Impact of Intercontinental Pollution Transport on North American Ozone Air Pollution: An HTAP Phase II Multi-model Study

Min Huang1,2, Gregory R. Carmichael3, R. Bradley Pierce4, Duseong S. Jo5, Rokjin J. Park5, Johannes Flemming6, Louisa K. Emmons7, Kevin W. Bowman8, Daven K. Henze9, Yanko Davila9, Kengo Sudo10, Jan Eiof Jonson11, Marianne Tronstad Lund12, Greet Janssens-Maenhout13, Frank J. Dentener13, Terry J. Keating14, Hilke Oetjen8, Vivienne H. Payne8

1George Mason University, Fairfax, VA, USA
2University of Maryland, College Park, MD, USA
3University of Iowa, Iowa City, IA, USA
4NOAA National Environmental Satellite, Data, and Information Service, Madison, WI, USA
5Seoul National University, Seoul, Korea
6European Center for Medium range Weather Forecasting, Reading, UK
7National Center for Atmospheric Research, Boulder, CO, USA
8Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA, USA
9University of Colorado-Boulder, Boulder, CO, USA
10Nagoya University, Furocho, Chigusa-ku, Nagoya, Japan
11Norwegian Meteorological Institute, Oslo, Norway
12Center for International Climate and Environmental Research, Oslo, Norway
13European Commission, Joint Research Center, Ispra, Italy
14US Environmental Protection Agency, Research Triangle Park, NC, USA

*Now at: University of Leicester, Leicester, UK

Correspondence to: Min Huang (mhuang10@gmu.edu)

Abstract

The recent update on the US National Ambient Air Quality Standards of the ground-level ozone (O₃) can benefit from a better understanding of its source contributions in different US regions during recent years. In the Hemispheric Transport of Air Pollution experiment Phase 1 (HTAP1), various global models were used to determine the O₃ source-receptor relationships among three continents in the North Hemisphere in 2001. In support of the HTAP Phase 2 (HTAP2) experiment that studies more recent years and involves higher-resolution global models and regional models’ participation, we conduct a number of regional scale Sulfur Transport and dEposition Model (STEM) air quality base and sensitivity simulations over North America during May-June 2010. The STEM top and lateral chemical boundary conditions were downscalped from three global chemical transport models’ (i.e., GEOS-Chem, RAQMS, and ECMWF C-IIFS) base and sensitivity simulations (in which the East Asian anthropogenic emissions were reduced by 20%). We perform analyses not only on large spatial/temporal scales relative to the HTAP1, but also on subcontinental- and event-scale that are more relevant to the US air quality management. The differences between STEM surface O₃ sensitivities (including the 24h mean and the US policy-relevant maximum daily 8h average (MDA8) metric averaged throughout the study period and during a selected pollution transport event) and its corresponding boundary condition model’s are often smaller than those among its boundary condition models. The STEM sensitivities are also compared with the mean sensitivities estimated by multi- global models, which are higher.
than the HTAP1 reported 2001 conditions, as well as the 2001-2005 conditions studied using the
tagged tracer approach. This indicates the increasing impacts of the East Asian anthropogenic
pollution on North America during 2001-2010. The GEOS-Chem sensitivities indicate that the
East Asian anthropogenic NOx emissions matter more than the other East Asian O3 precursors to
the North American O3, qualitatively consistent with previous adjoint sensitivity calculations. An
additional STEM simulation was performed in which the boundary conditions were downscaled
from a global RAQMS simulation without East Asian anthropogenic emissions, to assess the
scalability of O3 sensitivities to the size of the emission perturbation. The scalability is spatially
varying, and the full source contribution obtained by linearly scaling the North American regional
mean O3 sensitivity to the 20% reduction in the East Asian emissions may be underestimated.

Satellite NO2 (KNMI OMI) and O3 (TES, JPL-IASI, OMI, MLS, and AIRS) products help
detect pollution episodes, quantify or/and reduce the uncertainties in the bottom-up NOx emissions
and the model transported background O3. Based on model calculations and satellite/surface
observations during a selected event, we show the different influences from stratospheric O3
intrusions along with the transported East Asian pollution on O3 in the western and the eastern US.
Future directions of using satellite data in air quality research are also suggested.

1. Introduction

Tropospheric ozone (O3), a short-lived trace gas with a lifetime ranging from hours in the
boundary layer to weeks in the free troposphere, affects tropospheric chemistry, harms human and
ecosystem health, and induces climate change on local, regional and global scales (Jerrold et al.,
2009; Smith et al., 2009; Anenberg et al., 2010; Mauzerall and Wang, 2001; Avnery et al., 2011a,
b; Shindell et al., 2009, 2013; Bowman and Henze, 2012; Stevenson et al., 2006, 2013; Monks et
al., 2013). It has been recognized that the uneven distributions of tropospheric O3 can be attributed
to the stratosphere as well as local, regional and distant emission sources, through complicated
processes that occur on synoptic, meso- and micro-scales (Task Force on Hemispheric Transport
of Air Pollution (HTAP), 2010; National Research Council (NRC), 2009; Maas and Grennfelt,
2016). The mitigation of O3’s climate and health impacts would benefit from efforts to control the
emissions of its precursors from the various emission sources (United Nations Environment
Programme (UNEP) and World Meteorological Organization (WMO), 2011), such as nitrogen
oxides (NOx), carbon monoxide (CO), methane (CH4), and non-methane volatile organic
compounds (NMVOCs).

Ground-level O3 is one of the six criteria air pollutants regulated by the US Environmental
Protection Agency (EPA), and the US National Ambient Air Quality Standards (NAAQS) has
recently been lowered to 70 ppbv to better protect Americans’ health and the environment. Issues
regarding making accurate estimates of the total O3 as well as the background O3 level (defined as
the concentration that is not affected by recent locally-emitted or produced anthropogenic pollution
(e.g., McDonald-Buller et al., 2011; Zhang et al., 2011; Fiore et al., 2014; Huang et al., 2015),
have been recently discussed as part of the implementation of the new US O3 standard (US EPA,
2016a, b). This includes assessing the impacts of various components of the background O3, such
as stratospheric O3, local natural sources such as biogenic, lightning and wildfire emissions, as
well as the long-range transport (LRT) of pollution. The impact of the trans-Pacific pollution
transport on US air quality has been evaluated in numerous studies over the past decades (e.g.,
Fiore et al., 2009; Reidmiller et al., 2009; Zhang et al., 2008, 2009; Huang et al., 2010, 2013a; Lin
et al., 2012a; US EPA, 2016a). It has been found that the increasing trends of pollution in the
upwind continents, especially the populated East Asia (e.g., Zhang et al., 2014; Susaya et al., 2013;
Wang et al., 2012), may partially offset the US air quality improvements in recent decades due to
the regional and local emission controls (e.g., Verstraeten et al., 2015; Ambrose et al., 2011;
Wigder et al., 2013; Cooper et al., 2010; Parrish et al., 2009, 2012; Gratz et al., 2014). A better
understanding of the processes that determine the O₃ pollution levels, as well as an improved
capability of attributing the air pollution to nearby or distant sources is needed to assist with
designing and implementing effective local emission control strategies to comply with the tighter
air quality standards.

Chemical transport models are often used to reproduce and attribute the observed O₃ levels,
including assessing the impacts of the internationally transported O₃ on the US air quality. In the
HTAP modeling experiment Phase I (HTAP1), various global models with horizontal resolutions
ranging from 1°×1° to 5°×5°, only around half of which are finer than 3°×3°, were used to
determine the O₃ source-receptor (SR) relationships among three continents in the Northern
Hemisphere in 2001 (Chapter 4 in HTAP, 2010). The global model based SR relationships in
HTAP1 determined using the emission perturbation approach (i.e., calculating the changes of O₃
at the receptor regions in response to a 20% reduction in the emission inputs in a given source
region) were reported as either monthly 24h mean values or policy-relevant metrics such as the
maximum daily 8h average (MDA8) for the US (e.g., Fiore et al., 2009; Reidmiller et al., 2009).
Large intermodel diversity was found in the simulated total O₃ and the intercontinentally
transported pollution for the chosen SR pairs in the northern midlatitudes, indicating the challenges
with model simulations to accurately represent the key atmospheric processes. Multi-model mean
results were the foci of in these studies with the assumption that this approach can reduce the
uncertainty from the single model estimates for monthly or seasonal means. “Ensemble” model
analyses have been suggested by some US stakeholders as one of the methods for helping with the
characterization of the background O₃ components (US EPA, 2016b). Although the multi-model
approach can help identify some of the weaknesses of the individual models and may produce
more reliable estimates, it is necessary to well understand the uncertainties inherent in using the
same set of anthropogenic emissions in all these model simulations. Satellite observations over the
regions with limited in-situ measurements such as the East Asia can be particularly helpful for
quantifying such uncertainties.

The 20% emission perturbation in the HTAP1 modeling experiment was chosen to produce
a sizeable (i.e., larger than numerical noise) and realistic impact, but small enough in the assumed
near-linear atmospheric chemistry regime. The scalability of the modeled O₃ sensitivities to the
size of the emission perturbations has been assessed on continental scale (Wu et al., 2009; Fiore et
al., 2009; HTAP, 2010; Wild et al., 2012; Emmons et al., 2012). The receptor O₃ responses to the
source-region emission perturbations are found to be fairly linear within ~50% of the perturbations.
However, due to the chemical non-linearity, the full source contribution obtained by linearly
scaling the receptor regional mean O₃ sensitivity to the 20% reduction in the source region
emissions may be underestimated, and the scalability depended on seasons and the perturbed
emission species. Huang et al. (2013b) investigated the scalability of the O₃ sensitivity between
the southern California-US intermountain west SR pair for May 2010, in which study the southern
California anthropogenic emissions were perturbed by multiple amounts of +50%, -50%, -100%.
They reported that the scalability of the O₃ sensitivities changed with the distance from the source
regions. Further analyses on the scalability of these modeled O3 sensitivities during recent years especially for the East Asia-NAM SR pair, as well as their spatial variability, are still needed.

Furthermore, results generated using the emission perturbation approach need to be compared with those based on the other methods (e.g., tagged tracers, adjoint sensitivity).

Previous studies have demonstrated the advantages of high resolution chemical transport modeling for understanding SR relationships (e.g., Lin et al., 2010 for Europe and the East Asia; Lin et al., 2012a; Huang et al., 2010, 2013a for Asia and NAM). Using observations (satellite, sondes, aircraft) along with single model simulations, a few studies have reported that the US O3 sensitivities to extra-regional sources is region-dependent (e.g., Lin et al., 2012a; Langford et al., 2011; Ott et al., 2016), and therefore the necessity of evaluating the extra-regional source impacts on event scale has been emphasized in these studies as well as in US EPA (2016a, b). The HTAP Phase 2 (HTAP2) multi-model experiment, initiated in 2012, is designed to advance the understanding of the impact of intercontinental pollution transport during more recent years (i.e., 2008-2010) involving a number of global and regional models’ participation (Galmarini et al., 2016; Koffi et al., 2016). The regional models are anticipated to help connect the analyses over global and regional scales and enable discussions on small spatial (e.g., subcontinental) and temporal scales (i.e., event based analyses). The use of satellite products for identifying the transport events as well as for quantitative model evaluation is also encouraged in the work plan.

The HTAP2 modeling experiment was sequentially conducted in two steps. First, similar to the HTAP1 experiment, a group of global models with different resolutions conducted base and emission perturbation sensitivity simulations to determine the pollutants’ SR relationships. All models in their base simulations used the same set of harmonized sector-based global anthropogenic emissions developed specifically for the HTAP2 modeling experiment (Janssens-Maenhout et al., 2015). Most of these global models recorded only key chemical species from their base and sensitivity simulations in varied temporal frequencies. Several global models saved the three-dimensional (3D) chemical fields of more species with a 3- or 6-hour interval, which are suitable for being used as regional models’ chemical boundary conditions. In the second step, regional models conducted base and sensitivity simulations to analyze the pollutants’ SR relationships in greater detail. The regional model simulations used the same set of anthropogenic emissions as the global models within their simulation domains, and the chemical boundary conditions in these regional simulations were downscaled from the base and sensitivity simulations from the selected boundary condition model outputs. For regional simulations over the North America and Europe, boundary conditions were mostly taken from a single model such as the ECMWF C-IFS or GEOS-Chem.

In this study, we performed a number of regional scale STEM (Sulfur Transport and dEposition Model) base and sensitivity simulations over the NAM during May-June 2010, during which period strong trans-Pacific pollution transport were shown to episodically impact the US (Lin et al., 2012a). Extending the HTAP2 regional simulations’ basic setup, the STEM top and lateral chemical boundary conditions were downscaled from three global models’ (i.e., the Seoul National University (SNU) GEOS-Chem, RAQMS, and the ECMWF C-IFS) base and sensitivity simulations in which the East Asian anthropogenic emissions were reduced. The STEM surface O3 sensitivities over the NAM region based on different boundary condition models were inter-compared, in terms of the regional averages and the spatial patterns on monthly basis and during a selected event identified by satellite O3 and CO products. These were also compared with the
sensitivities estimated by their corresponding boundary condition models as well as all HTAP2 participating global models. An additional regional simulation was performed in which the STEM boundary conditions were downscaled from one global model simulation without the East Asian anthropogenic emissions, and the nonlinear relationship between the O₃ sensitivity and the size of the emission perturbation is discussed. In the discussion, we emphasize: 1) the differences in O₃ sensitivities generated from the HTAP2 and HTAP1 experiments to help address how the LRT impacts on NAM changed through time; 2) how the multi-model approach, as well as the refined model experiment design in HTAP2 can help advance our understanding of the LRT impacts, especially the benefits of increasing the global models’ resolutions and involving the regional models; 3) the usefulness of satellite observations for better understanding the sources of uncertainties in the modeled total O₃ (e.g., from the emission and regional models’ boundary condition inputs) as well as for reducing the uncertainties in some of these model inputs via chemical data assimilation.

2. Methods

2.1. Anthropogenic emission inputs

Identical anthropogenic emissions were used in all global and regional chemical transport models’ base and sensitivity simulations. This monthly-varying harmonized sectoral (i.e., power, industry, transportation, residential, shipping, aircraft, agriculture) emission inventory was provided on a gridded 0.1°×0.1° resolution for the years of 2008 and 2010, by compiling the officially reported emissions at the national scale (Janssens-Maenhout et al., 2015; http://edgar.jrc.ec.europa.eu/htap_v2). The temporal profiles for developing the monthly-varying emissions differ by region and sector. The amount of emissions of key O₃ precursors (CO, NOₓ, NMVOCs) from both years are summarized in Table S1 for the four major emissions sectors, over the NAM (US+Canada), based on data from the US EPA and the Environmental Canada, which shows lower emissions from the previous years as also discussed in Pouliot et al., 2015), MICS-Asia regions (south, southeast, and east Asia, based on country inventory for China and from the Clean Air Policy System and the Regional Emission inventory in ASia 2.1, more information also in Li et al., 2015), and for over the world. For all of these species, global total emissions in 2008 and 2010 are similar. The NOₓ, NMVOC, and CO emissions decreased from 2008 to 2010 over the NAM by 10.7%, 9.4%, and 15.7%, respectively. For 2010, the transportation sector contributed more than the other sectors to NAM anthropogenic NOₓ and CO emissions; industrial sector contributed more than the other sectors to NMVOCs emissions. Over East Asian countries, these emissions are ~2-5 times higher than the US emissions, and the NOₓ, NMVOC and CO emissions increased over Asia by 7.3%, 7.2% and 1.0%, with the dominant emission sectors in 2010 of transportation, industry, and residential, respectively. For both years, the emissions over the MICS-Asia regions contribute to over 40% of the global emissions. For these key O₃ precursors, the East Asian countries contribute to 45% (NMVOCs)-70% (NOₓ) of the emissions in the MICS-Asia domain in both years, and the south Asian countries contribute to ~22% (NOₓ)-34% (NMVOCs) of the MICS-Asia emissions. Non-anthropogenic emission inputs used in different models’ simulations may differ. As this paper focuses on the impact of anthropogenic emissions, we do not introduce this information in detail.

2.2. Region definitions for the SR study and the model base and sensitivity simulations

2.2.1. Base and 20% emission perturbation simulations from global and regional models
The HTAP2 simulations from eight global models, used in this study, are listed in Table 231. Horizontal resolutions of these models range from finer than 1° to coarser than 2.5°, and overall these resolutions are higher than the HTAP1 participating models’. Relevant references for the RAQMS model and the SNU GEOS-Chem are Pierce et al. (2007, 2009) and Park et al. (2004) (with additional descriptions on its HTAP simulation configurations at: http://iek8wikis.iek.fz-juelich.de/HTAPWiki/WP2.3?action=AttachFile&do=view&target=_README_GEOS-Chem.pdf), respectively. The descriptions of the remaining models can be found in published HTAP2 works such as in Stjern et al. (2016). Figure 1 defines the source regions used in the HTAP2 SR relationship study and we will focus in this study on assessing the East Asia (EAS), South Asia (SAS), Europe (EUR), and non-NAM anthropogenic source (interchangeable in this paper with “(all) foreign”) impacts on the NAM O₃ levels in 2010. Specifically, each model performed a base simulation and a number of sensitivity simulations in which the original HTAP2 anthropogenic emissions for all species and sectors in a defined source region were perturbed by a certain amount (referring to 20% as in most cases, unless specified differently) and these cases are defined in Table 1 as *source region*ALL (where “ALL” refers to “all species and sectors”, consistent with HTAP1 and HTAP2’s naming convention). The O₃ differences R(O₃, *source region*, *perturbation*) over the NAM were then calculated between each model’s base and sensitivity simulations:

\[
R(O₃, \text{EAS, 20%}) = \text{BASE O₃-EASALL O₃} \quad (1a)
\]

\[
R(O₃, \text{SAS, 20%}) = \text{BASE O₃-SASALL O₃} \quad (1b)
\]

\[
R(O₃, \text{EUR, 20%}) = \text{BASE O₃-EURALL O₃} \quad (1c)
\]

\[
R(O₃, \text{non-NAM, 20%}) = \text{NAMALL O₃-GLOALL O₃} \quad (1d)
\]

The monthly-mean R(O₃, *source region*, 20%) values were averaged over the NAM region for the analysis and compared with the findings in the HTAP1 study (e.g., Fiore et al., 2009). It is worth mentioning that the rectangular source regions defined in HTAP1 were modified in HTAP2 to align with the geo-political borders. For EAS and SAS, the regions not overlapped by HTAP1 and HTAP2 are mostly in the less populated/polluted regions such as the northwestern China, according to the HTAP2 emission maps (http://edgar.jrc.ec.europa.eu/htap_v2/index.php).

HTAP2’s EUR domain excludes certain regions in Russia/Belarus/Kazakhstan, Middle East and North Africa that are included in HTAP1’s EUR domain. The impact of emissions over these regions on comparing the NAM R(O₃, EUR, 20%) values in HTAP1 and HTAP2 will be discussed in Section 3.2.1.

A unitless “Response to Extra-Regional Emission Reductions (RERER)” metric (Galmarini et al., 2016), as defined in eq. (2), was also calculated to measure the importance of local versus non-local sources to NAM’s O₃ levels:

\[
\text{RERER (O₃, NAM)} = \frac{R_{O₃, \text{non-NAM, 20%}}}{R_{O₃, \text{global, 20%}}} = \frac{(\text{NAMALL O₃-GLOALL O₃})}{(\text{BASE O₃-GLOALL O₃})} \quad (2)
\]

The denominator and numerator terms of RERER represent the impacts of global and non-NAM anthropogenic emissions on NAM O₃, respectively. The higher the NAM RERER value is, the stronger impact from non-local sources on NAM is indicated. The RERER value can exceed 1, when emission reductions led to increasing concentrations (e.g. O₃ titration by nitrogen monoxide (NO)).
The STEM (version 2K3) regional simulations were then performed on a 60 km×60 km horizontal resolution (a typical coarse regional model resolution) grid over NAM within the domain defined in Figure 2b during May-June 2010. The meteorological conditions in Spring 2010 were compared with the climatology from the NCEP/NCAR reanalysis data for the 1981-2010 period (Kalnay et al., 1996) in Huang et al. (2013b), concluding that this spring represents a period of stronger-than-climatological average spring trans-Pacific transport (based on a stronger meridional gradient in the North Pacific and higher Pacific/North American (PNA) indexes) and stratospheric intrusion conditions (based on the tropopause pressure and the UTLS relative humidity). The mean near-surface air temperatures in the western US in this spring were lower than the climatology, with larger anomalies in the mountain states, which may have led to weaker local \( \text{O}_3 \) production and decomposition of the transported peroxyacetyl nitrates (PAN). In contrast, higher-than-normal temperatures were found in the eastern US that favored anomalously strong local \( \text{O}_3 \) production.

STEM has been used to interpret the observations collected by satellites and during aircraft campaigns in the past decade (e.g., Carmichael et al., 2003a, b; Huang et al., 2010, 2013a, b, 2014, 2015). STEM calculates gas-phase chemistry reactions based on the SAPRC 99 gaseous chemical mechanism (Carter, 2000) with thirty photolysis rates calculated online by the Tropospheric Ultraviolet-Visible radiation model (Madronich et al., 2002). Most of the key configurations of the 60 km base simulations are the same as those described in Lapina et al. (2014), i.e., meteorological fields were pre-calculated by the Advanced Research Weather Research and Forecasting Model (WRF-ARW, Skamarock et al., 2008) version 3.3.1 forced by the North American Regional Reanalysis data (Mesinger et al., 2006), using a similar set of the physics configuration to those in Huang et al. (2013a). Biomass burning emissions are from the Fire INventory from NCAR (FINN) inventory version 1.0 (Wiedinmyer et al., 2011). Biogenic emissions were calculated by the Model of Emissions of Gases and Aerosols from Nature (MEGAN) version 2.1 (Guenther et al., 2012), driven by the WRF meteorology. Lightning NO\(_3\) emissions are generated following the method in Allen et al. (2012), with the flash rates determined by the WRF convective precipitation and scaled to the National Lightning Detection Network flash rates. A major difference of the STEM simulations in this study from the Lapina (2014) study is that the anthropogenic emissions were replaced with the monthly-mean HTAP2 inventory with no weekday-weekend variability applied, rather than the earlier National Emission Inventory (NEI) 2005 in which the weekday-weekend variability exists. This change can introduce uncertainty, but was done to ensure consistency with the HTAP2 global model simulations, that also didn’t use daily variable emissions for any regions.

The VOC speciation for the SPRAC 99 chemical mechanism in the NEI 2005 (ftp://aftp.fsl.noaa.gov/divisions/taq/emissions_data_2005) were applied to break down the total NMVOC emissions provided in the HTAP2 inventory. The VOC speciation based on the year of 2005 can be unrealistic for 2005 as well as 2010 as studies have reported variable temporal changes of different VOC species in some US cities (e.g., Warneke et al., 2012). The time-varying lateral and top boundary conditions in the STEM base simulations were downscaled from three global models (i.e., 3 hourly SNU GEOS-Chem, 3 hourly ECMWF C-IFS, and 6 hourly RAQMS) base simulations. In support of the SR relationship study to quantify the East Asia anthropogenic impacts on the NAM, three STEM sensitivity simulations were also conducted in which the STEM boundary conditions were downscaled from the EASALL sensitivity simulations by these three global models (Table 1b). All STEM simulated 3D chemical fields were saved hourly for the convenience of calculating the US primary \( \text{O}_3 \) standard metric MDA8 as well as the quantitative
comparisons against the satellite Level 2 (L2) O₃ products. The STEM base case surface O₃ performance and its O₃ sensitivities were also compared with those of its boundary condition models as well as the multi- global model means. The latitude/longitude ranges (20-50°N/130-
65°W) of NAM for the global and regional model based sensitivity calculations were selected to mainly account for the coverage of the STEM domain, which are slightly different from the definition of North America in HTAP1.

2.2.2. Additional base and sensitivity simulations from selected models

In addition to the base and 20% EAS all-category emission perturbation simulations, the global RAQMS model conducted a sensitivity simulation in which the East Asian anthropogenic emissions were zeroed out, which was also used as STEM’s boundary conditions (Table 1b). We calculate the “S₀₃” metric (eq. (3)) using the O₃ sensitivities in STEM and RAQMS at the receptor regions in response to both 20% and 100% of emission reductions, to explore the relationships between the O₃ sensitivity and the size of the emission perturbation. A closer-to-one “S₀₃” value indicates higher scalability of the sensitivity based on the 20% emission perturbation method for obtaining the full “contribution” of the East Asian anthropogenic emissions on the NAM O₃.

\[
S₀₃ = R(O₃, \text{EAS, } 100\%) / R(O₃, \text{EAS, } 20\%) / 5
\] (3)

The RAQMS model also provided a base simulation that assimilated satellite O₃ products from the Ozone Monitoring Instrument (OMI, Levelt et al., 2006) and Microwave Limb Sounder (MLS, Livesey et al., 2008) (Pierce et al., 2007), which was used to help better understand the regional model base run error sources, as well as for demonstrating the use of satellite observations to help improve the representation of the trans-boundary pollution.

We also used a number of sensitivity simulations produced by the GEOS-Chem adjoint model v35f in which the emissions from selected anthropogenic emission sectors (power&industry, transportation, residential) or individual O₃ precursor chemical species (NOₓ, VOC, CO) over the East Asia were reduced by 20%. Additional simulations for the 2008-2009 periods by the SNU GEOS-Chem were also utilized to quantify the East Asia and non-NAM anthropogenic source impacts in comparison with the 2010 conditions that we mainly focus on in this study.

2.3. In-situ and satellite observations

2.3.1. In-situ observations

The hourly O₃ observations at the Clean Air Status and Trends Network (CASTNET, http://epa.gov/castnet/javaweb/index.html) sites were used to evaluate the global and regional models’ base simulations in four subregions: western US (i.e., the EPA regions 8, 9, 10); southern US (i.e., the EPA regions 4 and 6), the Midwest (i.e., the EPA regions 5 and 7), and the northeast (i.e., the EPA regions 1-3). The numbers of sites used in global and regional models’ evaluation in each US subregion are summarized in Tables 2-3. The locations of these sites and the subregions they belong to are indicated in Figure 2b, overlaid on a model-based terrain height map. A majority of the CASTNET sites in the western US are located at high elevation (>1 km) remote or rural regions, more susceptible to the trans-boundary pollution (e.g., Jaffé, 2011). Most of the sites in the other three subregions are located in low elevation regions, mainly affected by local and
regional pollution. The model-based terrain heights fairly well represent the reality on subregional scale – the differences between the actual and model-based subregional mean terrain heights at the CASTNET sites are smaller than 0.1 km (Table 3).

During May-June 2010, intense ozonesonde measurements were made at multiple California locations (Cooper et al., 2011), in support of the NOAA “California Nexus (CalNex): Research at the Nexus of Air Quality and Climate Change” field experiment (Ryerson et al., 2013). They have been used to evaluate the simulated $O_3$ vertical profiles by the HTAP2 participating models. The detailed evaluation results have been shown by Cooper et al. (2016), and will be covered by subsequent publications.

2.3.2. Satellite products

In a case study L2 and L3 $O_3$ and CO retrievals from several satellite instruments were used to assess the impacts of trans-Pacific pollution transport and stratospheric $O_3$ intrusions on NAM $O_3$ levels in early May. These include: 1) the early afternoon $O_3$ and CO profiles version 5 from the Tropospheric Emission Spectrometer (TES) (Beer et al., 2001; Beer, 2006) on the Aura satellite; 2) the mid-morning $O_3$ profiles from the METOP-Infrared Atmospheric Sounding Interferometer (IASI), which were retrieved using the Jet Propulsion Laboratory (JPL) TES optimal estimation retrieval algorithm (Bowman et al., 2006) for selected areas including the western US (Oetjen et al., 2014, 2016); as well as 3) the early afternoon L3 $O_3$ and CO maps (version 6, 1°×1°) from the Aqua Atmospheric Infrared Sounder (AIRS) instrument. The TES tropospheric $O_3$ retrieval is often sensitive to the mid- to lower free troposphere, and $O_3$ at these altitudes in the Eastern Pacific is known to possibly impact the downwind US surface air quality at later times (Huang et al., 2010; Parrish et al., 2010). TES $O_3$ is generally positively biased by $<15\%$ relative to high accuracy/precision reference datasets (e.g., Verstraeten et al., 2013). Although IASI is in general less sensitive than TES due to its coarse spectral resolution, the 681–316 hPa partial column-averaged $O_3$ mixing ratios in the JPL product agree well with TES $O_3$ for the 2008–2011 period with a $-3.9$ ppbv offset (Oetjen et al., 2016). Note that IASI $O_3$ data are processed operationally in Europe using a different algorithm. For this work we used $O_3$ profiles from TES and IASI processed using a consistent algorithm at JPL, although the latter set of data represents only a small subset of the full set of the IASI radiance measurements. The IASI and TES L2 $O_3$ profiles (screened by the retrieval quality and the C-Curve flags) were used to evaluate the STEM $O_3$ vertical distributions in the different base simulations, and the satellite observation operators were applied in these comparisons. Taking TES as an example, its observation operator $h_c$ for $O_3$ is written in (4):

$$ h_c = z_c + A_{TES} \left( \ln(F_{TES}(c)) - z_c \right) \quad (4) $$

where $z_c$ is the natural log form of the TES constraint vector (a priori) in volume mixing ratio. $A_{TES}$ is the averaging kernel matrix reflecting the sensitivity of retrieval to changes in the true state (Rodgers, 2000). $F_{TES}$ projects the modeled $O_3$ concentration fields $c$ to the TES grid using spatial and temporal interpolation. The exponential of $h_c$ is then used to compute the mismatches between the model and TES $O_3$ retrievals as the model evaluation. A small mismatch between model with the satellite observation operators and the satellite retrievals may indicate either good model performance or may be the low sensitivity of the retrievals to the true $O_3$ profile. AIRS $O_3$ is sensitive to the altitudes near the tropopause, with positive biases over the ozonesondes in the upper troposphere (e.g., Bian et al., 2007); AIRS CO is most sensitive to 300–600 hPa (Warner et
al., 2007) and is frequently used together with the AIRS O\textsubscript{3} to distinguish the stratospheric O\textsubscript{3} intrusions from long-range transported anthropogenic or biomass burning pollution. We use the L3 AIRS products in this study to get a broad overview of the areas that are strongly impacted by the stratospheric O\textsubscript{3} intrusions or/and LRT of pollution.

The bottom-up NO\textsubscript{x} emissions from the HTAP inventory were assessed on a monthly base by comparing the GEOS-Chem NO\textsubscript{2} columns with the de-stripped KNMI (Royal Netherlands Meteorological Institute) OMI column NO\textsubscript{2} product version 2.0 (Boersma et al., 2011a, b). For this model evaluation against the OMI L2 products, the NO\textsubscript{2} fields calculated by the GEOS-Chem adjoint model were saved daily at 13:30 local solar time, roughly coinciding with the Aura and Aqua overpassing times. Other parameters used in the model column calculations came from the GEOS-5/GEOS-Chem monthly mean conditions. The OMI data that passed the tropospheric quality flag at 13-14 local time were selected based on the following screening criteria: surface albedo<0.3; cloud fraction<0.2; solar zenith angle <75°; and viewing zenith angle <45°. The averaging kernels (Eskes and Boersma, 2003) and Air Mass Factors (AMFs) in the KNMI product were used to calculate the modeled tropospheric NO\textsubscript{2} vertical columns comparable to the OMI’s. Details of the method to compare the model-based NO\textsubscript{2} columns with the KNMI OMI’s can be found in Huang et al. (2014).

3. Results and Discussions

3.1. Evaluation of model base simulations and the HTAP2 bottom-up NO\textsubscript{x} emissions

3.1.1. Evaluation of the global model O\textsubscript{3} ensembles and the bottom-up NO\textsubscript{x} emissions

The monthly-mean surface O\textsubscript{3} from multiple global models’ free runs was evaluated with the CASTNET observations, at the stations with 95% of the hourly O\textsubscript{3} observation completeness for the 1 May-30 June 2010 period, and the mean biases and RMSEs for these two months were summarized in Table 2 by US subregions. The three boundary condition-model as well as the eight-model ensembles overall underpredicted O\textsubscript{3} in the western US (by ~3-6 ppbv), similar to the HTAP1 model performance over these regions for May-June 2001. This can be due to the underestimated trans-boundary pollution (as indicated by the evaluation of modeled O\textsubscript{3} profiles with ozonesondes and satellite O\textsubscript{3} products). In addition, the coarser model resolutions are less capable of resolving the local features that influence the pollutants’ import processes, chemical transformation, as well as regional processes such as the cross-state pollution transport over complex terrains. The global RAQMS base simulation with satellite assimilation improved the free tropospheric O\textsubscript{3} structure as its comparisons with the ozonesondes shows, which also enhanced the simulated monthly-mean surface O\textsubscript{3} by up to over 10 ppbv in the western US and some coastal areas in the southeastern US (Figure S1, left).

The global models overall significantly overestimated O\textsubscript{3} in the other subregions (by 8-12 ppbv), close to HTAP1 model performance for May-June 2001 over the similar areas (Fiore et al., 2009) and in the Lapina et al. (2014) study for 2010, in large part due to the uncertainties in the bottom-up emissions which will be discussed further in the following paragraphs. Satellite assimilation led to 2-6 ppbv higher RAQMS surface O\textsubscript{3} in the central/southern/eastern US than in its free simulation, which are associated with higher positive biases. Except in the northeastern US, the eight-model ensembles show better agreement with the CASTNET O\textsubscript{3} observations than the three boundary condition-model ensemble, suggesting that using a larger number of models in the
ensemble calculations in this case may result in better overall model performance. The representation of land use/land cover, boundary layer mixing and chemistry in certain global model (i.e., GEOS-Chem) can be sources of uncertainty as reported in the literature (e.g., Geddes et al., 2016; Travis et al., 2016), but how serious these issues were in the other models need to be investigated further. Future work should emphasize on evaluating and comparing the models on process level to better understand the multi-model results.

The comparison of the GEOS-Chem adjoint NO$_2$ columns with the OMI product was used to help assess the bottom-up HTAP2 NO$_2$ emissions. Figure 3 shows that NO$_2$ columns from GEOS-Chem’s base simulations over the US are overall overestimated, and larger disagreement was found over the central/eastern US during June 2010. While grid-scale differences in NO$_2$ columns may not be directly indicative of emissions biases (Qu et al., 2016), overall there does appear to be a positive bias in the bottom-up emissions, mainly from the anthropogenic sources, consistent with the findings of Anderson et al. (2014) and Travis et al. (2016). The NO$_2$ columns in the GEOS-Chem base simulation were overestimated in many northern China rural areas and underpredicted in a few urban areas in the East Asia as well as a broad area in the southwestern China. The mismatches between model and OMI NO$_2$ fell within the ranges of the comparison between the GOME2 NO$_2$ column product and six models’ simulations over China in summer 2008 (Quennehen et al., 2016). It is likely that other O$_3$ precursors’ co-emitted with NO$_2$ from the same sources are estimated with similar levels of uncertainties. Also, the use of monthly-mean anthropogenic emissions as well as the overall rough treatment of emission height and temporal profiles can be sources of uncertainty. These global model evaluation results suggest that the EAS-NAM SR relationships analyzed using this inventory may overall overestimate the NAM local contribution and underestimate the EAS contribution—This statement would also rely on the quality of other O$_3$ precursors in the HTAP2 emission inventory, so careful assessment of other key O$_3$ precursors’ emissions in the inventory is also needed. Note that this comparison does not account for the biases in the used OMI data, and would be further validated by using other OMI NO$_2$ products as well as the bias-corrected (if applicable) in-situ NO$_2$ measurements. We recommend more global models to save their calculations more frequently, at least near the satellite overpassing times, for a more comprehensive assessment of the emission inventory.

3.1.2. Evaluation of the STEM regional base simulations w/ three sets of boundary conditions

The three STEM base simulations using different boundary conditions were averaged and evaluated with the hourly O$_3$ observations at the CASTNET sites in the four US subregions. The evaluation included the 8 May-30 June 2010 period to exclude the results during the one-week spin-up period. The evaluation statistics is summarized in Table 3. The time series of observed and modeled O$_3$ at the western US CASTNET sites are shown in Figure 2a where the model overall simulated the surface O$_3$ fairly well, with a much smaller mean bias (~1.6 ppbv) than the global model ensembles. However, this good performance can be a net effect of incorrect partitioning between the trans-boundary and local source contributions, with the former being underestimated and offsetting the overestimation of the latter. Switching the STEM chemical boundary conditions to the assimilated RAQMS base simulation led to increases in the simulated surface O$_3$ concentrations by >9 ppbv in the western US (Figure S1, right), associated with higher positive biases (due to several factors discussed in the next paragraph). Regional-scale assimilation could further reduce uncertainties introduced from regional meteorological and emission inputs to obtain
better modeled total $O_3$ and the partitioning of trans-boundary versus US contributions (e.g., Huang et al., 2015).

The model significantly overpredicted $O_3$ over the rest of the US mainly due to the overall overestimated NOx emissions. These positive biases are higher than the global model ensembles’, which can partially result from the possible unrealistic VOC speciation of the emission inventory and the SAPRC 99 chemical mechanism: Although SAPRC mechanisms have been used in air quality modeling for regulatory applications in some US states such as California, they usually produced higher $O_3$ than other mechanisms such as the CB05 over the US, and the comparisons between SAPRC 99 and SAPRC 2007 are still in progress (e.g., Luecken et al., 2008; Zhang et al., 2012; Cai et al., 2011). It is important to timely update the chemical mechanisms in the chemistry models, and we also suggest to timely upgrade the VOC speciation in the bottom-up emission inventories in the US to benefit the air quality modeling. Additionally, the uncertainty from non-anthropogenic emissions, such as the biogenic VOC emissions from WRF/MEGAN which is known to often have positive biases, can be another cause: As Hogrefe et al. (2011) presented, the MEGAN emissions resulted in a higher $O_3$ response to hypothetical anthropogenic NOx emission reductions compared with another set of biogenic emission input. Some factors that caused the overpredicted MEGAN emissions, such as positively-biased temperature in WRF, can also be important sources of uncertainty in the STEM modeled $O_3$. We anticipate that the results from the Air Quality Model Evaluation International Initiative (AQMEII) experiment (e.g., Schere et al., 2012; Solazzo et al., 2012; Galmarini et al., 2015, 2016), which involves other regional model simulations over the US with the similar set of boundary conditions but different chemical mechanisms and non-anthropogenic emission inputs, can help better understand the causes of errors in the simulated total $O_3$.

3.2. The NAM surface $O_3$ sensitivity to extra-regional anthropogenic pollutants

3.2.1. Global model ensembles

The impact of all foreign (i.e. non-NAM) anthropogenic sources on NAM surface $O_3$ was first explored, including the spatial distributions of the RERER metric (eq. (2)) based on various global models’ simulations (Figure 4), and the domain wide mean sensitivities $R(O_3, \text{non-NAM}, 20\%)$ (eq. (1d)) (Figure 5). Across the NAM, the strongest impacts were found in spring time (March-April-May, larger than 1.5 ppbv in average over the domain) and the weakest impacts are shown during the summertime (June-July-August, 1.0-1.3 ppbv), consistent with the existing knowledge on the seasonal variability of the non-local pollution impacts on NAM for other years (e.g., Fiore et al., 2009; Reidmiller et al., 2009; Brown-Steiner and Hess, 2011). All global models indicate strong non-NAM anthropogenic source impacts on the western US mainly due to the impact of its high elevation, and also near the US-Mexico border areas, especially southern Texas, due to their vicinity to the Mexican emission sources. Over the western states, stronger non-local impacts were reflected from the results based on higher-resolution global models (e.g., the >0.6 RERER values from the half degree EMEP model, corresponding to its higher $R(O_3, \text{non-NAM}, 20\%)$ values than the other models’), similar to the findings in previous modeling studies (Lin et al., 2010, 2012a). Larger-than-1 RERER values are often seen near the urban areas and large point sources due to the titration, especially evident from the higher resolution model results. The $R(O_3, \text{EAS}, 20\%)$ values are larger than 1/3 of the $R(O_3, \text{non-NAM}, 20\%)$ (0.2-0.5 ppbv from April to June), more than 3-4 times higher than $R(O_3, \text{EUR}, 20\%)$ and $R(O_3, \text{SAS}, 20\%)$. Note that all eight
models contributed to the R(O₃, EAS, 20%) calculations, but one or two models did not provide all necessary sensitivity runs to compute the RERER, R(O₃, non-NAM, 20%), R(O₃, EUR, 20%), or R(O₃, SAS, 20%).

Comparing to the HTAP1 modeling results, the magnitudes of R(O₃, EUR, 20%) are smaller by a factor of 2-3, as a result of the substantial improvement in the European air quality over the past decades (Crippa et al., 2016; Poulid et al., 2015), and also possibly due to the changes in the HTAP2 experiment setup from HTAP1 (e.g., EUR by HTAP1’s definition includes regions in Russia/Belarussia/Ukraine, Middle East and North Africa that are excluded from the HTAP2’s EUR domain). In contrast, the R(O₃, non-NAM, 20%) and R(O₃, EAS, 20%) values are >50% higher than the HTAP1 modeling results. The R(O₃, EAS, 20%) based on the emission perturbation approach are also larger than 1/5 of the original estimates (i.e., 1.94 ppbv and 0.79 ppbv for March-April-May and June-July-August, respectively) for 2001-2005 using the tagged tracers (Brown-Steiner and Hess, 2011) a method that produces a higher “contribution” than the emission perturbation approach (e.g., Grewe et al., 2012; Emmons et al., 2012; Brown-Steiner and Hess, 2011). These results reflect the impact of the growing anthropogenic emissions from the East Asia and other developing countries during 2001-2010. The SNU GEOS-Chem-based R(O₃, EAS, 20%) and R(O₃, non-NAM, 20%) in Figure S2 both show that the NAM was a little more strongly affected by foreign anthropogenic pollution in 2010 than in 2008-2009, especially in April-May. This can be in part due to the higher O₃ precursors’ emissions in 2010 from extra-regions including the East Asia (Table S1). Such interannual variability can also be due to the Spring 2010 meteorological conditions that favored the trans-Pacific pollution transport, as introduced in Section 2.2.1.

Similar to the findings from the HTAP1 studies, the large intermodel variability (as indicated in Table 4) in the estimates of intercontinental SR relationships indicates the uncertainties of these models in representing the key atmospheric processes which needs more investigations in the future. Figure 5b compares the R (O₃, EAS, 20%) estimated by individual boundary condition models, their ensemble mean sensitivities, and the eight-global model mean. The averaged R(O₃, EAS, 20%) from the boundary condition model results are smaller than the eight-global model mean, and GEOS-Chem gives higher R (O₃, EAS, 20%) than RAQMS and C-IFS except for July-October 2010. The R (O₃, EAS, 20%) and the intermodel differences are overall much smaller than the modeled total O₃ (<<5%) and their biases in NAM, as the impact of the EAS anthropogenic sources has been diluted following transport over the great distances. Other factors contribute more significantly to the biases in the modeled total O₃, such as the stratospheric O₃ intrusions and the local O₃ formation, and controlling the local and regional emissions would still be more effective for complying with the tighter air quality standard.

The O₃ sensitivities in response to the perturbations of individual species or sector emissions in East Asia, estimated by the GEOS-Chem adjoint model, were also analyzed (Figure S2). The EAS anthropogenic NOₓ emissions more strongly impacted the NAM surface O₃ than the other major O₃ precursors, similar to the findings in Fiore et al. (2009) and Reidmiller et al. (2009) using the perturbation approach, as well as the conclusions in Lapina et al. (2014) based on the adjoint sensitivity analyses. Emissions from the power&industrial sectors are higher in East Asia than the other sectors (Table S1), resulting in its stronger influences on the NAM surface O₃. As the observed NOₓ columns started to drop since 2010 due to the effective denitrification devices
implemented at the Chinese power and industrial plants (e.g., Liu et al., 2016), depending on the changes in the VOC emissions, it is anticipated to see different R(O$_3$, EAS, 20%) values for the years after 2010. Therefore, continued studies to assess the East Asian anthropogenic pollution impacts on NAM during more recent years is needed. As emissions from various source sectors can differ by their emitted altitudes and temporal (from diurnal to seasonal) profiles. Efforts should also be placed to have the models timely update the heights and temporal profiles of the emissions from those various sectors.

3.2.2. Regional model sensitivities and their connections with the boundary condition models

The monthly-mean STEM surface R(O$_3$, EAS, 20%) sensitivities based on different boundary condition models were inter-compared, and also compared with the R(O$_3$, EAS, 20%) estimated by their boundary condition models as well as the global model ensemble mean (Figure 6). For both May and June 2010, the domain-wide mean R(O$_3$, EAS, 20%) values from STEM/RAQMS were higher than the estimates from RAQMS; the STEM/GEOS-Chem R(O$_3$, EAS, 20%) values are lower than those of GEOS-Chem, and the STEM/CIFS R(O$_3$, EAS, 20%) is higher than C-IFS’s in June but slightly lower in May. These differences are overall smaller than the inter-global model differences, and can be due to various factors including the uncertainties in boundary condition chemical species mapping, and the different meteorological/terrain fields/chemistry in the global and regional models. The STEM R(O$_3$, EAS, 20%) ensemble mean values, however, are less than 0.02 ppbv different from its boundary condition model’s ensemble mean for both months. The STEM R(O$_3$, EAS, 20%) ensemble mean value in June is also close to the eight-global model ensemble mean, but is ~0.05 ppbv lower than the eight-model mean in May.

Choosing other/more global model outputs as STEM’s boundary conditions may lead to different STEM ensemble mean R(O$_3$, EAS, 20%) estimates. We also found that the period mean R(O$_3$, EAS, 20%) of ~0.2 ppbv sampled only at the CASTNET sites (Table 3) are smaller than those averaged in all model grids. This indicates that currently the sparsely distributed surface network (especially over the western US that is more strongly affected by the extra-regional sources than the other US regions) may miss many LRT episodes that impact the NAM. The planned geostationary satellites with ~2-5 km footprint sizes and hourly sampling frequency (Hilsenrath and Chance, 2013) will help better capture the LRT episodes in these regions.

The spatial patterns of the monthly-mean STEM surface R(O$_3$, EAS, 20%) sensitivities based on the three boundary condition models are notably different, but overall resemble what’s estimated by the corresponding boundary condition model, and the STEM sensitivities show more local details in certain high elevation regions in the US west (Figure 7 shows the June 2010 conditions as an example). These different sensitivities were investigated further, by examining the R(O$_3$, EAS, 20%) values near the source regions (i.e., East Asia) as well as near the receptor regions (Figure 8). More East Asian anthropogenic O$_3$ seems to be transported at the upper troposphere in RAQMS than in the other two models. GEOS-Chem and RAQMS R(O$_3$, EAS, 20%) sensitivities are similar over the EAS as well as the 500-900 hPa near the receptor in the eastern Pacific (at ~135°W), the altitudes US surface O$_3$ are most strongly sensitive to during the summertime as concluded from previous studies (e.g., Huang et al., 2010, 2013a; Parrish et al., 2010). Despite the close NAM domain-wide mean values from the STEM/GEOS-Chem and STEM/RAQMS, the spatial patterns of R(O$_3$, EAS, 20%) over NAM differ in these two cases, with the latter case showing sharper gradients especially in the western US, partially due to the
impact of its higher horizontal resolution. The \( \text{R}(\text{O}_3, \text{EAS, 20\%}) \) values from STEM/CIFS are lower than from the other two cases both near the sources and at (near) NAM. The STEM surface (also near surface, not shown in figures) \( \text{R}(\text{O}_3, \text{EAS, 20\%}) \) does not spatially correlate well with the column \( \text{R}(\text{O}_3, \text{EAS, 20\%}) \), the latter of which contributed more to the base case \( \text{O}_3 \) columns, indicating that a good portion of the transported East Asian pollution did not descend to the lower altitudes to impact the boundary layer/ground level air quality. An additional regional simulation was performed in which the STEM boundary conditions were downscaled from a RAQMS simulation without the East Asian anthropogenic emissions. The non-linear emission perturbation-\( \text{O}_3 \) response relationships, as the larger-than-1 \( \text{S}_{\text{O}_3} \) metric (eq. (3)) indicate, are seen across the domain, for both the surface and column \( \text{O}_3 \) (Figure 7). Therefore, the full source contribution obtained by linearly scaling the receptor regional mean \( \text{O}_3 \) sensitivity to the 20% reduction in the source region emissions may be underestimated.

The temporal variability of the STEM \( \text{R}(\text{O}_3, \text{EAS, 20\%}) \) ensemble sensitivities were also studied. For most US subregions, 3-6 LRT episodes (defined as when the sensitivities are above the period mean) were identified during May-June. Throughout this period, the hourly \( \text{R}(\text{O}_3, \text{EAS, 20\%}) \) and the observed \( \text{O}_3 \) at the surface CASTNET sites are weakly correlated (Table 3), but they display similar diurnal cycles (e.g., Figure 2a for the western US sites), possibly because the deeper boundary layer depth during the daytime enhanced entrainment down-mixing of the extra-regional pollutants to the surface. The identified diurnal variability of the \( \text{R}(\text{O}_3, \text{EAS, 20\%}) \) can cause differences in the calculated MDA8 and all-hour mean \( \text{R}(\text{O}_3, \text{EAS, 20\%}) \) values. Figure S3 shows that the mean \( \text{R}(\text{MDA8, EAS, 20\%}) \) values, usually at daytimes, are higher than the all-hour averaged \( \text{R}(\text{O}_3, \text{EAS, 20\%}) \) in most STEM model grids during both months. Therefore, it is important for more HTAP2 participating models to save their outputs hourly in order to conveniently compute the policy-relevant metrics for the \( \text{O}_3 \) sensitivities. Also, the hourly sampling frequency of the planned geostationary satellites is anticipated to be more helpful for evaluating the impacts of the LRT episodes.

The STEM \( \text{R}(\text{MDA8, EAS, 20\%}) \) for May-June 2010 in four US subregions were averaged on all days and only on the days when the simulated total MDA8 \( \text{O}_3 \) is over 70 ppbv (Figure 9). These sensitivities also show appreciable spatial variability: from 0.35-0.58 ppbv in the western US (also with the largest standard deviations, not shown), which is slightly higher than the HTAP1 results reported by Reidmiller et al. (2009) for Spring 2001, to \( \sim \)0.1-0.25 ppbv in the rest three subregions, which is close to the Reidmiller et al. (2009) results. Qualitatively consistent with the findings in Reidmiller et al. (2009), \( \text{R}(\text{MDA8, EAS, 20\%}) \) is smaller during the high \( \text{O}_3 \) total days in all subregions. Note that the STEM base simulations overall substantially overpredicted the total \( \text{O}_3 \) in non-western US regions, so the \( \text{R}(\text{MDA8, EAS, 20\%}) \) calculated during the days of \( \text{O}_3 \) exceedances can actually represent the sensitivities during the non-exceedances.

### 3.3. Case study of the 9 May 2010 LRT event mixed with stratospheric \( \text{O}_3 \) intrusions

Lin et al. (2012a, b) and Neuman et al. (2012) showed that the trans-Pacific pollution transport intensely impacted the western US during 8-10 May, 2010, intermingled with a stratospheric intrusion. This episode is indeed indicated by the \( \text{O}_3 \) and CO products from AIRS and TES at \( \sim \)500 hPa over the E Pacific (Figure 10), and the observed TES and IASI \( \text{O}_3 \) profiles over the western US indicated elevated \( \text{O}_3 \) levels (>80 ppbv) at 700-900 hPa. Huang et al. (2013b)
found that the meteorological conditions during this period (i.e., a strong jet at ~700 hPa with wind speed >20 m/s) shifted southwesterly when passing the southern California and continued to travel towards the mountain states), along with the orographic lifting, efficiently exported the southern California anthropogenic pollution, which was chemically coupled with the extra-regional pollution and significantly enhanced the O₃ levels in the US intermountain west.

We selected this episode to compare the STEM surface total O₃ concentrations as well as the R(O₃, EAS, 20%) sensitivities based on the different HTAP boundary condition models. Figure 11 evaluates the simulated O₃ profiles in the western US from several STEM base simulations against the TES and IASI O₃ retrievals, and Figures 12a-d indicate the performance of the daily surface total MDA8 O₃ from these simulations. We found that the underestimated free tropospheric O₃ from the STEM simulations that used any single free-running chemical boundary conditions contributed to the underestimated STEM surface O₃ in the high elevation mountain states. The STEM base simulation using the RAQMS assimilated fields as the boundary conditions, agrees most with the observed O₃ at the CASTNET sites, as well as the TES and IASI O₃ profiles in the western states. Similar to the conclusions drawn in Huang et al. (2010, 2015) for summer 2008, we again demonstrated the robustness of satellite chemical data assimilation for improving the boundary condition models’ O₃ performance. As the enhancement of O₃ due to the assimilation is much larger than the O₃ sensitivities to the EAS anthropogenic emissions, the assimilation mainly improved the contributions from other sources, such as the stratospheric O₃.

The quality of the model boundary conditions only indicates how well the total “transported background” component is represented, and can not be directly connected with the accuracy of the model estimated R(O₃, EAS, 20%) sensitivities, which also show notable intermodel differences: The estimated R(MDA8, EAS, 20%) in the different STEM cases range from <1.0 ppbv to ~1.3 ppbv, at least 40% higher than the May-June period mean in Figure 9. Strong nonlinear emission perturbation-O₃ response relationships are also shown during this period (Figure 12, lower). The R(MDA8, EAS, 100%) from the STEM/RAQMS case is as high as >7 ppbv over the high terrain regions. These are of smaller magnitudes than the estimates in Lin et al. (2012a), possibly due to the differences in the used models and the bottom-up emission inputs.

A stratospheric O₃ intrusion also affected the NE US on the same day, as revealed by the satellite free tropospheric O₃ and CO observations (Figure 10). This intrusion was mixed with LRT East Asian pollution (Figure 12 and Figure S4). However, this intrusion did not enhance the NE boundary layer/surface O₃ concentrations, which were actually anomalously low (MDA8<40 ppbv) as indicated by the model base simulations and the CASTNET observations (Figure 12a-d). Similar characteristics during summertime stratospheric O₃ intrusion events around this region have been discussed by Ott et al. (2016). The East Asian pollution less intensely (<50%) affected the surface O₃ levels in these regions than in the US west, due to the greater transport distances as well as the impact of the overall flat terrain in the US east.

4. Conclusions and suggestions on future directions

In support of the HTAP Phase 2 experiment that involved high-resolution global models and regional models’ participation to advance the understanding of the pollutants’ SR relationships in the Northern Hemisphere, we conducted a number of regional scale STEM base and forward...
sensitivity simulations over North America during May-June 2010. The STEM top and lateral
chemical boundary conditions were downscaled from three global models’ (i.e., GEOS-Chem,
RAQMS, and ECMWF CIFS) base and sensitivity simulations (in which the East Asian
anthropogenic emissions were reduced by 20%). The STEM surface O₃ sensitivities (including the
24h mean and the policy-relevant MDA8 metric averaged throughout the study period and during
a selected transport event) over North America overall resembled those from the corresponding
boundary condition model, but can be quantitatively different from the mean sensitivities estimated
by all global model ensembles. Therefore, choosing other/more global model outputs as STEM’s
boundary conditions may lead to different STEM ensemble mean O₃ sensitivities. Overall, the
monthly-based US O₃ sensitivities to the 20% reduction of the East Asian anthropogenic emissions
contributed to <<5% of the total O₃ and are of smaller magnitudes than the biases in the modeled
total O₃. Better quantifying the contributions from other factors, such as the stratospheric O₃
intrusion and the local O₃ formation, would still be the most effective way to help reduce the North
American pollution levels and the model uncertainties. The US O₃ sensitivities to the East Asian
anthropogenic emissions were episodically strong, contributing to the O₃ exceedances in some high
terrain areas. Assessing the sources of intermodel differences are particularly important for better
evaluating the East Asian pollution impacts during these episodes. The STEM O₃ sensitivities
followed similar diurnal cycles as the total O₃, emphasizing the importance of saving model results
hourly for continently calculate policy-relevant metrics, as well as the usefulness of hourly
sampling frequency of the planned geostationary satellites for better evaluating the impacts of the
LRT events.

The ensemble mean O₃ sensitivities in 2010 were higher than the HTAP1 reported 2001
conditions, as well as 1/5 of the original estimates for 2001-2005 using the tagged tracers. This
indicates the increasing impacts of the East Asian anthropogenic pollution on North America. The
GEOS-Chem O₃ sensitivities in 2010 were also higher than the 2008-2009 conditions due to the
increasing Asian emissions and the Spring 2010 meteorological conditions that favored the trans-
Pacific pollution transport. The GEOS-Chem sensitivity calculations indicate that the East Asian
anthropogenic NOₓ emissions mattered more than the other East Asian O₃ precursors to the North
American O₃, qualitatively consistent with previous adjoint sensitivity calculations. Continued
research is needed on temporal changes of emissions for different species and sectors in North
America and foreign countries as well as their impacts on the SR relationships. As emissions from
various source sectors can differ by emitted altitudes and temporal profiles. Efforts should also be
placed to have the models timely update the height and temporal profiles of the emissions from
various sectors.

An additional STEM simulation was performed in which the boundary conditions were
downscaled from a RAQMS simulation without East Asian anthropogenic emissions (i.e., a 100%
emission reduction), to assess the scalability of the mean O₃ sensitivities to the size of the emission
perturbation. The scalability was found to be spatially varying, and the full source contribution
obtained by linearly scaling the NAM regional mean O₃ sensitivity to the 20% reduction in the
East Asian emissions may be underestimated. Motivated by Lapina et al. (2014), additional
calculations will be conducted in future to explore the scalability of different O₃ metrics in these
cases. For future source attribution analysis, in general it is recommended to directly choose the
suitable size of the emission perturbation based on the specific questions to address, and to avoid
linearly scaling O₃ sensitivities that are based on other amounts of the perturbations.
Satellite NO\textsubscript{2} (KNMI OMI) and O\textsubscript{3} (TES, JPL-IASI, OMI, MLS, and AIRS) products helped detect pollution episodes, quantify or/and reduce the uncertainties in the bottom-up NO\textsubscript{x} emissions and the model transported background O\textsubscript{3}. Based on model calculations and satellite/surface observations on a selected day of 9 May 2010, we showed the different influences from stratospheric O\textsubscript{3} intrusions along with the transported East Asian pollution on O\textsubscript{3} in the western and the eastern US. Continued studies on exceptional events during other seasons are in progress. As chemical data assimilation techniques keep developing (Bocquet et al., 2015), several HTAP2 participating global models have already been able to assimilate single- or multi-constitute satellite atmospheric composition data (e.g., Miyazaki et al., 2012; Parrington et al., 2008, 2009; Huang et al., 2015; Inness et al., 2015; Flemming et al., 2016). Comparing the performance of the assimilated fields from different models, and making the global model assimilated chemical fields in the suitable format for being used as boundary conditions would be very beneficial for future regional modeling, as well as for better interpreting the pollutants’ distributions especially during the exceptional events. Meanwhile, efforts should also be devoted to advancing and applying higher-resolution regional scale modeling and chemical data assimilation. Furthermore, although satellite observations have been applied for improving the estimated US background O\textsubscript{3} (e.g., Huang et al., 2015), using satellite (and/or other types of) observations to improve SR relationship studies also needs to be explored. Some of the possible methods include: 1) The combination of data assimilation and the tagging approach; 2) Introducing observation-constrained emission estimates in the emission perturbation analyses.

**Acknowledgements**

The global and regional modeling results used in this study have been submitted to the AeroCom database following the HTAP2 data submission guidelines (http://iek8wikis.iek.fz-juelich.de/HTAPWiki/HTAP-2-data-submission), or can be made available upon request. Technical support from Anna Carlin Benedictow, Brigitte Koffi, Jan Griesfeller, and Michael Schulz regarding formatting and submitting the data to the AeroCom is acknowledged. MH thanks the research resources at the University of Iowa and JPL/Caltech that supported this study, as well as the travel funding from the US EPA for attending the related HTAP2 workshops. DKH and YD recognize support from NASA AQAST. Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract to the National Aeronautics and Space Administration. Reference herein to any specific commercial product, process or service by trade name, trademark, manufacturer or otherwise does not constitute or imply its endorsement by the United States Government or the Jet Propulsion Laboratory, California Institute of Technology. The views, opinions, and findings contained in this report are those of the author(s) and should not be construed as an official National Oceanic and Atmospheric Administration or U.S. Government position, policy, or decision.
References


Cooper, O., et al. (2016), Western NA Performance Evaluation for HTAP2, HTAP2 workshop, Potsdam, Germany, 2016.


Figure 1. Definitions of the 16 source regions used in HTAP2 SR relationship study. The map is plotted based on data on a 0.1°×0.1° resolution grid. We focus in this study on the impact of anthropogenic pollution from selected non-North American source regions (i.e., EAS, SAS, and EUR), as highlighted in italic.

Figure 2. (a) Evaluation of the STEM modeled (averaged from the three base simulations using the GEOS-Chem, ECMWF C-IFS, and RAQMS base runs as the chemical boundary conditions) hourly O₃ at the western US (i.e., EPA regions 8, 9, and 10) CASTNET sites. Observations, modeled base O₃ and the modeled R(O₃, EAS, 20%) are in orange, black, and purple lines, respectively. The horizontal dashed lines indicate the period mean values. (b) The 60 km STEM NAM domain, colored by the model topography. The CASTNET sites used in the STEM base O₃ evaluation are marked as triangles in different colors that identify the subregions they belong to (red: western US; grey: southern US; purple: Midwest; blue: northeastern US).

Figure 3. Evaluation of the GEOS-Chem adjoint base NO₂ product (recorded at near the satellite overpassing time) with the OMI NO₂ columns. The differences between OMI and GEOS-Chem (OMI-modeled) tropospheric NO₂ columns (×10⁵ molec./cm²) are shown for (a) May and (b) June 2010. Details of the comparison are included in Section 2.3.2.
Figure 4. The RERER maps in May (left) and June (right) 2010 over the continental US, calculated based on the monthly mean $\text{O}_3$ from multiple global models’ base and emission sensitivity simulations. The RERER metric (unitless) was defined in eq. (2) in the text. Values larger than 1 and smaller than 0 are shown in purple and grey, respectively. The US (including continental US as well as Hawaii which is not shown in the plots) mean values are indicated for each panel at the lower right corner. The 7-model mean RERER values for May and June 2010 are ~0.5 and ~0.4, respectively.
Figure 5. (a) North American (130-65°W; 20-50°N) mean O₃ sensitivity to 20% anthropogenic emission reductions in various non-North American regions, averaged from multiple (six-eight, see details in text) global models. (b) North American surface R(O₃, EAS, 20%) values, as estimated by single (the three STEM boundary condition models) or multi-global model means. The “Multi-model” and “Three-model” in the legend indicate the mean sensitivities of all eight global models and only of the three boundary condition models, respectively.

Figure 6. Monthly-mean North American (130-65°W; 20-50°N) surface R(O₃, EAS, 20%) values from multiple global and regional model simulations for May (left) and June (right) 2010. STEM model mean values were calculated from its hourly output from 8 May and on. The “Multi-model” and “Three-model” in the legend indicate the mean sensitivities of all eight global models and only of the three boundary condition models, respectively.
Figure 7. The monthly-mean $R(O_3, \text{EAS}, 20\%)$ in June 2010 for: (a) surface $O_3$ (ppbv) from the three boundary condition models, (b) STEM surface $O_3$ (ppbv), and (c) STEM column $O_3$ ($\times 10^{16}$ molecules/cm$^2$). Columns 1-3 show $R(O_3, \text{EAS}, 20\%)$ from the simulations associated with GEOS-Chem, ECMWF C-IFS, and RAQMS, respectively. Column 4 shows 1/5 of the $R(O_3, \text{EAS}, 100\%)$ from the simulations related to RAQMS. (d) The STEM/RAQMS-based “Scalability” $S_{O_3}$ (eq. (3)) of the NAM surface (left) and column (right) in June 2010.

Figure 8. The monthly-mean $R(O_3, \text{EAS}, 20\%)$ in ppbv in June 2010 from the three boundary condition models at the source and near the receptor regions: (a-c) surface $O_3$ in the East Asia; and (d) $O_3$ (GEOS-Chem) or (e-f) $O_3$ (ECMWF C-IFS and RAQMS) along the cross section of 135°W (near the west boundary of the STEM model domain as defined in Figure 2b).
Figure 9. STEM R(MDA8, EAS, 20%) for May-June 2010 in four US subregions (defined in the inset of panel, also consistent with the definitions in Figures 2/S3 and Tables 2-3), averaged on all days (bars with solid fill) and only on the days when the simulated total MDA8 O$_3$ concentrations were over 70 ppbv (bars with grid pattern fill). The results from the STEM runs using GEOS-Chem, ECMWF C-IFS and RAQMS boundary conditions are shown separately.

Figure 10. Case study of 9 May 2010: (a-b) Ozone (ppbv) and (c-d) CO (ppbv) at ~500 hPa from the L2 (a,c) TES retrievals (circles) and (b,d) L3 AIRS products at early afternoon local time. The L2 IASI O$_3$ (ppbv) at ~500 hPa retrieved using the TES algorithm (details in Section 2.3.2) at the mid-morning local times is shown on panel (b) as triangles. The O$_3$ profiles within the purple box in panel (a) were used in the model evaluation shown in Figure 11.
Figure 11. Case study of 9 May 2010: The comparisons between (a) IASI and (b) TES O\textsubscript{3} in the western US with the simulated O\textsubscript{3} in the STEM runs using the GEOS-Chem (green), C-IFS (blue), RAQMS (purple), and assimilated RAQMS (red) boundary conditions. The O\textsubscript{3} profiles within the purple box in Figure 10a were used in the evaluation. Observation operators were applied in the comparisons (details in Section 2.3.2). Solid and open dots are TES/IASI data at the TES retrieval reporting levels and at the variable surface pressure levels, respectively. Solid lines are median O\textsubscript{3} profiles from the satellite observations and the different STEM simulations, calculated only on the TES retrieval reporting levels.

Figure 12. Case study of 9 May 2010: (a-d) Surface MDA8 total O\textsubscript{3} and (e-h) surface R(MDA8, EAS, 20\%) from the STEM simulations using the (a,e) GEOS-Chem, (b,f) ECMWF C-IFS, and (c,g) RAQMS free run as the boundary conditions. (d) Surface MDA8 total O\textsubscript{3} in a STEM base simulation using the RAQMS assimilation run as the boundary conditions. CASTNET observations are overlaid in filled circles in panels (a-d). (h) 1/5 of the surface R(MDA8, EAS, 100\%) from STEM/RAQMS simulations. The conditions at ~400-500 hPa are shown in Figure S4.
Table 1a. HTAP2 base and sensitivity simulations by various global models. Relevant references for the RAQMS model and the SNU GEOS-Chem are Pierce et al. (2007, 2009) and Park et al. (2004) (with additional descriptions on its HTAP2 simulation configurations at: http://iek8wikis.iek.fz-juelich.de/HTAPWiki/WP2.3?action=AttachFile&do=view&target=_README_GEOS-Chem.pdf), respectively, and the descriptions of the rest six global models can be found in published HTAP2 works such as in Stjern et al. (2016). The STEM boundary condition models are highlighted in bold.

<table>
<thead>
<tr>
<th>Global model and horizontal resolution</th>
<th>BASE</th>
<th>EASA LL-20%</th>
<th>EASA LL-100%</th>
<th>GLOA LL-20%</th>
<th>NAMA LL-20%</th>
<th>EURA LL-20%</th>
<th>SASA LL-20%</th>
</tr>
</thead>
<tbody>
<tr>
<td>CAM-Chem, 2.5°×1.9°</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>CHASER T42, ~2.8°×2.8°</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>EMEP rv48, 0.5°×0.5°</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>SNU GEOS-Chem v9-01-03, 2°×2.5°</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>CU-Boulder GEOS-Chem adjoint v35f, 2°×2.5°</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>RAQMS, 1°×1°, free running</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>RAQMS, 1°×1°, w/ satellite assimilation</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>Oslo, ~2.8°×2.8°</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
<tr>
<td>ECMWF CIFS, ~0.7°×0.7°/1.125°×1.125° (used as the STEM chemical boundary conditions)</td>
<td>✓</td>
<td>✓</td>
<td></td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
<td>✓</td>
</tr>
</tbody>
</table>

Table 1b. STEM regional simulations for HTAP2

<table>
<thead>
<tr>
<th>Boundary condition model</th>
<th>BASE</th>
<th>EASALL -20%</th>
<th>EASALL -100%</th>
</tr>
</thead>
<tbody>
<tr>
<td>SNU GEOS-Chem v9-01-03, 2°×2.5°</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>RAQMS, 1°×1°, free running</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>RAQMS, 1°×1°, w/ satellite assimilation</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
<tr>
<td>ECMWF CIFS, 1.125°×1.125°</td>
<td>✓</td>
<td>✓</td>
<td></td>
</tr>
</tbody>
</table>
Table 2. Evaluation of the period mean (1 May-30 June, 2010) multi-global model free simulations against the CASTNET observations, only at the sites where 95% of the hourly O$_3$ observations are available.

<table>
<thead>
<tr>
<th>Subregion</th>
<th>US EPA regions contained</th>
<th>Number of sites</th>
<th>Mean bias (ppbv)</th>
<th>RMSE (ppbv)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>3 BC$^a$ models</td>
<td>8 global models</td>
</tr>
<tr>
<td>Western US</td>
<td>8, 9, 10</td>
<td>19</td>
<td>-5.68</td>
<td>-2.52</td>
</tr>
<tr>
<td>Southern US</td>
<td>4, 6</td>
<td>18</td>
<td>11.61</td>
<td>10.24</td>
</tr>
<tr>
<td>Midwest</td>
<td>5, 7</td>
<td>13</td>
<td>8.03</td>
<td>7.66</td>
</tr>
<tr>
<td>Northeast</td>
<td>1, 2, 3</td>
<td>17</td>
<td>9.55</td>
<td>10.63</td>
</tr>
<tr>
<td>All</td>
<td>1-10</td>
<td>67</td>
<td>5.49</td>
<td>6.22</td>
</tr>
</tbody>
</table>

$^a$BC: Boundary Conditions

Table 3. Evaluation of the hourly STEM simulated total O$_3$ (averaged from the three base simulations that used the different free-running boundary conditions) against the CASTNET surface observations for 8 May-30 June, 2010. The subregional mean R(O$_3$, EAS, 100%), as well as its correlation coefficient with the observed O$_3$ are also shown.

<table>
<thead>
<tr>
<th>Subregion</th>
<th>US EPA regions contained</th>
<th>Number of sites</th>
<th>Mean elevation (km): actual/model</th>
<th>Mean bias (ppbv)</th>
<th>RMSE (ppbv)</th>
<th>Correlation (model base; obs)</th>
<th>Correlation (obs; modeled EAS)</th>
<th>Mean EAS sensitivity (ppbv)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Western US</td>
<td>8, 9, 10</td>
<td>22</td>
<td>1.75/1.71</td>
<td>1.60</td>
<td>4.86</td>
<td>0.76</td>
<td>0.34</td>
<td>0.48</td>
</tr>
<tr>
<td>Southern US</td>
<td>4, 6</td>
<td>22</td>
<td>0.38/0.31</td>
<td>20.33</td>
<td>22.13</td>
<td>0.58</td>
<td>0.27</td>
<td>0.15</td>
</tr>
<tr>
<td>Midwest</td>
<td>5, 7</td>
<td>16</td>
<td>0.29/0.28</td>
<td>15.64</td>
<td>17.97</td>
<td>0.70</td>
<td>0.15</td>
<td>0.17</td>
</tr>
<tr>
<td>Northeast</td>
<td>1, 2, 3</td>
<td>20</td>
<td>0.36/0.26</td>
<td>20.94</td>
<td>24.16</td>
<td>0.47</td>
<td>0.17</td>
<td>0.21</td>
</tr>
<tr>
<td>All</td>
<td>1-10</td>
<td>80</td>
<td>0.73/0.68</td>
<td>16.17</td>
<td>18.30</td>
<td>0.66</td>
<td>0.13</td>
<td>0.20</td>
</tr>
</tbody>
</table>
Table 4. The ranges and standard deviations (ppbv, separated by “;”) of R(O\textsubscript{3}, source region, 20%) by 6-8 global models (defined in eq. (1a-d)), summarized by months in 2010. The monthly multi-model mean values are shown in Figures 5-6.

<table>
<thead>
<tr>
<th>Month/Source region</th>
<th>All Foreign/Non-NAM (ppbv)</th>
<th>EUR (ppbv)</th>
<th>EAS (ppbv)</th>
<th>SAS (ppbv)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Jan</td>
<td>0.38-1.69; 0.41</td>
<td>0.002-0.12; 0.05</td>
<td>0.02-0.72; 0.24</td>
<td>0.001-0.11; 0.04</td>
</tr>
<tr>
<td>Feb</td>
<td>0.92-2.07; 0.37</td>
<td>0.02-0.15; 0.05</td>
<td>0.16-0.91; 0.28</td>
<td>0.02-0.12; 0.04</td>
</tr>
<tr>
<td>Mar</td>
<td>1.30-2.37; 0.38</td>
<td>0.07-0.21; 0.06</td>
<td>0.24-1.03; 0.30</td>
<td>0.03-0.12; 0.03</td>
</tr>
<tr>
<td>Apr</td>
<td>1.42-2.46; 0.33</td>
<td>0.09-0.23; 0.05</td>
<td>0.33-1.07; 0.28</td>
<td>0.04-0.12; 0.03</td>
</tr>
<tr>
<td>May</td>
<td>1.24-1.91; 0.21</td>
<td>0.06-0.17; 0.04</td>
<td>0.24-0.75; 0.19</td>
<td>0.05-0.11; 0.02</td>
</tr>
<tr>
<td>Jun</td>
<td>1.03-1.41; 0.13</td>
<td>0.03-0.07; 0.02</td>
<td>0.14-0.39; 0.09</td>
<td>0.04-0.07; 0.01</td>
</tr>
<tr>
<td>Jul</td>
<td>0.86-1.18; 0.13</td>
<td>0.02-0.04; 0.01</td>
<td>0.08-0.22; 0.06</td>
<td>0.01-0.04; 0.01</td>
</tr>
<tr>
<td>Aug</td>
<td>0.80-1.19; 0.13</td>
<td>0.01-0.04; 0.01</td>
<td>0.07-0.20; 0.05</td>
<td>0.02-0.04; 0.01</td>
</tr>
<tr>
<td>Sep</td>
<td>0.85-1.18; 0.13</td>
<td>0.03-0.05; 0.01</td>
<td>0.10-0.25; 0.06</td>
<td>0.02-0.06; 0.01</td>
</tr>
<tr>
<td>Oct</td>
<td>0.96-1.31; 0.14</td>
<td>0.04-0.10; 0.02</td>
<td>0.17-0.42; 0.09</td>
<td>0.03-0.08; 0.02</td>
</tr>
<tr>
<td>Nov</td>
<td>0.90-1.48; 0.19</td>
<td>0.05-0.15; 0.04</td>
<td>0.17-0.54; 0.14</td>
<td>0.04-0.10; 0.02</td>
</tr>
<tr>
<td>Dec</td>
<td>0.73-1.67; 0.29</td>
<td>0.03-0.18; 0.05</td>
<td>0.14-0.66; 0.19</td>
<td>0.04-0.12; 0.03</td>
</tr>
</tbody>
</table>