Estimating North American background ozone in U.S. surface air with two independent global models: Variability, uncertainties, and recommendations

A.M. Fiore[a*], J.T. Oberman[b], M.Y. Lin[c], L. Zhang[d,e], O.E. Clifton[a], D.J. Jacob[d], V. Naik[f], L.W. Horowitz[c], J.P. Pinto[g], G.P. Milly[h]

[a]Department of Earth and Environmental Sciences and Lamont-Doherty Earth Observatory of Columbia University, 61 Route 9W, Palisades, NY, USA; amfiore@ldeo.columbia.edu; oclifton@ldeo.columbia.edu

[b]Nelson Institute Center for Sustainability and the Global Environment (SAGE), University of Wisconsin-Madison, Madison, WI, USA; oberman.wisc@gmail.com

[c]NOAA Geophysical Fluid Dynamics Laboratory and Atmospheric and Oceanic Sciences, Princeton University, 201 Forrestal Road, Princeton, NJ, USA; meiyun.lin@noaa.gov; larry.horowitz@noaa.gov

[d]School of Engineering and Applied Sciences, Harvard University, 29 Oxford Street, Cambridge, MA, USA; djacob@fas.harvard.edu

[e]Department of Atmospheric and Oceanic Sciences & Laboratory for Climate and Ocean-Atmosphere Studies, School of Physics, Peking University, China; zhanglg@pku.edu.cn

[f]UCAR/NOAA Geophysical Fluid Dynamics Laboratory, Princeton, NJ, USA; Vaishali.Naik@noaa.gov

[g]U.S. EPA, National Center for Environmental Assessment, Research Triangle Park, NC, USA; Pinto.Joseph@epa.gov

[h]Lamont-Doherty Earth Observatory of Columbia University, 61 Route 9W, Palisades, NY, USA; gpm2109@columbia.edu

*Corresponding author; phone 1-845-365-8580; fax 1-845-365-8157

Submitted to *Atmospheric Environment*

December 26, 2013

Revised, July 24, 2014

Abstract

Accurate estimates for North American background (NAB) ozone (O_3) in surface air over the United States are needed for setting and implementing an attainable national O_3 standard. These estimates rely on simulations with atmospheric chemistry-transport models that set North American anthropogenic emissions to zero, and to date have relied heavily on one global model. We examine NAB estimates for spring and summer 2006 with two independent global models (GEOS-Chem and GFDL AM3). We evaluate the base simulations, which include North American anthropogenic emissions, with mid-tropospheric O_3 retrieved from space and ground-level O_3 measurements. The models often bracket the observed values, implying value in developing a multi-model approach to estimate NAB O_3. Consistent with earlier studies, the models robustly simulate the largest nation-wide NAB levels at high-altitude western U.S. sites (seasonal average maximum daily 8-hour values of ~40-50 ppb in spring and ~25-40 ppb in summer) where it correlates with observed O_3. At these sites, a 27-year GFDL AM3 simulation simulates observed O_3 events above 60 ppb and indicates that year-to-year variations in NAB O_3 influence their annual frequency (with NAB contributing 50-60 ppb or more during individual events). During summer over the eastern United States (EUS), when photochemical production from regional anthropogenic emissions peaks, NAB is largely uncorrelated with observed values and it is lower than at high-altitude sites (average values of ~20-30 ppb). Four processes contribute substantially to model differences in specific regions and seasons: lightning NO_x, biogenic isoprene emissions and chemistry, wildfires, and stratosphere-to-troposphere transport. Differences in the representations of these processes within the GFDL AM3 and GEOS-Chem models contribute more to uncertainty in NAB estimates, particularly in spring when NAB is highest, than the choice of horizontal resolution within a single model (GEOS-Chem). We propose that future efforts seek to constrain these processes with targeted analysis of multi-model simulations evaluated with observations of O_3 and related species from multiple platforms, and thereby reduce the error on NAB estimates needed for air quality planning.

1. Introduction

The United States Environmental Protection Agency (U.S. EPA) sets National Ambient Air Quality Standards (NAAQS) to protect human health and welfare. Under the Clean Air Act, ground-level ozone (O_3) is regulated as a criteria air pollutant, reviewed every five years to

63 assess and incorporate the best available scientific evidence. Following these reviews, the level
64 for the O_3 NAAQS has been lowered over the past decade, from 0.08 ppm in 1997 to the current
65 level of 0.075 ppm (75 ppb) in 2008, with proposals calling for even lower levels, within a range
66 of 60-70 ppb on the basis of the latest health evidence (Federal Register, 2010). A location is
67 considered to be in violation of the O_3 NAAQS when the three-year-average of the fourth highest
68 daily maximum 8-h average O_3 (MDA8) exceeds the current 75 ppb level. In order to better
69 understand how the O_3 NAAQS can be attained most effectively, a fundamental, quantitative
70 understanding of the background O_3 – both its magnitude and variability- over the United States
71 is needed.
72 The first draft of the current U.S. EPA Policy Assessment (EPA, 2013) and McDonald-
73 Bueller et al. (2011) describe the relevance of background O_3 in the U.S. national O_3 standard-
74 setting process. Here we review recent model estimates for background O_3 (Table 1) and
75 compare simulations from two independent models (GEOS-Chem and GFDL AM3) in the
76 context of observational constraints with a focus on spatial, seasonal, and daily variability.
77 Differences between the models provide a first estimate of the error in our quantitative
78 understanding. A process-oriented multi-model approach, tied closely to *in situ* and space-based
79 observations, can harness the strengths of individual models to provide information requested by
80 air quality managers during both the standard-setting and implementation processes.
81 The term "background" is ambiguous, with several definitions used in practice to
82 estimate it from observations and models (e.g., see discussion in Fiore et al., 2003). In the
83 context of a review of the NAAQS, it is useful to define background O_3 concentrations in a way
84 that distinguishes the O_3 produced from precursor emissions that are relatively less controllable
85 versus from precursor emissions that are relatively more controllable through U.S. policies. The
86 U.S. EPA thus defines a North American Background (NAB) as the O_3 levels that would exist in
87 the absence of continental North American (i.e., Canadian, U.S., and Mexican) anthropogenic
88 emissions (EPA, 2006). NAB includes contributions from natural sources (stratospheric
89 intrusions), emissions of precursors from natural sources (e.g., wildfires, lightning, biogenic)
90 throughout the globe, anthropogenic methane (CH_4), and emissions of anthropogenic pollutants
91 from countries outside North America that contribute to global O_3 abundances. This definition
92 restricts NAB to a model construct, estimated in simulations in which North American
93 anthropogenic emissions are set to zero. The desire to quantify the impact of Canadian and

94 Mexican emissions on NAB O_3 has led to the term "U.S. background", a parallel model construct
95 but estimated by setting only U.S. anthropogenic emissions to zero.
96 The development of effective State Implementation Plans (SIPs), by which states
97 demonstrate how non-attainment regions will reach compliance with the NAAQS, requires an
98 accurate assessment of the role of local, regional, and background sources in contributing to
99 individual high-O_3 events. The Clean Air Act includes a provision for 'exceptional events',
100 whereby high-O_3 events due to natural causes (such as wildfires or stratospheric intrusions) or
101 foreign influence (e.g., Asian pollution) can be exempted from counting towards non-attainment
102 status (Federal Register, 2007). Modeling the individual components of NAB can provide
103 information to aid in attributing such events to specific sources.
104 In the previous review cycle (EPA, 2006) of the O_3 NAAQS, the Air Quality Criteria
105 Document considered NAB estimates from the GEOS-Chem model for a single year (Fiore et al.,
106 2003), the only estimates documented in the published literature at that time. Recent work has
107 updated those estimates (Wang et al., 2009; Zhang et al., 2011), compared them with NAB in
108 regional models using GEOS-Chem boundary conditions (Emery et al., 2012; Mueller and
109 Mallard, 2011) and considered additional years. The first NAB estimates with an independent
110 global model, (GFDL AM3; hereafter AM3; Table 2) were found to episodically reach 60-75
111 ppb over the Western United States in spring (Lin et al., 2012a). By contrast, GEOS-Chem
112 estimated a maximum NAB of 65 ppb (Zhang et al., 2011) and the AM3 NAB was typically ~10
113 ppb higher than GEOS-Chem NAB on days when observations exceeded 70 ppb (Lin et al.,
114 2012a). Lin et al. (2012a) and Lin et al. (2012b) are discussed in the Integrated Science
115 Assessment (ISA) for the current O_3 NAAQS review cycle (EPA, 2013) but they focused on a
116 different simulation year (2010) from the other studies. The ISA supplemental material does
117 include a comparison of GEOS-Chem and GFDL AM3 for the same year (2006) at a
118 measurement site in Gothic, CO, U.S.A. Here we extend that initial analysis by examining, in a
119 consistent and process-oriented manner, the AM3 and GEOS-Chem NAB estimates during
120 March through August of 2006. We additionally draw on a multi-decadal AM3 simulation to
121 provide context for the single year inter-comparison. We include an evaluation of total surface
122 O_3 in the base simulations with ground-based and space-based observations during 2006 to
123 identify conclusions that are robust to the specific modeling system, as well as situations where
124 observation-based constraints can be most effective in reducing uncertainty.

2. Review of prior model estimates for NAB and its components

We focus here on model estimates for NAB using the U.S. EPA definition, which relies on simulations with North American anthropogenic emissions set to zero. Earlier reviews synthesize observations relevant for evaluating base model simulations at remote sites (McDonald-Buller et al., 2011; Reid et al., 2008; Vingarzan, 2004). Even with the same approach, model estimates will differ due to different representations of natural emissions and the choice of different years since meteorological variability alters the balance between transported versus regionally produced O_3. In Table 1, we summarize published model estimates of various statistics for NAB, along with estimates from individual NAB sources (wildfires, lightning, the stratosphere, global anthropogenic CH_4 plus international anthropogenic emissions, and the sum of all natural sources).

Despite quantitative differences, a basic consensus emerges that the highest NAB levels generally occur during springtime and over western U.S. (WUS) high-altitude regions, with lowest NAB levels over EUS low-altitude regions in summer. The summertime minimum reflects the peak in regional photochemistry, which leads to accumulation of O_3 generated from regional precursors at the same time as it shortens the lifetime of O_3 mixing downward into the photochemically active boundary layer (see e.g., Fiore et al., 2002). At high-altitude WUS sites, models consistently indicate a day-to-day correlation between NAB levels and total O_3 during spring (Emery et al., 2012; Fiore et al., 2003; Lin et al., 2012a; Lin et al., 2012b; Zhang et al., 2011), implying that enhanced NAB levels play a role in raising total O_3, including above the level of the NAAQS. While these results are qualitatively consistent across several modeling platforms, the models vary in their quantitative attributions for NAB and its specific sources.

A few studies report the annual fourth highest MDA8 NAB value, the metric used to assess compliance with the O_3 NAAQS. Consideration of different metrics and different years complicates using the ranges across different modeling systems in Table 1 as error estimates. For example, mean values of NAB are unlikely to be static from year to year due to trends and variability in global anthropogenic emissions of O_3 precursors, in natural sources of NAB, and in the dominant regional transport patterns (Lin et al., 2014). A multi-model parameterization indicates an increase of ~4 ppb between 1960 and 2000 due to rising global CH_4 plus

156 international anthropogenic emissions of non-CH_4 O_3 precursors (Wild et al., 2012). More
157 recent increases in Asian emissions may have additionally raised WUS NAB by up to 3 ppb in
158 spring between 2001 and 2006 (Zhang et al., 2008). The Asian, European, and global
159 anthropogenic CH_4 components of NAB have received particular attention under the UNECE
160 Task Force on Hemispheric Transport of Air Pollution (Fiore et al., 2009; Reidmiller et al., 2009;
161 TFHTAP, 2010; Wild et al., 2012). Recent studies have documented the mechanisms by which
162 Asian pollution can reach surface air over the WUS (e.g., Brown-Steiner and Hess, 2011; Lin et
163 al., 2012b).
164 Wang et al. (2009) additionally estimated summertime U.S. Background (USB), which
165 includes the influence of Canadian and Mexican anthropogenic emissions (but not CH_4) for 2001
166 conditions. They found that average USB is 4 ppb higher than NAB over the contiguous United
167 States, and up to 33 ppb higher during transport events at U.S. border sites directly downwind of
168 these sources. In the model, Canadian and Mexican sources often contributed more than 10 ppb
169 to total surface O_3 in excess of the 75 ppb NAAQS level in eastern Michigan, western New York,
170 New Jersey, and southern California (Wang et al., 2009).
171 The natural portion of NAB has been quantified in a few modeling studies and generally
172 follows the same patterns as total NAB, with maximum levels occurring during spring at high-
173 altitude regions of the WUS (Table 1). Natural sources of NAB can also contribute to high-O_3
174 events. Observational evidence indicates events mainly of stratospheric origin at high-altitude
175 sites in the WUS (e.g., Langford et al., 2009) but these efforts are hampered by a sparse
176 observational network. Models are useful for quantifying the frequency of these events and for
177 determining the contribution of these events to seasonal mean O_3 levels. For decades,
178 quantifying the stratospheric contribution to the troposphere, and particularly to surface air, has
179 been contentious, with controversy rooted in the imprecise methods for quantifying accurately
180 this component, as summarized in Lin et al. (2012a) (see their Section 2.3). Lin et al. (2012a)
181 demonstrate that stratospheric intrusions play an important role in driving variability, including
182 high-O_3 events, at high-altitude WUS sites during spring. High-altitude greatly increases
183 susceptibility to stratospheric influence; for days when observed O_3 exceeds 70 ppb at
184 monitoring sites in the Intermountain West during April-June of 2010, Lin et al. (2012a) find that
185 median values of stratospheric O_3 in the AM3 model are 10 ppb lower at the lower elevation
186 sites than at high-elevation sites. Episodic wildfires have also been shown to contribute to high-

187 O_3 events (e.g., Jaffe and Wigder, 2012; McKeen et al., 2002; Mueller and Mallard, 2011),
188 though Singh et al. (2010) found little O_3 production in wildfire plumes over California unless
189 mixing with an urban plume occurred.
190
191 **3. North American background estimates from two independent global models**
192 We compare background estimates for March through August of 2006 from two
193 independent global models: the GEOS-Chem global chemistry-transport model (CTM) and the
194 GFDL AM3 chemistry-climate model nudged to re-analysis winds. The models include different
195 representations of the processes contributing to the abundance and distributions of tropospheric
196 O_3 (Table 2). We evaluate the base O_3 simulations with hourly measurements from a ground-
197 based network of monitoring sites and with monthly averaged retrievals from satellite
198 instruments that are sensitive to O_3 in the mid-troposphere. We compare the models for March
199 through August of 2006, the period analyzed previously by Zhang et al. (2011), drawing on a 27-
200 year AM3 simulation to place the 2006 NAB estimates in the context of inter-annual variability.
201 We note that the inter-annual variability may be underestimated in AM3 over some regions due
202 its use of climatological inventories for soil NO_x and wildfire emissions.
203
204 **3.1. Model Simulations, Observations and Analysis Methods**
205 Table 2 describes the GEOS-Chem and AM3 model configurations for the base
206 simulations (meteorological year 2006). The GEOS-Chem CTM has been applied in various
207 configurations over the past decade to estimate NAB and its various components for the summer
208 of 1995 (Fiore et al., 2002), the 2001 O_3 season (Fiore et al., 2003; Wang et al., 2009), and the
209 2006-2008 O_3 seasons (Zhang et al., 2011; Zhang et al., 2014) including evaluation with *in situ*
210 and satellite observations. The AM3 model has previously been applied at ~50 km horizontal
211 resolution globally to estimate the impacts of Asian pollution and stratospheric intrusions on
212 surface O_3 over the WUS from March through June of 2010; evaluation with *in situ* and space-
213 based observations for that period shows it represents the subsidence of Asian and stratospheric
214 O_3 plumes over the WUS (Lin et al., 2012a; Lin et al., 2012b). The AM3 simulation used here is
215 ~200 km horizontal resolution and is multi-decadal (1980-2007; first year is discarded as
216 initialization), enabling us to place the year 2006 in the context of inter-annual variability (Table
217 2 and Section 4). Both models estimate NAB in U.S. surface air by setting North American

anthropogenic emissions of aerosol and O_3 precursors to zero. Anthropogenic sources include fossil and biofuel combustion (including aircraft and ship emissions within the domain), agricultural waste burning, and fertilizer application.

For anthropogenic emissions inventories, GEOS-Chem uses the 2005 National Emissions Inventory for the U.S., while AM3 uses the historical Atmospheric Chemistry-Climate Model Intercomparison Project (ACCMIP) emissions developed in support of the IPCC AR5 (Lamarque et al., 2011; Lamarque et al., 2010). Global, North American, and East Asian annual emissions for 2006 are provided in Table 2. Differences in the North American anthropogenic emissions inventories (5.58 and 6.67 Tg N a^{-1} in AM3 and GEOS-Chem, respectively; 4.85 and 5.32 Tg N a^{-1} for the United States), while crucial to the standard simulation for comparison with observations, should be irrelevant for the NAB simulations. Shortcomings in model representation of anthropogenic emissions or isoprene chemistry do not necessarily imply shortcomings in the NAB simulations, which may still be useful for quantifying daily to inter-annual variability driven by the transported components of NAB, such as O_3 associated with stratospheric intrusions, production from lightning NO_x, wildfires, or CH_4.

The ground-based U.S. EPA Clean Air Status and Trends Network (CASTNet) sites were situated in order to provide regionally representative measurements and to minimize the influence of polluted urban air (Baumgardner et al., 2002) and thus are useful for evaluating O_3 simulated by coarse grid models. Our evaluation focuses on MDA8 O_3 concentrations, the statistic currently used by the U.S. EPA to assess compliance with the O_3 NAAQS, at 77 CASTNet sites. Simulated MDA8 is calculated from archived hourly average O_3 concentrations in the model surface layer. The models use a terrain-following sigma-coordinate for near-surface layers, with the lowest layer centered at approximately 60m and 70m for a column where the lowest layer is at sea level in GFDL AM3 and GEOS-Chem, respectively. All statistics are calculated by sampling the models at the locations of CASTNet sites with bilinear interpolation from the four nearest model grid cells to the latitude and longitude at each station.

Columns retrieved from satellite instruments are sensitive to free tropospheric O_3 and enable an evaluation on a continuous spatial scale of the simulated background available to subside into surface air. We use here direct tropospheric O_3 retrievals from both the Ozone Monitoring Instrument (OMI) (Liu et al., 2010) and the Tropospheric Emission Spectrometer (TES) (Beer, 2006). All data are processed using a single fixed a priori profile as described in

Zhang et al. (2010). Previous validation of these retrievals against *in situ* and aircraft measurements indicate an accuracy to within 5 ppb at 500 hPa (Zhang et al., 2010 and references therein). We remove the average bias of the satellite columns as compared to sondes at northern mid-latitudes prior to comparing with the model mid-tropospheric O_3 distributions. Annual biases averaged over 2005-2007 relative to northern mid-latitude O_3 sondes (Zhang et al., 2010) have been uniformly subtracted from the retrieved products prior to evaluating the AM3 and GEOS-Chem base simulations. As evident from Zhang et al. (2010) (see their Figure 5), the bias is not uniform and thus the real model error may deviate at any particular location from the true O_3 abundance differently than implied by the comparison with the satellite products reported here. We apply the appropriate satellite averaging kernels to the model daily O_3 fields for direct comparison with the retrieved satellite O_3 columns (Zhang et al., 2010). While the averaging kernels for the 500 hPa retrieved product for both the TES and OMI instruments are most sensitive to the mid-troposphere, there is a broad vertical sensitivity throughout the troposphere, but generally very little information is retrieved from the boundary layer (see example averaging kernels in Figure 1 of Zhang et al. (2010)).

3.2 Regional and seasonal NAB estimates

Seasonal mean MDA8 NAB O_3 is consistently higher over the WUS than the EUS in both models (Figure 1). During spring, AM3 simulates higher NAB over the high-altitude WUS than GEOS-Chem, which we attribute at least partially to a larger stratospheric influence in AM3 (Lin et al., 2012a) than in GEOS-Chem (Zhang et al., 2011). The diagnostics necessary to determine whether AM3 actually simulates more stratosphere-to-troposphere exchange of O_3, or whether it mixes free tropospheric air (including the stratospheric component) into the planetary boundary layer more efficiently, were not archived from these simulations. However, an evaluation of AM3 O_3 profiles at sonde launch locations during the 2010 CalNex field campaign indicates that the model captures much of the observed vertical structure of O_3 throughout the troposphere and lower stratosphere, including its day-to-day variability (Lin et al., 2012a; Lin et al., 2012b). During summer, the different simulated spatial patterns for NAB over the WUS are influenced by differences in the lightning NO_x sources (see Section 3.5).

Figure 2 shows the spatial patterns of the fourth highest NAB value between March 1 and August 31. As the O_3 seasonal cycle is typically highest during the summer in polluted regions,

280 we expect the fourth highest during this six-month period to represent reasonably this statistic
281 over a full year. AM3 simulates the highest values over Colorado whereas GEOS-Chem
282 indicates that the highest values occur over New Mexico (Figure 2), reflecting excessive NAB
283 associated with lightning NO_x and subsequent O_3 production and transport (Zhang et al., 2014).
284 Due to different timing of these processes, AM3 simulates the fourth highest NAB values during
285 spring over much of Colorado but GEOS-Chem simulates peak values over much of New
286 Mexico during August (Figure 2). Over Minnesota and Wisconsin, GEOS-Chem generally
287 produces the fourth highest values in spring, but AM3 indicates that they occur in summer. Over
288 the northeastern states and west coast, the fourth highest MDA8 concentrations generally occur
289 during spring, though they occur later in the year over southeastern states, with occurrences
290 generally later in GEOS-Chem than AM3.
291 The fourth highest values often occur during months when model biases are largest
292 (Section 3.4), indicating that bias-correction techniques may be necessary for quantitatively
293 accurate NAB estimates at specific locations and times. In the following sections, we analyze the
294 model NAB estimates in the context of evaluating the total surface O_3 simulations with both
295 space- and ground-based observations, a first step towards developing the process-level
296 knowledge needed for accurate bias correction.
297
298 **3.3 Constraints from space-based observations**
299 With the exception of O_3 produced within the U.S. boundary layer from CH_4 or natural
300 NMVOC and natural NO_x, NAB in surface air mixes downward from the free troposphere. We
301 use 500 hPa products retrieved from both the OMI and TES instruments aboard the NASA Aura
302 satellite to evaluate the potential for space-based constraints on simulated mid-tropospheric total
303 O_3 distributions. Our comparison thus evaluates the reservoir of mid-tropospheric O_3, of any
304 origin, that can mix into the planetary boundary layer. For context, the ratio of NAB to total O_3
305 over North America at 500 hPa in AM3 varies spatially and seasonally, with highest values
306 generally in spring and a minimum contribution within any model grid cell of 70% (not shown).
307 During spring, AM3 estimates a stronger north-to-south O_3 gradient in the mid-
308 troposphere than GEOS-Chem (compare first and third rows of Figure 3). The satellite retrievals
309 from both instruments suggest a stronger gradient than simulated with GEOS-Chem, which
310 generally underestimates O_3 in the northern half of the United States compared both to TES (5-

311 15 ppb) and OMI (up to 10 ppb). In contrast, AM3 mid-tropospheric O_3 is higher than the
312 satellite products in the northern half of the domain, with a closer match to the OMI retrievals
313 (generally within 5 ppb over the United States) than TES (positive biases up to 10-20 ppb). Prior
314 direct evaluation of AM3 with O_3 sondes indicates biases of up to 10 ppb in AM3 at the high
315 northern latitude measurement sites of Alert and Resolute at 500 and 800 hPa with little bias in
316 spring at the mid-latitude North American sites of Edmonton, Trinidad Head, Boulder and
317 Wallops Island (Naik et al., 2013), roughly consistent with the biases relative to OMI.
318 Both satellite instruments indicate a general decrease from spring into summer over the
319 western and northern United States, but an increase over several southeastern states, northern
320 Mexico, and the Gulf of Mexico (compare Figures 3 and 4). The summertime spatial pattern of
321 U.S. O_3 observed from space is broadly consistent with that estimated by interpolating upper
322 tropospheric ozonesonde measurements during August of 2006 (Cooper et al., 2007). While the
323 increases from spring to summer in the mid-troposphere over the EUS may include a
324 contribution from lofting of regional anthropogenic O_3 production, there is likely also a
325 contribution from the larger lightning NO_x source in the free troposphere during summer. GEOS-
326 Chem estimates a summertime mid-tropospheric O_3 enhancement at mid-latitudes, centered over
327 the United States whereas AM3 simulates a gradient with O_3 generally increasing along the
328 southwest-to-northeast direction (Figure 4). Over Canada, the AM3 model tends to be higher
329 than both retrievals by up to 15-20 ppb during summer, as occurs during spring versus TES (but
330 springtime biases are smaller compared to OMI).
331 We expect discrepancies between AM3 and observations during summer over forested
332 boreal regions due to the use of a climatological wildfire inventory and the vertical distribution
333 used to prescribe those emissions (Dentener et al., 2006), which lofts fire effluents into the mid-
334 troposphere where they can efficiently produce O_3 and PAN (see also Section 3.5). GEOS-Chem
335 includes fire emissions representative of the year 2006 and emits them only the planetary
336 boundary layer. These model differences in wildfire treatment are likely reflected in the mid-
337 tropospheric O_3 distributions shown in Figures 3 and 4. They may also contribute to the different
338 spatial distributions of simulated NAB at the surface, specifically the higher NAB estimated with
339 AM3 over the northern United States and Canada relative to the NAB estimated with GEOS-
340 Chem (Figures 1 and 2).

In both Figures 3 and 4, the models are generally more consistent with the OMI retrievals, which may reflect differences in the vertical sensitivity of the TES and OMI instruments. While the satellite retrievals provide useful qualitative constraints on the simulated mid-tropospheric distributions, the disagreement between OMI and TES over many locations (grey boxes in Figures 3 and 4) hinders their quantitative utility. The higher sampling frequency possible from instruments on geostationary satellites such as TEMPO (Hilsenrath and Chance, 2013) should improve the potential for space-based constraints on free-tropospheric and near-surface distributions.

We can nevertheless glean additional insights into the model vertical distributions of O_3 by examining differences in the models sampled with the two different averaging kernels (Figures 3 and 4). For example, over Canada, GEOS-Chem indicates that OMI would measure higher O_3 than TES whereas AM3 indicates that TES should retrieve higher O_3 than OMI during both seasons. In the spring, the retrieved OMI product is generally higher than TES over this region. GEOS-Chem is generally within 10 ppb of the OMI product with a tendency to underestimate springtime mid-tropospheric O_3 over Canada, whereas AM3 is generally within 5 ppb of OMI over much of the United States and Canada, with a tendency towards a positive bias. During summer, TES is higher than OMI over Canada. The high O_3 bias over the EUS in AM3 is confined close to the surface (Figure 5) since AM3 tends to underestimate free tropospheric O_3, particularly over the convectively active Gulf of Mexico region where lightning NO_x is expected to be an important source of NAB O_3. We conclude that the estimates from the models could bracket the true NAB in many cases, but caution that the ability of the models to bracket the satellite measurements does not necessarily imply accuracy in their NAB estimates.

3.4 Constraints from ground-based measurements

We use the CASTNet MDA8 O_3 observations to further constrain the model NAB estimates through an evaluation of the base simulations, which include all anthropogenic emissions, to simulate total surface O_3. Since NAB depends strongly on altitude (Figure 1; references in Table 1), the remainder of our analysis separates the data by altitude to gain insight into the different processes shaping NAB distributions. Specifically, we divide the CASTNet sites into two groups: (1) below 1.5 km in elevation (low-altitude sites), primarily sites in the EUS, and (2) Intermountain West CASTNet sites with elevation greater than 1.5 km (high-

372 altitude sites). This second category includes all high-altitude CASTNet sites except for those in
373 California.
374
375 **3.4.1. Seasonal Variability**
376 Figure 5 shows the observed and simulated seasonal cycles at the CASTNet sites. At the
377 high-altitude sites, both models are generally within 5 ppb of the regional mean observed values
378 and usually fall within one standard deviation of the observed monthly mean values at the sites
379 within the region. Consistent with the evaluation in Section 3.3, the models tend to bracket the
380 observations, but with notably different seasonal cycles. AM3 peaks in early spring,
381 overestimating observed values in March but then declines to slightly underestimate observed
382 values in June and July. In contrast, GEOS-Chem underestimates observed values from March
383 through July but increases to overestimate observed values in August. The model differences are
384 amplified in the NAB estimates: AM3 simulates a large seasonal decline in NAB from
385 springtime (near 50 ppb) into summer (below 35 ppb) while GEOS-Chem estimates little
386 seasonality in NAB (monthly mean values around 40 ppb).
387 At the low-altitude sites, AM3 exhibits a large positive bias in total surface O_3 in all
388 months, most evident during summer. The exacerbation of the bias in summer at low-altitude
389 sites implies a problem with O_3 produced from regional emissions, with isoprene-NO_x-O_3
390 chemistry a likely culprit given its different treatment in the models (Table 2; see Section 3.5).
391 Both models show declining NAB levels from spring into summer, though the GEOS-Chem
392 amplitude of the NAB seasonal cycle is smaller than that of AM3. The AM3 discrepancy with
393 observations is much larger than the difference between the GEOS-Chem and AM3 NAB
394 estimates except for March and April. If we assume that the model biases during March and
395 April at both the high and low altitude sites are entirely due to NAB and correct the NAB
396 estimates accordingly, the model NAB estimates would become more similar. While we
397 conclude that the AM3 NAB at low-altitude sites is too high in March since we expect NAB to
398 be lower than the observed value, it is possible that NAB could actually be higher in an
399 atmosphere with lower NO_x than under current conditions, due to more efficient O_3 production
400 and slower chemical loss.
401 At the high-altitude sites in summer, the GEOS-Chem overestimate of observed O_3 has
402 been attributed previously to an overestimate of O_3 associated with production from lightning

403 NO_x plus subsequent transport when prescribing a higher production of NO_x from flashes at mid-
404 latitudes and spatially scaling the source to match LIS-OTD climatological flash counts (Murray
405 et al., 2012), which may lead to regional errors for a specific year (Zhang et al., 2014). The
406 larger difference between the NAB estimates from the two models in August than between the
407 simulated and observed total O_3 implies that the agreement with observations does not
408 sufficiently constrain the NAB estimates.
409
410 **3.4.2. Daily Variability**
411 Figure 6 shows probability density distributions constructed from observed and simulated
412 MDA8 O_3 in spring (top) and summer (bottom) sampled at the high-altitude (left) versus low-
413 altitude (right) CASTNet sites, and statistics are summarized for the AM3 and high-resolution
414 GEOS-Chem simulations in Table 3. We additionally include in Figure 6 estimates from a
415 coarse resolution version of the GEOS-Chem model (green) in order to examine the extent to
416 which differences in horizontal resolution contribute to the different NAB and total O_3 estimates
417 in AM3 versus GEOS-Chem. In all cases, the NAB (dotted lines) differ more between the
418 GEOS-Chem and AM3 models than between the high- versus low-resolution versions of GEOS-
419 Chem. This conclusion also holds for the total O_3 distributions in spring. In summer, however,
420 the total O_3 distributions in GEOS-Chem are more sensitive to the choice of horizontal resolution,
421 presumably reflecting the larger contributions from local-to-regional photochemical production
422 during this season and the importance of spatially resolving domestic anthropogenic and natural
423 emissions distributions. Emery et al. (2012) found that the higher resolution CAMx model
424 generally simulated higher WUS NAB than a coarse resolution version of GEOS-Chem, and
425 better agreement has been noted between CAMx and the higher resolution version of GEOS-
426 Chem (EPA, 2013). Simulation of higher WUS NAB by higher resolution models (Emery et al.,
427 2012; Lin et al., 2012a) likely reflects improved resolution of topography and mesoscale
428 meteorology at higher resolution and the damping of vertical eddy transport at coarser resolution
429 (Wang et al., 2004; Zhang et al., 2011).
430 AM3 simulates a wider NAB range than GEOS-Chem (Figure 6 and Table 3). This wider
431 range of NAB may contribute to the wider total surface O_3 distribution in the AM3 versus
432 GEOS-Chem standard simulations, which aligns more closely with the observed variability,
433 except for O_3 simulated with the high-resolution GEOS-Chem model in summer at high-altitude

434 sites. The relative skill of AM3 in capturing the variability of NAB despite its generally high
435 bias implies that AM3 is useful for process-level analysis and for quantifying day-to-day
436 variability.
437 In Table 3, we partition statistics for total and NAB O_3 in surface air into average versus
438 high-O_3 days. We use observed values, rather than simulated values used in Zhang et al. (2011),
439 to select for high-O_3 days in order to sample the same temporal subset from both models. Using
440 the simulated total O_3 values would lead to subsets of different sizes given the individual model
441 biases. During spring, the models robustly estimate NAB to be ~10 ppb higher on average at
442 high-altitude than at low-altitude CASTNet sites, but AM3 estimates higher NAB levels than
443 GEOS-Chem. During summer, the models also estimate higher NAB at high-altitude than at
444 low-altitude sites, and average NAB levels decrease from spring to summer at low-elevation sites.
445 GEOS-Chem suggests little change from spring to summer in average high-altitude NAB
446 whereas AM3 simulates a decrease of over 10 ppb. At the high-altitude sites, both models
447 suggest that NAB increases as total O_3 increases, although the sample size is small for events
448 above 75 ppb and the average values for the different data subsets all fall within one standard
449 deviation each other. At the low altitude sites, there is little change in the average NAB when
450 selecting for observed values exceeding 60, 70, or 75 ppb. The variability in NAB, as measured
451 by the standard deviation in Table 3, is similar in the two models at the low-elevation sites, but
452 AM3 simulates more variability in NAB at the high-altitude sites than GEOS-Chem, particularly
453 on high-O_3 days.
454 We next evaluate the ability of the models to capture observed day-to-day variability by
455 correlating observed and simulated MDA8 O_3 in surface air at each CASTNet site, separately for
456 the spring and summer seasons of 2006 (Figure 7). During spring, the correlation coefficients
457 between observations and total simulated surface O_3 over WUS sites are generally higher in
458 GEOS-Chem, but the level of correlation in summer is maintained or improved in AM3 whereas
459 it decreases in GEOS-Chem (Figure 7). Over the EUS, the models show similar spatial patterns
460 in their ability to reproduce the observed day-to-day variability. GEOS-Chem is generally better
461 than AM3 over northern sites, whereas AM3 captures more of the variability over the
462 southeastern sites in summer (Figure 7).
463 Figure 8 shows the correlation coefficients of the NAB estimates versus the simulated
464 total O_3, separately for spring and summer, for each model sampled at the CASTNet sites that

were used in Figure 7 to evaluate the simulated daily variability. Over the Intermountain WUS sites, the models robustly indicate that variability in NAB drives a substantial portion of the total surface O_3 variability in both seasons, but with a stronger influence (higher correlations) during spring. AM3 generally indicates a stronger role for NAB in contributing to variability in total surface O_3 at the sites in Florida and Texas. Over the central states and over the inland mid-Atlantic region, the NAB is more strongly correlated with total surface O_3 in GEOS-Chem than in AM3; the stronger correlation in GEOS-Chem in these regions may arise from soil NO_x emissions, which respond to meteorological variability in GEOS-Chem whereas AM3 uses a monthly climatology. At the EUS sites in Figure 8, the NAB in both models is poorly correlated, and in some cases, anti-correlated with the total simulated surface O_3. The highest total surface O_3 events over the EUS are thus generally decoupled from the highest NAB events, consistent with the current understanding that regional pollution is the dominant influence on total O_3 distributions in this region.

We select four sites, encircled in black in Figures 7 and 8, to probe more deeply the day-to-day variability in NAB and total surface O_3. The time series at these sites (Figures 9a and 9b) provide evidence at the local scale for our assessment of regional and seasonal biases. At the two western U.S sites (Gothic, CO and Grand Canyon NP, AZ) in Figure 9a, the 6-month average NAB is nearly the same in both models, but this reflects little seasonal variation in the GC NAB (thin blue line) versus a sharp seasonal decline from spring into summer in AM3 (thin red line). The standard deviation is twice as large in AM3 as in GEOS-Chem, consistent with the frequency distributions of NAB in Figure 6 (left side); AM3 also captures the observed variability. Despite the summertime high bias in AM3 at the two EUS sites (M.K. Goddard, PA and Georgia Station, GA), AM3 correlates at least as well with the observations as GEOS-Chem (Figures 7 and 9).

3.5 Processes contributing to inter-model differences in total and NAB surface O_3

We use the sites in Figure 9a to examine the role of specific processes in contributing to differences in the GEOS-Chem and AM3 Base and NAB simulations. Superimposed in Figures 9a and 9b are results from a separate simulation (Lin et al., 2014) in which a stratospheric O_3 tracer (O3Se90) was available (orange line), tagged relative to the e90 tropopause (Prather et al., 2011) as described in Lin et al. (2012a). The high summertime correlation of O3Se90 and NAB

496 at the WUS sites (Figure 9a) does not imply that stratospheric O_3 intrusions are the dominant
497 factor. The magnitude of the O3Se90 enhancements in summer do not account for all of the
498 episodic NAB enhancements. Rather, this result implies that other sources enhance NAB free
499 tropospheric O_3, which then mixes into the surface air alongside the O3Se90 tracer in the model.
500 We interpret the high correlations in both seasons at the EUS sites (Figure 9a) in a similar
501 manner: the O3Se90 indicates a larger influence of O_3 mixing down from the free troposphere.
502
503 ***Deep stratospheric intrusions over the WUS in spring***. As described by Lin et al. (2012a),
504 stratospheric O_3 drives a substantial portion of the daily variability in observed springtime O_3
505 over the WUS. Inspection of Figure 9a (top two panels) shows that the episodic enhancements
506 in the O3Se90 tracer can explain much of the episodic enhancements in NAB. A caveat is that
507 the magnitude of the stratospheric contribution is an upper limit due to the definition of the
508 O3Se90 tracer, which could be tagging O_3 in the lower stratosphere that originated in the
509 troposphere (see discussions in Lin et al. (2012b) and Zhang et al. (2014)). Nevertheless, the
510 strong correlation of the orange and red lines in Figure 9a implies a key role for mixing of free
511 tropospheric air into the planetary boundary layer in driving day-to-day variability in NAB O_3
512 levels.
513 For illustrative purposes, we focus on an event during late May of 2006 at the Gothic and
514 Grand Canyon sites, during which the AM3 model NAB (thin red line in Figures 9a and 9b)
515 spikes, with an associated increase in the simulated total O_3 (thick red line). In contrast, GEOS-
516 Chem NAB (thin blue line) decreases, as does total O_3 (thick blue line) during this event. The
517 opposing trends in the models during this event raise questions as to whether both models
518 simulate a mixing event but import different NAB levels, or whether the boundary layer in AM3
519 mixes more vigorously with the free troposphere than in GEOS-Chem. The observations (black)
520 increase during this period, as in AM3, but the model overestimates the observed values on May
521 27 and 28. The O3Se90 tracer (orange line in Figures 9a and 9b) suggests that the AM3 model is
522 simulating surface O_3 enhancements associated with a stratospheric intrusion; the spatial pattern
523 of these enhancements aligns with observed spatial pattern of enhanced ground-level O_3 at the
524 CASTNet sites (not shown). Figure 9a further suggests that these events drive much of the
525 variability in NAB at high-altitude western sites in spring, consistent with earlier findings for
526 April through June of 2010 (Lin et al., 2012a).

Wildfires over the EUS in spring and summer. There are several EUS events during spring and summer where AM3 simulates a localized spike in NAB that is not simulated by GEOS-Chem, which we attribute at least partially to the different treatment of wildfire emissions in the models (Table 2). For example, there is an extreme "NAB event" in the AM3 model on June 28, 2006 at the Pennsylvania CASTNet site in Figure 9a. AM3 estimates NAB values above 60 ppb, exceeding the total observed value of about 60 ppb, while GEOS-Chem simulates NAB below 20 ppb (Figure 9a, yellow highlight). We find that the use of a year-specific fire inventory (as is done in the GEOS-Chem simulations) versus a climatology leads to differences of 10 ppb for the June 28, 2006 event in AM3 (not shown). In AM3, the recommendations from Dentener et al. (2006) are applied to vertically distribute biomass burning emissions north of 25°N, placing 40% of the total emissions between 3 and 6 km (see their Table 4) over boreal North America. This lofting of fire effluents likely contributes to the summertime O_3 overestimates at 500 hPa over Canada (Figure 4). Vertical mixing of NAB O_3 from the free troposphere into surface air in the AM3 model is indicated by associated enhancements of the O3Se90 tracer on days with high NAB. In contrast, biomass burning is emitted only in the boundary layer in GEOS-Chem, likely resulting in less efficient O_3 production and subsequent long-range transport. The GEOS-Chem approach appears more consistent with the observations.

Lightning NO_x over the Southwestern United States in summer. During August at the two WUS sites in Figure 9a, the models reverse their relative rankings of simulated NAB relative to springtime, with the GEOS-Chem NAB as much as 10-20 ppb higher than AM3 NAB in summer. In notable contrast to the spring, GEOS-Chem overestimates the observed O_3 values. We attribute the summertime overestimate and poor correlations of GEOS-Chem with the observed values over WUS sites in Figures 7 and 9a to the lightning NO_x source and subsequent transport. GEOS-Chem produces approximately 10 times more lightning NO_x than AM3 over the southwestern states during summer (0.018 Tg N in AM3 versus 0.159 Tg N in GC within the region 26°N-42°N, 124°W-97°W) and the models further differ in their spatial distributions of the lightning NO_x source (Table 2). This source has been reduced in a newer version of GEOS-Chem, decreasing simulated NAB O_3 over these regions (Zhang et al., 2014).

Isoprene oxidation chemistry over the EUS in summer. Earlier work (e.g. Fiore et al. 2002, 2003) demonstrated that NAB is fundamentally different between the EUS and the WUS, with the EUS more strongly controlled by regional photochemistry, where the O_3 lifetime in the planetary boundary layer is as short as 1-2 days and isoprene-NO_x-O_3 chemistry dominates much of the region from May through September (Jacob et al., 1995). At the two EUS sites in Figure 9a (M.K. Goddard, PA and Georgia Station, GA), we attribute some of the differences in the summertime simulations to the isoprene oxidation mechanism (Table 2) that would tend to reduce O_3 production in GEOS-Chem relative to AM3 due to isoprene ozonolysis serving as a more important loss pathway for NAB in GEOS-Chem (Fiore et al., 2002; Mickley et al., 2001). These differences in isoprene oxidation chemistry could at least partially explain the higher NAB in AM3 during the isoprene emission season (*i.e.*, a longer O_3 lifetime in the AM3 boundary layer). The largest inter-model differences in NAB, however, occur in spring when transported sources are more important than regional production involving natural sources.

The isoprene oxidation chemistry likely also contributes to the large bias in AM3 total surface O_3. GEOS-Chem assumes a much higher yield of isoprene nitrates from the reaction of isoprene hydroxyperoxy radicals with NO and assumes that they are a permanent sink for NO_x (Table 2). In contrast, AM3 assumes an 8% isoprene nitrate yield and allows 40% of the products to recycle back to NO_x on the basis of observational constraints from field campaigns (Horowitz et al., 2007; Perring et al., 2009). Earlier work with predecessors of the models used here suggests that these differences may explain over 10 ppb of the high bias in AM3 relative to GEOS-Chem over the EUS in summer (Fiore et al., 2005). The fact that GEOS-Chem best captures the observations implies that the additional O_3 production from isoprene oxidation using the field-based constraints on isoprene nitrates must be offset by larger O_3 losses, such as may occur through additional HO_x uptake by aerosol (Mao et al., 2013) and halogen-induced O_3 destruction (Parrella et al., 2012).

4. Inter-annual variability in NAB MDA8 O_3 estimates in surface air

The 27-year AM3 NAB simulation (1981-2007) enables us to define spring and summer climatologies of seasonal mean NAB O_3 in surface air, and to quantify the year-to-year variability as the standard deviation of the annual seasonal mean values (Figure 10). The seasonal mean spatial patterns are similar to those in 2006 (Figure 1), with little year-to-year

589 variation over much of the country. Figure 10 also includes the climatological fourth highest
590 MDA8 value between March 1 and August 31 over the multi-decadal simulation. We emphasize
591 that these estimates are subject to the biases diagnosed above in comparison to observations. In
592 particular, NAB estimates over the EUS are probably too high in AM3. The variability over
593 central Texas and central Mexico in the fourth highest values may indicate year-to-year
594 variations in events involving NAB production from lightning NO_x and convective mixing.
595 Large variability in both mean NAB levels and the highest events is simulated over Western
596 Colorado in spring, with standard deviations of 2-3 ppb, likely reflecting variability in year-to-
597 year influence from stratospheric O_3 intrusions.
598 Jaffe (2011) noted regionally coherent year-to-year variability in the number of high-O_3
599 events at high-altitude western U.S. measurement sites in both spring and summer and we
600 examine here the potential contribution of NAB to this observed variability. Specifically, Jaffe
601 (2011; see their Figure 6) found that the number of O_3 events above levels of 65, 70, and 75 ppb
602 varied together, with the lowest and highest number of springtime events occurring in 1997 and
603 1999, respectively; for summer, the lowest and highest years were 1997 and 2002. We follow
604 the approach of Wang et al. (2009; see their Figure 5) to illustrate simultaneously the model skill
605 at capturing the observed values, and the simulated NAB contribution to observed levels within
606 specific ranges for total surface O_3. Figure 11 shows the AM3 NAB contributions throughout
607 the overall observed distributions for 2006 in comparison to a low versus high year for observed
608 high-O_3 events at the 11 high-altitude WUS sites in Figures 5 and 6. Note that the highest years
609 in Figure 11 differ for spring and summer, but the lowest year is 1997 in both seasons.
610 For observed O_3 events above 60 ppb, AM3 tends to overestimate observations during
611 spring but does not exhibit any systematic bias during summer. Furthermore, the model captures
612 events up to 80 ppb during spring of 1999, though in other years there is a general tendency to
613 underestimate events above 75 ppb. This finding is in contrast to those from higher-resolution
614 models including the GEOS-Chem version used here, which underestimates events above 60 ppb
615 (Zhang et al., 2011). During all years and both seasons shown in Figure 11, there is a tendency
616 for the median simulated NAB contribution to increase from observed values of 40 ppb to those
617 in the 70 ppb range, with 75th percentile values reaching 50-60 ppb for observed values above 60
618 ppb during 2006 and 1999, implying that enhanced NAB levels contribute to the higher observed
619 values. This interpretation is consistent with the findings of Lin et al. (2012a) that stratospheric

620 O_3 intrusions over the high-altitude WUS drive much of the observed day-to-day variability in
621 spring, as well as with Jaffe (2011) who suggests that a large-scale process drives coherent
622 variability at the monitoring sites in this region.
623 Consistent with earlier work (Fiore et al., 2003), Figure 11 shows that summertime NAB
624 levels are typically much lower than in spring, with maximum values nearly always below 60
625 ppb and 75th percentile values generally below 50 ppb. Jaffe (2011) suggested that summertime
626 inter-annual variability is strongly influenced by wildfire activity. The lack of year-to-year
627 variations wildfires in this version of the AM3 model may contribute to its underestimate of the
628 highest events in 2002 and 2006, which were the first and second highest fire activity years for
629 the 1997-2006 period analyzed by Jaffe (2011).
630
631 **5. Conclusions and Recommendations**
632 On the basis of health evidence, the level of the National Ambient Air Quality Standard
633 (NAAQS) for ground-level O_3 has been lowered in recent years, pushing closer to "background"
634 levels. In the past, the U.S. Environmental Protection Agency considered model-based estimates
635 of background O_3 as part of the process for setting the NAAQS. These model-based estimates,
636 previously called "Policy-Relevant Background", are now termed "North American Background"
637 (NAB), which is defined to be background levels that would exist in the absence of North
638 American anthropogenic emissions. Identifying high-background events is crucial for
639 determining whether an observation merits consideration for "exceptional event" status, which
640 exempts a particular observation from counting towards non-attainment if it can be shown that
641 the event occurred due to processes beyond the control of U.S. air quality management options.
642 The model simulations presented here can provide information on the frequency of such events
643 and the individual components contributing to NAB, including O_3 originating from international
644 pollution, wildfires, or the stratosphere.
645 As a first step towards assessing our understanding of NAB and its components, we
646 briefly reviewed recent model estimates (Table 1). We then evaluated total surface O_3 and NAB
647 estimates from two independent models (GEOS-Chem and AM3) for March through August of
648 2006, using comparisons between the base simulations and space-based and ground-based
649 measurements to place constraints on the model estimates. A 27-year NAB simulation in the
650 AM3 model provides context for our two-model analysis and indicates that 2006 is a typical year

651 in terms of its spatial and seasonal patterns in NAB, though 2006 NAB levels are generally
652 higher than the climatological averages (compare Figure 10 with 1 and 2). The largest
653 variability in mean NAB MDA8 estimated with AM3 occurs over Idaho, western Colorado and
654 Wyoming, and New Mexico, with standard deviations of over 2 ppb; the largest variability in the
655 fourth highest MDA8 NAB occurs over Colorado and Texas (Figure 10). A comparison of low-
656 versus high-O_3 years at high-altitude western U.S. (WUS) sites indicates a role for NAB in
657 driving year-to-year differences in the frequency of springtime high-O_3 events (Figure 11).
658 At high-altitude WUS sites, the GEOS-Chem and AM3 models consistently show higher
659 NAB than at low-altitude sites, but the magnitude and day-to-day variability often differs
660 (Figures 1,5,6,9a, Table 3). In some months (e.g., August), the larger differences between the
661 NAB estimates from the two models than between the simulated and observed total O_3, imply
662 that agreement with observations does not sufficiently constrain the NAB estimates. While AM3
663 indicates a seasonal decline of NAB into summer over this region, GEOS-Chem suggests a
664 relatively weak seasonal cycle associated with an increase of influence from lightning NO_x in
665 that model during the late summer (Figures 5 and 9). Higher stratosphere-troposphere exchange
666 in AM3 may explain the springtime NAB enhancement in the free troposphere relative to GEOS-
667 Chem (Figure 3), which, followed by more vigorous mixing between the free troposphere and
668 boundary layer, may explain the higher NAB in surface air during this season in AM3 (Figure 1).
669 At low-altitude sites, such as over the EUS, the models consistently show lower NAB
670 levels than at high-altitude sites, as in earlier work (Table 1). We find that the highest total
671 surface O_3 events over the EUS are often decoupled from the highest NAB events (Figure 8),
672 consistent with the understanding that regional pollution is the dominant influence on total O_3
673 distributions there. Over the EUS, uncertainties in isoprene-NO_x-O_3 chemistry (Table 2) likely
674 contribute to differences in simulated total O_3, and to a lesser extent, NAB estimates.
675 We find little evidence that horizontal resolution is a major contributor to differences in
676 mean NAB estimates in the models (Figure 6), consistent with EPA (2013). Higher resolution
677 refines spatially local NAB estimates, including at the tails of the distribution. It also resolves
678 better the impact from local and regional emissions, as evidenced by the larger differences
679 associated with resolution in summertime distributions when photochemical production peaks
680 over many U.S. regions (Figure 6). We conclude that simulated NAB distributions reflect large-
681 scale synoptic transport that is resolved sufficiently at the relatively coarse scale of global

682 models, with the NAB differences mainly stemming from different treatments of NAB sources
683 such as stratospheric O_3, boreal fires, and lightning NO_x. The regional and seasonal variability in
684 these driving processes further manifests as differences in the model timings of the fourth highest
685 NAB over many regions (Figure 2).
686 Future efforts to determine the processes contributing to model differences, and to the
687 biases in individual models versus observations, would benefit from evaluation with daily O_3
688 vertical profiles as measured by sondes, consistently defined tracers of stratospheric influence
689 (e.g., the O3Se90 tracer in AM3), as well as daily three-dimensional archival of other chemical
690 species (e.g., CO, PAN, H_2O) that can aid in disentangling tropospheric versus stratospheric
691 origins and from meteorological variables (e.g., mixing depth, mass fluxes) to diagnose the role
692 of mixing processes. The routine use of synthetic tracers could further aid in distinguishing
693 between model differences in transport, dilution, and mixing versus chemical evolution during
694 transport. Improved estimates of NAB in a given region and season will require better
695 constraints on, for example: lightning NO_x for central and southwestern states in summer;
696 transported stratospheric O_3 over the high-altitude WUS in spring; isoprene chemistry and its
697 impact on chemical processing and NAB lifetime over the EUS in summer; and wildfires, which
698 may influence NAB throughout the nation from late spring into summer.
699 We propose that future multi-model studies target limited time periods to enable process-
700 oriented analysis during field campaigns when ground-based and satellite observations are
701 supplemented with a broader suite of observations from intensive aircraft flights and balloon
702 launches. If combined with a thorough evaluation of O_3 precursors, such analysis should hasten
703 progress towards understanding the impact of specific sources on NAB O_3. We further
704 recommend developing bias-correction techniques, such as those routinely applied in numerical
705 weather prediction, to improve the accuracy of local NAB estimates. As a first step, simple
706 assumptions assuming the bias is entirely driven by one process (e.g., as applied to the
707 stratospheric O_3 estimates from the AM3 model by Lin et al. (2012a)) can be applied to
708 individual models and then used to generate a multi-model estimate with uncertainties. The two
709 models analyzed here often bracket the observations (Figures 3-6 and 9), thereby indicating
710 different sources of error, which leads us to conclude that a multi-model approach can harness
711 unique capabilities of different modeling systems and thus provide more accurate NAB estimates
712 than a single model.

Acknowledgments. We acknowledge support from the NASA Air Quality Applied Sciences Team (NNX12AF15G to AMF; NNX11AH93G to DJJ), the NOAA Ernest F. Hollings Scholarship program (JTO) and NOAA GFDL. We are grateful to P. Dolwick (U.S. EPA) for useful discussions and comments, and suggestions from anonymous review. The information in this article has been subjected to review by the National Center for Environmental Assessment, U.S. Environmental Protection Agency, and approved for publication. The views expressed in this manuscript are those of the authors and do not necessarily represent the views or policies of the U.S. Environmental Protection Agency.

References

Baumgardner, R.E., Lavery, T.F., Rogers, C.M., Isil, S.S., 2002. Estimates of the Atmospheric Deposition of Sulfur and Nitrogen Species: Clean Air Status and Trends Network, 1990-2000. Environmental Science & Technology 36, 2614-2629.

Beer, R., 2006. TES on the aura mission: scientific objectives, measurements, and analysis overview. Geoscience and Remote Sensing, IEEE Transactions on 44, 1102-1105.

Bey, I., Jacob, D.J., Yantosca, R.M., Logan, J.A., Field, B.D., Fiore, A.M., Li, Q., Liu, H.Y., Mickley, L.J., Schultz, M.G., 2001. Global modeling of tropospheric chemistry with assimilated meteorology: Model description and evaluation. J. Geophys. Res. 106, 23073-23095.

Brown-Steiner, B., Hess, P., 2011. Asian influence on surface ozone in the United States: A comparison of chemistry, seasonality, and transport mechanisms. J. Geophys. Res. 116, D17309.

Collins, W.J., Derwent, R.G., Garnier, B., Johnson, C.E., Sanderson, M.G., Stevenson, D.S., 2003. Effect of stratosphere-troposphere exchange on the future tropospheric ozone trend. Journal of Geophysical Research: Atmospheres 108, 8528.

Cooper, O.R., Trainer, M., Thompson, A.M., Oltmans, S.J., Tarasick, D.W., Witte, J.C., Stohl, A., Eckhardt, S., Lelieveld, J., Newchurch, M.J., Johnson, B.J., Portmann, R.W., Kalnajs, L., Dubey, M.K., Leblanc, T., McDermid, I.S., Forbes, G., Wolfe, D., Carey-Smith, T., Morris, G.A., Lefer, B., Rappenglück, B., Joseph, E., Schmidlin, F., Meagher, J., Fehsenfeld, F.C., Keating, T.J., Van Curen, R.A., Minschwaner, K., 2007. Evidence for a recurring eastern North America upper tropospheric ozone maximum during summer. Journal of Geophysical Research: Atmospheres 112, D23304.

Dentener, F., Kinne, S., Bond, T., Boucher, O., Cofala, J., Generoso, S., Ginoux, P., Gong, S., Hoelzemann, J.J., Ito, A., Marelli, L., Penner, J.E., Putaud, J.P., Textor, C., Schulz, M., van der Werf, G.R., Wilson, J., 2006. Emissions of primary aerosol and precursor gases in the years 2000 and 1750 prescribed data-sets for AeroCom. Atmos. Chem. Phys. 6, 4321-4344.

Donner, L.J., Wyman, B.L., Hemler, R.S., Horowitz, L.W., Ming, Y., Zhao, M., Golaz, J.-C., Ginoux, P., Lin, S.J., Schwarzkopf, M.D., Austin, J., Alaka, G., Cooke, W.F., Delworth, T.L., Freidenreich, S.M., Gordon, C.T., Griffies, S.M., Held, I.M., Hurlin, W.J., Klein, S.A., Knutson, T.R., Langenhorst, A.R., Lee, H.-C., Lin, Y., Magi, B.I., Malyshev, S.L., Milly, P.C.D., Naik, V., Nath, M.J., Pincus, R., Ploshay, J.J., Ramaswamy, V., Seman, C.J., Shevliakova, E., Sirutis, J.J., Stern, W.F., Stouffer, R.J., Wilson, R.J., Winton, M., Wittenberg, A.T., Zeng, F., 2011. The Dynamical Core, Physical Parameterizations, and Basic Simulation Characteristics of the Atmospheric Component AM3 of the GFDL Global Coupled Model CM3. Journal of Climate 24, 3484-3519.

Emery, C., Jung, J., Downey, N., Johnson, J., Jimenez, M., Yarwood, G., Morris, R., 2012. Regional and global modeling estimates of policy relevant background ozone over the United States. Atmospheric Environment 47, 206-217.

Emmons, L.K., Walters, S., Hess, P.G., Lamarque, J.F., Pfister, G.G., Fillmore, D., Granier, C., Guenther, A., Kinnison, D., Laepple, T., Orlando, J., Tie, X., Tyndall, G., Wiedinmyer, C., Baughcum, S.L., Kloster, S., 2010. Description and evaluation of the Model for Ozone and Related chemical Tracers, version 4 (MOZART-4). Geosci. Model Dev. 3, 43-67.

Environmental Protection Agency (EPA), 2006. Air Quality Criteria for Ozone and Related Photochemical Oxidants (2006 Final). in: Agency, U.S.E.P. (Ed.), Washington, DC.

770 EPA, 2013. Integrated Science Assessment of Ozone and Related Photochemical Oxidants
771 (Final Report). in: Agency, U.S.E.P. (Ed.), Washington, DC.
772 Federal Register, 2007. Environmental Protection Agency, 40 CFR Parts 50 and 51,Treatment
773 of Data Influenced by Exceptional Events; Final Rule, pp. 13560-13581.
774 Federal Register, 2010. National Ambient Air Quality Standards for Ozone, pp. 2938-3052.
775 Fiore, A.M., Dentener, F.J., Wild, O., Cuvelier, C., Schultz, M.G., Hess, P., Textor, C., Schulz,
776 M., Doherty, R.M., Horowitz, L.W., MacKenzie, I.A., Sanderson, M.G., Shindell, D.T.,
777 Stevenson, D.S., Szopa, S., Van Dingenen, R., Zeng, G., Atherton, C., Bergmann, D., Bey, I.,
778 Carmichael, G., Collins, W.J., Duncan, B.N., Faluvegi, G., Folberth, G., Gauss, M., Gong, S.,
779 Hauglustaine, D., Holloway, T., Isaksen, I.S.A., Jacob, D.J., Jonson, J.E., Kaminski, J.W.,
780 Keating, T.J., Lupu, A., Marmer, E., Montanaro, V., Park, R.J., Pitari, G., Pringle, K.J., Pyle,
781 J.A., Schroeder, S., Vivanco, M.G., Wind, P., Wojcik, G., Wu, S., Zuber, A., 2009.
782 Multimodel estimates of intercontinental source-receptor relationships for ozone pollution. J.
783 Geophys. Res. 114, D04301.
784 Fiore, A.M., Horowitz, L.W., Purves, D.W., Levy, H., II, Evans, M.J., Wang, Y., Li, Q.,
785 Yantosca, R.M., 2005. Evaluating the contribution of changes in isoprene emissions to
786 surface ozone trends over the eastern United States. J. Geophys. Res. 110, D12303.
787 Fiore, A.M., Jacob, D.J., Liu, H., Yantosca, R.M., Fairlie, T.D., Li, Q., 2003. Variability in
788 surface ozone background over the United States: Implications for air quality policy. J.
789 Geophys. Res. 108, 4787.
790 Fiore, A.M., Jacob, D.J., Bey, I., Yantosca, R.M., Field, B.D., Fusco, A.C., Wilkinson, J.G.,
791 2002. Background ozone over the United States in summer: Origin, trend, and contribution to
792 pollution episodes. J. Geophys. Res. 107, 4275.
793 Guenther, A., Karl, T., Harley, P., Wiedinmyer, C., Palmer, P., Geron, C., 2006. Estimates of
794 global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and
795 Aerosols from Nature). Atmospheric Chemistry and Physics 6, 3181-3210.
796 Hilsenrath, E., Chance, K., 2013. NASA Ups the TEMPO on Monitoring Air Pollution. The
797 Earth Observer 25, 10-15,35.
798 Horowitz, L.W., Fiore, A.M., Milly, G.P., Cohen, R.C., Perring, A., Wooldridge, P.J., Hess,
799 P.G., Emmons, L.K., Lamarque, J.-F., 2007. Observational constraints on the chemistry of
800 isoprene nitrates over the eastern United States. J. Geophys. Res. 112, D12S08.
801 Horowitz, L.W., Walters, S., Mauzerall, D.L., Emmons, L.K., Rasch, P.J., Granier, C., Tie, X.,
802 Lamarque, J.-F., Schultz, M.G., Tyndall, G.S., Orlando, J.J., Brasseur, G.P., 2003. A global
803 simulation of tropospheric ozone and related tracers: Description and evaluation of
804 MOZART, version 2. J. Geophys. Res. 108, 4784.
805 Hudman, R.C., Jacob, D.J., Turquety, S., Leibensperger, E.M., Murray, L.T., Wu, S., Gilliland,
806 A.B., Avery, M., Bertram, T.H., Brune, W., Cohen, R.C., Dibb, J.E., Flocke, F.M., Fried, A.,
807 Holloway, J., Neuman, J.A., Orville, R., Perring, A., Ren, X., Sachse, G.W., Singh, H.B.,
808 Swanson, A., Wooldridge, P.J., 2007. Surface and lightning sources of nitrogen oxides over
809 the United States: Magnitudes, chemical evolution, and outflow. J. Geophys. Res. 112,
810 D12S05.
811 Jacob, D.J., Horowitz, L.W., Munger, J.W., Heikes, B.G., Dickerson, R.R., Artz, R.S., Keene,
812 W.C., 1995. Seasonal transition from NOx- to hydrocarbon-limited conditions for ozone
813 production over the eastern United States in September. J. Geophys. Res. 100, 9315-9324.
814 Jaffe, D., 2011. Relationship between Surface and Free Tropospheric Ozone in the Western U.S.
815 Environmental Science & Technology 45, 432-438.

816 Jaffe, D.A., Wigder, N.L., 2012. Ozone production from wildfires: A critical review.
817 Atmospheric Environment 51, 1-10.
818 Kalnay, E., Kanamitsu, M., Kistler, R., Collins, W., Deaven, D., Gandin, L., Iredell, M., Saha,
819 S., White, G., Woollen, J., Zhu, Y., Leetmaa, A., Reynolds, R., Chelliah, M., Ebisuzaki, W.,
820 Higgins, W., Janowiak, J., Mo, K.C., Ropelewski, C., Wang, J., Jenne, R., Joseph, D., 1996.
821 The NCEP/NCAR 40-Year Reanalysis Project. Bulletin of the American Meteorological
822 Society 77, 437-471.
823 Kaynak, B., Hu, Y., Martin, R.V., Russell, A.G., Choi, Y., Wang, Y., 2008. The effect of
824 lightning NOx production on surface ozone in the continental United States. Atmos. Chem.
825 Phys. 8, 5151-5159.
826 Lamarque, J.-F., Kyle, G., Meinshausen, M., Riahi, K., Smith, S., van Vuuren, D., Conley, A.,
827 Vitt, F., 2011. Global and regional evolution of short-lived radiatively-active gases and
828 aerosols in the Representative Concentration Pathways. Climatic Change 109, 1-22.
829 Lamarque, J.F., Bond, T.C., Eyring, V., Granier, C., Heil, A., Klimont, Z., Lee, D., Liousse, C.,
830 Mieville, A., Owen, B., Schultz, M.G., Shindell, D., Smith, S.J., Stehfest, E., Van Aardenne,
831 J., Cooper, O.R., Kainuma, M., Mahowald, N., McConnell, J.R., Naik, V., Riahi, K., van
832 Vuuren, D.P., 2010. Historical (1850-2000) gridded anthropogenic and biomass burning
833 emissions of reactive gases and aerosols: methodology and application. Atmos. Chem. Phys.
834 10, 7017-7039.
835 Langford, A.O., Aikin, K.C., Eubank, C.S., Williams, E.J., 2009. Stratospheric contribution to
836 high surface ozone in Colorado during springtime. Geophys. Res. Lett. 36, L12801.
837 Lin, M., Fiore, A.M., Cooper, O.R., Horowitz, L.W., Langford, A.O., Levy, H., Johnson, B.J.,
838 Naik, V., Oltmans, S.J., Senff, C.J., 2012a. Springtime high surface ozone events over the
839 western United States: Quantifying the role of stratospheric intrusions. Journal of
840 Geophysical Research: Atmospheres 117, D00V22.
841 Lin, M., Fiore, A.M., Horowitz, L.W., Cooper, O.R., Naik, V., Holloway, J., Johnson, B.J.,
842 Middlebrook, A.M., Oltmans, S.J., Pollack, I.B., Ryerson, T.B., Warner, J.X., Wiedinmyer,
843 C., Wilson, J., Wyman, B., 2012b. Transport of Asian ozone pollution into surface air over
844 the western United States in spring. J. Geophys. Res. 117, D00V07.
845 Lin, M., Horowitz, L.W., Oltmans, S.J., Fiore, A.M., Fan, S., 2014. Tropospheric ozone trends
846 at Mauna Loa Observatory tied to decadal climate variability. Nature Geosci 7, 136-143.
847 Liu, X., Bhartia, P.K., Chance, K., Spurr, R.J.D., Kurosu, T.P., 2010. Ozone profile retrievals
848 from the Ozone Monitoring Instrument. Atmos. Chem. Phys. 10, 2521-2537.
849 Mao, J., Paulot, F., Jacob, D.J., Cohen, R.C., Crounse, J.D., Wennberg, P.O., Keller, C.A.,
850 Hudman, R.C., Barkley, M.P., Horowitz, L.W., 2013. Ozone and organic nitrates over the
851 eastern United States: Sensitivity to isoprene chemistry. Journal of Geophysical Research:
852 Atmospheres, 118, 1–13, doi:10.1002/jgrd.50817.
853 Martin, R.V., Sauvage, B., Folkins, I., Sioris, C.E., Boone, C., Bernath, P., Ziemke, J., 2007.
854 Space-based constraints on the production of nitric oxide by lightning. J. Geophys. Res. 112,
855 D09309.
856 McDonald-Buller, E.C., Allen, D.T., Brown, N., Jacob, D.J., Jaffe, D., Kolb, C.E., Lefohn, A.S.,
857 Oltmans, S., Parrish, D.D., Yarwood, G., Zhang, L., 2011. Establishing Policy Relevant
858 Background (PRB) Ozone Concentrations in the United States. Environmental Science &
859 Technology 45, 9484-9497.

McKeen, S.A., Wotawa, G., Parrish, D.D., Holloway, J.S., Buhr, M.P., Hübler, G., Fehsenfeld, F.C., Meagher, J.F., 2002. Ozone production from Canadian wildfires during June and July of 1995. Journal of Geophysical Research: Atmospheres 107, ACH 7-1-ACH 7-25.

McLinden, C.A., Olsen, S.C., Hannegan, B., Wild, O., Prather, M.J., Sundet, J., 2000. Stratospheric ozone in 3-D models: A simple chemistry and the cross-tropopause flux. J. Geophys. Res. 105, 14653-14665.

Mickley, L.J., Jacob, D.J., Rind, D., 2001. Uncertainty in preindustrial abundance of tropospheric ozone: Implications for radiative forcing calculations. J. Geophys. Res. 106, 3389-3399.

Mueller, S.F., Mallard, J.W., 2011. Contributions of Natural Emissions to Ozone and PM2.5 as Simulated by the Community Multiscale Air Quality (CMAQ) Model. Environmental Science & Technology 45, 4817-4823.

Murray, L.T., Jacob, D.J., Logan, J.A., Hudman, R.C., Koshak, W.J., 2012. Optimized regional and interannual variability of lightning in a global chemical transport model constrained by LIS/OTD satellite data. Journal of Geophysical Research: Atmospheres 117, D20307.

Naik, V., Horowitz, L.W., Fiore, A.M., Ginoux, P., Mao, J., Aghedo, A.M., Levy, H., 2013. Impact of preindustrial to present-day changes in short-lived pollutant emissions on atmospheric composition and climate forcing. Journal of Geophysical Research: Atmospheres 118, 8086-8110.

Olivier, J.G.J., Berdowski, J.J.M., 2001. Global emissions sources and sinks, in: The Climate System, in: Berdowski, J., Guicherit, R., Heij, B.J. (Eds.), Lisse, The Netherlands, pp. 33-78.

Park, R.J., Jacob, D.J., Field, B.D., Yantosca, R.M., Chin, M., 2004. Natural and transboundary pollution influences on sulfate-nitrate-ammonium aerosols in the United States: Implications for policy. Journal of Geophysical Research: Atmospheres 109, D15204.

Parrella, J.P., Jacob, D.J., Liang, Q., Zhang, Y., Mickley, L.J., Miller, B., Evans, M.J., Yang, X., Pyle, J.A., Theys, N., Van Roozendael, M., 2012. Tropospheric bromine chemistry: implications for present and pre-industrial ozone and mercury. Atmos. Chem. Phys. 12, 6723-6740.

Perring, A.E., Bertram, T.H., Wooldridge, P.J., Fried, A., Heikes, B.G., Dibb, J., Crounse, J.D., Wennberg, P.O., Blake, N.J., Blake, D.R., Brune, W.H., Singh, H.B., Cohen, R.C., 2009. Airborne observations of total RONO2: New contraints on the yield and lifetime of isoprene nitrates. Atmospheric Chemistry and Physics 9, 1451-1463.

Prather, M.J., Zhu, X., Tang, Q., Hsu, J., Neu, J.L., 2011. An atmospheric chemist in search of the tropopause. J. Geophys. Res. 116, D04306.

Price, C., Rind, D., 1992. A simple lightning parameterization for calculating global lightning distributions. Journal of Geophysical Research: Atmospheres 97, 9919-9933.

Rasmussen, D.J., Fiore, A.M., Naik, V., Horowitz, L.W., McGinnis, S.J., Schultz, M.G., 2012. Surface ozone-temperature relationships in the eastern US: A monthly climatology for evaluating chemistry-climate models. Atmospheric Environment 47, 142-153.

Reid, N., Yap, D., Bloxam, R., 2008. The potential role of background ozone on current and emerging air issues: An overview. Air Qual Atmos Health 1, 19-29.

Reidmiller, D.R., Fiore, A.M., Jaffe, D.A., Bergmann, D., Cuvelier, C., Dentener, F.J., Duncan, B.N., Folberth, G., Gauss, M., Gong, S., Hess, P., Jonson, J.E., Keating, T., Lupu, A., Marmer, E., Park, R., Schultz, M.G., Shindell, D.T., Szopa, S., Vivanco, M.G., Wild, O., Zuber, A., 2009. The influence of foreign vs. North American emissions on surface ozone in the US. Atmos. Chem. Phys. 9, 5027-5042.

Singh, H.B., Anderson, B.E., Brune, W.H., Cai, C., Cohen, R.C., Crawford, J.H., Cubison, M.J., Czech, E.P., Emmons, L., Fuelberg, H.E., Huey, G., Jacob, D.J., Jimenez, J.L., Kaduwela, A., Kondo, Y., Mao, J., Olson, J.R., Sachse, G.W., Vay, S.A., Weinheimer, A., Wennberg, P.O., Wisthaler, A., 2010. Pollution influences on atmospheric composition and chemistry at high northern latitudes: Boreal and California forest fire emissions. Atmospheric Environment 44, 4553-4564.

TFHTAP, 2010. HEMISPHERIC TRANSPORT OF AIR POLLUTION 2010 PART A: OZONE AND PARTICULATE MATTER, Air Pollution Studies No. 17, in: Dentener, F., Keating, T., Akmoto, H. (Eds.), AIR POLLUTION STUDIES No. 17. UNITED NATIONS, New York.

van der Werf, G.R., Randerson, J.T., Giglio, L., Collatz, G.J., Kasibhatla, P.S., Arellano Jr, A.F., 2006. Interannual variability in global biomass burning emissions from 1997 to 2004. Atmos. Chem. Phys. 6, 3423-3441.

Vingarzan, R., 2004. A review of surface ozone background levels and trends. Atmospheric Environment 38, 3431-3442.

Wang, H., Jacob, D.J., Le Sager, P., Streets, D.G., Park, R.J., Gilliland, A.B., van Donkelaar, A., 2009. Surface ozone background in the United States: Canadian and Mexican pollution influences. Atmospheric Environment 43, 1310-1319.

Wang, Y.X., McElroy, M.B., Jacob, D.J., Yantosca, R.M., 2004. A nested grid formulation for chemical transport over Asia: Applications to CO. Journal of Geophysical Research: Atmospheres 109, D22307.

Wild, O., Fiore, A.M., Shindell, D.T., Doherty, R.M., Collins, W.J., Dentener, F.J., Schultz, M.G., Gong, S., MacKenzie, I.A., Zeng, G., Hess, P., Duncan, B.N., Bergmann, D.J., Szopa, S., Jonson, J.E., Keating, T.J., Zuber, A., 2012. Modelling future changes in surface ozone: a parameterized approach. Atmos. Chem. Phys. 12, 2037-2054.

Wu, S., Mickley, L.J., Jacob, D.J., Rind, D., Streets, D.G., 2008. Effects of 2000-2050 changes in climate and emissions on global tropospheric ozone and the policy-relevant background surface ozone in the United States. J. Geophys. Res. 113, D18312.

Zhang, L., Jacob, D.J., Boersma, K.F., Jaffe, D.A., Olson, J.R., Bowman, K.W., Worden, J.R., Thompson, A.M., Avery, M.A., Cohen, R.C., Dibb, J.E., Flocke, F.M., Fuelberg, H.E., Huey, L.G., McMillian, W.W., Singh, H.B., Weinheimer, A.J., 2008. Transpacific transport of ozone pollution and the effect of recent Asian emission increases on air quality in North America: an integrated analysis using satellite, aircraft, ozonesonde, and surface observations. Atmospheric Chemistry and Physics 8, 6117-6136.

Zhang, L., Jacob, D.J., Downey, N.V., Wood, D.A., Blewitt, D., Carouge, C.C., van Donkelaar, A., Jones, D.B.A., Murray, L.T., Wang, Y., 2011. Improved estimate of the policy-relevant background ozone in the United States using the GEOS-Chem global model with ½°x⅔° horizontal resolution over North America. Atmospheric Environment 45, 6769-6776.

Zhang, L., Jacob, D.J., Liu, X., Logan, J.A., Chance, K., Eldering, A., Bojkov, B.R., 2010. Intercomparison methods for satellite measurements of atmospheric composition: application to tropospheric ozone from TES and OMI. Atmos. Chem. Phys. 10, 4725-4739.

Zhang, L., Jacob, D.J., Yue, X., Downey, N.V., Wood, D.A., Blewitt, D., 2014. Sources contributing to background surface ozone in the US Intermountain West. Atmos. Chem. Phys., 14, 5295-5309, doi:10.5194/acp-14-5295-2014.

952
953 **FIGURE CAPTIONS**
954
955 **Figure 1**. Mean MDA8 values of North American Background (NAB) in the lowest model layer
956 for the GFDL AM3 (left; ~2°x2° horizontal resolution) and GEOS-Chem (right; ½°x⅔°)
957 simulations for spring (MAM; top row) and summer (JJA; bottom row) of 2006. NAB is
958 estimated with simulations in which North American anthropogenic emissions are set to zero.
959 See Table 1 for model configurations.
960
961 **Figure 2.** Fourth highest MDA8 NAB O_3 between March 1 and August 31 2006 in the lowest
962 model layer (top) and day of occurrence (bottom) for GFDL AM3 (left; ~2°x2° horizontal
963 resolution) and GEOS-Chem (right; ½°x⅔°) simulations.
964
965 **Figure 3**. Springtime (March-April-May average) mid-tropospheric O_3 as retrieved (second row)
966 from the TES (left column) and OMI (right column) satellite instruments and as simulated with
967 the GFDL AM3 (top row) and GEOS-Chem (third row) with the appropriate averaging kernels
968 applied to daily average O_3 fields archived from the base model simulations. Grey boxes denote
969 locations where no coincident TES and OMI data points meet the retrieval quality criteria. The
970 simulations evaluated here are at coarse horizontal resolution (~2°x2°) in both models. The
971 2005-2007 annual average bias of the satellite products (5.7 ppb for TES; 3.1 ppb for OMI)
972 relative to ozone sondes between 20-60°N determined by Zhang et al. (2010) has been removed
973 from the satellite products. The third and fourth rows show the difference of the simulated mid-
974 tropospheric O_3 with each satellite product; grey boxes denote places where the OMI and TES
975 retrievals disagree by over 10 ppb.
976
977 **Figure 4.** As in Figure 3 but for summer (June-July-August).
978
979 **Figure 5.** Monthly mean MDA8 O_3 in surface air for March through August of 2006 in
980 CASTNet observations (black) and the base simulations (thick lines) for the GEOS-Chem (GC
981 ½°x⅔° horizontal resolution; blue) and GFDL AM3 (~2°x2° horizontal resolution red)
982 simulations sampled at the CASTNet sites (using bilinear interpolation of the nearest four grid
983 cells and sampling only on days with valid measurements) over the CONtiguous United States
984 (CONUS) at altitudes a) above 1.5km excluding California to focus on the InterMountain West
985 region and b) below 1.5 km in altitude. Also shown are NAB estimates (thin lines) with GC
986 (blue) and AM3 (red). The grey band delineates the one standard deviation range about the
987 observed regional mean monthly values.
988
989 **Figure 6**. Probability density curves calculated via kernel (Gaussian) density estimation with a
990 bandwidth of 2 ppbv from surface MDA8 O_3 data over the CONtiguous United States (CONUS)
991 during spring (top) and summer (bottom) and at high (left, excluding California sites) and low
992 (right) elevation CASTNet sites: observed (black) and GFDL AM3 (red) and GEOS-Chem at
993 low (green; 2°x2.5°) and high (blue; ½° x ⅔°) horizontal resolution models sampled at the
994 CASTNet sites for total (solid lines) and NAB (dashed lines) O_3.
995
996 **Figure 7.** Correlation coefficients (r) of observed MDA8 O_3 in surface air with that simulated
997 by the GFDL AM3 (left) and GEOS-Chem (right) models during spring (MAM; top row) and

summer (JJA; bottom row). Black circles mark the sites used in Figure 9 and discussed in Section 3.5. The color bar saturates to emphasize regional patterns.

Figure 8. As in Figure 7 but for simulated NAB versus simulated total O_3 in surface air.

Figure 9a. MDA8 O_3 in surface air observed (black) at four CASTNet sites for March through August 2006 (highlighted by black circles in Figures 7 and 8), and simulated with the GEOS-Chem (blue thick lines) and GFDL AM3 (red thick lines) models. Also shown are NAB estimates with GEOS-Chem (blue thin lines) and GFDL AM3 (red thin lines) and an estimate of stratospheric O_3 influence in the AM3 model (orange lines) following the method described in Lin et al. (2012b) from a simulation described in Lin et al. (2014). Statistics in the upper left corner of each panel are for the entire March through August period: the mean and standard deviation (in parentheses) of total surface O_3 as observed (black) and simulated with GEOS-Chem (blue) and GFDL AM3 (red); coefficients of determination for each model versus the observations (see also correlation coefficients in Figures 7 and 8 for separate correlations during spring and summer); the mean and standard deviation of the MDA8 NAB O_3 simulated with each model. The stratospheric O_3 tracer (orange), which can be interpreted more broadly as a diagnostic of mixing between the free troposphere and the surface, correlates strongly at all sites (r^2 ranges from 0.72 to 0.86) with the NAB (red). Yellow highlighted days at the western U.S. sites and PA site are discussed in Section 3.5. The Gothic, CO panel is adapted from Figure 3-75 of the U.S. EPA Integrated Science Assessment for O_3 (EPA, 2013).

Figure 9b. As in Figure 9a but for mid-April through mid-June at the two high-altitude Western U.S. sites in order to focus in on the stratospheric intrusion event discussed in Section 3.5.

Figure 10. Climatological (1981-2007) average (left) and standard deviation (right) of spring (top) and summer (middle) seasonal mean MDA8 NAB O_3, and of the fourth highest value between March 1 and August 31 (bottom) as estimated with the GFDL AM3 model simulation in which North American anthropogenic emissions are set to zero.

Figure 11. GFDL AM3 simulated daily maximum 8-hour (MDA8) surface O_3 versus observed values (black) and AM3 NAB statistics (green) at 11 Intermountain Western U.S. CASTNet sites above 1.5 km altitude for a "low-O_3" year (left column) and "high-O_3" year (right column) to provide context for the year 2006 (middle column) during spring (top panel) and summer (bottom panel), following the approach of Wang et al. (2009; see their Figure 5). The 1:1 line (solid black) and a 60 ppb threshold (dashed line) are shown. Box and whisker plots show the median (triangle), 25^{th}-75^{th} range (box) and minimum and maximum NAB values (vertical lines) for 10 ppb bins of observed O_3 values. The "low" and "high" years are selected from Figure 6 of Jaffe (2011).

Table 1. Model estimates for North American Background (NAB) ozone using the current U.S. EPA definition (North American anthropogenic emissions set to zero) and for specific components of NAB (ppb)

Study **Model**	**Study period;** **Metric**	**NAB**	**Components**
Fiore et al. (2003) GC (2°x2.5°)	Mar-Oct 2001; 1-5pm mean	Typically 15-35; up to 40-50 (highest in spring and WUS)	*Natural*: 18-23 (NW), 18-27 (SW), 13-20 (NE), 15-21 (SE) *Strat*: always < 10 *CH4+intercontinental transport (ICT)*: 5-12
Wu et al. (2008) ; GC with winds from GISS GCM (4°x5°)	2000 climatology; 1-5pm mean	12-30 (summer); 22-40 (April); highest in WUS	*Natural*: 10-15 (EUS, summer); 15-25 (WUS, summer)
Wang et al. (2009) GC (1°x1°)	summer 2001; MDA8	26±8 *USB*: 30±8; up to 33 ppb during events	
Zhang et al. (2011) GC (½°x⅔°)	Mar-Aug 2006-2008; MDA8	39-44 (spring); 35-45(summer); low-alt 27±8; high-alt 40±7; 51-59 (4th highest)	*Natura*l: 18±6 (low-alt); 27±6 (high-alt); 34-45 (4th highest). *CH4+ICT*: 13-16 (spring) 11-13 (summer); 13 (high alt); 9 (low alt)
Emery et al. (2012) CAMx (12 km^2), GC boundary conditions	Mar-Aug 2006; MDA8	25-50 ppb (20-45 in GC); 35-100 (4th highest; 65 max without fires; 55 max in GC)	*Fires*: 10-50 ppb (events)
Lin et al. (2012a) GFDL AM3 (~50km^2)	Apr-Jun 2010; MDA8	15 WUS high-alt sites: 50±11 (mean); 55±11 (days when obs exceed 60 ppb)	*Strat*: *15 WUS high-alt sites*: 22±12 (mean); 15-25 for obs O_3 @ 60-70; 17-40 for obs O_3 @ 70-85 *Median, bias-corrected*: 10-22 (W); 8-13(NE); 3-8 (SE) *Max, bias-corrected*: 35-55 (W); 30-45 (EUS)
McKeen et al. (2002); 3D regional model (60 km^2)	Jun-Jul 1995; 1-4pm mean		*Fires*: 10-30 ppb (event, Central and EUS)
Collins et al. (2003);	March 1991-		*Strat*: 5-15 ppb (highest in WUS)

STOCHEM driven by UM HadAM4 GCM)	1994 monthly mean		
Kaynak et al. (2008); CMAQ (36km^2)	Jul-Aug 2004; MDA8		*Lightning*: up to 10 ppb; 14 ppb 4th highest; < 2 ppb 71% of the time
Mueller and Mallard (2011); CMAQ, GC boundary conditions (36 km^2)	2002; MDA8		*Fires*: 30-50 (WUS, events) *Lightning*: 10-30 (Southern US, events)
Zhang et al. (2014); GC (½°x⅔°)	Mar-Aug 2006; MDA8		*Lightning:* 6-10 ppbv (summer); *Fires*: ~20 (local events), 1-3 (WUS summer mean); *Strat*: 8-10 (WUS spring mean), up to 15 (events)

Table 2. GFDL AM3 and GEOS-Chem model configurations

Model	GFDL AM3 (Donner et al., 2011) (Rasmussen et al., 2012) (Naik et al., 2013) (Lin et al., 2012b)	GEOS-Chem http://acmg.seas.harvard.edu/geos/ (Zhang et al., 2011) (Bey et al., 2001) (Park et al., 2004)
Grid	Cubed sphere with 48x48 cell faces, approximately 2°x2° horizontal resolution. Vertical coordinate is a 48-level hybrid sigma grid, with the top level at 0.01 hPa; lowest 5 layers extend to 60, 130, 220, 330, and 470 m for surface pressure of 1013.25 hPa and scale height of 7.5km.	Continental North American nested (Wang et al., 2004) simulation at ½° latitude by ⅔° longitude using boundary conditions from boundary conditions from a 2°x2.5° global simulation. Vertical grid has 47 levels to 0.01 hPa, with lowest 5 layers centered at 70, 200, 330, 470, 600 m for a column at sea level.
Meteorology	Online, nudged to NCEP *u* and *v* (Kalnay et al., 1996). The nudging timescale is inversely proportional to pressure (Lin et al., 2012b)	Assimilated from NASA GEOS-5
Stratospheric ozone	Stratospheric chemistry and dynamics seamlessly coupled to the troposphere (Naik et al., 2013)	Linoz parameterization (McLinden et al., 2000)
Isoprene nitrate yield and fate	Observationally-constrained 8% yield with 40% NO_x recycling (Horowitz et al., 2007 and references therein)	18% yield with no NO_x recycling (permanent sink for NO_x)
Lightning NO_x distribution	Parameterized based on convective cloud top height (Price and Rind, 1992), and described in Horowitz et al. (2003); global source in 2006 is 4.9 Tg N a^{-1}; range over 1981-2007 is	Scaled to match a top-down global constraint of 6 Tg N a^{-1} (Martin et al., 2007) and spatially redistributed based on the LIS/OTD flash climatology (Murray et al., 2012) and includes a

	4.4-4.9 Tg N a^{-1}.	yield of 500 mol N $flash^{-1}$ at northern mid-latitudes and 125 mol N $flash^{-1}$ elsewhere (Hudman et al., 2007)
Anthropogenic emissions	ACCMIP (Lamarque et al., 2010) with annual interpolation after 2000 to RCP4.5 2010 value (Lamarque et al., 2011). NO_x emissions are 27 (global), 5.6 (North America[a]), 5.9 (East Asia[b]) Tg N a^{-1}. CO are 640 (global), 100 (North America), 160 (East Asia) Tg CO a^{-1}. C_3H_8[c] are 2.9 (global), 0.3 (North America), 0.5 (East Asia) Tg C a^{-1}.	EDGAR (Olivier and Berdowski, 2001) with U.S. emissions from 2005 National Emissions Inventory (NEI-05). NO_x emissions are 27.9 (global), 6.4 (North America[a]), 7.9 (East Asia[b]) Tg N a^{-1}. CO are 582 (global), 91 (North America), 192 (East Asia) Tg CO a^{-1}. C_3H_8[c] are 27 (global), 0.45 (North America), 4.1 (East Asia) Tg C a^{-1}.
Biogenic emissions	Model of Emissions of Gases and Aerosols from Nature (MEGAN) 2.1 (Guenther et al., 2006), implemented as described by Emmons et al. (2010) and Rasmussen et al. (2012) Emissions are 385 (global) and 32 (North America) Tg C a^{-1}.	MEGAN 2.0 (Guenther et al., 2006) Emissions are 404 (global) and 30 (North America) Tg C a^{-1}.
Biomass burning emissions	As for anthropogenic emissions but distributed vertically as recommended by Dentener et al. (2006). NO_x emissions are 5.3 (global), 0.18 (North America[a]), 0.14 (East Asia[b]) Tg N a^{-1}. CO are 440 (global), 18 (North America), 14 (East Asia) Tg CO a^{-1}. C_3H_8[c] are 1.5 (global), 0.09 (North America), 0.09 (East Asia) Tg C a^{-1}.	GFEDv2 year-specific monthly fires (van der Werf et al., 2006), emitted at the surface. NO_x emissions are 5.1 (global), 0.14 (North America[a]), 0.12 (East Asia[b]) Tg N a^{-1}. CO are 410 (global), 12 (North America), 11 (East Asia) Tg CO a^{-1}. C_3H_8[c] are 1.7 (global), 0.05 (North America), 0.06 (East Asia) Tg C a^{-1}

[a]North American domain is 15-55 °N and 60-125 °W

[b]East Asian domain is 15-50 °N, 95-160 °E

[c]C_3H_8 is provided as an example non-CH_4 volatile organic compound (NMVOC).

Table 3: Summary statistics of seasonal mean MDA8 total and NAB O_3 in surface air (ppb) as observed and estimated with the GFDL AM3 and GEOS-Chem (GC) models, segregated by altitude, season, and observed values. California is excluded to focus on the Intermountain West.

Season	Filter	N	OBS	AM3 Base	GC Base	AM3 NAB	GC NAB
Above 1.5 km (excluding CA)							
MAM	None	993	57±7	60±7	54±6	48±8	42±5
MAM	Obs ≥ 60	300	64±4	63±7	58±6	52±8	45±5
MAM	Obs ≥ 70	33	73±4	66±6	62±4	55±7	47±5
MAM	Obs ≥ 75	7	80±4	65±7	61±2	56±8	50±3
JJA	None	899	58±7	55±6	57±8	35±8	40±7
JJA	Obs ≥ 60	344	65±4	58±5	59±7	38±8	41±6
JJA	Obs ≥ 70	38	73±5	61±4	62±7	43±8	42±6
JJA	Obs ≥ 75	9	80±6	64±4	64±6	47±6	42±3
Below 1.5 km							
MAM	None	5769	49±11	57±8	48±8	39±8	29±7
MAM	Obs ≥ 60	969	65±6	64±8	57±8	37±8	29±7
MAM	Obs ≥ 70	175	75±6	69±8	63±10	36±10	31±8
MAM	Obs ≥ 75	58	82±6	71±10	68±12	36±11	34±9
JJA	None	5583	51±15	69±15	54±14	29±9	24±8
JJA	Obs ≥ 60	1509	69±9	76±13	63±11	30±9	25±9
JJA	Obs ≥ 70	537	78±8	77±13	67±12	30±9	27±10
JJA	Obs ≥ 75	294	83±8	76±14	69±14	31±9	28±10

Figures.

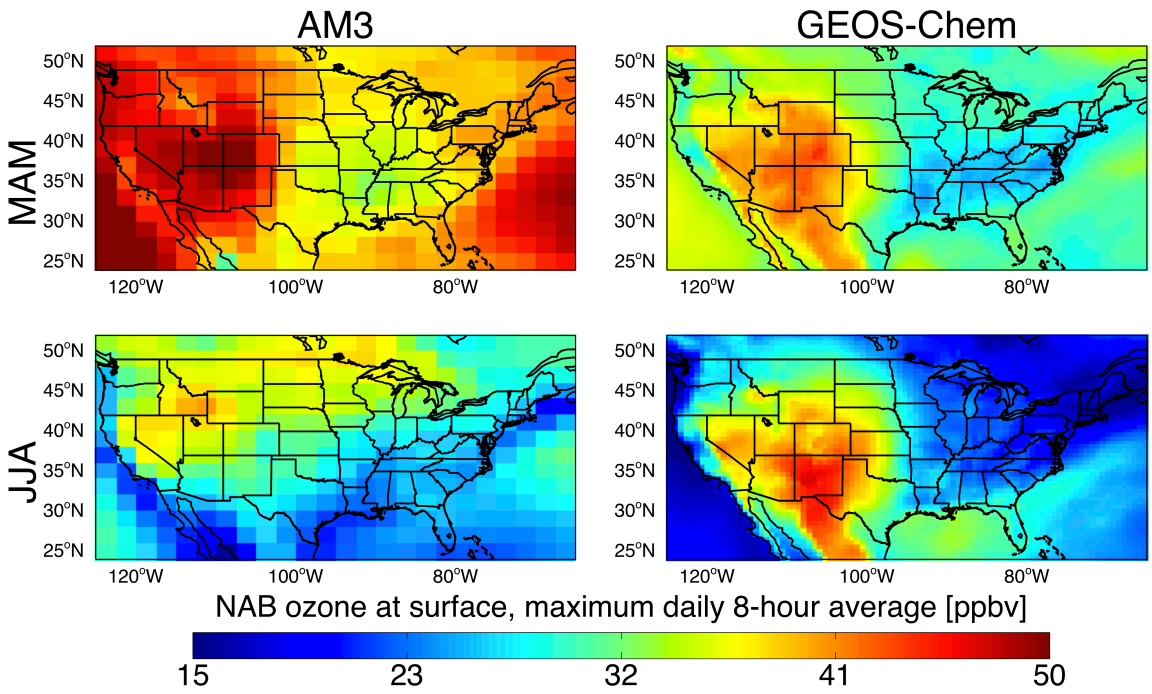

Figure 1. Mean MDA8 values of North American Background (NAB) in the lowest model layer for the GFDL AM3 (left; ~2°x2° horizontal resolution) and GEOS-Chem (right; ½°x⅔°) simulations for spring (MAM; top row) and summer (JJA; bottom row) of 2006. NAB is estimated with simulations in which North American anthropogenic emissions are set to zero. See Table 1 for model configurations.

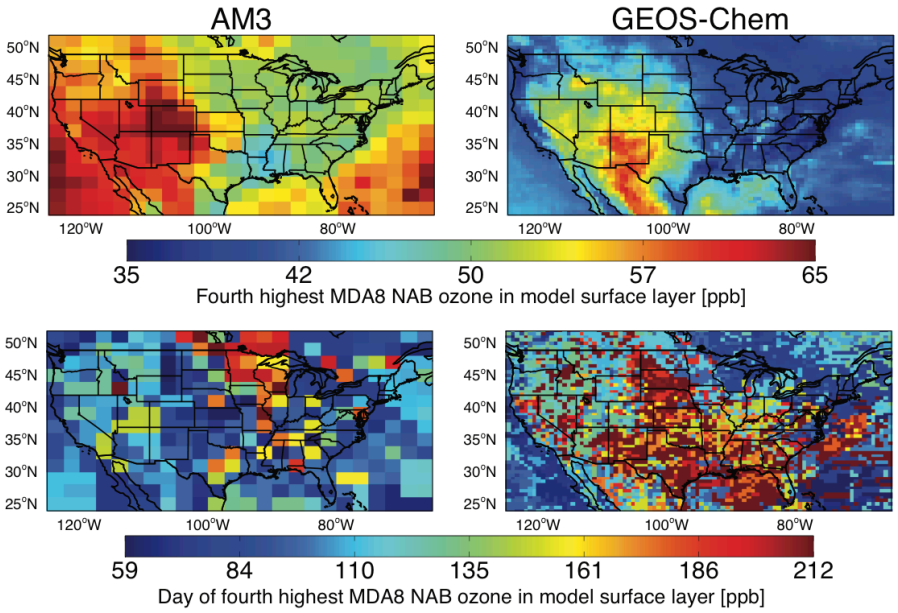

Figure 2. Fourth highest MDA8 NAB O_3 between March 1 and August 31 2006 in the lowest model layer (top) and day of occurrence (bottom) for GFDL AM3 (left; ~2°x2° horizontal resolution) and GEOS-Chem (right; ½°x⅔°) simulations.

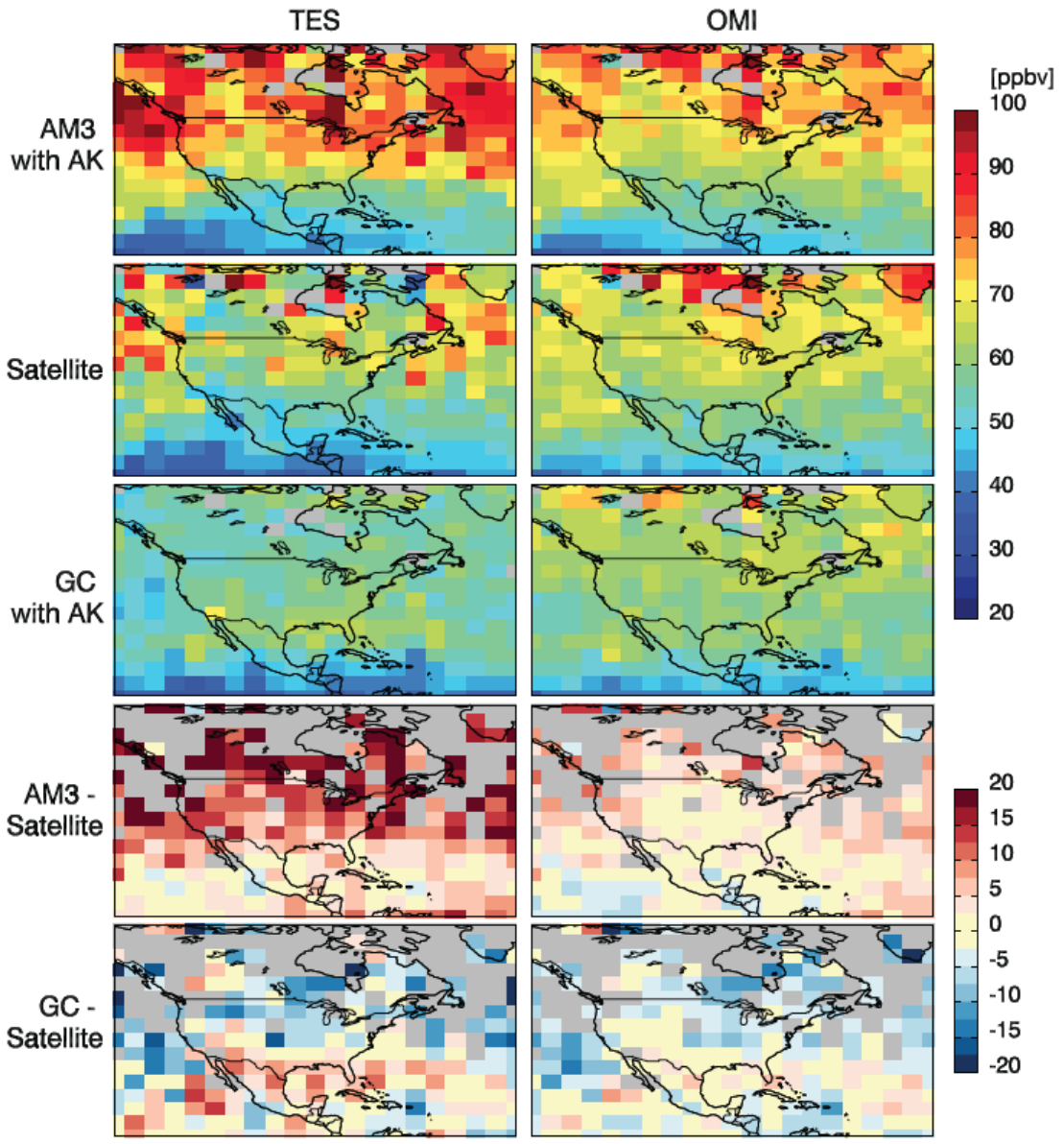

Figure 3. Springtime (March-April-May average) mid-tropospheric O_3 as retrieved (second row) from the TES (left column) and OMI (right column) satellite instruments and as simulated with the GFDL AM3 (top row) and GEOS-Chem (third row) with the appropriate averaging kernels applied to daily average O_3 fields archived from the base model simulations. Grey boxes denote locations where no coincident TES and OMI data points meet the retrieval quality criteria. The simulations evaluated here are at coarse horizontal resolution (~2°x2°) in both models. The 2005-2007 annual average bias of the satellite products (5.7 ppb for TES; 3.1 ppb for OMI) relative to ozone sondes between 20-60°N determined by Zhang et al. (2010) has been removed from the satellite products. The third and fourth rows show the difference of the simulated mid-tropospheric O_3 with each satellite product; grey boxes denote places where the OMI and TES retrievals disagree by over 10 ppb.

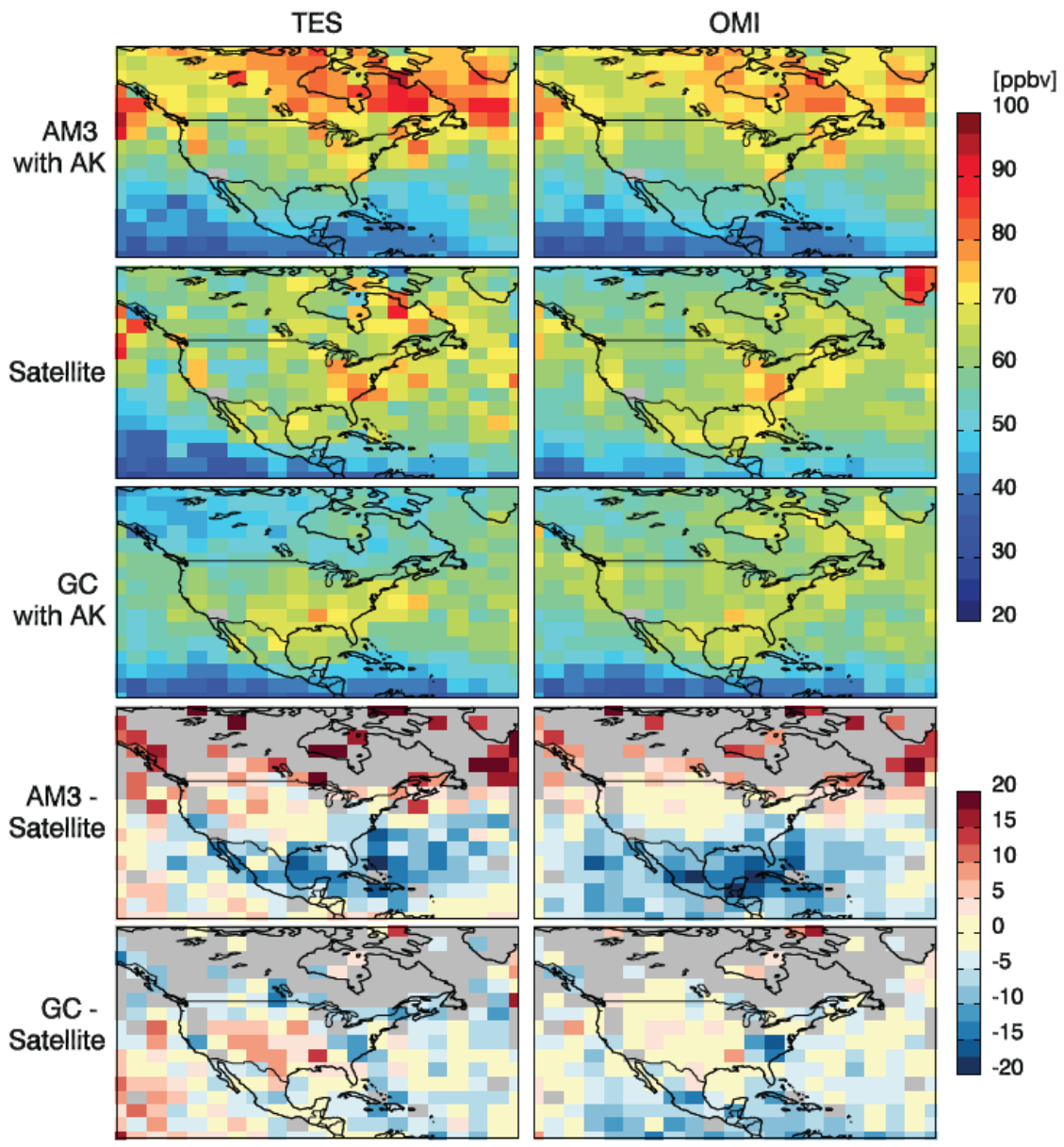

Figure 4. As in Figure 3 but for summer (June-July-August).

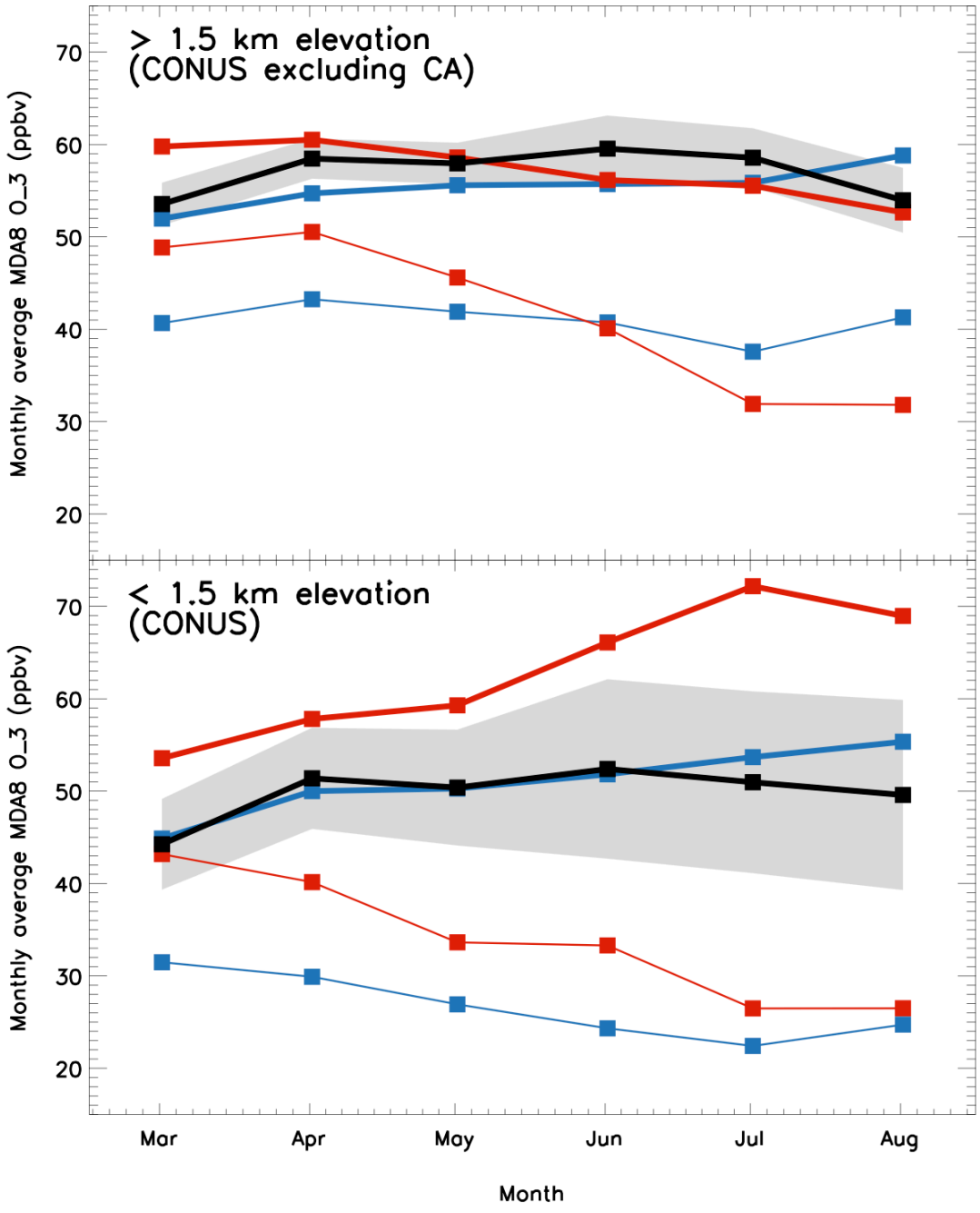

Figure 5. Monthly mean MDA8 O_3 in surface air for March through August of 2006 in CASTNet observations (black) and the base simulations (thick lines) for the GEOS-Chem (GC ½°x⅔° horizontal resolution; blue) and GFDL AM3 (~2°x2° horizontal resolution red) simulations sampled at the CASTNet sites (using bilinear interpolation of the nearest four grid cells and sampling only on days with valid measurements) over the CONtiguous United States (CONUS) at altitudes a) above 1.5km excluding California to focus on the InterMountain West region and b) below 1.5km in altitude. Also shown are NAB estimates (thin lines) with GC (blue) and AM3 (red). The grey band delineates the one standard deviation range about the observed regional mean monthly values.

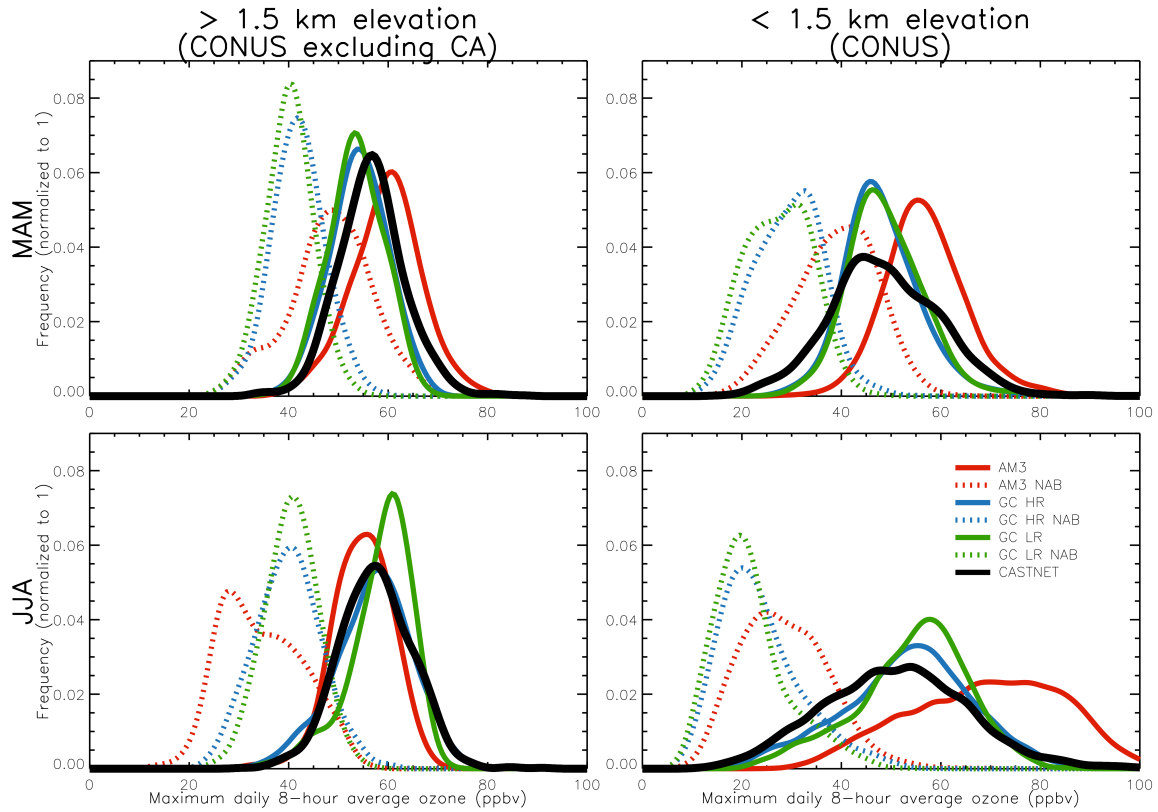

Figure 6. Probability density curves calculated via kernel (Gaussian) density estimation with a bandwidth of 2 ppbv from surface MDA8 O_3 data over the CONtiguous United States (CONUS) during spring (top) and summer (bottom) and at high (left, excluding California sites) and low (right) elevation CASTNet sites: observed (black) and GFDL AM3 (red) and GEOS-Chem at low (green; 2°x2.5°) and high (blue; ½° x ⅔°) horizontal resolution models sampled at the CASTNet sites for total (solid lines) and NAB (dashed lines) O_3.

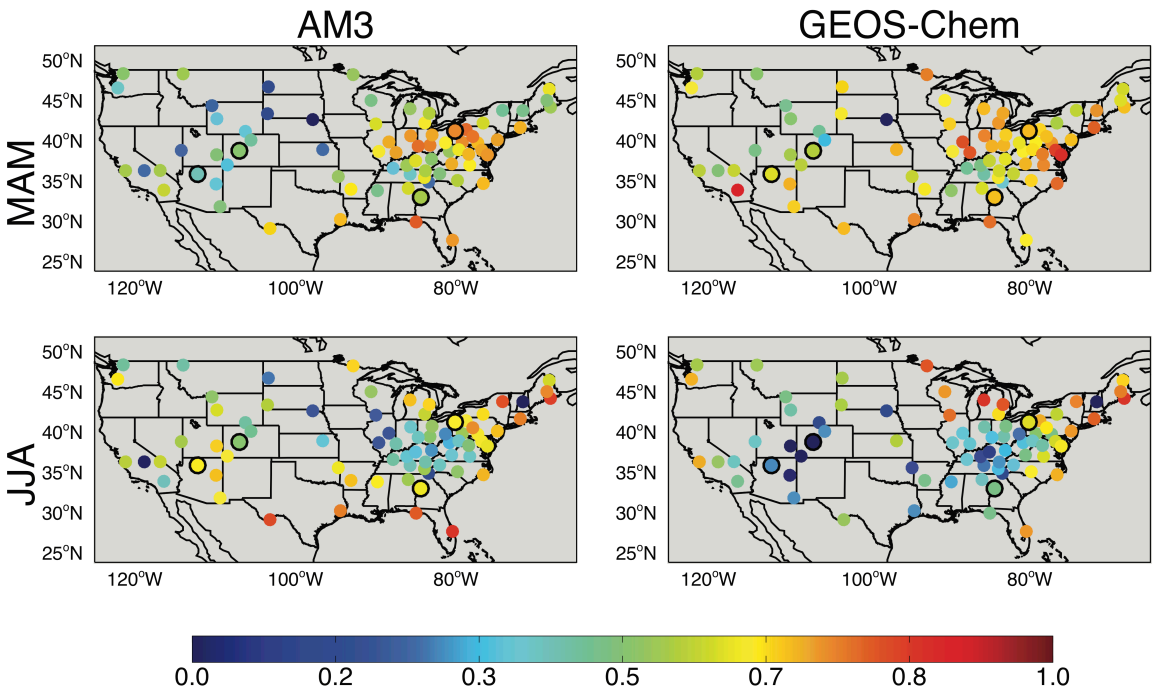

Figure 7. Correlation coefficients (r) of observed MDA8 O_3 in surface air with that simulated by the GFDL AM3 (left) and GEOS-Chem (right) models during spring (MAM; top row) and summer (JJA; bottom row). Black circles mark the sites used in Figure 9 and discussed in Section 3.5. The color bar saturates to emphasize regional patterns.

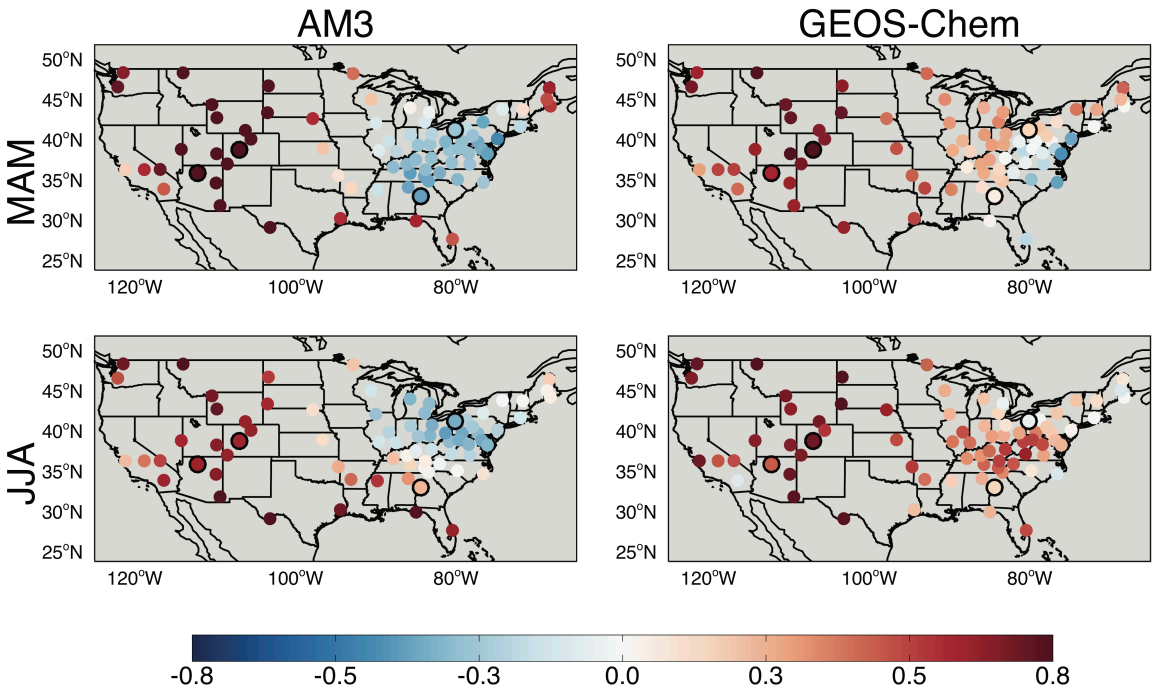

Figure 8. As in Figure 7 but for simulated NAB versus simulated total O_3 in surface air.

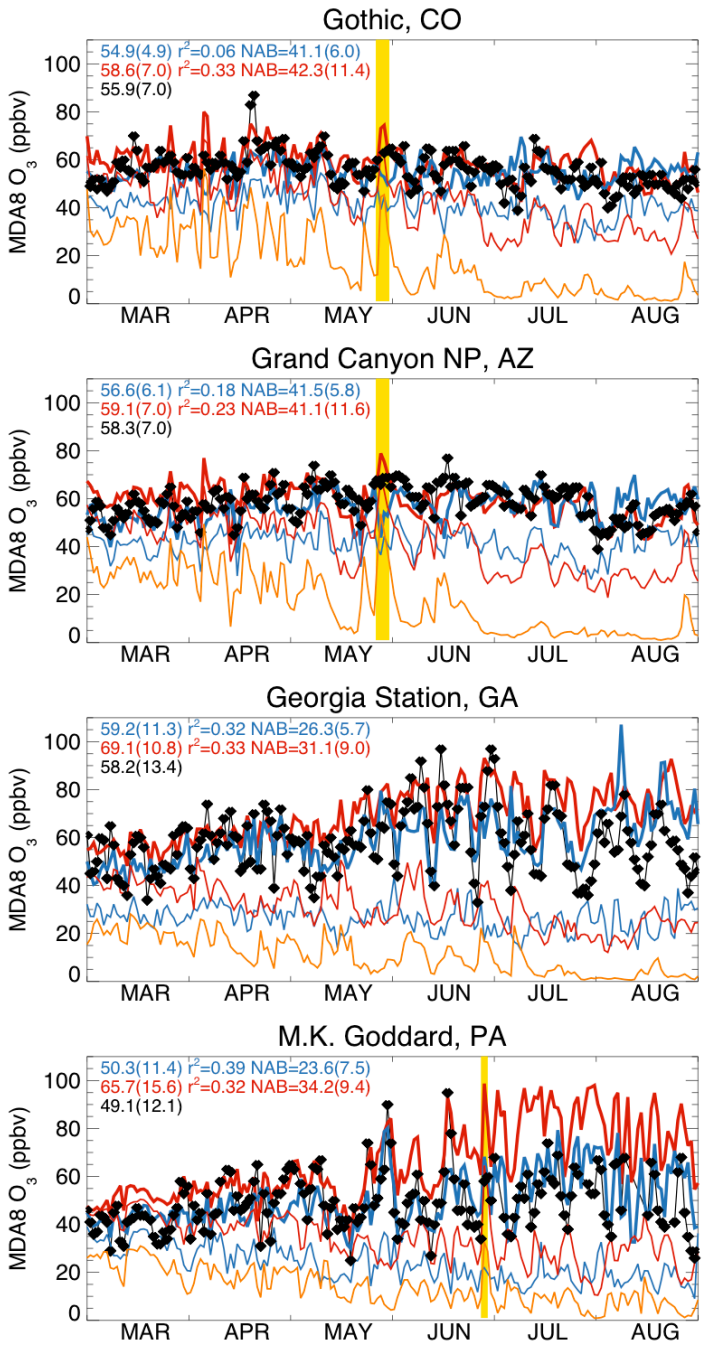

Figure 9a. MDA8 O_3 in surface air observed (black) at four CASTNet sites for March through August 2006 (highlighted by black circles in Figures 7 and 8), and simulated with the GEOS-Chem (blue thick lines) and GFDL AM3 (red thick lines) models. Also shown are NAB estimates with GEOS-Chem (blue thin lines) and GFDL AM3 (red thin lines) and an estimate of stratospheric O_3 influence in the AM3 model (orange lines) following the method described in Lin et al. (2012b) from a simulation described in Lin et al. (2014). Statistics in the upper left corner of each panel are for the entire March through August period: the mean and standard deviation (in parentheses) of total surface O_3 as observed (black) and simulated with GEOS-Chem (blue) and GFDL AM3 (red); coefficients of determination for each model versus the

observations (see also correlation coefficients in Figures 7 and 8 for separate correlations during spring and summer); the mean and standard deviation of the MDA8 NAB O_3 simulated with each model. The stratospheric O_3 tracer (orange), which can be interpreted more broadly as a diagnostic of mixing between the free troposphere and the surface, correlates strongly at all sites (r^2 ranges from 0.72 to 0.86) with the NAB (red). Yellow highlighted days at the western U.S. sites and PA site are discussed in Section 3.5. The Gothic, CO panel is adapted from Figure 3-75 of the U.S. EPA Integrated Science Assessment for O_3 (EPA, 2013).

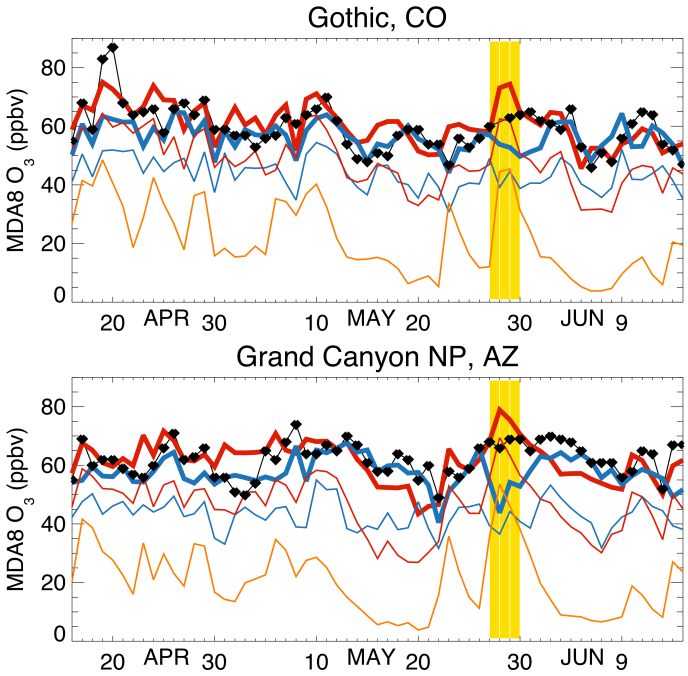

Figure 9b. As in Figure 9a but for mid-April through mid-June at the two high-altitude Western U.S. sites in order to focus in on the stratospheric intrusion event discussed in Section 3.5.

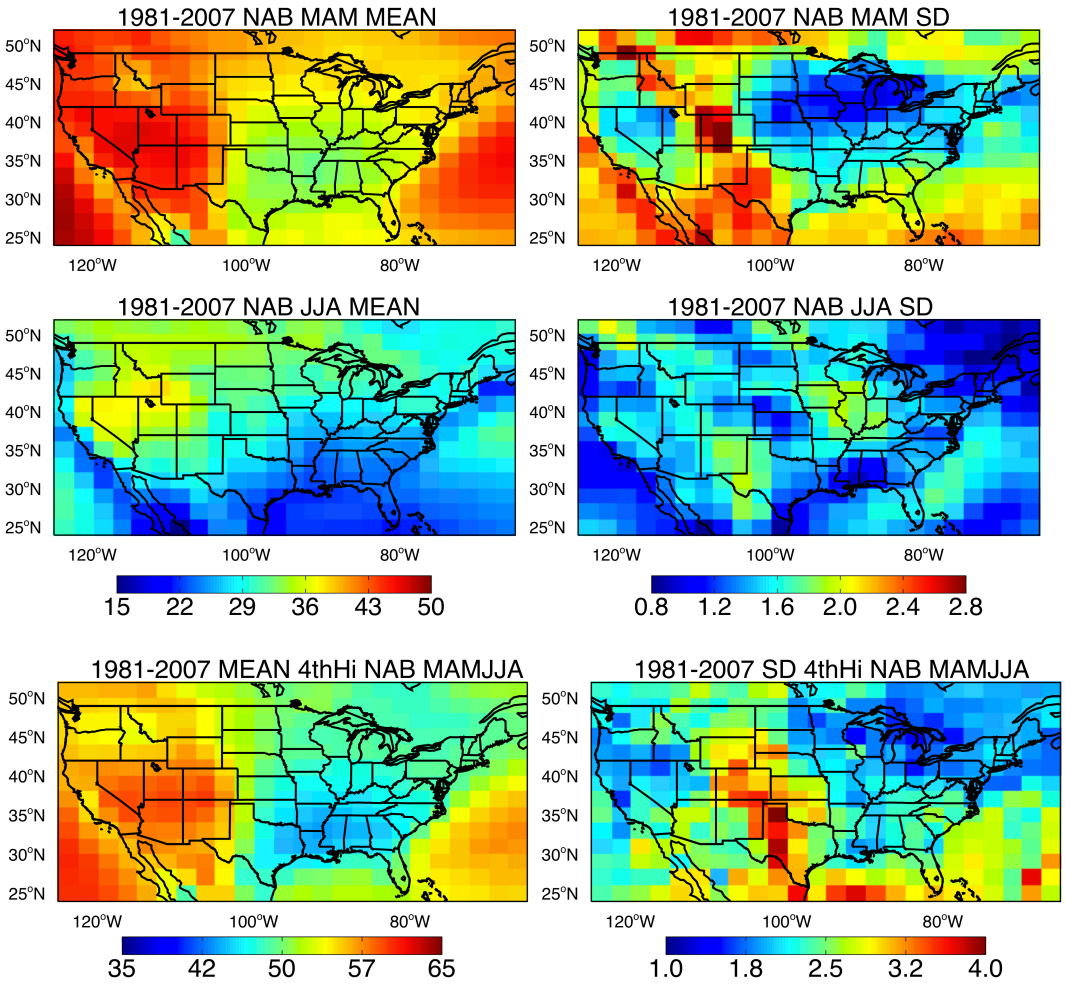

Figure 10. Climatological (1981-2007) average (left) and standard deviation (right) of spring (top) and summer (middle) seasonal mean MDA8 NAB O_3, and of the fourth highest value between March 1 and August 31 (bottom) as estimated with the GFDL AM3 model simulation in which North American anthropogenic emissions are set to zero.

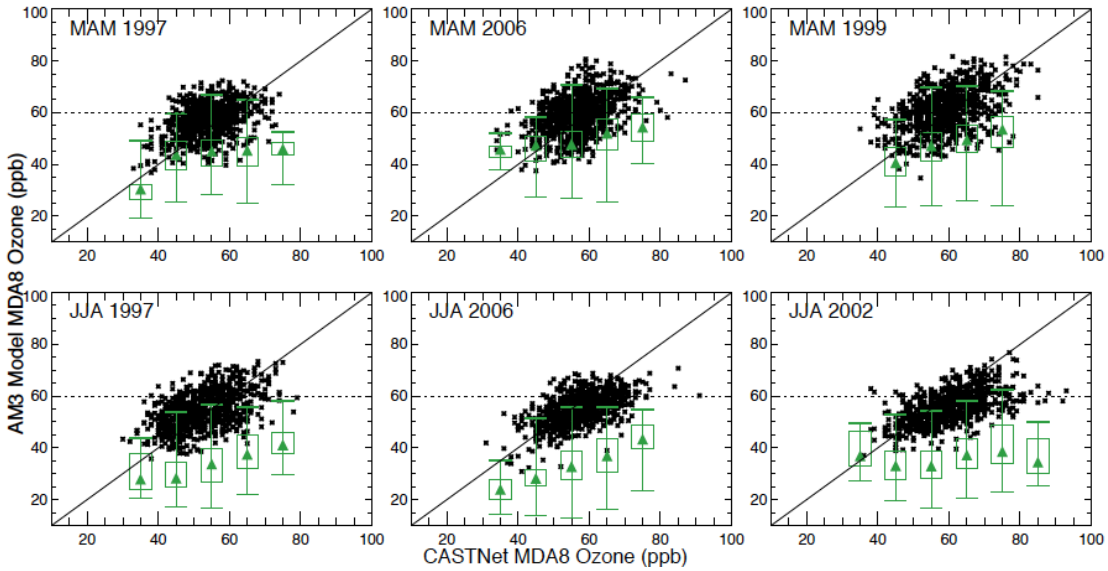

Figure 11. GFDL AM3 simulated daily maximum 8-hour (MDA8) surface O_3 versus observed values (black) and AM3 NAB statistics (green) at 11 Intermountain Western U.S. CASTNet sites above 1.5 km altitude for a "low-O_3" year (left column) and "high-O_3" year (right column) to provide context for the year 2006 (middle column) during spring (top panel) and summer (bottom panel), following the approach of Wang et al. (2009; see their Figure 5). The 1:1 line (solid black) and a 60 ppb threshold (dashed line) are shown. Box and whisker plots show the median (triangle), 25th-75th range (box) and minimum and maximum NAB values (vertical lines) for 10 ppb bins of observed O_3 values. The "low" and "high" years are selected from Figure 6 of Jaffe (2011).