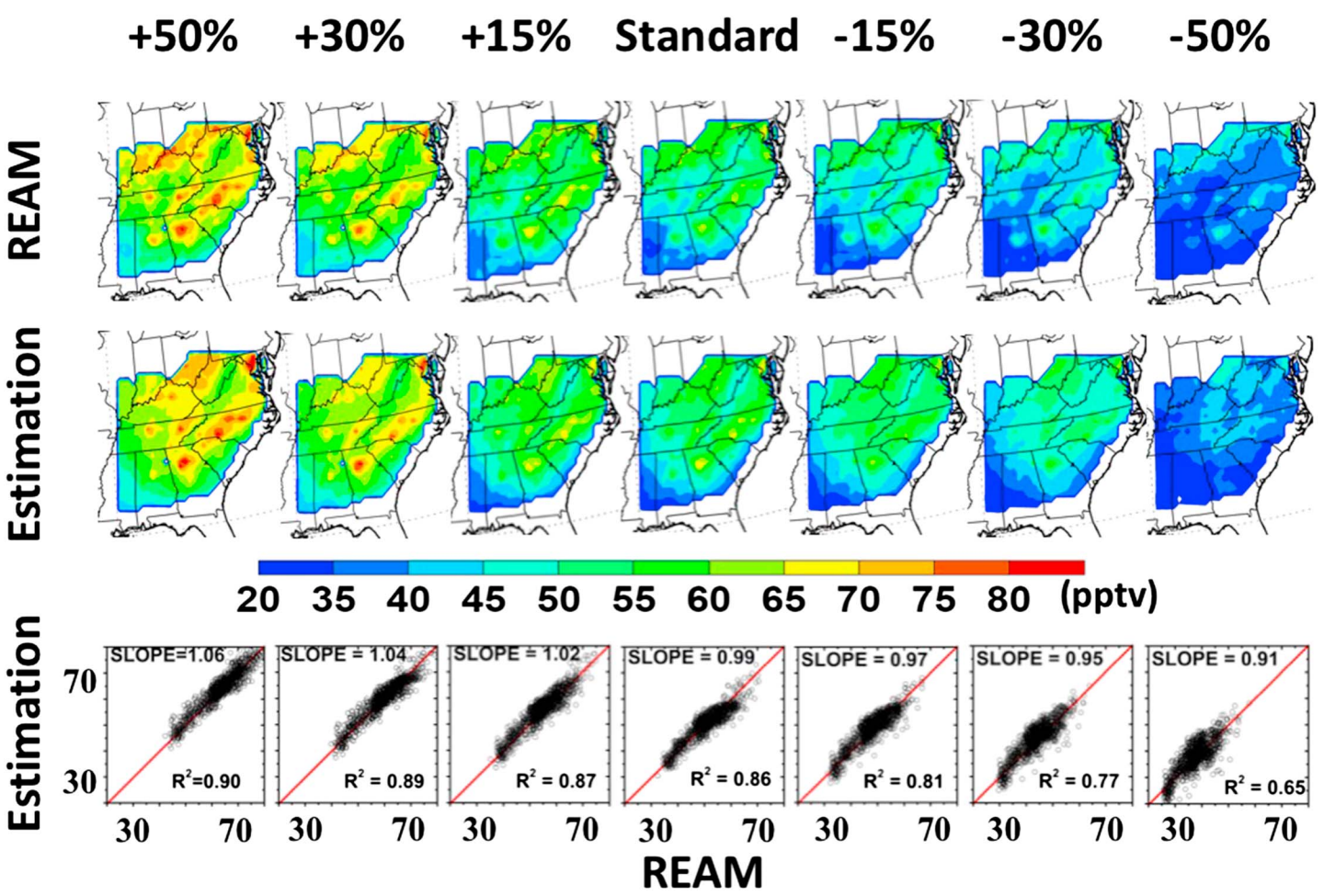

Figure 4. Distributions of monthly mean (11:00 a.m. to 4:00 p.m.) REAM and LOOCV estimation of surface O_3 concentrations for July 2011 under different anthropogenic emission scenarios. In LOOVC hourly estimation, the regional parameters in equation (7) are estimated using data not including the day of estimation. Scatterplots of corresponding grid-cell hourly REAM and LOOCV estimation data are shown in the third row; the 1:1 line is shown in red. Seven emission scenarios are presented. *+50%*, *+30%*, *+15%*, *Standard*, *−15%*, *−30%*, and *−50%* on the top of columns denote 150%, 130%, 115%, original, 85%, 70%, and 50% of anthropogenic CO and NO_x emissions of the 2011 National Emission Inventory. The comparison statistics are listed in Table 1. REAM = Regional chEmical trAnsport Model; LOOCV = leave-one-out cross validation

surface primary emissions. However, we also test the correlation of surface $CO_{biogenic}$ and CH_2O, which shows an *R* value in a range of 0.85 to 0.95 in the areas where more than 90% monthly mean CH_2O is from biogenic isoprene oxidation. It implies that CH_2O and the correlation of CO-CH_2O can be used to separate anthropogenic and biogenic CO. Above 2.5 km, the *R* value decreases from the boundary layer top to the free troposphere. This is because above the boundary layer CO is mainly from transport from lateral and upper boundaries, which does not contribute to CH_2O as much due to its short chemical lifetime (Figure S3). The slope of CO to CH_2O also increases from the boundary layer to the free troposphere due to low concentrations of CH_2O in free troposphere. As a function of time, the *R* value does not show a significant variation. The regression slope of CO-CH_2O remains at ~20 ppbv/ppbv in daytime except higher values in the morning (before 8: 00 a.m.) when photochemistry is weak.

3.2. Surface Ozone Estimation Using Equation (7)

For July 2011, we estimate the LOOCV surface O_3 distribution using equation (7). The averaged parameters of equation (7) for the 30 validations are listed in Table S2 in the supporting information. To examine the sensitivity of the estimation to emissions, we also increase (or decrease) anthropogenic emissions by 15%, 30%, and 50%, respectively. The full REAM model is run with different emissions. For the LOOCV estimation

Table 1
Comparison of Hourly REAM and LOOCV Estimation of Surface O_3 Concentrations Under Different Anthropogenic Emissions Scenarios

Emission scenarios	REAM mean ±Std (ppbv)	Estimation mean ±Std (ppbv)	Mean Bias	MSE	R^2	Least squares slope (ppbv/ppbv)
+50%	63.0 ± 8.2	64.1 ± 8.1	1.1	7.6	0.90	1.06
+30%	59.0 ± 7.8	59.9 ± 7.6	0.9	7.1	0.89	1.04
+15%	53.9 ± 7.5	54.6 ±7.5	0.7	8.5	0.87	1.02
Standard	51.4 ±7.0	51.8 ± 6.9	0.4	7.8	0.86	0.99
−15%	48.2 ± 6.8	47.2 ±6.8	−1.0	9.5	0.81	0.97
−30%	44.3 ± 6.6	43.2 ± 6.3	−1.1	11.0	0.77	0.95
−50%	38.8 ± 6.2	37.3 ± 5.5	−1.5	13.6	0.65	0.91

Note. All grids in the domain in all certain hours (11:00 a.m. to 4:00 p.m.) are grouped together. REAM = Regional chEmical trAnsport Model; LOOCV = leave-one-out cross validation; Std = standard deviation; MSE = mean squared error.

using equation (7), the $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{anthro}}}$, $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}$, and $\frac{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}{\Delta[\mathbf{CH_2O}]_{\mathbf{bioISOP}}}$ values are the same as in the standard simulation. The only changes are for the CO and CH_2O surface concentrations used. The premise is that the regression slopes of $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{anthro}}}$, $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}$, and $\frac{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}{\Delta[\mathbf{CH_2O}]_{\mathbf{bioISOP}}}$ are relatively stable with respect to emission changes and therefore the estimation model using equation (7) provides a robust means to estimate surface O_3 distribution when the distributions of CO and CH_2O are known.

Since equations (4)–(7) are more accurate when CH_2O is dominated by oxidation of biogenic isoprene, the estimation evaluation is only for the eastern U.S. regions where monthly mean biogenic CH_2O is >90%. We compare the averaged estimation results with REAM results for these regions under different emission scenarios in Figure 4. In the standard simulation, the estimation shows a similar distribution and explains 86% of the variance of the full REAM results (R^2 = 0.86) with no significant overall bias (Table 1), although the scatterplot shows a slight tendency of low biases for O_3 above 60 ppbv with an average of −1.7 ppbv. As the anthropogenic emissions decrease by 15–50%, the REAM model shows a decrease from 3.0 to 12.4 ppbv on average in these regions. The estimation model using equation (7) overestimates the O_3 decrease by 1.0–1.5 ppbv on average and the explained variance decreases from 81% to 65%. When anthropogenic emissions increase by 15%–50%, the full REAM results show surface O_3 increases from 2.7 to 11.8 ppbv. As anthropogenic emissions increase, the O_3 *hot spots* due to urban emissions become more obvious. The estimation model using equation (7) shows similar features but overestimates the O_3 increase by 0.7–1.1 ppbv on average. However, the explained variance (R^2 value) increases from 0.87 to 0.90 due largely to good estimations of urban increases. We also tested the estimation model performance by increasing (or decreasing) biogenic isoprene emissions by 15%, 30%, and 50%, respectively. The full REAM model shows much lower dependence of surface O_3 to biogenic emissions than anthropogenic emissions (Figure S4 in the supporting information). The estimation model shows similar results with R^2 values ranging from 0.77 to 0.86 (Table S1 in the supporting information).

4. Discussion

The surface O_3 estimator (equation (7)) works very well for the regions shown in Figure 4, and, it is also quite robust with R^2 values >0.77 for anthropogenic emissions in the range of −30% to +50% and biogenic emissions of −50% to +50%. Therefore, it has the potential of being used for rapid O_3 distribution assessment if the distributions of surface CO and CH_2O are known. Given the current lack of usable surface CO and CH_2O observations, we cannot test the estimator using observed data. In this analysis, the key parameters of the estimator, that is, the regression slopes of $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{anthro}}}$, $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}$, and $\frac{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}{\Delta[\mathbf{CH_2O}]_{\mathbf{bioISOP}}}$ are based on model simulations, although we show that model simulations are in good agreement with DISCOVER-AQ observations (Figures 1 and 3; Cheng et al., 2017). If the distributions of O_3, CO, and CH_2O are known, equation (6) can be used to obtain the observation-based regression slopes using least squares regression. These parameters can provide insights in understanding of biases of air quality model simulations and be applied to improve the model.

Surface CO measurements are readily obtained if existing instruments are calibrated and the reporting limit is lowered. In situ observation of CH_2O is more complex and expensive than CO. For both species, the high concentrations in the boundary layer and very low concentrations in the stratosphere imply that satellite instruments have better sensitivity to derive their near-surface concentrations than O_3. The relatively large uncertainties of the current generation instruments on Sun-synchronous satellites can be greatly reduced (due in part to the large increase of observation frequency) by instruments on board geostationary satellites such as TEMPO over North America (Chance et al., 2013), SENTINEL-4 over Europe (Ingmann et al., 2012), and GEMS over East Asia (Bak et al., 2013). High-quality CH_2O measurements over the United States will be available from TEMPO. Deployment of near-IR and thermal instruments on geostationary satellites will be needed for improved satellite measurements of near-surface CO. More complex statistical methods can be applied to combine such derived O_3 estimation with in situ surface O_3 observations (Y. Zhang et al., 2018).

The estimator is valid only in the regions where CH_2O is dominated by oxidation of biogenic isoprene during the period of the year when isoprene emissions are large. As anthropogenic emissions are expected to continue decreasing (Cheng et al., 2017), the regions where biogenic CH_2O dominates and the estimator can be applied will increase. There are limitations in the estimator of equation (7). The relatively good performance of the estimation model with changing anthropogenic emissions is not because surface O_3 is insensitive to anthropogenic NO_x emissions. In fact, most of the changes shown in Table 1 are due to NO_x. However, surface NO_x is not as useful a predictor as CO and CH_2O in the formulation of equation (7). O_3 is a secondary pollutant while, NO_x is mostly a primary pollutant in our study region. In comparison, biogenic CO and CH_2O are secondary. As NO_x is oxidized and its concentration decreases, O_3, biogenic CO, and CH_2O are produced and their concentrations increase. The anthropogenic CO is a better tracer for the cumulative effect of O_3 production by anthropogenic NO_x because of its much longer lifetime than NO_x and fast-reacting VOCs (e.g., Chin et al., 1994).

In REAM simulations, we make an implicit assumption that NO_x and CO emission ratios of anthropogenic sources do not change. This assumption is not always valid; for example, NO_x emission reduction from the electric generating utility sources in the past two decades did not reduce CO emissions significantly (https://www.epa.gov/air-emissions-inventories/air-pollutant-emissions-trends-data). When the emission changes of anthropogenic NO_x and CO are sufficiently different, the estimator needs to be reconstructed with updated regression slopes of $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{anthro}}}$, $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}$, and $\frac{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}{\Delta[\mathbf{CH_2O}]_{\mathbf{bioISOP}}}$. Furthermore, if a region is affected by pollutant sources such as fires that are not included in the model, the estimation results will be biased. Another uncertainty of this estimator could come from isoprene chemistry. If a specific isoprene oxidation pathway significantly affects the regression slopes of $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{anthro}}}$, $\frac{\Delta[\mathbf{O_3}]}{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}$, and $\frac{\Delta[\mathbf{CO}]_{\mathbf{bioISOP}}}{\Delta[\mathbf{CH_2O}]_{\mathbf{bioISOP}}}$, equation (6) can be applied to derive these parameters on the basis of the observations, which can then be applied to evaluate the chemical pathway representation in the model.

5. Conclusions

Extensive measurement of O_3, CO, and CH_2O were conducted during the DISCOVER-AQ aircraft experiment in July 2011 over the Washington-Baltimore area. We find strong correlation and stable linear regression slopes of O_3-CH_2O, O_3-CO, and CO-CH_2O with no significant variation with time (11 a.m. to 4 p.m.) or altitude in the boundary layer. The concentrations, correlations, and regression slopes of these tracers are reproduced well by the REAM model. We find that biogenic isoprene oxidation makes most of the contribution to the regression slopes of CH_2O-O_3 in large regions of the eastern United States using the slope decomposition method by tracing separately three different CH_2O sources, including primary anthropogenic emissions and oxidation of anthropogenic VOCs, oxidation of biogenic isoprene, and transport from the lateral and upper model boundaries.

Making use of the robust regression slopes, we construct a surface ozone estimation model using the distributions of CH_2O and CO as input parameters. In this model, CH_2O is used as a proxy to calculate O_3 and CO produced by the oxidation of biogenic VOCs. The estimator can explain >77% of the surface O_3 variance simulated by the full 3-D model in the range of 70% to 150% of the anthropogenic emissions. It provides a fast regional surface O_3 estimation in most regions of the eastern United States in summer where CH_2O is

dominated by oxidation of biogenic isoprene. With high-quality geostationary satellite observations of CO and CH_2O, the estimator could be applied to improve surface O_3 distribution, which is challenging to measure directly from space. The function of equation (6) relates surface concentrations of O_3 to those of CO and CH_2O. Using a sizable observation data set of these concentrations of a given region, regional metrics, such as the regression slopes of $\frac{\Delta[O_3]}{\Delta[CO]_{anthro}}$, $\frac{\Delta[O_3]}{\Delta[CO]_{biolSOP}}$, and $\frac{\Delta[CO]_{biolSOP}}{\Delta[CH_2O]_{biolSOP}}$, can be empirically determined and applied to investigate model performance and biases. Therefore, the estimation model provides the means of using observations to evaluate and improve full-fledged 3-D air quality models.

Appendix A

CH_2O concentration is the sum of that from primary anthropogenic emissions and the oxidation of anthropogenic VOCs ($CH_2O_{anthroVOCs}$), the oxidation of biogenic isoprene ($CH_2O_{biolSOP}$), and transport from model lateral and upper boundaries (CH_2O_{BC}):

$$[CH_2O]_{total} = [CH_2O]_{anthroVOCs} + [CH_2O]_{BC} + [CH_2O]_{biolSOP}. \tag{A1}$$

The slope of O_3 to CH_2O in a least squares regression is thus

$$\text{Least squares regression slope of } O_3\text{-}CH_2O = \frac{\mathrm{Cov}(CH_2O_{totals}, O_3)}{\mathrm{Var}(CH_2O_{total})} \tag{A2}$$

$$= \frac{\overline{\left([CH_2O]_{total} - \overline{[CH_2O]_{total}}\right)\left([O_3] - \overline{[O_3]}\right)}}{\mathrm{Var}(CH_2O_{total})} \tag{A3}$$

$$= \frac{\overline{\left([CH_2O]_{anthroVOCs} + [CH_2O]_{BC} + [CH_2O]_{biolSOP} - \overline{([CH_2O]_{anthroVOCs} + [CH_2O]_{BC} + [CH_2O]_{biolSOP})}\right)\left([O_3] - \overline{[O_3]}\right)}}{\mathrm{Var}(CH_2O_{total})} \tag{A4}$$

$$= \frac{\begin{array}{c}\overline{\left([CH_2O]_{anthroVOCs} - \overline{[CH_2O]_{anthroVOCs}}\right)\left([O_3] - \overline{[O_3]}\right)} \\ +\overline{\left([CH_2O]_{BC} - \overline{[CH_2O]_{BC}}\right)\left([O_3] - \overline{[O_3]}\right)} + \overline{\left([CH_2O]_{biolSOP} - \overline{[CH_2O]_{biolSOP}}\right)\left([O_3] - \overline{[O_3]}\right)}\end{array}}{\mathrm{Var}(CH_2O_{total})} \tag{A5}$$

$$= \frac{\mathrm{Cov}(CH_2O_{anthroVOCs}, O_3)}{\mathrm{Var}(CH_2O_{total})} + \frac{\mathrm{Cov}(CH_2O_{BC}, O_3)}{\mathrm{Var}(CH_2O_{total})} + \frac{\mathrm{Cov}(CH_2O_{biolSOP}, O_3)}{\mathrm{Var}(CH_2O_{total})}, \tag{A6}$$

where $\overline{X}$ donates the average value of X.

Acknowledgments
The research was supported by the NASA ACMAP program. The data for this paper are available at the DISCOVER-AQ data archive (http://www-air.larc.nasa.gov/missions/discover-aq/discover-aq.html).

References

Atkinson, R., & Arey, J. (1998). Atmospheric chemistry of biogenic organic compounds. *Accounts of Chemical Research*, *31*(9), 574–583. https://doi.org/10.1021/ar970143z

Aumann, H. H., Chahine, M. T., Gautier, C., Goldberg, M. D., Kalnay, E., McMillin, L. M., et al. (2003). AIRS/AMSU/HSB on the Aqua mission: Design, science objectives, data products, and processing systems. *IEEE Transactions on Geoscience and Remote Sensing*, *41*(2), 253–264. https://doi.org/10.1109/TGRS.2002.808356

Bak, J., Kim, J. H., Liu, X., Chance, K., & Kim, J. (2013). Evaluation of ozone profile and tropospheric ozone retrievals from GEMS and OMI spectra. *Atmospheric Measurement Techniques*, *6*(2), 239. https://doi.org/10.5194/amtd-5-6733-2012

Beer, R. (2006). TES on the Aura mission: Scientific objectives, measurements, and analysis overview. *IEEE Transactions on Geoscience and Remote Sensing*, *44*(5), 1102–1105. https://doi.org/10.1109/TGRS.2005.863716

Bey, I., Jacob, D. J., Yantosca, R. M., Logan, J. A., Field, B. D., Fiore, A. M., et al. (2001). Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. *Journal of Geophysical Research*, *106*(D19), 23,073–23,095. https://doi.org/10.1029/2001JD000807

Biancofiore, F., Verdecchia, M., Di Carlo, P., Tomassetti, B., Aruffo, E., Busilacchio, M., et al. (2015). Analysis of surface ozone using a recurrent neural network. *Science of the Total Environment*, *514*, 379–387. https://doi.org/10.1016/j.scitotenv.2015.01.106

Buchwitz, M., Khlystova, I., Bovensmann, H., & Burrows, J. P. (2007). Three years of global carbon monoxide from SCIAMACHY: Comparison with MOPITT and first results related to the detection of enhanced CO over cities. *Atmospheric Chemistry and Physics*, *7*(9), 2399–2411. https://doi.org/10.5194/acp-7-2399-2007

Buhr, M., Sueper, D., Trainer, M., Goldan, P., Kuster, B., Fehsenfeld, F., et al. (1996). Trace gas and aerosol measurements using aircraft data from the North Atlantic Regional Experiment (NARE 1993). *Journal of Geophysical Research*, *101*(D22), 29,013–29,027. https://doi.org/10.1029/96JD01159

Burrows, W. R., Benjamin, M., Beauchamp, S., Lord, E. R., McCollor, D., & Thomson, B. (1995). CART decision-tree statistical analysis and prediction of summer season maximum surface ozone for the Vancouver, Montreal, and Atlantic regions of Canada. *Journal of Applied Meteorology*, *34*(8), 1848–1862. https://doi.org/10.1175/1520-0450(1995)034%3C1848:CDTSAA%3E2.0.CO;2

Cardenas, L., Austin, J., Burgess, R., Clemitshaw, K., Dorling, S., Penkett, S., & Harrison, R. (1998). Correlations between CO, NO_y, O_3 and non-methane hydrocarbons and their relationships with meteorology during winter 1993 on the North Norfolk coast, UK. *Atmospheric Environment*, *32*(19), 3339–3351. https://doi.org/10.1016/S1352-2310(97)00445-7

Chance, K., Liu, X., Suleiman, R. M., Flittner, D. E., Al-Saadi, J., & Janz, S. J. (2013). Tropospheric Emissions: Monitoring of Pollution (TEMPO). Earth Observing Systems XVIII (Vol. 8866, p. 88660D). International Society for Optics and Photonics. https://doi.org/10.1117/12.2024479

Chance, K., Palmer, P. I., Spurr, R. J. D., Martin, R. V., Kurosu, T. P., & Jacob, D. J. (2000). Satellite observations of formaldehyde over North America from GOME. *Geophysical Research Letters*, *27*(21), 3461–3464. https://doi.org/10.1029/2000GL011857

Cheng, Y., Wang, Y., Zhang, Y., Chen, G., Crawford, J. H., Kleb, M. M., et al. (2017). Large biogenic contribution to boundary layer O_3-CO regression slope in summer. *Geophysical Research Letters*, *44*, 7061–7068. https://doi.org/10.1002/2017GL074405

Chin, M., Jacob, D. J., Munger, J. W., Parrish, D. D., & Doddridge, B. G. (1994). Relationship of ozone and carbon monoxide over North America. *Journal of Geophysical Research*, *99*(D7), 14,565–14,573. https://doi.org/10.1029/94JD00907

Choi, Y., Osterman, G., Eldering, A., Wang, Y., & Edgerton, E. (2010). Understanding the contributions of anthropogenic and biogenic sources to CO enhancements and outflow observed over North America and the western Atlantic Ocean by TES and MOPITT. *Atmospheric Environment*, *44*(16), 2033–2042. https://doi.org/10.1016/j.atmosenv.2010.01.029

Choi, Y., Wang, Y., Yang, Q., Cunnold, D., Zeng, T., Shim, C., et al. (2008). Spring to summer northward migration of high O_3 over the western North Atlantic. *Geophysical Research Letters*, *35*, L04818. https://doi.org/10.1029/2007GL032276

Choi, Y., Wang, Y., Zeng, T., Cunnold, D., Yang, E. S., Martin, R., et al. (2008). Springtime transitions of NO_2, CO, and O_3 over North America: Model evaluation and analysis. *Journal of Geophysical Research*, *113*, D20311. https://doi.org/10.1029/2007JD009632

Choi, Y., Wang, Y., Zeng, T., Martin, R. V., Kurosu, T. P., & Chance, K. (2005). Evidence of lightning NO_x and convective transport of pollutants in satellite observations over North America. *Geophysical Research Letters*, *32*, L02805. https://doi.org/10.1029/2004GL021436

Cobourn, W. G. (2007). Accuracy and reliability of an automated air quality forecast system for ozone in seven Kentucky metropolitan areas. *Atmospheric Environment*, *41*(28), 5863–5875. https://doi.org/10.1016/j.atmosenv.2007.03.024

Cooper, O., Moody, J., Parrish, D., Trainer, M., Holloway, J., Hübler, G., et al. (2002). Trace gas composition of midlatitude cyclones over the western North Atlantic Ocean: A seasonal comparison of O_3 and CO. *Journal of Geophysical Research*, *107*(D7), 4057. https://doi.org/10.1029/2001JD000902

Cooper, O., Moody, J., Parrish, D., Trainer, M., Ryerson, T., Holloway, J., et al. (2002). Trace gas composition of midlatitude cyclones over the western North Atlantic Ocean: A conceptual model. *Journal of Geophysical Research*, *107*(D7), 4056. https://doi.org/10.1029/2001JD000901

Cuesta, J., Eremenko, M., Liu, X., Dufour, G., Cai, Z., Höpfner, M., et al. (2013). Satellite observation of lowermost tropospheric ozone by multispectral synergism of IASI thermal infrared and GOME-2 ultraviolet measurements over Europe. *Atmospheric Chemistry and Physics*, *13*(19), 9675–9693. https://doi.org/10.5194/acp-13-9675-2013

Dabberdt, W. F., Carroll, M. A., Baumgardner, D., Carmichael, G., Cohen, R., Dye, T., et al. (2004). Meteorological research needs for improved air quality forecasting: Report of the 11th prospectus development team of the US weather research program. *Bulletin of the American Meteorological Society*, *85*(4), 563–586. https://doi.org/10.1175/BAMS-85-4-563

De Laat, A. T. J., Dijkstra, R., Schrijver, H., Nédélec, P., & Aben, I. (2012). Validation of six years of SCIAMACHY carbon monoxide observations using MOZAIC CO profile measurements. *Atmospheric Measurement Techniques*, *5*(9), 2133–2142. https://doi.org/10.5194/amt-5-2133-2012

De Smedt, I., Müller, J. F., Stavrakou, T., Van Der, A. R., Eskes, H., & Van Roozendael, M. (2008). Twelve years of global observations of formaldehyde in the troposphere using GOME and SCIAMACHY sensors. *Atmospheric Chemistry and Physics*, *8*(16), 4947–4963. https://doi.org/10.5194/acp-8-4947-2008

De Smedt, I., Van Roozendael, M., Stavrakou, T., Müller, J.-F., Lerot, C., Theys, N., et al. (2012). Improved retrieval of global tropospheric formaldehyde columns from GOME-2/MetOp-A addressing noise reduction and instrumental degradation issues. *Atmospheric Measurement Techniques*, *5*(11), 2933–2949. https://doi.org/10.5194/amt-5-2933-2012

Delle Monache, L., Deng, X., Zhou, Y., & Stull, R. (2006). Ozone ensemble forecasts: 1. A new ensemble design. *Journal of Geophysical Research*, *111*, D05307. https://doi.org/10.1029/2005JD006310

Delle Monache, L., Hacker, J. P., Zhou, Y., Deng, X., & Stull, R. B. (2006). Probabilistic aspects of meteorological and ozone regional ensemble forecasts. *Journal of Geophysical Research*, *111*, D24307. https://doi.org/10.1029/2005JD006917

Duane, M., Poma, B., Rembges, D., Astorga, C., & Larsen, B. (2002). Isoprene and its degradation products as strong ozone precursors in Insubria, northern Italy. *Atmospheric Environment*, *36*(24), 3867–3879. https://doi.org/10.1016/S1352-2310(02)00359-X

Emmons, L. K., Deeter, M. N., Gille, J. C., Edwards, D. P., Attié, J. L., Warner, J., et al. (2004). Validation of Measurements of Pollution in the Troposphere (MOPITT) CO retrievals with aircraft in situ profiles. *Journal of Geophysical Research*, *109*, D03308. https://doi.org/10.1029/2003JD003970

Fishman, J., & Seiler, W. (1983). Correlative nature of ozone and carbon monoxide in the troposphere: Implications for the tropospheric ozone budget. *Journal of Geophysical Research*, *88*(C6), 3662–3670. https://doi.org/10.1029/JC088iC06p03662

Fried, A., Cantrell, C., Olson, J., Crawford, J. H., Weibring, P., Walega, J., et al. (2011). Detailed comparisons of airborne formaldehyde measurements with box models during the 2006 INTEX-B and MILAGRO campaigns: Potential evidence for significant impacts of unmeasured and multi-generation volatile organic carbon compounds. *Atmospheric Chemistry and Physics*, *11*(22), 11,867–11,894. https://doi.org/10.5194/acp-11-11867-2011

Geng, F., Tie, X., Guenther, A., Li, G., Cao, J., & Harley, P. (2011). Effect of isoprene emissions from major forests on ozone formation in the city of Shanghai, China. *Atmospheric Chemistry and Physics*, *11*(20), 10,449–10,459. https://doi.org/10.5194/acp-11-10449-2011

Gloudemans, A. M. S., Schrijver, H., Kleipool, Q., Van den Broek, M. M. P., Straume, A. G., Lichtenberg, G., et al. (2005). The impact of SCIAMACHY near-infrared instrument calibration on CH_4 and CO total columns. *Atmospheric Chemistry and Physics*, *5*(9), 2369–2383. https://doi.org/10.5194/acp-5-2369-2005

González Abad, G., Liu, X., Chance, K., Wang, H., Kurosu, T. P., & Suleiman, R. (2015). Updated Smithsonian Astrophysical Observatory Ozone Monitoring Instrument (SAO OMI) formaldehyde retrieval. *Atmospheric Measurement Techniques*, *8*(1), 19–32. https://doi.org/10.5194/amt-8-19-2015

González Abad, G., Vasilkov, A., Seftor, C., Liu, X., & Chance, K. (2016). Smithsonian Astrophysical Observatory Ozone Mapping and Profiler Suite (SAO OMPS) formaldehyde retrieval. *Atmospheric Measurement Techniques*, *9*(7), 2797–2812. https://doi.org/10.5194/amt-9-2797-2016

Gu, D., Wang, Y., Smeltzer, C., & Boersma, K. F. (2014). Anthropogenic emissions of NO_x over China: Reconciling the difference of inverse modeling results using GOME-2 and OMI measurements. *Journal of Geophysical Research: Atmospheres*, *119*, 7732–7740. https://doi.org/10.1002/2014JD021644

Gu, D., Wang, Y., Smeltzer, C., & Liu, Z. (2013). Reduction in NO_x emission trends over China: Regional and seasonal variations. *Environmental Science & Technology*, *47*(22), 12,912–12,919. https://doi.org/10.1021/es401727e

Gu, D., Wang, Y., Yin, R., Zhang, Y., & Smeltzer, C. (2016). Inverse modelling of NO_x emissions over eastern China: Uncertainties due to chemical non-linearity. *Atmospheric Measurement Techniques*, *9*(10), 5193–5201. https://doi.org/10.5194/amt-9-5193-2016

Guenther, A., Hewitt, C. N., Erickson, D., Fall, R., Geron, C., Graedel, T., et al. (1995). A global model of natural volatile organic compound emissions. *Journal of Geophysical Research*, *100*(D5), 8873–8892. https://doi.org/10.1029/94JD02950

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. https://doi.org/10.5194/gmd-5-1471-2012

Guillas, S., Bao, J., Choi, Y., & Wang, Y. (2008). Statistical correction and downscaling of chemical transport model ozone forecasts over Atlanta. *Atmospheric Environment*, *42*(6), 1338–1348. https://doi.org/10.1016/j.atmosenv.2007.10.027

Honrath, R., Owen, R. C., Val Martin, M., Reid, J., Lapina, K., Fialho, P., et al. (2004). Regional and hemispheric impacts of anthropogenic and biomass burning emissions on summertime CO and O_3 in the North Atlantic lower free troposphere. *Journal of Geophysical Research*, *109*, D25310. https://doi.org/10.1029/2004JD005147

Hudman, R. C., Murray, L. T., Jacob, D. J., Millet, D., Turquety, S., Wu, S., et al. (2008). Biogenic versus anthropogenic sources of CO in the United States. *Geophysical Research Letters*, *35*, L04801. https://doi.org/10.1029/2007GL032393

Huntrieser, H., Heland, J., Schlager, H., Forster, C., Stohl, A., Aufmhoff, H., et al. (2005). Intercontinental air pollution transport from North America to Europe: Experimental evidence from airborne measurements and surface observations. *Journal of Geophysical Research*, *110*, D01305. https://doi.org/10.1029/2004JD005045

Ingmann, P., Veihelmann, B., Langen, J., Lamarre, D., Stark, H., & Courrèges-Lacoste, G. B. (2012). Requirements for the GMES atmosphere service and ESA's implementation concept: Sentinels-4/−5 and-5p. *Remote Sensing of Environment*, *120*, 58–69. https://doi.org/10.1016/j.rse.2012.01.023

Jing, P., Cunnold, D., Choi, Y., & Wang, Y. (2006). Summertime tropospheric ozone columns from Aura OMI/MLS measurements versus regional model results over the United States. *Geophysical Research Letters*, *33*, L17817. https://doi.org/10.1029/2006GL026473

Kesselmeier, J., & Staudt, M. (1999). Biogenic volatile organic compounds (VOC): An overview on emission, physiology and ecology. *Journal of Atmospheric Chemistry*, *33*(1), 23–88. https://doi.org/10.1023/A:1006127516791

Koo, J.-H., Wang, Y., Kurosu, T., Chance, K., Rozanov, A., Richter, A., et al. (2012). Characteristics of tropospheric ozone depletion events in the Arctic spring: Analysis of the ARCTAS, ARCPAC, and ARCIONS measurements and satellite BrO observations. *Atmospheric Chemistry and Physics*, *12*(20), 9909–9922. https://doi.org/10.5194/acp-12-9909-2012

Kurosu, T. P., Chance, K., & Sioris, C. E. (2004). December. Preliminary results for HCHO and BrO from the EOS-aura ozone monitoring instrument. In *Passive optical remote sensing of the atmosphere and clouds IV* (Vol. 5652, pp. 116–124). Bellingham, Washington: International Society for Optics and Photonics. https://doi.org/10.1117/12.578606

Lathiere, J., Hauglustaine, D., Friend, A., Noblet-Ducoudré, N. D., Viovy, N., & Folberth, G. (2006). Impact of climate variability and land use changes on global biogenic volatile organic compound emissions. *Atmospheric Chemistry and Physics*, *6*(8), 2129–2146. https://doi.org/10.5194/acp-6-2129-2006

Lee, K.-Y., Kwak, K.-H., Ryu, Y.-H., Lee, S.-H., & Baik, J.-J. (2014). Impacts of biogenic isoprene emission on ozone air quality in the Seoul metropolitan area. *Atmospheric Environment*, *96*, 209–219. https://doi.org/10.1016/j.atmosenv.2014.07.036

Lelieveld, J., & Dentener, F. J. (2000). What controls tropospheric ozone? *Journal of Geophysical Research*, *105*(D3), 3531–3551. https://doi.org/10.1029/1999JD901011

Li, C., Joiner, J., Krotkov, N. A., & Dunlap, L. (2015). A new method for global retrievals of HCHO total columns from the Suomi National Polar-orbiting Partnership Ozone Mapping and Profiler Suite. *Geophysical Research Letters*, *42*, 2515–2522. https://doi.org/10.1002/2015GL063204

Li, Q., Jacob, D. J., Bey, I., Palmer, P. I., Duncan, B. N., Field, B. D., et al. (2002). Transatlantic transport of pollution and its effects on surface ozone in Europe and North America. *Journal of Geophysical Research*, *107*(D13), 4166. https://doi.org/10.1029/2001JD001422

Liu, X., Bhartia, P. K., Chance, K., Spurr, R. J. D., & Kurosu, T. P. (2010). Ozone profile retrievals from the Ozone Monitoring Instrument. *Atmospheric Chemistry and Physics*, *10*(5), 2521–2537. https://doi.org/10.5194/acp-10-2521-2010

Liu, Z., Wang, Y., Costabile, F., Amoroso, A., Zhao, C., Huey, L. G., et al. (2014). Evidence of aerosols as a media for rapid daytime HONO production over China. *Environmental Science & Technology*, *48*(24), 14,386–14,391. https://doi.org/10.1021/es504163z

Liu, Z., Wang, Y., Gu, D., Zhao, C., Huey, L. G., Stickel, R., et al. (2010). Evidence of reactive aromatics as a major source of peroxy acetyl nitrate over China. *Environmental Science & Technology*, *44*(18), 7017–7022. https://doi.org/10.1021/es1007966

Liu, Z., Wang, Y., Gu, D., Zhao, C., Huey, L. G., Stickel, R., et al. (2012). Summertime photochemistry during CAREBeijing-2007: RO_x budgets and O_3 formation. *Atmospheric Chemistry and Physics*, *12*(16), 7737–7752. https://doi.org/10.5194/acp-12-7737-2012

Liu, Z., Wang, Y., Vrekoussis, M., Richter, A., Wittrock, F., Burrows, J. P., et al. (2012). Exploring the missing source of glyoxal (CHOCHO) over China. *Geophysical Research Letters*, *39*, L10812. https://doi.org/10.1029/2012GL051645

Logan, J. A., Prather, M. J., Wofsy, S. C., & McElroy, M. B. (1981). Tropospheric chemistry: A global perspective. *Journal of Geophysical Research*, *86*(C8), 7210–7254. https://doi.org/10.1029/JC086iC08p07210

Mao, H., & Talbot, R. (2004). O_3 and CO in New England: Temporal variations and relationships. *Journal of Geophysical Research*, *109*, D21304. https://doi.org/10.1029/2004JD004913

Pang, X., Mu, Y., Zhang, Y., Lee, X., & Yuan, J. (2009). Contribution of isoprene to formaldehyde and ozone formation based on its oxidation products measurement in Beijing, China. *Atmospheric Environment*, *43*(13), 2142–2147. https://doi.org/10.1016/j.atmosenv.2009.01.022

Parrish, D. D., Holloway, J. S., Trainer, M., Murphy, P. C., Fehsenfeld, F. C., & Forbes, G. L. (1993). Export of North American ozone pollution to the North Atlantic Ocean. *Science*, *259*(5100), 1436–1439. https://doi.org/10.1126/science.259.5100.1436

Parrish, D. D., Trainer, M., Holloway, J., Yee, J., Warshawsky, M., Fehsenfeld, F., et al. (1998). Relationships between ozone and carbon monoxide at surface sites in the North Atlantic region. *Journal of Geophysical Research*, *103*(D11), 13,357–13,376. https://doi.org/10.1029/98JD00376

Perez, P., & Reyes, J. (2006). An integrated neural network model for PM_{10} forecasting. *Atmospheric Environment*, *40*(16), 2845–2851. https://doi.org/10.1016/j.atmosenv.2006.01.010

Pierce, T., Geron, C., Bender, L., Dennis, R., Tonnesen, G., & Guenther, A. (1998). Influence of increased isoprene emissions on regional ozone modeling. *Journal of Geophysical Research*, *103*(D19), 25,611–25,629. https://doi.org/10.1029/98JD01804

Polonsky, I. N., O'Brien, D. M., Kumer, J. B., & O'Dell, C. W. (2014). Performance of a geostationary mission, geoCARB, to measure CO_2, CH_4 and CO column-averaged concentrations. *Atmospheric Measurement Techniques*, *7*(4), 959–981. https://doi.org/10.5194/amt-7-959-2014

Sachse, G. W., Hill, G. F., Wade, L. O., & Perry, M. G. (1987). Fast-response, high-precision carbon monoxide sensor using a tunable diode laser absorption technique. *Journal of Geophysical Research: Atmospheres*, *92*(D2), 2071–2081. https://doi.org/10.1029/JD092iD02p02071

Shad, R., Mesgari, M. S., & Shad, A. (2009). Predicting air pollution using fuzzy genetic linear membership kriging in GIS, Computers. *Environment and Urban Systems*, *33*(6), 472–481. https://doi.org/10.1016/j.compenvurbsys.2009.10.004

Sindelarova, K., Granier, C., Bouarar, I., Guenther, A., Tilmes, S., Stavrakou, T., et al. (2014). Global data set of biogenic VOC emissions calculated by the MEGAN model over the last 30 years. *Atmospheric Chemistry and Physics*, *14*(17), 9317–9341. https://doi.org/10.5194/acp-14-9317-2014

Straume, A. G., Schrijver, H., Gloudemans, A. M. S., Houweling, S., Aben, I., Maurellis, A. N., et al. (2005). The global variation of CH_4 and CO as seen by SCIAMACHY. *Advances in Space Research*, *36*(5), 821–827. https://doi.org/10.1016/j.asr.2005.03.027

Van der Wal, J., & Janssen, L. (2000). Analysis of spatial and temporal variations of PM_{10} concentrations in the Netherlands using Kalman filtering. *Atmospheric Environment*, *34*(22), 3675–3687. https://doi.org/10.1016/S1352-2310(00)00085-6

Vautard, R., Schaap, M., Bergström, R., Bessagnet, B., Brandt, J., Builtjes, P., et al. (2009). Skill and uncertainty of a regional air quality model ensemble. *Atmospheric Environment*, *43*(31), 4822–4832. https://doi.org/10.1016/j.atmosenv.2008.09.083

Wang, Y., Choi, Y., Zeng, T., Davis, D., Buhr, M., Huey, L. G., & Neff, W. (2007). Assessing the photochemical impact of snow NO_x emissions over Antarctica during ANTCI 2003. *Atmospheric Environment*, *41*(19), 3944–3958. https://doi.org/10.1016/j.atmosenv.2007.01.056

Wang, Y., Choi, Y., Zeng, T., Ridley, B., Blake, N., Blake, D., & Flocke, F. (2006). Late-spring increase of trans-Pacific pollution transport in the upper troposphere. *Geophysical Research Letters*, *33*, L01811. https://doi.org/10.1029/2005GL024975

Wang, Y., & Jacob, D. J. (1998). Anthropogenic forcing on tropospheric ozone and OH since preindustrial times. *Journal of Geophysical Research*, *103*(D23), 31,123–31,135. https://doi.org/10.1029/1998JD100004

Wang, Y., Jacob, D. J., & Logan, J. A. (1998). Global simulation of tropospheric O_3-NO_x-hydrocarbon chemistry: 1. Model formulation. *Journal of Geophysical Research*, *103*(D9), 10,713–10,725. https://doi.org/10.1029/98JD00158

Wang, Y., & Zeng, T. (2004). On tracer correlations in the troposphere: The case of ethane and propane. *Journal of Geophysical Research*, *109*, D24306. https://doi.org/10.1029/2004JD005023

Weibring, P., Richter, D., Walega, J., Rippe, L., & Fried, A. (2010). Difference frequency generation spectrometer for simultaneous multispecies detection. *Optics Express*, *18*(26), 27,670–27,681. https://doi.org/10.1364/OE.18.027670

Wiedinmyer, C., Friedfeld, S., Baugh, W., Greenberg, J., Guenther, A., Fraser, M., & Allen, D. (2001). Measurement and analysis of atmospheric concentrations of isoprene and its reaction products in central Texas. *Atmospheric Environment*, *35*(6), 1001–1013. https://doi.org/10.1016/S1352-2310(00)00406-4

Wittrock, F., Richter, A., Oetjen, H., Burrows, J. P., Kanakidou, M., Myriokefalitakis, S., et al. (2006). Simultaneous global observations of glyoxal and formaldehyde from space. *Geophysical Research Letters*, *33*, L16804. https://doi.org/10.1029/2006GL026310

Yang, Q., Wang, Y., Zhao, C., Liu, Z., Gustafson, W. I. Jr., & Shao, M. (2011). NO_x emission reduction and its effects on ozone during the 2008 Olympic Games. *Environmental Science & Technology*, *45*(15), 6404–6410. https://doi.org/10.1021/es200675v

Zeng, T., & Wang, Y. (2011). Biomass burning induced nationwide summer peaks of OC/EC ratios in the continental United States. *Atmospheric Environment*, *45*(3), 578–586. https://doi.org/10.1016/j.atmosenv.2010.10.038

Zeng, T., Wang, Y., Chance, K., Blake, N., Blake, D., & Ridley, B. (2006). Halogen-driven low-altitude O_3 and hydrocarbon losses in spring at northern high latitudes. *Journal of Geophysical Research*, *111*, D17313. https://doi.org/10.1029/2005JD006706

Zeng, T., Wang, Y., Chance, K., Browell, E. V., Ridley, B. A., & Atlas, E. L. (2003). Widespread persistent near-surface ozone depletion at northern high latitudes in spring. *Geophysical Research Letters*, *30*(24), 2298. https://doi.org/10.1029/2003GL018587

Zhang, L., Jiang, H., Lu, X., & Jin, J. (2016). Comparison analysis of global carbon monoxide concentration derived from SCIAMACHY, AIRS, and MOPITT. *International Journal of Remote Sensing*, *37*(21), 5155–5175. https://doi.org/10.1080/01431161.2016.1230282

Zhang, R., Wang, Y., He, Q., Chen, L., Zhang, Y., Qu, H., et al. (2017). Enhanced trans-Himalaya pollution transport to the Tibetan Plateau by cut-off low systems. *Atmospheric Chemistry and Physics*, *17*(4), 3083–3095. https://doi.org/10.5194/acp-17-3083-2017

Zhang, Y., Bocquet, M., Mallet, V., Seigneur, C., & Baklanov, A. (2012). Real-time air quality forecasting, part II: State of the science, current research needs, and future prospects. *Atmospheric Environment*, *60*, 656–676. https://doi.org/10.1016/j.atmosenv.2012.02.041

Zhang, Y., & Wang, Y. (2016). Climate-driven ground-level ozone extreme in the fall over the southeast United States. *Proceedings of the National Academy of Sciences of the United States of America*, *113*(36), 10,025–10,030. https://doi.org/10.1073/pnas.1602563113

Zhang, Y., Wang, Y., Chen, G., Smeltzer, C., Crawford, J., Olson, J., et al. (2016). Large vertical gradient of reactive nitrogen oxides in the boundary layer: Modeling analysis of DISCOVER-AQ 2011 observations. *Journal of Geophysical Research: Atmospheres*, *121*, 1922–1934. https://doi.org/10.1002/2015JD024203

Zhang, Y., Wang, Y., Crawford, J., Cheng, Y., & Li, J. (2018). Improve observation-based ground-level ozone spatial distribution by compositing satellite and surface observations: A simulation experiment. *Atmospheric Environment*, *180*, 226–233. https://doi.org/10.1016/j.atmosenv.2018.02.044

Zhao, C., & Wang, Y. (2009). Assimilated inversion of NO_x emissions over East Asia using OMI NO_2 column measurements. *Geophysical Research Letters*, *36*, L06805. https://doi.org/10.1029/2008GL037123

Zhao, C., Wang, Y., Choi, Y., & Zeng, T. (2009). Summertime impact of convective transport and lightning NO_x production over North America: Modeling dependence on meteorological simulations. *Atmospheric Chemistry and Physics*, *9*(13), 4315–4327. https://doi.org/10.5194/acp-9-4315-2009

Zhao, C., Wang, Y., Yang, Q., Fu, R., Cunnold, D., & Choi, Y. (2010). Impact of East Asian summer monsoon on the air quality over China: View from space. *Journal of Geophysical Research*, *115*, D09301. https://doi.org/10.1029/2009JD012745

Zhao, C., Wang, Y., & Zeng, T. (2009). East China plains: A "basin" of ozone pollution. *Environmental Science & Technology*, *43*(6), 1911–1915. https://doi.org/10.1021/es8027764

Zhu, L., Jacob, D. J., Kim, P. S., Fisher, J. A., Yu, K., Travis, K. R., et al. (2016). Observing atmospheric formaldehyde (HCHO) from space: Validation and intercomparison of six retrievals from four satellites (OMI, GOME2A, GOME2B, OMPS) with SEAC 4 RS aircraft observations over the southeast US. *Atmospheric Chemistry and Physics*, *16*(21), 13,477–13,490. https://doi.org/10.5194/acp-16-13477-2016