Sensitivity of the WRF-Chem v4.4 ozone, formaldehyde, and precursor simulations to multiple bottom-up emission inventories over East Asia during the KORUS-AQ 2016 field campaign

Kyoung-Min Kim\textsuperscript{1}, Si-Wan Kim\textsuperscript{2*}, Seunghwan Seo\textsuperscript{1}, Donald R. Blake\textsuperscript{3}, Seogju Cho\textsuperscript{4}, James H. Crawford\textsuperscript{5}, Louisa Emmons\textsuperscript{6}, Alan Fried\textsuperscript{7}, Jay R. Herman\textsuperscript{8,9}, Jinkyu Hong\textsuperscript{1}, Jinsang Jung\textsuperscript{10} Gabriele Pfister\textsuperscript{6}, Andrew J. Weinheimer\textsuperscript{6}, Jung-Hun Woo\textsuperscript{11}, and Qiang Zhang\textsuperscript{12}

\textsuperscript{1}Department of Atmospheric Sciences, Yonsei University, Seoul, South Korea
\textsuperscript{2}Irreversible Climate Change Research Center, Yonsei University, Seoul, South Korea
\textsuperscript{3}Department of Chemistry, University of California at Irvine, Irvine, CA, US
\textsuperscript{4}Seoul Metropolitan Government Research Institute of Public Health and Environment, Gyeonggi-do, South Korea
\textsuperscript{5}NASA Langley Research Center, Hampton, VA, US
\textsuperscript{6}National Center for Atmospheric Research, Boulder, CO, US
\textsuperscript{7}Institute of Arctic and Alpine Research, University of Colorado, Boulder, CO, US
\textsuperscript{8}NASA Goddard Space Flight Center, Greenbelt, MD, US
\textsuperscript{9}University of Maryland Baltimore County, Baltimore, MD, USA
\textsuperscript{10}Korea Research Institute of Standards and Science, Daejeon, South Korea
\textsuperscript{11}Department of Advanced Technology Fusion, Konkuk University, Seoul, South Korea
\textsuperscript{12}Department of Earth System Science, Tsinghua University, Beijing, China

*To whom correspondence should be addressed. E-mail: siwan.kim@yonsei.ac.kr

Date: 6/21/2023
Abstract

In this study, the WRF-Chem v4.4 model was utilized to evaluate three bottom-up emission inventories (EDGAR-HTAP v2, v3, and KORUS v5) using surface and aircraft data in East Asia during the Korea-United States Air Quality (KORUS-AQ) campaign period in 2016. All emission inventories were found to reproduce the diurnal variations of O₃ and NO₂ as compared to the surface monitor data. However, the spatial distributions of the daily maximum 8-hour average (MDA8) O₃ in the model do not completely align with the observations. The model MDA8 O₃ had a negative (positive) bias north (south) of 30°N over China. All simulations underestimated the observed CO by 50-60% over China and South Korea. In the Seoul Metropolitan Area (SMA), EDGAR-HTAP v2, v3, and KORUS v5 simulated the vertical shapes and diurnal patterns of O₃ and other precursors effectively, but the model underestimated the observed O₃, CO and HCHO concentrations. Notably, the model aromatic VOCs were significantly underestimated with the three bottom-up emission inventories, although the KORUS v5 shows improvements. The model isoprene estimations had a positive bias relative to the observations, suggesting that the Model of Emissions of Gases and Aerosols from Nature (MEGAN) version 2.04 overestimated isoprene emissions. Additional model simulations were conducted by doubling CO and VOC emissions over China and South Korea to investigate the causes of the model O₃ biases and the effects of the long-range transport on the O₃ over South Korea. The doubled CO and VOC emission simulations improved the model O₃ simulations for the local emission dominant case, but led to the model O₃ overestimations for the transport dominant case, which emphasizes the need for accurate representations of the local VOC emissions over South Korea.
1. Introduction

Air pollutants not only harm human health but also affect radiative balance, resulting in climate change (Anenberg et al., 2018; Franklin et al., 2015; Lee et al., 2014; Manning and von Tiedemann, 1995; Rosenzweig et al., 2008; Wild et al., 2001). Anthropogenic activities are the primary source of air pollutant emissions, which have significant temporal and spatial variability. Chemical transport models (CTMs) use bottom-up emission data to simulate ambient concentrations of air pollutants. CTMs then process these emissions, tracking their impact through chemistry, transport, and loss through deposition (Zhong et al., 2016). Therefore, sensitivity evaluations of CTMs to anthropogenic emission data are an essential part of atmospheric modeling research.

Several bottom-up emission inventories are available for chemical modeling of Asia, including the Multi-resolution Emission Inventory for China (MEIC), Regional Emission inventory in Asia (REAS), and Emissions Database for Global Atmospheric Research-Hemispheric Transport of Air Pollution (EDGAR-HTAP). Since 2010, Tsinghua University has developed the high-resolution MEIC emission inventory for China and updated the data to the v1.3, providing anthropogenic emissions by sector and species from 2008 to 2017 (Zheng et al., 2018). REAS provides emission data in Asia from 1950 to 2015 (Kurokawa and Ohara, 2020). In Europe, EDGAR-HTAP has been developed and widely used for CTM simulations from global to regional scale (Kim et al., 2021; Sharma et al., 2017; Sicard et al., 2021). Recently, EDGAR-HTAP v3 has been published, covering 19 years from 2000 to 2018 compared to only two years (2008 and 2010) in the version 2 data (Crippa et al., 2023). Zhong et al. (2016) compared REAS with EDGAR in July, 2007 over China, while Saikawa et al. (2017) compared 5 emission inventories including REAS, EDGAR, MEIC in China, without

https://doi.org/10.5194/gmd-2023-132
Preprint. Discussion started: 29 August 2023
© Author(s) 2023. CC BY 4.0 License.
validation. As bottom-up emission inventories are continuously updated for recent years, there is an ongoing need to evaluate new emissions data.

The Ministry of Environment (MOE) in South Korea and National Aeronautics and Space Administration (NASA) in the U.S. conducted the Korea-United States Air Quality (KORUS-AQ) campaign in May-June 2016. The campaign provided a variety of data sets, including ground-based and airborne observations, useful for the validation of model simulations. The KORUS emissions, developed by Konkuk University, were used by many modeling teams to simulate the air pollutant concentrations during the campaign period. In this study, we selected the EDGAR-HTAP versions 2, v3, and KORUS version 5 emission data and used the Weather Research and Forecasting model coupled with Chemistry (WRF-Chem) version 4.4 for intercomparison of the three emissions data sets. O3 and its major precursors were selected for model evaluation and the model results were validated with surface observation data in China and South Korea and aircraft data acquired over the South Korean peninsula and surrounding waters.

2. Data and Methods

2.1. WRF-Chem model configurations

In this study, we utilized the WRF-Chem v4.4, which was developed by the National Oceanic and Atmospheric Administration (NOAA) and National Center for Atmospheric Research (NCAR), to simulate meteorological variables and chemical species in the atmosphere (Grell et al., 2005). The WRF-Chem v4.4 includes N2O5 heterogeneous chemistry that consists of several chemical reactions related with ClNO2.
and N\textsubscript{2}O\textsubscript{5} reactions, resulting in nitrate aerosol. The reactions are incorporated in Secondary Organic Aerosol-Volatility Basis Set (SOA-VBS) with Regional Atmospheric Chemistry Mechanism (RACM) chemistry option (chem = 108) in WRF-Chem (Li et al., 2016).

We set 59 vertically customized \(\eta\) levels as vertical layers. The model’s first layer height is approximately 40 m above ground level for the entire domain. The model’s vertical layers are designed to include about 17 layers under 1.5 km to simulate planetary boundary layer chemistry and near surface vertical distribution in detail. The horizontal resolution is 28 x 28 km\(^2\). The simulations in this study start at 12 UTC on April 24 and end at 12 UTC on June 11. The model meteorology restarts every 12 UTC (9 PM local time in South Korea) to minimize numerical errors. After the first 7 days of model initiation (spin-up), we analyzed the model results from May 1 to June 10. We used China standard time (+8 UTC) and Korea standard time (+9 UTC) for evaluations with observations. The model physics, chemistry, and aerosol schemes are summarized in Table S1 with corresponding references. The Global Forecast System (GFS) Final (FNLI analysis data are used for meteorological input and boundary conditions. The Community Atmosphere Model with Chemistry (CAM-Chem) output is used for chemical boundary conditions (https://www.acom.ucar.edu/cam-chem/cam-chem.html) (Buchholz et al., 2019; Emmons et al., 2020). We used the Model of Emissions of Gases and Aerosols from Nature (MEGAN) v2.04 to calculate biogenic emissions (Guenther et al., 2006). We did not account for fire emissions because of small impact on air quality simulations during the KORUS-AQ campaign period (Park et al., 2021).
2.2. The model simulations using different anthropogenic emissions

2.2.1. Bottom-up emission data

EDGAR-HTAP v2, v3, and KORUS v5 emissions are compared with respect to their spatial distribution and total amount in Figure 1 and Table S3. We applied the same diurnal factor for all three emissions data by species, following the diurnal patterns for the Los Angeles Basin as in Kim et al. (2016) (also see Figure S1).

EDGAR-HTAP v2 provides 2-dimensional emissions of CH₄, CO, SO₂, NOx (NO + NO₂), non-methane volatile organic compound (NMVOC), NH₃, PM₁₀, PM₂.₅, BC, and OC in 2008 and 2010 with a horizontal resolution of 0.1° x 0.1°. We used 2010 data since it is the most recent data available. The data are partitioned by each sector and its sources such as air, ships, energy, industry, transport, residential, and agriculture (https://edgar.jrc.ec.europa.eu/dataset_htap_v2). For East Asia, it included data from the Model Inter-Comparison Study for Asia (MICS-Asia) and REAS v2.1. In South Korea, it adopted data from the Clean Air Policy Support System (CAPSS) (Janssens-Maenhout et al., 2015), and the underlying emission data had an original horizontal resolution of 0.25° x 0.25° over East Asia, which is resampled to 0.1° x 0.1° resolution by raster resampling and aggregation. The specifically mapped EDGAR-HTAP v2 data were obtained through the WRF-Chem site (https://www.acom.ucar.edu/wrf-chem/download.shtml) in the anthro_emiss program with the Model for Ozone and Related chemical Tracers (MOZART) species. The anthro_emiss program converts the EDGAR-HTAP v2 data into 28 x 28 km² grid by the RACM chemical species. It mapped the MOZART volatile organic compounds (VOC) species into the RACM VOC species (See the detailed equations in Supporting Information, Table S4) (Li et al., 2014; Emmons et al., 2010).
The EDGAR-HTAP v3 is extended to much longer time scale than the previous version EDGAR-HTAP (v2). The EDGAR-HTAP v3 covers 2000 to 2018 with a more detailed horizontal resolution (https://edgar.jrc.ec.europa.eu/dataset_htap_v3) (Crippa et al. 2023). While EDGAR-HTAP v2 uses MICS-Asia, only the REAS data are used in China and India in the EDGAR-HTAP v3. It adopts the CAPSS-Konkuk University (CAPSS-KU) data for South Korea and emission data provided by the Japanese government for Japan. We chose the data for 2016, according to the KORUS-AQ campaign period. Because the original EDGAR-HTAP v3 data provide VOC as total NMVOC with the unit of ton/month, we distributed the total NMVOC to MOZART VOC species with the ratio of each VOC species to total NMVOC from EDGAR-HTAP v2 in anthro_emiss program. Then, the assigned EDGAR-HTAP v3 data were again converted to the RACM.

The KORUS v5 emission data represent 2016 in China and 2015 in other regions. The Comprehensive Regional Emissions Inventory for Atmospheric Transport Experiment (CREATE) v2.3 data from 2015 were used and the ship emissions from CAPSS were added near the coastal region in South Korea (Jang et al., 2020; Woo et al., 2012). The CREATE is originally developed by combining REAS, MEIC, Japan Auto-Oil Program emission inventory (JATOP), and Korean Clear Air Policy Support System (CAPSS). The NMVOC species from KORUS v5 were mapped following the Statewide Air Pollution Research Center (SAPRC-99) mechanism, and we also assigned the SAPRC-99 species to RACM (Carter, 2000) (Supporting information, Table S4).

Figure 1 shows the spatial distribution of NO, CO, and toluene emissions in May for each inventory. The NOx emissions were assumed to be emitted as NO. The major
cities in China and South Korea had relatively high NO, CO, and TOL (toluene and less aromatics defined in RACM, see Table S2) emissions, which are major precursors affecting O₃ formation. We define three boxes representing Eastern China, South Korea, and the Seoul metropolitan area (SMA) and calculated the emissions (see Table S3). In South Korea including SMA, EDGAR-HTAP v3 had the largest NOₓ emission among the emission inventories. The KORUS v5 has lower NOₓ emissions in Eastern China by 46% and 39% compared to EDGAR-HTAP v2 and v3, respectively. The CO emission was the lowest in EDGAR-HTAP v2 in South Korea, being 56% (69%) lower than that in KORUS v5 (EDGAR-HTAP v3). KORUS v5 showed the highest CO emissions in SMA though EDGAR-HTAP v3 showed more CO emissions in South Korea. However, KORUS v5 had the smallest CO emissions in China, being 7% (9%) lower than that in EDGAR-HTAP v2 (v3). The TOL emission from KORUS v5 is higher than those from EDGAR-HTAP v2 (EDGAR-HTAP v3) by 176% (98%) in China. The relative difference between KORUS v5 and EDGAR-HTAP v2 (EDGAR-HTAP v3) is larger in South Korea by 263%. These discrepancies of VOC emissions may lead to a change in the NOₓ/VOC-sensitive regime and O₃ production efficiency. The sensitivity of O₃ formation to NOₓ emission has discrepancies by its regime, which will be further discussed in section 3.2.

2.2.2. The model experiments

The model experiments are summarized in Table 1. The simulations using EDGAR-HTAP v2, v3, and KORUS v5 emissions are named as EDV2, EDV3, and KOV5, respectively. In this study, we found consistent underestimation of CO and VOC for all
emissions by -40% (± 2%) and -25% (± 1%) (HCHO) compared to DC-8 in South Korea. This is in line with the results reported by Park et al. (2021), who found that almost every model underestimated CO. Underestimation of CO in East Asia is a well-known feature revealed by many studies. For example, Gaubert et al. (2020) mentioned that CAM-Chem underestimates CO during the KORUS-AQ campaign period and presented a CO compensation method utilizing data assimilation with CO observations. Wada et al. (2012) pointed out that EDGAR v4.1 underestimates anthropogenic CO emissions in China by 45% compared to observation-based estimations of CO emissions. Moreover, underestimation of VOC is also found for all anthropogenic emission inventories. Kwon et al. (2021) estimated top-down emissions of anthropogenic VOCs utilizing Geostationary Trace gas and Aerosol Sensor Optimization spectrometer (GeoTASO). They found that top-down VOC emissions were up to 6.9 times higher than bottom-up emissions (KORUS v5). With all emission, O₃ is underestimated at most ground-based observation sites in South Korea. Therefore, we conducted two additional model simulations using EDGAR-HTAP v3 that shows lowest bias of O₃ concentrations compared to DC-8 (-14.2 ppb) than EDGAR-HTAP v2 (-16.9 ppb) and KORUS v5 (-18.1 ppb) over the SMA: one is with twice the anthropogenic CO and VOC emissions in China (EDV3_Ch2) and the other simulation uses double CO and VOC emissions in both China and South Korea (EDV3_ChKo2) to investigate possible improvements in the simulated O₃ and CO from these emission changes.
2.3. Observations

2.3.1. Meteorological data

The meteorological field that WRF-Chem reproduced is validated with the surface synoptic observation (SYNOP) data operated by the World Meteorological Organization (WMO) (http://www.meteomanz.com). Surface temperature, relative humidity, and surface wind speed are adopted for model validation. As the SYNOP data are provided every 3 or 6-hourly, we selected model data when the observation data are available. There were 271 sites in China-Taiwan-Hongkong and 48 sites in South Korea.

2.3.2. Ground-based observations

The surface observation network used in this study was obtained from Airkorea in South Korea and the China Ministry of Ecology and Environment (MEE) in China. The Airkorea observation network provides 1-hourly measurements of NO$_2$, SO$_2$, CO, O$_3$, PM$_{10}$, and PM$_{2.5}$ at suburban, background, roadside, city, and port sites (www.airkorea.or.kr). The concentrations of NO$_2$, CO, and O$_3$ are measured using the chemiluminescent, non-dispersive infrared, and ultraviolet photometric methods, respectively. The model data with 28 x 28 km$^2$ horizontal resolution were linearly interpolated to the 365 sites in South Korea, and we selected NO$_2$, O$_3$, and CO for model validation. The Chinese observations were provided by MEE through the website (beijingair.sinaapp.com). Surface NO$_2$ over China was measured using a molybdenum converter, which has the potential for positive biases due to other NO$_2$-related oxidation products (Dunlea et al., 2007). In South Korea, the positive biases exist regarding NO$_2$ surface observations, which could to overestimations of 28.9% at suburban sites.
spring (Jung et al., 2017). CO was measured using infrared absorption (Zhang and Cao., 2015), and there were 1454 stations in China during the campaign period.

For validation of NO₂ and HCHO vertical column density, data from the Pandora spectrometer were used, which the model reproduced with emission inventories at the Olympic Park site (37.5232°N, 127.126°E). The HCHO data from Pandora is corrected because of internal off-gasing to avoid positive biases (Spinei et al., 2021). At the same observation site, surface NO₂ was also measured by a KEN TEK NOx analyzer with photolytic method, and surface O₃ was measured using the same instrument. Ground-based HCHO was measured using Aerodyne QCL. We compared the observed diurnal cycle of vertical column and surface concentrations of NO₂ and HCHO with the model simulations utilizing EDV2, EDV3, and KOV5. We also used ground-based VOC data from gas chromatography flame ionization detector (GC-FID) operated by the Seoul Research Institute of Public Health and Environment (SIHE).

### 2.3.3. Aircraft data

The DC-8 research aircraft, operated by NASA, performed multiple flight measurements with a variety of measuring instruments. We utilized 1 minute interval merged data of O₃, NO₂, CO, HCHO, and VOC along the 20 flight paths (Figure 7). The nearest WRF-Chem grid is selected and then temporally and vertically interpolated to the aircraft data to fully utilize the observations. Atmospheric NO₂ and O₃ concentrations were measured using a 4-channel chemiluminescence instrument, with an uncertainty of 100 pptv + 30% and 5 ppbv + 10%, respectively. CO concentrations were observed using a diode laser spectrometer, with an uncertainty of 2% or 2 ppbv.
The Compact Atmospheric Multi-species Spectrometer (CAMS) was used to measure HCHO concentration, with a possible 3% systematic error (Richter et al., 2015). We also utilized data from the Whole Air Sampler (WAS) to analyze VOC species from different emission inventories (Colman et al., 2001). In this study, we focused on DC-8 observations below a height of 2 km to concentrate on planetary boundary layer (PBL) chemistry. The observation height was determined by GPS altitude above ground level.

3. Results

3.1. The model meteorology simulations

The model temperature and relative humidity were compared with surface observations in China and South Korea. The model-simulated temperature had a slight negative mean bias of -0.91 °C (correlation coefficient R = 0.90) in China, with the largest negative bias in southwestern China. In South Korea, the mean bias was -1.71 °C (R = 0.88). The simulated relative humidity showed a negative bias of -20 to -10% in the North China Plain (NCP) area and a positive bias of 10 to 20% in southwestern China. There was a negative bias of relative humidity over the west coastal area and a positive bias of 10 to 20% at most observation stations in South Korea. The correlation coefficients between the model relative humidity and observations were 0.85 and 0.76 for China and South Korea, respectively. Overall, the comparisons showed decent model simulations of meteorology. During the KORUS-AQ campaign period, WRF-Chem accurately simulated the daytime PBL height from a laser ceilometer (CL-31, Vaisla Inc., Finland) observed at Yonsei University in Seoul, South Korea (Lee et al., 2019). But, Travis et al. (2022) has indicated the possibility of PBL height underestimations by CTM. Furthermore, due to limitations of the instrument, the ceilometer has potential to
inadequately estimate nighttime PBL height. It is primarily attributed to the method
based on aerosol gradients (Jordan et al., 2020). Therefore, the interpretation of
simulated nighttime concentrations of air pollutants should be approached with caution.
More analysis of meteorological fields, including PBL height, can be found in the
Supporting Information (Table S5 and Figure S2-S3).

3.2. Evaluations with routine surface chemical observational data
The study compared simulated concentrations of O₃, NO₂, and CO with data from
routine surface observational networks (Table 2 and Figure 2-6). First, the diurnal
variations of the model O₃ using different emissions inventories were compared with
observations for each subregion (Table 2 and Figure 2). Overall, all emissions
successfully reproduced diurnal variations and absolute values of O₃ for most regions,
but there were notable discrepancies in several regions. In the North China Plain (NCP)
region, EDV2 led to a negative model O₃ bias (-12 ppb) with R=0.65, while EDV3 and
KOV5 simulated O₃ better with reduced biases and increased correlations (R=0.68-
0.71). Similarly, EDV2 had a negative O₃ bias (-17 ppb) with R=0.62 in the Yangtze
River Delta (YRD) area, but EDV3 and KOV5 much improved the simulations, which
was also observed in the Northeastern China (NEC) area. However, the model O₃
concentrations based on the three emission inventories were overestimated in the
Sichuan-Chongqing-Guizhou (SCG) and Southeastern China (SEC) area. In Pearl River
Delta (PRD), EDV2 showed the lowest bias (-0.3 ppb) compared to EDV3 and KOV5.
In the suburban area of Northern China (NOC), all emission inventories reasonably
simulated hourly O₃ concentrations. Averaged O₃ was well simulated in South Korea
(KOR) with low biases (-1 to 0.7 ppb), but a negative bias appears over the Seoul
metropolitan area (SMA) with all emissions (-5.5 to -3.5 ppb) (Table 2).

The study also analyzed the mean values of daily maximum 8-hour average (MDA8) O₃ concentration at each site and their spatial distributions for the entire campaign period (Figure 3). The spatial distributions of the model MDA8 O₃ were not well correlated with those of the observations. But, notable disparities were observed in simulating MDA8 O₃ when the different emissions were used. For the north and eastern part of China including Beijing and Shanghai, large negative biases disappear when using EDV3 and KOV5. KOV5 only shows a significant correlation with the surface MDA8 O₃ observations (including 929 sites) than EDV2 and EDV3 in China (0.43 versus 0.01, 0.20). The correlations between the time series of the model MDA8 O₃ and observations varied at each site, with about 40-60% of sites (depending on the emission inventories) showing a correlation coefficient greater than 0.6 (see Supporting Information, Figure S4), and the locations of these sites were scattered. The correlation slightly improved with hourly O₃ concentrations instead of MDA8 O₃, with about 50-60% of sites having a correlation coefficient greater than 0.6 (Supporting Information, Figure S4). For this metric, high correlations occurred in pollution hot spots north of 30°N and the South Coast of China, in which the ratio of HCHO to NO₂ (FNR) was much less than 1, suggesting VOC-limited/NOx-saturated chemical regime (Supporting Information, Figure S5). The model MDA8 O₃ were underestimated for the pollution hot spots with a low HCHO to NO₂ ratio located north of 30°N, suggesting a possibility of model underestimations of anthropogenic VOC emissions causing model MDA8 O₃ biases at these sites. In contrast, the simulated MDA8 O₃ was generally overestimated for sites south of 30°N in which HCHO concentrations were high (Supporting Information, Figure S5). Zhang et al. (2020) reported that simulated biogenic isoprene
from MEGAN was overestimated compared to observation sites under 35°N in China.

The EDV2 and EDV3 showed a positive NO\textsubscript{2} bias over the YRD, NCP, and PRD regions, which include large cities in China (Table 2 and Figure 4-5). On the other hand, EDV2 and EDV3 had low negative NO\textsubscript{2} biases in the NEC and NOC regions (Figure 4). All models demonstrated reasonable NO\textsubscript{2} model performance in the SCG region, where MDA\textsubscript{8} O\textsubscript{3} was overestimated (Figure 2 and 4). In the YRD region, there were large positive NO\textsubscript{2} biases with EDV2, EDV3, and KOV5 (ranging from 6.4 to 22.7 ppb).

Liu et al. (2021) reported that YRD is in a VOC-limited regime when using EDV2. The findings indicated that a reduction in NO\textsubscript{x} emissions led to an increase in O\textsubscript{3} concentrations, while a reduction in VOC emissions resulted in lower O\textsubscript{3} concentrations. The lower bias of O\textsubscript{3} in YRD can be attributed to the combined influence of higher anthropogenic NO\textsubscript{x} emissions and VOC originated from both anthropogenic and biogenic sources (Figure S5). In contrast, KOV5 underestimated NO\textsubscript{2} in the NCP region, while EDV2 and EDV3 did not. All emissions showed significant discrepancies compared to NO\textsubscript{2} observations in the SEC area, with a low correlation coefficient (0.19 to 0.26). EDV3 showed the lowest bias of -1.9 ppb (-0.8 ppb) compared to EDV2 and KOV5 in South Korea (SMA). The daily averaged NO\textsubscript{2} exhibited spatial distributions similar to MDA\textsubscript{8} O\textsubscript{3} and CO (Figure 5). The slopes of regression between the three model simulations and observations were 1.31, 1.03, and 0.8 for EDV2, EDV3, and KOV5, respectively, in China. The correlation coefficients between the simulated NO\textsubscript{2} utilizing EDV2, EDV3, and KOV5 and surface data were around 0.6 in China. EDV2, EDV3, and KOV5 demonstrated good correlations with observations in South Korea (R = 0.69-0.74). Correlation coefficient (R) was the highest with KOV5 in South Korea (R=0.74).
The simulated CO was averaged at each site and compared with observations during the KORUS-AQ campaign period (Figure 6). The three model results showed similar spatial distributions to observations, indicating higher CO concentrations in the NCP, YRD, and PRD regions than their surrounding areas. However, all simulations failed to reproduce the abundance of CO, indicating large negative biases throughout the country. The bias was larger in South Korea than in China.

3.3. Evaluations with the airborne and special surface chemical observations during KORUS-AQ

3.3.1. The aircraft observations

Figure 7 shows the flight paths flown by the DC-8 during the KORUS-AQ campaign period. In Table 3, we compare the model results for O₃, NO₂, CO, HCHO, TOL, XYL, ETE (Ethene or OL2), and ISO with the corresponding observed values for all flight tracks under 2 km height in South Korea (Table 3). On average, the model underestimated O₃ by 15-18 ppb, with EDV3 exhibiting the lowest O₃ bias (-15.1 ppb) compared to EDV2 and KOV5 (-16.8 and -17.5 ppb, respectively). All emissions showed positive biases for NO₂ (0.64 to 1.72 ppb), ETE (0.08 to 0.14 ppb), and ISO (0.1 to 0.11 ppb). However, the model significantly underestimated CO, HCHO, TOL, and XYL for all three emissions. Given the large spatial variability of air pollutants in South Korea, we also sampled aircraft data from six regions (see Figure 7) and compared the three model results with the aircraft observations under 2 km height (Figure 8).

The flight tracks that surveyed large power plants and factories in the Chungnam region on a daily basis are shown in Figure S6 in the Supporting Information. The
The largest negative model O₃ bias was observed over the Chungnam region, with a difference of 38-41 ppb. Emission estimation uncertainties can be significant over this region, where there are large point sources such as coal-burning power plants and petrochemical industries. The model NO₂ agreed with the aircraft observations in SMA, but it tends to overestimate the measurements in the other areas. There were substantial model overestimations of NO₂ with EDV3 over the Chungnam and Busan areas, while KOV5 showed the most reasonable model NO₂ simulations. The model CO near the surface was underestimated in the entire domain, resulting in high negative model CO biases relative to the aircraft observations across the six regions (Figure 8). Additionally, the model HCHO was underestimated by all emission inventories for all subregions, with negative biases being evident in the SMA, Yellow Sea, and Chungnam regions. Other model VOC species, such as TOL, XYL, ETE, and ISO, were also analyzed. These VOC species are classified by their chemical structures and reactivities in the RACM (Stockwell et al., 1997) (Table S2). For example, TOL includes toluene and relatively less reactive aromatics, while XYL includes xylene and more reactive aromatics. The WAS data from DC-8 were lumped into RACM (Supporting Information Table S6, Lu et al., 2013) and were compared with aircraft observations. When the model TOL or XYL was compared with the observed toluene and xylene, the model using KOV5 reasonably reproduced the observed concentrations (light gray bars in Figure 8). However, the model TOL (even using KOV5) underestimated the observed lumped TOL for most of the regions except for Busan (bars including the dark gray part in Figure 8). The model using KOV5 reasonably reproduced the observed xylene or XYL, except for the Chungnam and Busan regions. The observed ethene (or ETE) concentrations were low (< 0.5 ppb), except for the Chungnam region, where the
average of measurements was 2.1 ppb. The model ethene concentration was higher than the observations for the SMA, Kyungbuk, and Busan regions, while it had a large negative bias (-1.6 ~ -1.3 ppb) for the Chungnam region. Regarding isoprene (ISO), one of the most important biogenic VOCs, the model values were larger than the observations by a factor of 2. In summary, underestimated CO and aromatic VOCs are the main features, along with underestimated ozone and HCHO. The largest discrepancies occur over the Chungnam region, where large point sources are located on the west coast of South Korea. The detailed statistics over the SMA and Chungnam area can be obtained from the Supporting Information (Table S7-S8).

Figure 9 displays the vertical distributions of observed and simulated O3 and related species over SMA. The shapes of the simulated profile were in agreement with the observations. Particularly, the model accurately reproduced the observed NO2 profiles though the surface NO2 is underestimated by -4.2 to -0.8 ppb in SMA (Table 2 and Figure 9b). The underestimation of simulated surface NO2 is explained by the overestimation of molybdenum converter method; surface concentrations of NO2 from molybdenum converter is larger than photolytic converter by 13.6% on average and 64% at 4 pm (Figure 10). However, the simulated O3 and HCHO had negative biases of 16.4 ppb and 0.73 ppb, respectively, persisting from the surface to 2 km. Additionally, the simulated CO underestimated the observations by 40% throughout the vertical layer. While the model TOL and XYL, utilizing KOV5, agreed well with the observations below 1 km, the results using EDV2 and EDV3 substantially underestimated the observations throughout the layer. On the other hand, the model simulated ETE and ISO overestimated the observations below 1 km over SMA.
3.3.2. The ground-based observations

During the KORUS-AQ campaign, Pandora and surface measurements were co-located at the Olympic Park. Figure 10 compares the observed diurnal cycle of Pandora vertical columns and surface concentrations of NO$_2$ and HCHO with the model simulations. The photolytic converter was used to measure surface NO$_2$ to minimize positive bias from the molybdenum converter. All emissions reasonably simulated the diurnal patterns of vertical column and surface NO$_2$ and HCHO concentrations. The surface NO$_2$ peak appeared at 07 LT in the model and 08 LT in the observations, associated with the increase of traffic and the under-developed convective boundary layer. On the other hand, the Pandora NO$_2$ column amount increased from 06 LT to 12 LT and stayed at that value throughout the afternoon, indicating the increase of NO$_x$ emissions from morning to afternoon. The model-simulated NO$_2$ columns agreed with those from Pandora in terms of absolute values and diurnal variations. The opposite patterns between surface and column NO$_2$ were also shown in Crawford et al. (2020). The simulated and observed HCHO show similar diurnal variations, but all three emissions underestimated both column and surface HCHO values by up to $-8.5 \times 10^{15}$ molecules cm$^{-2}$ (-46%) at 7 LT and -0.9 ppbv (-26%) at the surface on average. The underestimations of the model HCHO relative to the Pandora and surface observations are similar to findings from comparisons of the model results with the aircraft data (Figure 9). Therefore, the model VOC performance needs to be investigated at the Olympic Park.

The diurnal variations of the model O$_3$, CO, TOL, and XYL were evaluated against the surface observations at the Olympic Park acquired during the KORUS-AQ campaign (Figure 11). The diurnal pattern and hourly averaged mixing ratio of O$_3$ were
well simulated with the three emission inventories with slight model negative biases. The observed CO was 2.7 times higher than the model on average. Considering the diurnal profile of observed TOL and XYL, KOV5 reduced the model negative biases from EDV2 and EDV3, but it still showed negative biases. The model TOL and XYL showed peak concentrations at 08 LT, but the observation had a maximum value at 06 LT. The model biases of XYL (-3.7 to -0.6 ppb, -89 to -20%) were much larger than those in TOL at the surface. Our study demonstrates that the improvement of VOC emission/chemistry representations in the model is necessary for better simulations of air quality over SMA and South Korea.

3.4. The model performances for the Local and Transport Cases

Previous studies have used meteorological conditions to classify synoptic patterns that affect air pollutant concentrations (Park et al. 2021; Peterson et al. 2019). In contrast, we defined the Transport and Local cases by comparing model results that used the EDV3 base emission and the EDV3 zero-out-Chinese emission (see Figure 12). The Local case comprises May 4, May 20, June 2, and June 3 (Supporting Information, Figure S7), while the Transport case includes May 25, May 26, and May 31 (Supporting Information, Figure S8). The Local (Transport) case in this study generally aligns with the Stagnant and Blocking (Transport) cases in Peterson et al. (2019). The Local case has a Chinese contribution to O₃ of under 11%, whereas the Transport case has a Chinese contribution to O₃ of over 46%. EDV3 performed better in simulating O₃ for the Transport case compared to EDV2 and KOV5, with a bias of only 2.7 ppb in comparison with the DC-8 airborne observations. In contrast, for the Local case, all emissions had a negative bias ranging from 15.5-18.2 ppb. See the Table S9 in
Supporting Information to obtain detailed information of model performances against DC-8 measurements for different cases. Surface concentrations of O$_3$ at Olympic Park also exhibited enhanced contributions from Chinese anthropogenic emissions for Transport case (Figure S9). This section focuses on the model simulations using EDV3 and its modified versions, EDV3_Ch2 and EDV3_ChKo2 (doubling Chinese and South Korean CO and VOC emissions).

Figure 13 illustrates the biases in the model O$_3$, CO, and HCHO using EDV3 and its variants relative to DC-8 observations over SMA. The plot highlights differences in biases for the Local and Transport cases. The model O$_3$ biases were negative, and the absolute values of biases were larger in the Local case than in the Transport case (-20% versus -6%). The model CO biases were also negative, and the absolute values of biases were larger in the Transport case than in the Local case. The model HCHO biases were negative and similar for the two cases, except for a larger discrepancy between model and observation in the Local case than in the Transport case. Doubling Chinese CO and VOC emissions (EDV3_Ch2) only slightly reduced biases in the Local case, whereas doubling South Korean CO and VOC emissions, as well as Chinese CO and VOC emissions (EDV3_ChKo2), were necessary to substantially reduce the model biases for the Local case. For the Transport case, doubling Chinese CO and VOC emissions reduced biases to almost zero for CO and HCHO, but the model O$_3$ was much overestimated, with 14% positive biases (from an original bias of -6%). Further increasing South Korean CO and VOC emissions led to overestimations of O$_3$ (20%) and HCHO (33%). These sensitivity tests modifying EDV3 indicate that increases in CO and VOC emissions over South Korea improve the model O$_3$, CO, and VOC simulations. However, increasing Chinese VOC (and CO) emissions may overestimate
the model O₃ for the studied period.

4. Summary and conclusions

We conducted sensitivity tests using WRF-Chem with three different bottom-up emission inventories (EDGAR-HTAP v2, v3, and KORUS v5) to investigate the impacts of different emissions on the simulation of O₃ and precursors in East Asia. This study is the first to use EDGAR-HTAP v3 with WRF-Chem v4.4 and extends the validation domain to the whole of China during the KORUS-AQ campaign period. We extensively validated these emission inventories using both ground and aircraft observations in East Asia.

The three emission inventories accurately reproduced the diurnal profiles and absolute values of surface O₃ for most subregions in China, except for the SCG and SEC areas. However, discrepancies were observed in the model performance for the MDA8 O₃ concentrations, with poor correlations observed over regions with high HCHO concentrations (south of 30°N) and relatively low ratios of FNR (north of 30°N). The emission inventories reasonably reproduced the spatial distribution of daily surface NO₂ concentrations. However, we found that CO was considerably underestimated by the emission inventories over both China and South Korea.

We evaluated the model simulations against vertical profile measurements of O₃, NO₂, CO, HCHO, TOL, XYL, ETE, and ISO from the DC-8 aircraft, as well as surface observations over South Korea. The simulated vertical shapes of O₃, NO₂, CO, HCHO, TOL, XYL, ETE, and ISO agreed well with the DC-8 measurements in the SMA, although negative biases were observed for O₃, CO, TOL, XYL, and HCHO, with the
largest discrepancy between the model results and observations in the Chungnam area. When we compared the simulations with the surface in-situ measurements and PANDORA observations at the Olympic Park in Seoul, the model accurately reproduced the diurnal patterns of surface and vertical columns of NO₂ and HCHO. However, we found that the model underestimated TOL and XYL. This underestimation of TOL and XYL is one of the reasons why the model underestimates O₃ concentrations, as VOCs contribute to NO to NO₂ conversions resulting in O₃ production via photochemistry.

We also classified the flight tracks into two categories: Local and Transport cases. We found that the negative bias of O₃ was much larger under the Local case than the Transport case. When the increment of CO and VOC emissions in South Korea is taken into account, the biases of O₃ are significantly reduced, indicating the need for local emission adjustments to decrease O₃ bias in South Korea.

Our study revealed a consistent overestimation of isoprene over SMA. The uncertainty of biogenic VOC emissions from MEGAN can affect the model O₃ performance. Therefore, to achieve more accurate simulations of O₃ in East Asia, it is essential to explore precise representations of both anthropogenic and biogenic VOC emissions.

**Code and data availability**

WRF-Chem source codes are distributed by NCAR (https://doi:10.5065/D6MK6B4K). WRF-Chem v4.4 is available in the GitHub (wrf-model, 2022). The exact version of WRF-Chem codes and configuration files are archived at

Author contribution

KMK conducted simulations, analysis and wrote the paper. SWK designed this study, secured funding, performed analysis and wrote the paper. SS supported model set-up and contributed to refining the paper. DRB measured VOC data from DC-8. SC acquired ground-based in-situ VOC data at Olympic Park. JHC performed analysis and wrote the paper. LE and GF assisted in setting up the model emissions and discussed about the model performance. AF measured HCHO data from DC-8. JRH measured Pandora data (NO2, HCHO). JH retrieved PBL height and discussed about the results. JJ acquired NO2 data at Olympic Park with different methods. AJW acquired NO2 and O3 data from DC-8. JHW and QZ provided emissions inventories and related information. All authors reviewed the manuscript.
Competing interests

At least one of the (co-)authors is a member of the editorial board of Geoscientific Model Development.

Acknowledgements

This work was supported by the National Research Foundation of Korea (NRF) grant funded by the Korea government (MSIT) (No. 2020R1A2C2014131). S.-W. Kim also acknowledges support from NRF-2018R1A5A1024958. All the computing resources are provided by National Center for Meteorological Supercomputer. The National Center for Atmospheric Research (NCAR) is sponsored by the National Science Foundation (NSF) (NNX16AD96G). We would like to express our gratitude to Glen Diskin for generously providing the CO data from the DC-8 aircraft. We also thanks to Andrew Whitehill and Russell Long for providing the HCHO data from Olympic Park. We would also like to thank Meng Li and Brian McDonald for their valuable discussions, which greatly enhanced our understandings.

References


Table List

Table 1. The model experiments with different emissions.

Table 2. Comparison of the ground-based hourly O₃, NO₂, and CO observations with the simulations utilizing EDGAR-HTAP v2 (EDV2) and v3 (EDV3) and KORUS v5 (KOV5) in each regional box (unit = ppb). N is the number of samples. R is correlation coefficient.

Table 3. Comparison of aircraft-based 1-minute-interval O₃, NO₂, CO, HCHO, TOL, XYL, ETE, and ISO observations with EDV2, EDV3, and KOV5 for all flight cases under 2 km height (unit = ppb). N is the number of samples. R is correlation coefficient.
Figure List

Figure 1. The averaged spatial distribution map of the NO, CO, and TOL emissions from EDGAR-HTAP v2, v3, and KORUS v5 in May.

Figure 2. Averaged O₃ concentrations from ground-based observations and model simulations over the areas that distinguish urban (red box) and non-urban (green box) region (central plot). Box-averaged diurnal cycle (solid lines) of O₃ and 1/4 of standard deviations (filled area) from observations (black), EDV2 (green), EDV3 (blue), and KOV5 (red) by local time are shown. The results are shown for Northern China (NOC, 38-42°N/106-110°E), Sichuan-Chongqing-Guizhou (SCG, 27-33°N/103-109°E), Pearl River Delta (PRD, 21.5-24°N/112-115.5°E), Southeastern China (SEC, 24-28°N/116-120°E), Yangtze River Delta (YRD, 30-33°N/119-122°E), South Korea (KOR, 34.5-38°N/126-130°E), North China Plain (NCP, 34-41°N/113-119°E), and Northeastern China (NEC, 43-47°N/124-130°E).

Figure 3. Comparison of (a) the campaign averaged ground-based maximum daily average of 8-hour O₃ (MDA8 O₃) (unit: ppb) observations and WRF-Chem simulations with (d) EDGAR-HTAP v2 (EDV2), (e) v3 (EDV3), (f) KORUS v5 (KOV5) and (g, h, i) the differences between the observations and model results. The sub-regions are presented with red (urban) and green (non-urban) boxes. The scatter plots comparing averaged observations and the three-emission-based WRF-Chem simulations (green; EDV2, blue; EDV3, red; KOV5) are shown in (b) and (c) for Eastern China and South Korea, respectively. (a, d-e) Color-filled circles in (a), (d), (e), and (f) represent the averaged MDA8 O₃ for the whole campaign period (1st May to 10th June).

Figure 4. The same as Figure 2 except NO₂.

Figure 5. The same as Figure 3 except daily NO₂ (unit: ppb).

Figure 6. The same as Figure 3 except daily CO (unit: ppm).

Figure 7. The DC-8 flight paths during the KORUS-AQ campaign period (black) and 6 regional boxes (1: Seoul Metropolitan Area (SMA); 2: Yellow Sea; 3: Chungnam; 4: Kyungbuk; 5: Gwangju; 6: Busan) (red).
Figure 8. The mean (bars) and 1/4 of standard deviations (whiskers) of (a) O$_3$, (b) NO$_2$, (c) CO, (d) HCHO, (e) TOL, (f) XYL, (g) ethene (ETE), and (h) isoprene (ISO) (unit = ppb) from EDV2 (green), EDV3 (blue), and KOV5 (red) for each box are shown, respectively. TOL and XYL are calculated based on Table S6 (Supporting Information). The contribution of toluene to TOL and m/p-Xylene + o-Xylene to XYL is represented with light grey bars (e, f). The sampling numbers are represented with magenta color above the plots.

Figure 9. Vertically averaged (a) O$_3$, (b) NO$_2$, (c) CO, (d) HCHO, (e) TOL, (f) XYL, (g) ETE, and (h) ISO from DC-8 (black), EDV2 (green), EDV3 (blue), and KOV5 (red) in SMA under 2 km height above ground level. The 1/2 of standard deviations are represented with black whiskers in each 200m layer. The sample number is presented with magenta color on the right side of the plots.

Figure 10. The diurnal cycles of vertical columns and surface concentrations of (a) NO$_2$ and (b) HCHO from Pandora spectrometer (column), and ground-based instruments (TEI 42i NOx analyzer and Aerodyne QCL) at the Olympic Park site (37.5232°N, 127.126°E). EDV2 (green), EDV3 (blue), and KOV5 (red) are compared with observations. The WRF-Chem vertical column concentrations are produced by summing all vertical layers.

Figure 11. Diurnal cycles of surface (a) O$_3$, (b) CO, (c) TOL, and (d) XYL at the Olympic Park site. EDV2 (green), EDV3 (blue), and KOV5 (red) are compared with the observations. 1/4 of standard deviations are represented with grey shades. The average period is from the 11th May to the 10th June.

Figure 12. Averaged O$_3$ (bars) and 1/4 of standard deviations (whiskers) (unit: ppbv) for the 20 DC8 flights (under 2 km height). The observations (grey) are compared with the model results utilizing EDV2 (green), EDV3 (blue), and KOV5 (red). White hatch-filled bars over blue bars are the contribution of Chinese emissions to O$_3$ concentrations obtained from the default and sensitivity model runs with/without Chinese anthropogenic emissions. The Local (5/4,20 and 6/2,3) and Transport (5/25,26,31) cases are shaded with light blue and orange, respectively.

Figure 13. The biases in (a) the model O$_3$, (b) CO, and (c) HCHO concentrations (bars) relative to the DC-8 observations under 2 km height over SMA (dark gray: EDV3, red: EDV3 Ch2, blue: EDV3 ChKo2): (left panel) Local and (right panel) Transport case. Fractional differences (%) are shown in the white boxes.
Table 1. The model experiments with different emissions

<table>
<thead>
<tr>
<th>Experiments</th>
<th>Emissions</th>
</tr>
</thead>
<tbody>
<tr>
<td>EDV2</td>
<td>EDGAR-HTAP v2</td>
</tr>
<tr>
<td>EDV3</td>
<td>EDGAR-HTAP v3</td>
</tr>
<tr>
<td>KOV5</td>
<td>KORUS v5</td>
</tr>
<tr>
<td>EDV3_Ch2</td>
<td>EDGAR-HTAP v3 with double CO, VOC emission in China</td>
</tr>
<tr>
<td>EDV3_ChKo2</td>
<td>EDGAR-HTAP v3 with double CO, VOC emission in China &amp; South Korea</td>
</tr>
</tbody>
</table>
Table 2. Comparison of the ground-based hourly O₃, NO₂, and CO observations with the simulations utilizing EDGAR-HTAP v2 (EDV2) and v3 (EDV3) and KORUS v5 (KOV5) in each regional box (unit = ppb). N is the number of samples. R is correlation coefficient.

<table>
<thead>
<tr>
<th>Region</th>
<th>NCP</th>
<th>SCG</th>
<th>YRD</th>
<th>PRD</th>
<th>KOR (SMA)</th>
<th>NEC</th>
<th>NOC</th>
<th>SEC</th>
</tr>
</thead>
<tbody>
<tr>
<td>OBS Mean</td>
<td>44.5</td>
<td>34.6</td>
<td>38.2</td>
<td>27.9</td>
<td>41.5 (36.6)</td>
<td>40.9</td>
<td>44.3</td>
<td>43</td>
</tr>
<tr>
<td>EDV2 Mean</td>
<td>32.2</td>
<td>53.5</td>
<td>21.6</td>
<td>27.6</td>
<td>40.5 (31.1)</td>
<td>28.6</td>
<td>39.4</td>
<td>40.8</td>
</tr>
<tr>
<td>R</td>
<td>0.65</td>
<td>0.53</td>
<td>0.62</td>
<td>0.61</td>
<td>0.59 (0.60)</td>
<td>0.48</td>
<td>0.63</td>
<td>0.52</td>
</tr>
<tr>
<td>O₃ Bias</td>
<td>-12.3</td>
<td>18.9</td>
<td>-16.6</td>
<td>-0.3</td>
<td>-1.0 (-5.5)</td>
<td>-12.3</td>
<td>-4.9</td>
<td>14.7</td>
</tr>
<tr>
<td>KOV5 Mean</td>
<td>43.4</td>
<td>57.5</td>
<td>35.7</td>
<td>34.7</td>
<td>41.0 (32.6)</td>
<td>35.2</td>
<td>43.7</td>
<td>45.5</td>
</tr>
<tr>
<td>R</td>
<td>0.68</td>
<td>0.55</td>
<td>0.66</td>
<td>0.65</td>
<td>0.56 (0.57)</td>
<td>0.63</td>
<td>0.67</td>
<td>0.55</td>
</tr>
<tr>
<td>NO₂ Obs Mean</td>
<td>17.5</td>
<td>13.8</td>
<td>17.1</td>
<td>12.9</td>
<td>23.2 (32.5)</td>
<td>13.5</td>
<td>11.9</td>
<td>9.6</td>
</tr>
<tr>
<td>EDV2 Mean</td>
<td>25.8</td>
<td>12.7</td>
<td>39.8</td>
<td>22.0</td>
<td>18.8 (29.6)</td>
<td>13.7</td>
<td>12.9</td>
<td>11.0</td>
</tr>
<tr>
<td>R</td>
<td>0.45</td>
<td>0.37</td>
<td>0.38</td>
<td>0.54</td>
<td>0.51 (0.34)</td>
<td>0.49</td>
<td>0.47</td>
<td>0.19</td>
</tr>
<tr>
<td>O₃ Bias</td>
<td>-8.3</td>
<td>-1.0</td>
<td>22.7</td>
<td>9.1</td>
<td>-4.3 (-3.0)</td>
<td>0.2</td>
<td>1.0</td>
<td>1.5</td>
</tr>
<tr>
<td>KOV5 Mean</td>
<td>4.5</td>
<td>20.7</td>
<td>2.8</td>
<td>7.8</td>
<td>0.7 (-3.5)</td>
<td>-3.8</td>
<td>-0.5</td>
<td>16.3</td>
</tr>
<tr>
<td>R</td>
<td>0.71</td>
<td>0.53</td>
<td>0.65</td>
<td>0.70</td>
<td>0.62 (0.64)</td>
<td>0.62</td>
<td>0.67</td>
<td>0.54</td>
</tr>
<tr>
<td>NO₂ Obs Mean</td>
<td>835</td>
<td>597</td>
<td>694</td>
<td>636</td>
<td>443 (493)</td>
<td>527</td>
<td>579</td>
<td>655</td>
</tr>
<tr>
<td>EDV2 Mean</td>
<td>373</td>
<td>389</td>
<td>455</td>
<td>282</td>
<td>175 (210)</td>
<td>206</td>
<td>162</td>
<td>258</td>
</tr>
<tr>
<td>R</td>
<td>0.24</td>
<td>0.20</td>
<td>0.42</td>
<td>0.30</td>
<td>0.31 (0.30)</td>
<td>0.21</td>
<td>0.09</td>
<td>0.18</td>
</tr>
<tr>
<td>NO₂ Bias</td>
<td>-462</td>
<td>-208</td>
<td>-239</td>
<td>-354</td>
<td>-267 (-283)</td>
<td>-321</td>
<td>-417</td>
<td>-397</td>
</tr>
<tr>
<td>KOV5 Mean</td>
<td>374</td>
<td>359</td>
<td>535</td>
<td>282</td>
<td>196 (208)</td>
<td>221</td>
<td>162</td>
<td>256</td>
</tr>
<tr>
<td>R</td>
<td>0.22</td>
<td>0.19</td>
<td>0.35</td>
<td>0.31</td>
<td>0.26 (0.33)</td>
<td>0.24</td>
<td>0.10</td>
<td>0.25</td>
</tr>
<tr>
<td>CO Obs Mean</td>
<td>190</td>
<td>104</td>
<td>93</td>
<td>68</td>
<td>358 (125)</td>
<td>45</td>
<td>28</td>
<td>43</td>
</tr>
<tr>
<td>EDV2 Mean</td>
<td>375</td>
<td>358</td>
<td>475</td>
<td>305</td>
<td>190 (217)</td>
<td>231</td>
<td>176</td>
<td>266</td>
</tr>
<tr>
<td>R</td>
<td>0.27</td>
<td>0.21</td>
<td>0.48</td>
<td>0.29</td>
<td>0.35 (0.36)</td>
<td>0.15</td>
<td>0.10</td>
<td>0.13</td>
</tr>
</tbody>
</table>

1) Urban area, 2) Non-urban area
5 a) Sichuan-Chongqing-Guizhou, b) South Korea, c) Northeastern China, d) Northern China, e) Southeastern China
Table 3. Comparison of aircraft-based 1-minute-interval O₃, NO₂, CO, HCHO, TOL, XYL, ETE, and ISO observations with EDV2, EDV3, and KOV5 for all flight cases under 2 km height (unit = ppb). N is the number of samples. R is correlation coefficient.

<table>
<thead>
<tr>
<th>Species</th>
<th>Type</th>
<th>N</th>
<th>Mean</th>
<th>Bias</th>
<th>σ</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>O₃</td>
<td>OBS</td>
<td>5191</td>
<td>84.4</td>
<td></td>
<td>19.9</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EDV2</td>
<td></td>
<td>67.5</td>
<td>-16.8</td>
<td>16.7</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>EDV3</td>
<td></td>
<td>69.3</td>
<td>-15.1</td>
<td>17.8</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>KOV5</td>
<td></td>
<td>66.9</td>
<td>-17.5</td>
<td>15.8</td>
<td>0.50</td>
</tr>
<tr>
<td>NO₂</td>
<td>OBS</td>
<td>5047</td>
<td>2.19</td>
<td></td>
<td>4.49</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EDV2</td>
<td></td>
<td>3.06</td>
<td>0.87</td>
<td>4.60</td>
<td>0.71</td>
</tr>
<tr>
<td></td>
<td>EDV3</td>
<td></td>
<td>3.91</td>
<td>1.72</td>
<td>5.34</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>KOV5</td>
<td></td>
<td>2.83</td>
<td>0.64</td>
<td>4.73</td>
<td>0.73</td>
</tr>
<tr>
<td>CO</td>
<td>OBS</td>
<td>5575</td>
<td>253</td>
<td></td>
<td>100</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EDV2</td>
<td></td>
<td>148</td>
<td>-105</td>
<td>48</td>
<td>0.60</td>
</tr>
<tr>
<td></td>
<td>EDV3</td>
<td></td>
<td>156</td>
<td>-97</td>
<td>47</td>
<td>0.59</td>
</tr>
<tr>
<td></td>
<td>KOV5</td>
<td></td>
<td>146</td>
<td>-107</td>
<td>43</td>
<td>0.62</td>
</tr>
<tr>
<td>HCHO</td>
<td>OBS</td>
<td>5365</td>
<td>2.37</td>
<td></td>
<td>1.64</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EDV2</td>
<td></td>
<td>1.75</td>
<td>-0.62</td>
<td>1.01</td>
<td>0.69</td>
</tr>
<tr>
<td></td>
<td>EDV3</td>
<td></td>
<td>1.78</td>
<td>-0.59</td>
<td>1.02</td>
<td>0.67</td>
</tr>
<tr>
<td></td>
<td>KOV5</td>
<td></td>
<td>1.80</td>
<td>-0.57</td>
<td>1.10</td>
<td>0.71</td>
</tr>
<tr>
<td>TOL</td>
<td>OBS</td>
<td>730</td>
<td>2.60</td>
<td></td>
<td>2.02</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EDV2</td>
<td></td>
<td>0.47</td>
<td>-2.13</td>
<td>0.38</td>
<td>0.39</td>
</tr>
<tr>
<td></td>
<td>EDV3</td>
<td></td>
<td>0.55</td>
<td>-2.05</td>
<td>0.48</td>
<td>0.38</td>
</tr>
<tr>
<td></td>
<td>KOV5</td>
<td></td>
<td>1.58</td>
<td>-1.01</td>
<td>1.30</td>
<td>0.37</td>
</tr>
<tr>
<td>XYL</td>
<td>OBS</td>
<td>289</td>
<td>0.73</td>
<td></td>
<td>0.65</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EDV2</td>
<td></td>
<td>0.23</td>
<td>-0.50</td>
<td>0.23</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>EDV3</td>
<td></td>
<td>0.30</td>
<td>-0.43</td>
<td>0.31</td>
<td>0.30</td>
</tr>
<tr>
<td></td>
<td>KOV5</td>
<td></td>
<td>0.49</td>
<td>-0.24</td>
<td>0.47</td>
<td>0.27</td>
</tr>
<tr>
<td>ETE</td>
<td>OBS</td>
<td>2573</td>
<td>0.42</td>
<td></td>
<td>1.59</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EDV2</td>
<td></td>
<td>0.51</td>
<td>0.09</td>
<td>0.65</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>EDV3</td>
<td></td>
<td>0.56</td>
<td>0.14</td>
<td>0.76</td>
<td>0.15</td>
</tr>
<tr>
<td></td>
<td>KOV5</td>
<td></td>
<td>0.51</td>
<td>0.08</td>
<td>0.58</td>
<td>0.20</td>
</tr>
<tr>
<td>ISO</td>
<td>OBS</td>
<td>1294</td>
<td>0.08</td>
<td></td>
<td>0.09</td>
<td></td>
</tr>
<tr>
<td></td>
<td>EDV2</td>
<td></td>
<td>0.18</td>
<td>0.10</td>
<td>0.21</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>EDV3</td>
<td></td>
<td>0.19</td>
<td>0.11</td>
<td>0.20</td>
<td>0.41</td>
</tr>
<tr>
<td></td>
<td>KOV5</td>
<td></td>
<td>0.17</td>
<td>0.10</td>
<td>0.20</td>
<td>0.42</td>
</tr>
</tbody>
</table>
Figure 1. The averaged spatial distribution map of the NO, CO, and TOL emissions from EDGAR-HTAP v2, v3, and KORUS v5 in May.
Figure 2. Averaged O₃ concentrations from ground-based observations and model simulations over the areas that distinguish urban (red box) and non-urban (green box) region (central plot). Box-averaged diurnal cycle (solid lines) of O₃ and 1/4 of standard deviations (filled area) from observations (black), EDV2 (green), EDV3 (blue), and KOV5 (red) by local time are shown. The results are shown for Northern China (NOC, 38-42°N/106-110°E), Sichuan-Chongqing-Guizhou (SCG, 27-33°N/103-109°E), Pearl River Delta (PRD, 21.5-24°N/112-115.5°E), Southeastern China (SEC, 24-28°N/116-120°E), Yangtze River Delta (YRD, 30-33°N/119-122°E), South Korea (KOR, 34.5-38°N/126-130°E), North China Plain (NCP, 34-41°N/113-119°E), and Northeastern China (NEC, 43-47°N/124-130°E).
Figure 3. Comparison of (a) the campaign averaged ground-based maximum daily average of 8-hour O₃ (MDA8 O₃) (unit: ppb) observations and WRF-Chem simulations with (d) EDGAR-HTAP v2 (EDV2), (e) v3 (EDV3), (f) KORUS v5 (KOV5) and (g, h, i) the differences between the observations and model results. The sub-regions are presented with red (urban) and green (non-urban) boxes. The scatter plots comparing averaged observations and the three-emission-based WRF-Chem simulations (green; EDV2, blue; EDV3, red; KOV5) are shown in (b) and (c) for Eastern China and South Korea, respectively. (a, d-e) Color-filled circles in (a), (d), (e), and (f) represent the averaged MDA8 O₃ for the whole campaign period (1st May to 10th June).
Figure 4. The same as Figure 2 except NO₂.
Figure 5. The same as Figure 3 except daily NO$_2$ (unit: ppb).
Figure 6. The same as Figure 3 except daily CO (unit: ppm).
Figure 7. The DC-8 flight paths during the KORUS-AQ campaign period (black) and 6 regional boxes (1: Seoul Metropolitan Area (SMA); 2: Yellow Sea; 3: Chungnam; 4: Kyungbuk; 5: Gwangju; 6: Busan) (red).
Figure 8. The mean (bars) and 1/4 of standard deviations (whiskers) of (a) O3, (b) NO2, (c) CO, (d) HCHO, (e) TOL, (f) XYL, (g) ethene (ETE), and (h) isoprene (ISO) (unit = ppb) from EDV2 (green), EDV3 (blue), and KOV5 (red) for each box are shown, respectively. TOL and XYL are calculated based on Table S6 (Supporting Information). The contribution of toluene to TOL and m/p-Xylene + o-Xylene to XYL is represented with light grey bars (e, f). The sampling numbers are represented with magenta color above the plots.
Figure 9. Vertically averaged (a) O₃, (b) NO₂, (c) CO, (d) HCHO, (e) TOL, (f) XYL, (g) ETE, and (h) ISO from DC-8 (black), EDV2 (green), EDV3 (blue), and KOV5 (red) in SMA under 2 km height above ground level. The 1/2 of standard deviations are represented with black whiskers in each 200m layer. The sample number is presented with magenta color on the right side of the plots.
Figure 10. The diurnal cycles of vertical columns and surface concentrations of (a) NO$_2$ and (b) HCHO from Pandora spectrometer (column), and ground-based instruments (TEI 42i NO$_x$ analyzer and Aerodyne QCL) at the Olympic Park site (37.5232°N, 127.126°E). Surface concentrations of NO$_2$ are obtained by the two methods: molybdenum converter and photolytic method. EDV2 (green), EDV3 (blue), and KOV5 (red) are compared with observations. The WRF-Chem vertical column concentrations are produced by summing all vertical layers.
Figure 11. Diurnal cycles of surface (a) O₃, (b) CO, (c) TOL, and (d) XYL at the Olympic Park site. EDV2 (green), EDV3 (blue), and KOV5 (red) are compared with the observations. 1/4 of standard deviations are represented with grey shades. The average period is from the 11th May to the 10th June.
Figure 12. Averaged O₃ (bars) and 1/4 of standard deviations (whiskers) (unit: ppbv) for the 20 DC8 flights (under 2 km height). The observations (grey) are compared with the model results utilizing EDV2 (green), EDV3 (blue), and KOV5 (red). White hatch-filled bars over blue bars are the contribution of Chinese emissions to O₃ concentrations obtained from the default and sensitivity model runs with/without Chinese anthropogenic emissions. The Local (5/4,20 and 6/2,3) and Transport (5/25,26,31) cases are shaded with light blue and orange, respectively.
Figure 13. The biases in (a) the model O₃, (b) CO, and (c) HCHO concentrations (bars) relative to the DC-8 observations under 2 km height over SMA (dark gray: EDV3, red: EDV3 Ch2, blue: EDV3 ChKo2): (left panel) Local and (right panel) Transport case. Fractional differences (%) are shown in the white boxes.