DISSERTATION

SATELLITE CONSTRAINTS ON SURFACE CONCENTRATIONS OF PARTICULATE MATTER

Submitted by

Bonne Ford Hotmann

Department of Atmospheric Science

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, Colorado

Spring 2015

Doctoral Committee:

Advisor: Colette L. Heald

Sonia M. Kreidenweis
Jennifer L. Peel
Jeffrey L. Collett, Jr.

ABSTRACT

SATELLITE CONSTRAINTS ON SURFACE CONCENTRATIONS OF PARTICULATE MATTER

Because of the increasing evidence of the widespread adverse effects on human health from exposure to poor air quality and the recommendations of the World Health Organization to significantly reduce $PM_{2.5}$ in order to reduce these risks, better estimates of surface air quality globally are required. However, surface measurements useful for monitoring particulate exposure are scarce, especially in developing countries which often experience the worst air pollution. Therefore, other methods are necessary to augment estimates in regions with limited surface observations.

The prospect of using satellite observations to infer surface air quality is attractive; however, it requires knowledge of the complicated relationship between satellite-observed aerosol optical depth (AOD) and surface concentrations. This dissertation explores how satellite observations can be used in conjunction with a chemical transport model (GEOS-Chem) to better understand this relationship.

First, we investigate the seasonality in aerosols over the Southeastern United States using observations from several satellite instruments (MODIS, MISR, CALIOP) and surface network sites (IMPROVE, SEARCH, AERONET). We find that the strong summertime enhancement in satellite-observed aerosol optical depth (factor 2-3 enhancement over wintertime AOD) is not present in surface mass concentrations (25-55% summertime enhancement). *Goldstein et al*. [2009] previously attributed this seasonality in AOD to biogenic organic aerosol; however, surface observations show that organic aerosol only accounts for ~35% of $PM_{2.5}$ mass and exhibits similar seasonality to total surface $PM_{2.5}$. The GEOS-Chem model generally reproduces these surface aerosol measurements, but under represents the AOD seasonality observed by

satellites. We show that seasonal differences in water uptake cannot sufficiently explain the magnitude of AOD increase. As CALIOP profiles indicate the presence of additional aerosol in the lower troposphere (below 700 hPa), which cannot be explained by vertical mixing; we conclude that the discrepancy is due to a missing source of aerosols above the surface layer in summer.

Next, we examine the usefulness of deriving premature mortality estimates from "satellite-based" $PM_{2.5}$ concentrations. In particular, we examine how uncertainties in the model AOD-to-surface-$PM_{2.5}$ relationship, satellite retrieved AOD, and particulars of the concentration-response function can impact these mortality estimates. We find that the satellite-based estimates suggest premature mortality due to chronic $PM_{2.5}$ exposure is 2-16% higher in the U.S. and 4-13% lower in China compared to model-based estimates. However, this difference is overshadowed by the uncertainty in the methodology, which we quantify to be on order of 20% for the model-to- surface-$PM_{2.5}$ relationship, 10% for the satellite AOD and 30-60% or greater with regards to the application of concentration response functions.

Because there is a desire for acute exposure estimates, especially with regards to extreme events, we also examine how premature mortality due to acute exposure can be estimated from global models and satellite-observations. We find similar differences between model and satellite-based mortality estimates as with chronic exposure. However the range of uncertainty is much larger on these shorter timescales. This work suggests that although satellites can be useful for constraining model estimates of $PM_{2.5}$, national mortality estimates from the two methods are not significantly different. In order to improve the efficacy of satellite-based $PM_{2.5}$ mortality estimates, future work will need to focus on improving the model representation of the regional

AOD-to-surface-$PM_{2.5}$ relationship, reducing biases in satellite-retrieved AOD and advancing our understanding of personal and population-level responses to $PM_{2.5}$ exposure.

ACKNOWLEDGEMENTS

I would like to thank my advisor, Colette Heald. I have appreciated her patience, encouragement and genuine goal to train me to be a better scientist. I feel really blessed to have gotten to work with you and be privy to your insight.

Thanks to my committee members Dr. Peel, Dr. Kreidenweis, and Dr. Collett for challenging me and providing extra advice and support during my time at CSU.

I would also like to thank past and present members of the Heald group for their camaraderie, insightful questions and interest in my work. Special thanks to my former officemate, Luke Schiferl, for being my brother-in-arms as we crawled through the trenches of GEOS-Chem and to Dr. David Ridley for his honest professional and personal advice (and music).

Next I would like to thank Dr. Jeff Pierce and his group for adopting me, with special thanks to Stephen D'Andrea.

Deepest gratitude also goes to Shaila Parasher and ENS. Thank you for your patience and help in maintaining our servers. I promise to never again ask you to restart the daemons.

I also want to give thanks to the original Dr. Hotmann, my wonderful Ford and Hotmann families, and all my awesome, long-suffering friends.

Acknowledgements for Chapter 2:

This work was supported by the National Science Foundation (ATM-0929282). We thank the satellite (MODIS, MISR, CALIOP, CloudSat) and in situ measurement teams (AERONET, IMPROVE, SEARCH).

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1 Introduction

Atmospheric aerosols, or small particles suspended in the air (or other gas), are important for two main reasons: (1) they can impact climate through interactions with radiation and clouds and (2) they can have negative impacts on human health. Additionally, aerosols can degrade visibility and impact ecosystem health. Further discussion on the health impacts is given in Chapter 3, but because of the strong associations of premature mortality and exposure to particulate matter (PM, specifically those with radii of less than 2.5μm, $PM_{2.5}$), the World Health Organization (WHO) guidelines suggest that annual (24-hour) concentrations should not exceed 10 μgm^{-3} (25 μgm^{-3}). In response to the harmful impacts on health, many countries have enacted their own regulations, although typically not as strict as the WHO guidelines.

Particles can vary by size (orders of magnitude), shape, and composition. They have both natural and anthropogenic sources and can be directly emitted or formed in the atmosphere. Although anthropogenic emissions have increased dramatically over the last century, natural sources still dominate aerosol mass loading globally. Tropospheric aerosols typically have lifetimes on the order of days to a week [*Roelofs*, 2013; *Tasić et al*., 2006] and can be removed both through wet deposition (washout or forming ice/cloud condensation nuclei) and dry deposition (impaction or settling).

Aerosols can absorb and/or scatter solar radiation (large particles can also interact with terrestrial radiation). The specifics of this process depend on particle size and optical properties, which are functions of the chemical composition and formation mechanism. The amount of radiation that is depleted by aerosols along a path can be described by the extinction coefficient (b_{ext}), which captures the combined effect of scattering and absorption and is given in Equation 1.1.

$$b_{ext}(\lambda) = \int_{D^{min}}^{D^{max}} \frac{\pi D_p^2}{4} Q_{ext}(m, x) n(D_p) dD_p \quad (1.1)$$

The aerosol extinction has units of inverse length and is a function of the wavelength (λ), particle size (D), number of particles (n), and the extinction efficiency (Q_{ext}, which is a function of the size parameter x and refractive index, m). The column integrated extinction (over depth z) is referred to as the aerosol optical depth (AOD, τ), which is a unitless measure given by Equation 1.2. It is the most common parameter determined by satellite retrievals of aerosols.

$$\tau = \int_{z_0}^{Z_{TOA}} b_{ext}\, dz \quad (1.2)$$

Because of the aforementioned impacts on human health, climate and the environment, monitoring of atmospheric aerosols has become more prevalent. Surface-based mass measurements are preferable for health studies, but for climate effects, we are also interested in aerosols throughout the atmospheric column. Because of the short lifetime of aerosols, to accurately represent the temporal and spatial variability requires dense surface coverage. Information on the vertical distribution can also be useful in regards to health impacts by indicating how and where pollutants are transported. Airborne measurements are extremely useful for this purpose, but are not a viable option for long-term widespread monitoring. Therefore, there has been increased reliance on exploiting aerosols' interactions with radiation through remote sensing to provide information on atmospheric loading. Space-borne instruments in particular can provide global coverage on multi-year to multi-decadal timescales that can help us better understand climate impacts. Additionally, the potential to translate these observations into information on surface air quality can be beneficial for studying health impacts associated with particulate exposure in regions where surface-based measurements are not available.

However, because aerosol properties can vary dramatically (temporally, spatially, and vertically), it can be difficult to use a remotely sensed column-integrated observation as a proxy for surface air quality. This issue of relating satellite optical measurements to surface mass concentrations is the main focus of this dissertation. One of the methods employed herein is to use a chemical transport model (CTM, GEOS-Chem) to help interpret these satellite observations by examining different processes and variables that can impact the mass loading and the simulated AOD. CTMs such as GEOS-Chem can represent the temporally and spatially varying physics and chemistry of the atmosphere; however, they rely on meteorological and emission inventory datasets as input. Therefore, the satellite observation can provide insight into missing sources or processes in the model. This method however requires a delicate balance to use the information from both the satellites and the model, while being cognizant of the limitations of both.

This dissertation is divided into two topics that reflect the two-fold motivation for studying atmospheric aerosols, climate and air quality. In the first part, satellite observations are used in conjunction with a global model (GEOS-Chem) in order to investigate aerosol loading in the Southeastern United States, a region that has experienced cooling over recent decades, potentially driven by changes in atmospheric aerosol loading. In the second part, PM exposure and the associated premature mortality are estimated using model simulated and satellite-based $PM_{2.5}$. The goal is to better understand how satellite observations can be used for these types of investigations and how to temper interpretation by their limitations.

2 Aerosol loading in the Southeastern United States: Reconciling surface and satellite observations

Bonne Ford and Colette L. Heald

This chapter is a published article in Atmospheric Chemistry and Physics, 13, 9269–9283, 2013.

2.1 Introduction

Portmann et al. [2009] suggest that increases in atmospheric aerosols of biogenic origin associated with regional reforestation may have caused cooling over the Southeastern United States (SEUS) in recent decades. This theory is supported by the strong winter to summer seasonality in satellite-derived aerosol optical depth (AOD) that spatially and temporally matches biogenic volatile organic compound (BVOC) emissions in the region [*Goldstein et al.*, 2009]. A potential source of this summertime aerosol could be enhanced production of secondary organic aerosols (SOA) formed from the oxidation of volatile organic compounds (VOC) emitted from vegetation in the presence of anthropogenic pollutants from urban areas [*Volkamer et al.*, 2006; *Hoyle et al.*, 2011]. The SEUS could be particularly susceptible to such an effect [*Weber et al.*, 2007], which could augment summertime aerosol loading in the region.

The SEUS is densely forested and primarily a rural environment, although there are also several major urban centers in the region. Previous studies have shown that $PM_{2.5}$ concentrations in the region are dominated by ammonium sulfate and organic matter (OM), which together account for 60-90% of the surface $PM_{2.5}$ concentrations [*Edgerton et al.*, 2005; *Weber et al.*, 2007; *Hand et al.*, 2012; *Zhang et al.*, 2012a]. Throughout most of the year, organic carbon is produced from wood combustion and diesel exhaust; secondary production dominates in the summertime [*Zheng et al.*, 2002]. *Lim and Turpin* [2002] suggest that SOA generally makes up half of the measured organic carbon. Although urban centers often have higher $PM_{2.5}$

concentrations and their emissions can have a regional impact; water soluble organic carbon concentrations, which are often used as a marker for SOA, appear to have a more widespread homogenous source over the region [*Peltier et al.*, 2007].

Several studies have also suggested that aerosol loading over the Southeastern U.S. has decreased over the last decade. *Edgerton et al.* [2005] note a 15-20% decrease in surface $PM_{2.5}$ mass in the region over the five-year period from 1999-2003, mainly attributable to declines in sulfate and organic matter. Using satellite measurements of AOD and surface $PM_{2.5}$ measurements over Georgia, *Alston et al.* [2012] also suggest that aerosol loading over the Southeastern U.S. declined from 2000-2009. Additionally, *Leibensperger et al.* [2012] suggest that anthropogenic aerosols are responsible for regional cooling over the Eastern U.S. over the last century, but that this radiative forcing has declined since 1990 mainly due to decreases in domestic emissions of sulfur dioxide. However, given limitations in both our measurement and understanding of BVOC emissions and SOA formation [*Hallquist et al.*, 2012], it is unclear whether biogenic emissions, and the aerosols produced upon oxidation of these emissions, have also changed over this same time period. The evolution of these biogenic emissions is difficult to predict [e.g. *Heald et al.*, 2009], representing a significant hurdle for future air quality management efforts and the prediction of climate forcing.

In this study, we use a suite of satellite and surface observations with a global model to explore the origin of the observed enhancement of summertime AOD in the SEUS. We aim to provide insight relevant to the Southeastern Atmosphere Study (SAS) campaign in 2013, whose primary objective is to investigate the impact of biogenic aerosol on regional climate and air quality.

2.2 Description of Observations and Model

2.2.1 Satellite Observations

For this study, we use a variety of satellite instruments and products to analyze aerosol and cloud distributions and variability along with fire activity.

The Multi-imaging Spectroradiometer (MISR) instrument was launched into sun-synchronous orbit aboard the EOS-Terra satellite in 1999 and provides global measurements of AOD with an equator crossing of ~10:30am local time [*Diner et al.*, 2005; *Martonchik et al.*, 2009]. MISR employs nine different cameras to make multi-angle radiance measurements in four spectral bands (visible to near infrared). Here we use the version 22 Level 3 (gridded) global aerosol product which provides daily averaged AOD at 555 nm gridded and filter to remove any grid boxes wherein the standard deviation of the averaged Level 2 (ungridded) AOD data are greater than 2.5 [*Ridley et al.*, 2012].

The Moderate Resolution Imaging Spectroradiometer (MODIS) measures radiances at 36 wavelengths to characterize a variety of land and atmospheric properties. We use observations here from the MODIS instrument launched aboard the EOS-Aqua platform in 2002, which flies as part of the A-Train constellation, making simultaneous measurements at an equator crossing time of ~13:30 local time. AOD from MODIS is retrieved separately over the ocean and land to account for differences in surface properties [e.g. *Remer et al.*, 2005]. While some studies have found a high AOD bias in the western U.S. due to the use of an estimated surface reflectance over the bright land surface [*Drury et al.*, 2010], the SEUS has dense vegetation that provides good dark targets and greater confidence in the MODIS aerosol retrieval [*Roy et al.*, 2007]. For this work, we use Collection 5, Level 3 daily measurements and combine land and ocean optical depth retrievals. We filter the MODIS data to include only grid boxes with cloud fractions below

0.5 and aerosol optical depths less than 1.5. We note that the magnitude of AOD observed by MODIS is sensitive to the cloud fraction filtering [*Zhang et al.*, 2005]; however, the spatial distribution and relative increase from winter to summer remain the same when cloud fraction filtering is varied from 0.1 to 0.8. To investigate the impact of biomass burning on aerosol loading, we also examined MODIS fire counts. For this, we use V005 MODIS Aqua 1°x1° monthly gridded active fire counts, which have frequently been used as an estimate of biomass burning activity [*Duncan et al.*, 2003; *Zeng et al.*, 2008; *Zhang et al.*, 2010].

The Cloud-Aerosol Lidar with Orthogonal Polarization (CALIOP) was launched aboard the CALIPSO satellite in 2006 as part of the A-Train constellation. The instrument detects the intensity and orthogonally polarized components of backscattered radiation at two wavelengths, 532 nm and 1064 nm [*Winker*, 2003]. Through extensive comparisons between CALIOP and the airborne NASA Langley Research Center High Spectral Resolution Lidar (HSRL), *Rogers et al.* [2011] have demonstrated the high accuracy of CALIOP's 532 nm attenuated backscatter calibration, finding that total attenuated backscatter from the two instruments agrees within 2.7% ± 2.1% at night and 2.9% ± 3.9% during the day. Other studies have also found good agreement between CALIOP and ground-based lidar measurements [e.g. *Mamouri et al.*, 2009; *Mona et al.*, 2009].

A vertical profile of aerosol extinction is estimated from the measurement of backscattered radiation, but relies on a lidar ratio for the conversion [*Young and Vaughan*, 2009]. This value can be derived from layer transmittance, or the aerosol classification scheme (which relies on the two wavelength backscatter measurements, approximate volume depolarization ratios, surface type, geographic location and layer altitude) can specify a lidar ratio based on the assumed aerosol type [*Omar et al.*, 2009]. The six aerosol types used are defined from cluster analysis of

AERONET datasets [*Omar et al.*, 2005]. While the CALIOP algorithm uses the mean lidar ratio for each aerosol type, the associated standard deviations suggest that the uncertainty in these values could be 30-50%. The correct classification of cloud and aerosols and selection of an appropriate lidar ratio is the largest source of the uncertainty in the retrieved extinction profile [*Young et al.*, 2013].

We use the Level 2 Version 3.01 5-km Aerosol Profiles and filter the CALIOP observations using cloud aerosol distinction (CAD) scores, extinction uncertainty values, atmospheric volume descriptors, extinction quality control (QC) flags and total column optical depths. We make the approximation that all extinction observations with a corresponding atmospheric volume descriptor that indicates clear air have zero aerosol extinction. For comparisons of simulated extinction profiles with observed profiles, we match clear sky CALIOP profiles with the corresponding grid box and apply a simple detection limit to the model profile following *Ford and Heald* [2012]. For further information on the impact of our filtering and sampling methods, we refer the reader to *Ford and Heald* [2012]. Although the nighttime data have a greater signal-to-noise ratio (SNR) due to the lack of noise from background solar illumination [*Hunt et al.*, 2009]; in Figure 2.1, we use daytime observations to coincide with the MODIS observations. For the rest of our analysis (Figure 2.5 and Figure 2.6), we use the more reliable nighttime profiles. Previous comparisons of seasonally averaged AOD from CALIOP show that daytime observations have only a slight low bias compared to the night observations in source and outflow regions, and a slight high bias over remote marine regions [*Ford and Heald,* 2012]. We re-grid the satellite AOD observations to a 2ºx2.5º resolution and calculate daily averages. In order to preserve the amount of data, we do not co-sample all the

data. However, co-sampling the data does not change the spatial distributions or the magnitude of seasonality reported in Section 3.

For seasonal cloud fraction over the SEUS, we compare observations from MODIS Aqua and MODIS Terra with observations from the CloudSat Cloud Profiling Radar. In particular, we use the CloudSat Level 2 Radar-Lidar GEOPROF Product for profiles of cloud fraction and the Cloudsat Auxiliary Data to convert above ground altitude to pressure coordinates for December 2008-2009. The GEOPROF product combines observations from both CALIOP, which is useful for observing thin cirrus clouds but is completely attenuated in deep clouds (optical depths >3), and the CloudSat Cloud Profiling Radar, which has a millimeter wavelength and is able to penetrate through most non-precipitating clouds [*Stephens et al.*, 2002, 2008]. To compare with two-dimensional spatial distributions of cloud fractions observed by MODIS, we use the maximum cloud fraction from each profile observed by CloudSat/CALIOP.

2.2.2 Ground-based Data

We use observations from the global AERosol Robotic NETwork (AEORONET) of sun photometers in the SEUS [*Holben et al.*, 1998]. AERONET sites record AOD and aerosol properties at several wavelengths in the visible and near IR, and have been used for validation studies of satellite measurements [e.g. *Remer et al.*, 2002a]. For this work, we use hourly Version 2, Level 2 measurements from the Walker Branch and UA Huntsville sites for months in 2008-2009.

We also use surface measurements of $PM_{2.5}$ concentrations from both the Interagency Monitoring of Protected Visual Environments (IMPROVE) and Southeastern Aerosol Research and Characterization (SEARCH) networks. Surface measurements of atmospheric composition

from the IMPROVE network are taken over a 24-hour period once in three days and are analyzed for the concentration of fine, total, and speciated particle mass [*Malm et al.*, 1994]. Ammonium mass is determined by assuming that sulfate and nitrate are fully neutralized, which means that this is an upper bound for the dry mass of ammonium. The IMPROVE network $PM_{2.5}$ values used in this study are the reconstructed fine mass (RCFM) determined by adding the values of ammonium sulfate, ammonium nitrate, soil, sea salt, elemental carbon and organic matter. We use 1.8 as the organic carbon to organic matter multiplier following *Hand et al.* [2012]; though we note that this could be too high or too low at specific sites as *Malm and Hand* [2007] have given a range of 1.2 to 2.6.

The SEARCH network is composed of eight sites consisting of pairs of urban and rural/suburban locations in four states (AL, FL, GA, MS) which all measure meteorological parameters, gas phase pollutants and major $PM_{2.5}$ components (organic carbon, elemental carbon, sulfate, nitrate, ammonium, and trace metals) [*Hansen et al.*, 2003; *Edgerton et al.*, 2005, 2006]. These sites have both continuous and 24-hour integrated filter-based measurements of $PM_{2.5}$, and the sample frequency is every 3 days, except for at the sites located in Atlanta and Birmingham which report daily. However, we sample the data from these two sites to the same measurement days as the other sites.

SEARCH network $PM_{2.5}$ concentrations are calculated as the sum of sulfate, nitrate, ammonium, organic matter (using 1.8[OC]), elemental carbon, and major metal oxides (MMO). MMO is the sum of aluminum, calcium, iron, potassium, silica, and titanium in the highest oxidation state and is almost equivalent to the soil concentrations reported by the IMPROVE network. However, the soil equation used for the IMPROVE network makes corrections to account for a lower oxidation state of iron, contributions from other elements, and potassium

from non-soil sources; so that [SOIL] = (2.20[Al] + 2.49[Si] + 1.63[Ca] + 2.42[Fe] + 1.94[Ti]) [*Malm et al.*, 1994]. We also note that this standard calculation of $PM_{2.5}$ for SEARCH sites does not include sea salt as included with the IMPROVE data (calculated as 1.8[Cl]), and chlorine has only been reported since 2009. To estimate species contributions to the total $PM_{2.5}$ from the SEARCH sites, we use the assumptions for the IMPROVE measurements to include a modified soil concentration and a value for sea salt using chlorine measurements from 2009. We also use hourly TEOM measurements of total $PM_{2.5}$ mass concentrations to characterize the diurnal variation.

All eight of the SEARCH sites and several IMPROVE sites are co-located with nephelometers, which provide measurements of ambient relative humidity (RH). For comparisons of vertical profiles of RH, we also use ground based soundings from 8 NOAA sites (http://weather.uwyo.edu/upperair/sounding.html). Locations of these IMPROVE, SEARCH, and NOAA sounding sites are shown in Figure 2.1 and Figure 2.4.

2.2.3 GEOS-Chem

We use v9.01.01 of the GEOS-Chem chemical transport model, driven by GEOS-5 meteorology, in the nested grid configuration over North America (0.5°x0.667° horizontal resolution). The GEOS-Chem aerosol simulation includes sulfate, nitrate, ammonium (Park et al., 2004), primary carbonaceous aerosols [*Park et al.*, 2003], dust [*Fairlie et al.*, 2007; *Ridley et al.*, 2012], sea salt [*Alexander et al.*, 2005], and secondary organic aerosols (SOA) [*Henze et al.*, 2008]. Aerosols and gases are removed by both wet and dry deposition in the model. The wet deposition scheme includes scavenging in convective updrafts, rainout and washout (Liu et al., 2001), while dry deposition of gases and aerosols is dependent on surface characteristics and

meteorological conditions [*Wesely*, 1989; *Wang et al.*, 1998]. The EPA NEI05 inventory inventory (scaled to be year-specific) is used for most anthropogenic and biofuel emissions over the USA [*Hudman et al.*, 2007, 2008]; however, anthropogenic emissions of black and organic carbon follow *Cooke et al.* [1999] with the seasonality from *Park et al.* [2003]. Biogenic VOC emissions are calculated interactively following MEGAN [*Guenther et al.*, 2006], while year-specific biomass burning is specified according to the GFED2 inventory [*van der Werf et al.*, 2006]. We implement a fix for artificially low nighttime boundary layer heights in the GEOS-5 product as well as a 25% reduction in the HNO_3 concentrations, both of which improve comparisons with surface nitrate observations in the United States as described by *Heald et al.* [2012].

We calculate surface $PM_{2.5}$ in the model by combining sulfate, nitrate, ammonium, elemental carbon, organic matter (organic carbon scaled by a factor of 1.8 to account for total organic matter, consistent with IMPROVE and SEARCH measurements), fine dust, and accumulation mode sea salt concentrations in the lowest grid box. For comparison with IMPROVE and SEARCH daily $PM_{2.5}$ measurements, we compute 24 hour averages and sample the data to site locations and measurement days.

Aerosol optical depth in the model is calculated for a specific wavelength using the extinction efficiency (Q_{ext}), the column mass loading (M), effective radius (r_{eff}), and particle mass density (ρ), such that the AOD (τ) is calculated by the following equation [*Tegen and Lacis*, 1996]:

$$\tau = \frac{3Q_{ext}M}{4\rho r_{eff}} = \alpha M \quad (2.1)$$

The aerosol mass extinction efficiency (α) is calculated with Mie code based on wavelength-resolved optical and size parameters at 7 relative humidity values (0, 50%, 70%, 80%, 90%, 95%, 99%) for various aerosol types from the Global Aerosol Data Set (GADS) [*Koepke et al.*, 1997], with recent updates based on *Drury et al.* [2010], *Jaeglé et al.* [2011], *Ridley et al.* [2012]. The extinction efficiency (α) for each grid box is calculated from local relative humidity conditions [*Martin et al.*, 2003]. In order to determine whether the observed seasonality in AOD can be attributed to changes in mass loading or mass extinction efficiency, we explore the sensitivity of the mass extinction efficiency to the observed seasonal changes in relative humidity (Section 3.2). For this purpose, we use the properties of sulfate aerosol (Table 2.1), which exhibits the strongest relationship with relative humidity (i.e. the largest hygroscopicity).

Table 2.1 Mass extinction efficiency values used in GEOS-Chem for sulfate aerosols at given relative humidity values.

Relative Humidity (%)	Mass Extinction Efficiency (m^2g^{-1})
0	2.24
50	6.12
70	8.90
80	11.90
90	19.09
95	32.75
99	104

2.3 Results

2.3.1 Seasonality in AOD and Surface Concentrations

We mimic the seasonal AOD comparisons of *Goldstein et al.* [2009] in Figure 2.1 for 2007-2009 using observations from MODIS Aqua, MISR, and CALIOP. Figure 2.1 also shows the seasonal mean AOD measured by ground-based sun photometers at two AERONET sites in the SEUS.

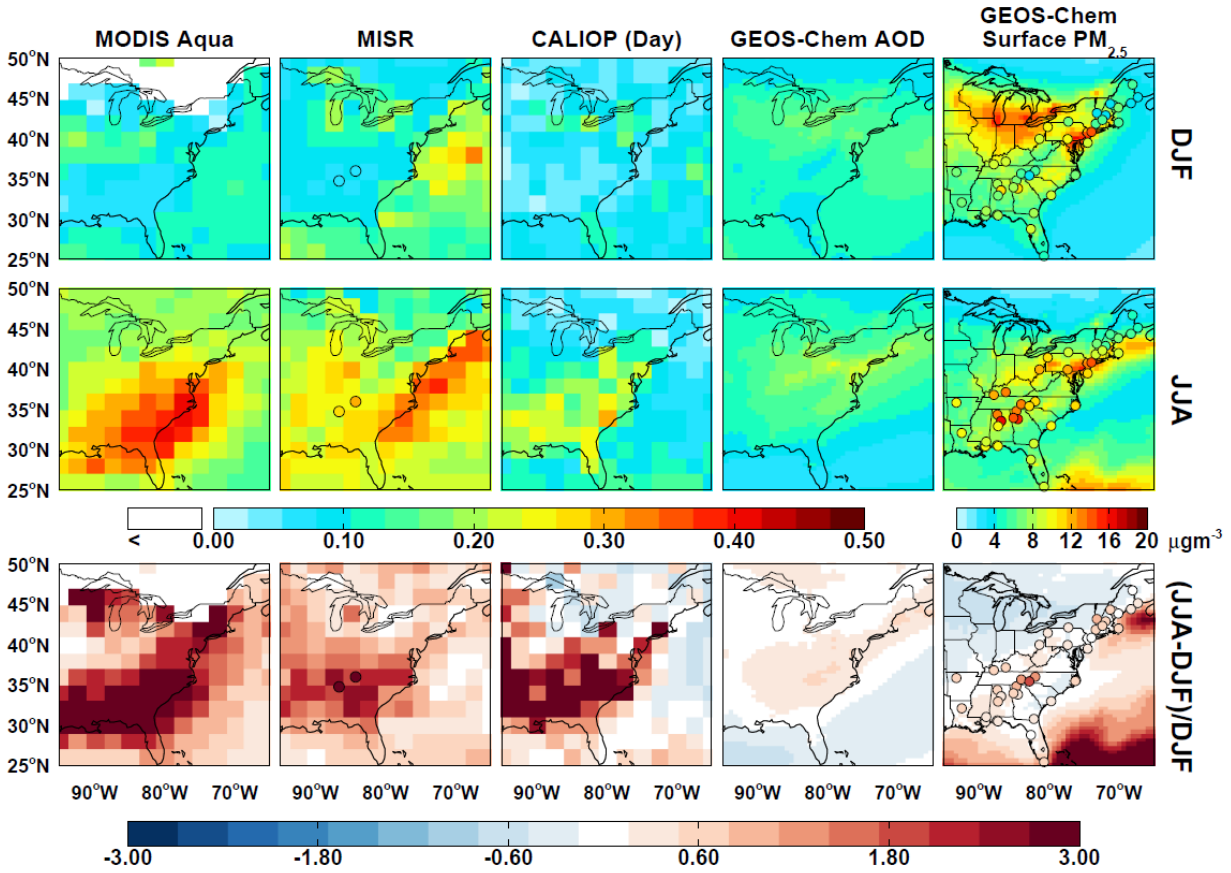

Figure 2.1 Seasonally averaged total column AOD for winter (DJF, top row) and summer (JJA, middle row) for December 2006-August 2009 as observed by MODIS (column 1), MISR (column 2), and CALIOP (daytime, column 3) gridded to 2x2.5° and compared to simulated AOD from GEOS-Chem (column 4). Concentrations of surface $PM_{2.5}$ simulated by GEOS-Chem are overlaid with concentrations measured at IMPROVE and SEARCH network sites (circles) in column 5. Bottom row shows the relative enhancement of summer over winter for observed and simulated AOD and surface concentrations. Average AOD observed at the University of Alabama, Huntsville and Walker Branch AERONET sites for 2008-2009 is overlaid (circles) on the MISR maps.

Figure 2.1 shows that observed summertime AOD is consistently a factor of 2-3 higher than wintertime values in the SEUS. This is consistent with *Alston et al.* [2012] who find a threefold increase in summertime AOD over wintertime, although they use a finer resolution product and therefore show greater spatial variability. The magnitude of both the seasonal mean and the relative enhancement differs among instruments as shown in the figure, but is spatially

consistent. Similar summertime enhancements are reported at AERONET sites in 2008-2009. Furthermore, AOD measurements at the two AERONET sites show little variability during daylight hours (<20% in summertime) and suggest that this enhancement is consistent throughout the day. Four independent observations, made with four different measurement techniques all indicate a large regional enhancement in summertime AOD.

We contrast these observations with the GEOS-Chem chemical transport model simulation. Simulated summertime enhancements in AOD through the Eastern U.S. range from 15 to 40%. While the model does show a summertime maximum in AOD over the Ohio River Valley and northeastern U.S. associated with increases in sulfate via SO_2 oxidation [*Chin et al.*, 2000] and stagnation events, it does not reproduce the strong observed seasonality in column AOD over the SEUS.

The fourth column of Figure 2.1 shows the GEOS-Chem simulated surface concentrations of $PM_{2.5}$ overlaid with observations from the SEARCH and IMPROVE networks. Comparison with surface observations indicates that GEOS-Chem generally captures the spatial, seasonal, and diurnal (Figure 2.2) variation of $PM_{2.5}$ in the Eastern U.S., as will be discussed in further sections and consistent with previous studies [e.g. *Heald et al.*, 2012; *Leibensperger et al.*, 2012]. However, Figure 2.1 shows that the measured surface concentrations exhibit only a fraction of the seasonality in the column AOD observed by the satellite instruments, suggesting that changes in surface concentrations do not dictate the seasonality observed in AOD. This is in agreement with *Alston et al.* [2012], who show that the summertime enhancement in surface $PM_{2.5}$ in Georgia is considerably less than the AOD enhancement over the region. This presents an intriguing disconnect between surface and column measurements of aerosol loading in the SEUS.

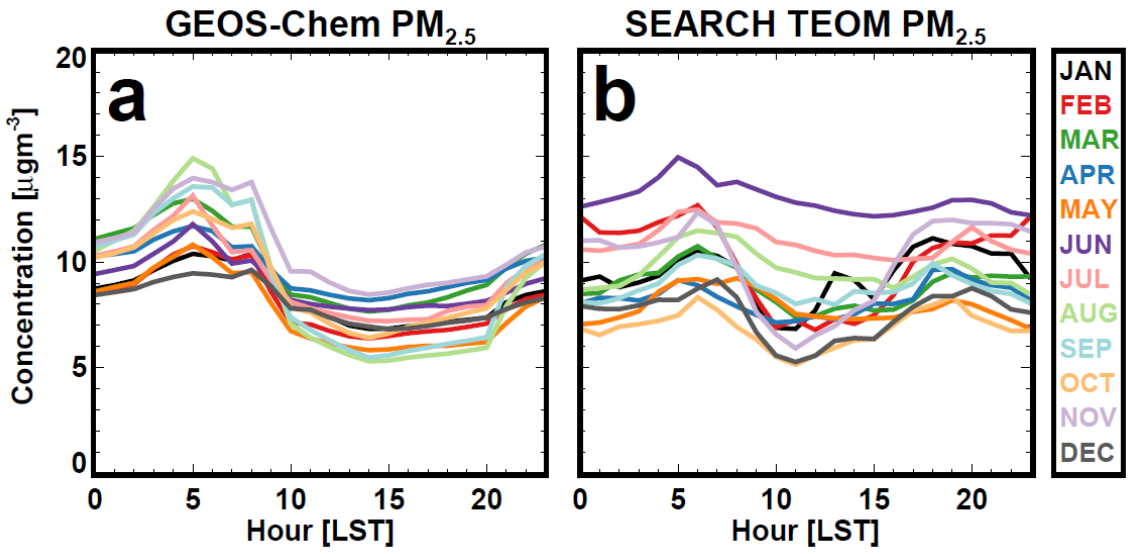

Figure 2.2 Diurnal cycles of monthly mean $PM_{2.5}$ concentrations as simulated by GEOS-Chem (a) and as measured by TEOM filters at SEARCH network sites (b), color coded by month in 2009 over which the averaging was done.

In Figure 2.3, we explore the chemical composition of surface $PM_{2.5}$ over the SEUS in winter and in summer. Overall, the model captures the chemical speciation and the seasonality of IMPROVE surface concentrations in the region (Figure 2.3a), but somewhat underestimates the observed summertime enhancement (mean $PM_{2.5}$ concentration is ~55% greater in the summertime). The observed surface seasonality is mainly due to inorganic species. Both observed and simulated nitrate concentrations are higher in the wintertime, consistent with more favorable formation of ammonium nitrate at cooler temperatures, while the increase in observed summertime sulfate is the result of enhanced SO_2 oxidation [*Chin et al.*, 2000]. The model simulation underestimates the doubling in sulfate concentrations from winter to summer seen at IMPROVE sites (specifically the Appalachian sites), but does capture the summertime enhancement associated with dust transport from North Africa [*Ridley et al.*, 2012].

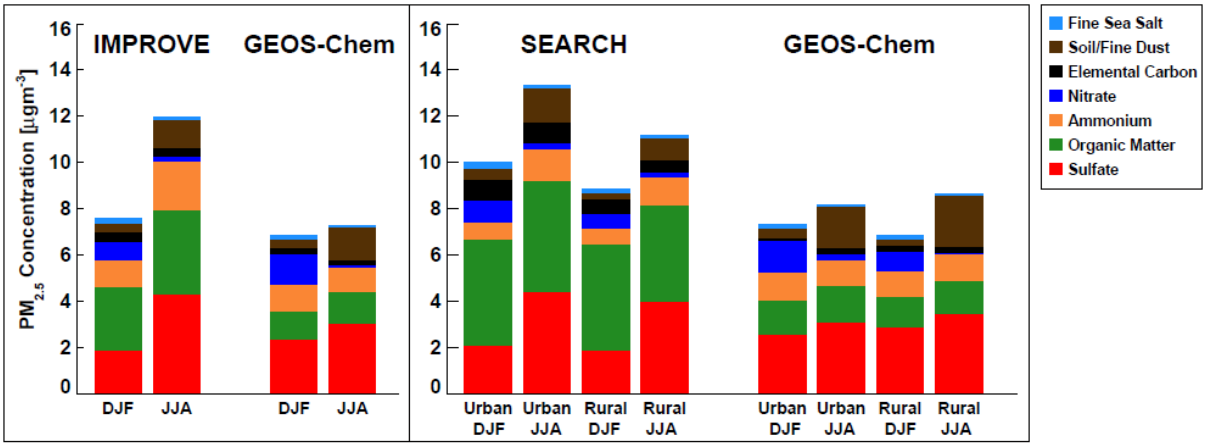

Figure 2.3 Mean seasonally averaged surface mass concentrations of $PM_{2.5}$ (a) observed at the 13 IMPROVE network sites in the SEUS and as simulated by GEOS-Chem (sampled to the IMPROVE sampling days and corresponding model grid boxes). (b) Same as (a) for SEARCH network sites. All data are for December 2006-2009, except for SEARCH sea salt, which only uses data from 2009.

Observed OM at IMPROVE sites in the SEUS is consistently ~2-4µgm^{-3}, making up less than 35% of mean observed $PM_{2.5}$, consistent with the analysis of *Hand et al.* [2012]. *Zhang et al.* [2012b] find slightly higher OM/$PM_{2.5}$ fractions (40-50%) in urban areas in the SEUS; however, when biomass burning events are removed, the fraction of water soluble OM/$PM_{2.5}$ that they observe is reduced to ~25%. GEOS-Chem under predicts OM concentrations by about a factor of two, consistent with other regions of the world [*Heald et al.*, 2011]. However, the magnitude of OM seasonality is similar between the observations and model with very little variation in OM throughout the year (a 35% increase in summer over winter), in agreement with values reported by *Zhang et al.* [2012b]. The seasonality in OM (and consequently $PM_{2.5}$) might be enhanced if we used a varying OM/OC ratio, as studies have shown that the ratio is greater in the summer than in the winter [e.g. *Simon et al.*, 2011]. However, the ratio would be applied to both the model simulation and observations; and, as GEOS-Chem is already able to simulate the

OM seasonality, would therefore not explain the model discrepancy. We also compare these OM concentrations and seasonality with IMPROVE sites in the Northeastern U.S., where there is less of an increase in summertime AOD, and find similar values at the surface.

Observations from the SEARCH network show less seasonal variability in $PM_{2.5}$ than the IMPROVE sites, as shown in Figure 2.3b, which also highlights the differences between rural and urban sites. As expected, $PM_{2.5}$ concentrations are greater at urban sites and in particular, sulfate and elemental carbon concentrations are higher. Additionally, although GEOS-Chem is able to simulate dust concentrations and seasonality at IMPROVE sites; it slightly overpredicts summertime dust at SEARCH sites, especially rural sites. There are also higher concentrations and more seasonality in the OM at urban sites, consistent with *Yan et al.* [2009]. Overall, the OM fraction at these SEARCH sites is greater than observed at the IMPROVE sites (Figure 2.3a). Thus the model underestimate of OM is also greater. However, even more so than the IMPROVE comparisons, both the SEARCH observations and the simulation of concentrations at these sites, show very little seasonal variation in the OM at both rural and urban sites. Additionally, there is little diurnal variability in $PM_{2.5}$ (Figure 2.2, relative standard deviation of ~10% throughout the year), which confirms that the seasonality in surface concentrations is consistent throughout the day and that using a 24-hour average $PM_{2.5}$ concentration rather than a one-hour average $PM_{2.5}$ concentration sampled to the satellite overpass times does not bias the comparison with satellite observations provided at two snapshots.

We conclude that the seasonality in satellite AOD over the SEUS does not match the surface concentrations, or more specifically, organic aerosol at the surface. This weaker correlation between column AOD and $PM_{2.5}$ (R=0.2-0.31, for daily matched pairs across the region) in the SEUS contrasts other regions where previous studies have found strong correlation

between satellite-observed column AOD and surface concentrations throughout the year (R=0.39-0.9, e.g. *Engel-Cox et al.*, 2004; *Al-Saadi et al.*, 2005; *van Donkelaar et al.*, 2006, 2010; *Paciorek and Liu*, 2009; *Zhang et al.*, 2009a). These previous studies note that discrepancies often arise in comparing surface concentrations and column AOD due to inaccurate assumptions or lack of information about the hygroscopicity of the aerosols, the composition, size distributions, the vertical distribution of aerosols, the presence of transported aerosols above the surface layer, and the meteorological environment, especially with regards to clouds. Correlations are highest when the aerosol is near the surface, uniformly mixed, relative humidity is moderate, and coarse mode aerosol fraction is small [*Al-Saadi et al.*, 2005; *van Donkelaar et al.*, 2006]. Therefore, in the following sections we investigate several of these factors in an attempt to determine what, other than surface concentrations, could be driving the seasonality in AOD.

2.3.2 Effect of Relative Humidity

Changes in aerosol water uptake could play a role in the seasonality of AOD in the SEUS. This is briefly examined by *Goldstein et al.* [2009], who find that the AOD and relative humidity (RH) at a single AERONET site (Walker Branch) are only weakly correlated across seasons. We investigate this further at several locations to better represent the effect of RH on a broader region of the SEUS. In addition, given that water uptake is already included in the simulation of AOD, we explore whether there is any evidence of a bias in RH in the GEOS-5 meteorology which could degrade the model simulation of AOD.

Hourly surface RH values used in GEOS-Chem are highly correlated with observations from the IMPROVE and SEARCH network nephelometer sites (R=0.6-0.83, across sites and

seasons) with a mean bias of less than 5%. Correlations are highest during the morning and afternoon hours and degrade during the nighttime and at coastal sites (where GEOS-Chem has a low RH bias). Figure 2.4a demonstrates that there is little seasonality in mean surface RH in the SEUS, and that the model reproduces both the magnitude and consistency of RH year-round. Monthly RH values averaged over all the sites vary less than 10% throughout the year at the 13:30 satellite overpass and less than 20% when all hours are used to construct monthly means.

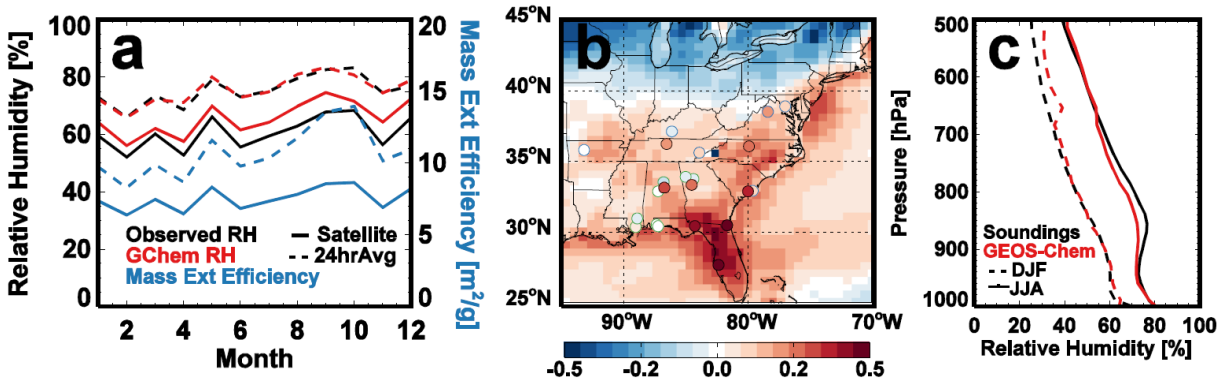

Figure 2.4 (a) Average 2009 monthly surface RH recorded at IMPROVE and SEARCH network nephelometer sites (black) using 24 hour averages (dashed line) and sampled to the afternoon overpass time (solid line). These are compared with values used in GEOS-Chem (red lines) sampled to observational site locations and converted to mass extinction efficiency values (for sulfate aerosols, blue lines). (b) Fractional increase in surface mass extinction efficiency of summer over winter simulated with GEOS-Chem and overlaid with values calculated from surface RH observed at IMPROVE nephelometer (blue circles), SEARCH (green circles) and sounding (black circles) sites for 2009. GEOS-Chem, SEARCH and IMPROVE data are sampled for the afternoon satellite overpass time (13-15LT) and the sounding data are an average from 0000 and 1200 UTC. (c) mean profile of RH measured at the 8 sounding sites (black lines) at 0000 and 1200 UTC and the corresponding RH in GEOS-Chem (red line) sampled to the site locations and times for winter (dashed) and summer (solid) 2009.

In order to examine the impact of these modest seasonal differences in RH on AOD, we use the optical properties applied to aerosols in GEOS-Chem to convert these differences in RH to differences in aerosol mass extinction efficiency. For simplicity, we use sulfate aerosol

properties, which have the highest hygroscopicity (other than NaCl which is not a significant contributor to aerosol mass in the region) [*Petters and Kreidenweis*, 2007], and would be the most affected by changes in RH. At the surface, seasonal differences in mass extinction efficiency due to water uptake account for less than a 25% increase in aerosol extinction from winter to summer over the SEUS (Figure 2.4b).

The seasonality in the vertical profile of RH measured at the 8 NOAA sounding sites in the region is larger, with about a 20% absolute increase in mean RH from winter to summer throughout the troposphere (Figure 2.4c). However, the seasonal difference in aerosol extinction resulting from water uptake (40% when integrated over the column assuming sulfate aerosol) can account for only a fraction of the 100-300% difference in AOD observed by the satellite instruments over the SEUS (Figure 2.1). While this increase in summertime RH is generally captured by the model (and thus reflected in the simulated AOD seasonality), RH values in the lower troposphere (900-700hPa) are slightly underestimated in summertime (<5%). This translates to a 5-12% underestimate in sulfate extinction efficiency in GEOS-Chem at these altitudes; however, the effect is likely to be more modest for the mix of ambient aerosol with lower hygroscopicity. Therefore, aerosol water uptake cannot explain the observed seasonality in AOD, nor can biases in RH explain the model underestimate of this seasonality.

Although we only show the sensitivity of the mass extinction efficiency to RH here, we also note that changes to the aerosol size distribution not accounted for in the model could impact the observed AOD. *Heald et al.* [2011] use Mie code to estimate that uncertainty in mass extinction efficiency associated with the size of OM (associated with a doubling or halving of the assumed mean geometric radius and varying the geometric standard deviation from values of 1.4 to 1.8) is ~50%. Therefore, a seasonal shift towards larger particles in summertime could account

for some of the model-measurement discrepancy. However, measurements of the geometric mean diameters for ambient SOA dominated aerosol and fresh smoke are similar [*Levin et al.*, 2009, 2010]. Additionally, the fraction of total AOD accounted for by organics in the model is on average less than 15% in the summertime, and substantial seasonal changes in the size of inorganic aerosol are less likely [e.g. *Stanier et al.*, 2004; *Zhang et al.*, 2008], so it is unlikely that shifts in the fine aerosol size distribution are a dominant source of model error.

2.3.3 The vertical profile of aerosol

The lack of strong seasonality in surface aerosol concentrations and aerosol water content throughout the vertical column suggests that summertime increases in AOD over the SEUS must be associated with an increase in aerosol mass above the surface layer that is not accounted for in GEOS-Chem, potentially consistent with the hypothesis of *Goldstein et al.* [2009] which suggests a missing source of aerosols aloft.

Winter and summer vertical distributions of all the nighttime aerosol extinction values reported by CALIOP for three years (2007-2009) over the SEUS are shown in Figure 2.5. Separating profiles based on aerosol type indicates broadly what sources are likely to contribute to the mass loading in each season, although the CALIOP algorithm is based on physical properties and does not distinguish aerosol by chemical composition.

Figure 2.5 shows that there are larger aerosol extinction observations at higher altitudes during the summer months. There is little seasonal difference in GEOS-5 nighttime boundary layer height (coincident with CALIOP measurements); but the deeper daytime mixed layers in summer may vertically distribute aerosol to higher altitudes. We investigate the impact of the mixing heights on the vertical profile in the next section.

In the final panels of Figure 2.5, we compare the average nighttime profiles of aerosol extinction observed over the region by CALIOP with the average profile simulated by GEOS-Chem. First, these profiles confirm that the mean observed extinction profile over the SEUS is higher in the summer than in winter. The integrated mean AOD from these profiles increases more than 3-fold from winter to summer (0.07 to 0.25), consistent with the picture presented in Figure 2.1. Second, these profiles demonstrate that the model underestimate of summertime AOD shown in Figure 2.1 is associated with above-surface aerosols. The simulated and observed profile shapes are similar in the wintertime, but the model greatly under predicts aerosol extinction above the surface layer in summer. While there might be a bias in the CALIOP measurements near the surface due to interference of clouds, lower sensitivity, and/or inclusion of clear air retrievals below aerosol layers [*Winker et al*., 2013], surface extinction values are generally reproduced by the model throughout the year, consistent with our $PM_{2.5}$ surface comparisons in Figures 2.1 and 2.3. Additionally, the retrieval sensitivity of CALIOP increases with altitude [*Winker et al*., 2009], and *Sheridan et al.* [2012] suggest that if there is a bias in the free troposphere, it is a low bias (particularly under cloud-free conditions, as shown here), which suggests that the discrepancy in extinction above the surface layer between the simulated and measured profile is unlikely to be due to systematic observational errors.

Mean summertime AOD calculated by integrating the model profile in Figure 2.5 is 0.12, less than half the mean CALIOP values. While these profiles do not suggest that there is a distinct lofted layer of aerosol, they do indicate that the seasonality in AOD is primarily associated with an increase in aerosol mass above the surface layer, in the lower troposphere (below 700hPa), which is not captured by the GEOS-Chem model. This discrepancy is inconsistent with *Heald et al*. [2011] who show that the GEOS-Chem model generally captures

the profiles of sulfate and organic aerosol in the Northern Hemisphere, even when concentrations are significantly underestimated.

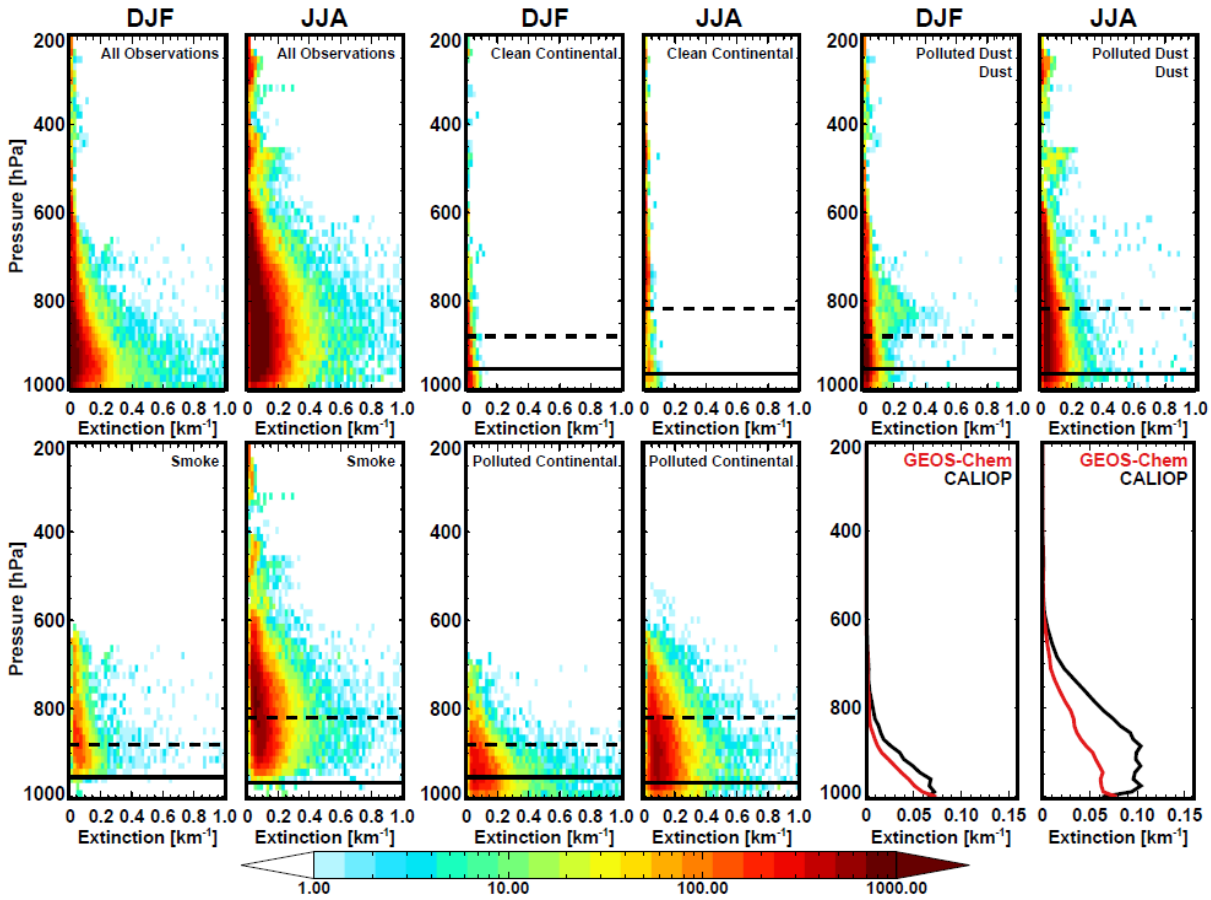

Figure 2.5 Density plots of all nighttime aerosol extinction values observed by CALIOP for winter and summer seasons of 2007-2009 over the SEUS (30.5°N to 37.5°N and 90°W to 81.5°W), classified by aerosol type. The color denotes the number of observations with given extinction values at a given altitude. There are ~ 70,000 extinction values for the winter and ~200,000 for the summer from December 2006-2009, including values below the detection limit. Panels 1 and 2 show observations with all aerosol types noted, and panels 3-10 separate values based on aerosol type with seasonally averaged boundary layer heights from GEOS-Chem for nighttime (solid black horizontal lines) and daytime (dashed lines) overlaid. Bottom right panels show average aerosol extinction profiles as observed by CALIOP (black) and the corresponding profiles simulated by GEOS-Chem and sampled to the CALIOP overpasses (red) for the two seasons applying detection limits as in *Ford and Heald* [2012].

We distinguish these CALIOP profiles by the observed aerosol types to understand the source of these aerosols, and Figure 2.5 shows that the majority of aerosols observed in the SEUS are classified as polluted dust, polluted continental, and smoke. The maximum in dust above the boundary layer in the winter is likely due to dust transport from the western U.S. while increases in the summer are associated with transport of African dust.

2.3.4 Effect of planetary boundary layer height

One possible explanation of a large seasonal enhancement in AOD with modest change to surface $PM_{2.5}$ is a summertime aerosol source with a coincident deepening of the mixed layer. As shown in Figure 2.5, this deepening of the daytime planetary boundary layer (PBL) is simulated in the GEOS-Chem model, with an increase of more than 60 hPa from winter to summer on average. The CALIOP average summer profile suggests that the higher concentrations above the surface are relatively uniform over a deep layer (up to 800 hPa), which could indicate that the GEOS-Chem simulation does not mix pollutants through a deep enough layer of the atmosphere, although we note that this altitude is similar to the mean summertime PBL depths used in the model (shown in Figure 2.5). If the PBL heights used in the model are too shallow, pollutants could be trapped and more easily removed before being mixed upward.

The diurnal cycle of surface concentrations can provide some information about vertical mixing. For relatively constant emissions, surface concentrations generally increase at night with a shallow boundary layer and then as the PBL height grows throughout the day, pollutants are diluted and surface concentrations decline. Thus, a shallow bias in summertime afternoon mixed layer would reduce the diurnal variability in simulated $PM_{2.5}$ concentrations. However, the diurnal cycle of simulated and observed $PM_{2.5}$ at SEARCH sites are relatively similar (Figure 2.2), with observations showing slightly more consistency throughout the day, especially in the

summer months. The observations shown here are consistent with *Weber et al.* [2003] who show that on average $PM_{2.5}$ concentrations in Atlanta during August generally vary less than 20% through the day due to sulfate peaking in the afternoon while OM, EC, and nitrate tend to peak in the early morning. Thus, while the contribution of the timing of different sources and mixing depths to the diurnal profile can be a challenge to untangle, the ability of the model to capture a relatively flat profile in surface $PM_{2.5}$, as well as the overall seasonality and composition of that $PM_{2.5}$ (Figure 2.3), provides evidence of a relatively unbiased simulation of mixing depth.

It remains a challenge to validate the PBL heights used in the model simulation as there are several different approaches used to estimate the height of the PBL, all of which can produce differing results [e.g. *Berman et al.*, 1997]. As shown in *Marsik et al.* [1995], various measurement systems can at some times differ on the height of the PBL over Atlanta by almost 1km. Therefore, we examine the sensitivity of the vertical profile and summertime AOD to changes in mixing depth in model simulations. We performed four simulations for summer 2009, in which (1) PBL heights are raised by 100 hPa at all hours of the day, (2) PBL heights are raised by 100 hPa only during the daytime hours of 7:00 AM and 7:00 PM LST, (3) SO_2 emissions in North America are doubled, and (4) SO_2 emissions are doubled and the daytime PBL heights were raised. The impact of these simulations on the average regional profile is shown in Figure 2.6.

Raising the PBL for all hours of the day increases extinction values throughout most of the profile, particularly at the top of the PBL, as concentrations are mixed throughout a deeper layer. However, values near the surface decrease as aerosols are mixed away from the surface into the deeper PBL, particularly at night. In the simulation where the PBL was only raised during the day, aerosol extinction values near the surface are similar to the original profile, but

greater above the nighttime PBL, in the daytime residual layer. However, both of these simulations continue to substantially under predict the extinction values compared to CALIOP (AOD increases by only 0.02 over the region), suggesting that a potential bias in the PBL height would not be enough to explain the discrepancy in AOD.

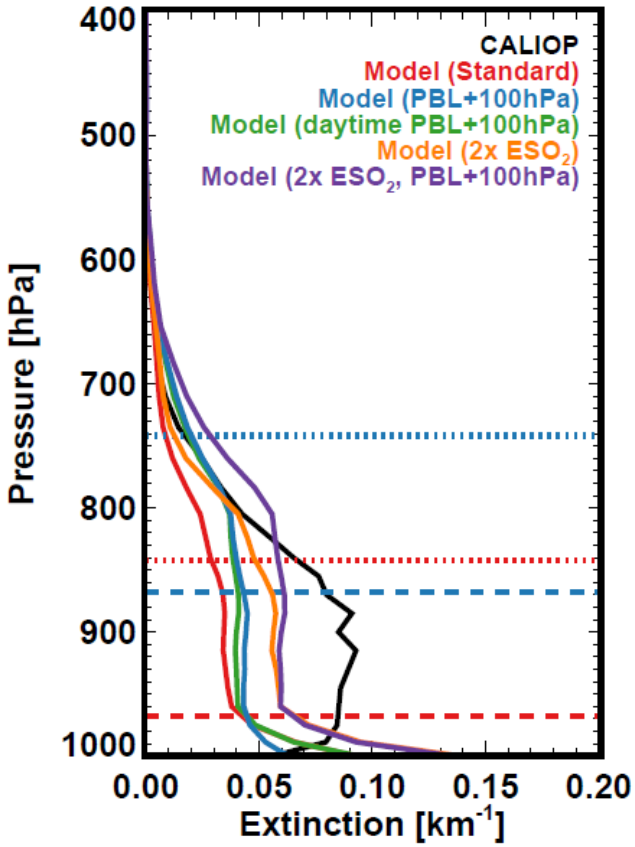

Figure 2.6 Aerosol extinction profiles over the SEUS for JJA 2009 as observed by CALIOP (black) and as simulated by GEOS-Chem with original PBL heights (red), PBL heights raised by 100 hPa at all hours (blue), raised by 100 hPa during daytime hours (green), doubled SO_2 emissions (orange), and doubled SO_2 emissions and PBL heights raised during daytime hours (purple). Average PBL heights are shown in red for original simulation and blue for simulations with raised PBL heights, with dashed lines for nighttime and dotted lines for daytime.

Finally, we verify whether the aerosol profile measured by CALIOP is consistent with an increase in existing sources. We test increasing the sources of sulfate, which would be consistent with both the surface underestimate shown in Figure 2.3, as well as the possibility of formation aloft via in-cloud processing. The profiles in Figure 2.6 demonstrate that doubling SO_2 emissions effectively scales up the entire simulated profile, but in doing so, significantly degrades the comparison with both CALIOP and the speciated surface concentrations (not shown). Simultaneously increasing the daytime PBL heights produces virtually the same profile, but produces an overestimate of observed aerosol extinction from 800 to 700 hPa. We therefore conclude that an increase in mixing depth, with or without a coincident increase in an existing aerosol source, cannot explain the observed CALIOP profile. This implies that the underestimate in aerosol mass aloft is due to an additional above-surface source. Using the average summertime mass extinction efficiency and the difference between the satellite estimated and simulated extinction profile, we roughly estimate that this additional source is equivalent to three times the current sulfate mass (above the surface only).

2.4 Discussion

Satellite observations show a strong summertime enhancement in AOD over the SEUS, previously linked with biogenic activity, which is not observed in surface $PM_{2.5}$ concentrations. By determining here that there is little agreement between surface concentrations and column AOD in the Southeastern U.S., we surmise that changes in surface mass concentrations do not control the seasonality of AOD in the region. Furthermore, the GEOS-Chem model generally captures observed surface concentrations (with the exception of a modest underestimate of summertime sulfate, and year-round underestimate of OA), but it does not reproduce the

observed AOD seasonality, indicating an underestimation of aerosol extinction above the surface layer in the model. We show that neither a bias in model RH (and hence aerosol water uptake) nor summertime mixing depths can explain this discrepancy. *Zhang et al.* [2012a] show that GEOS-Chem reproduces wet deposition measurements in the U.S. This suggests that this discrepancy is also not due to a bias in aerosol removal, but rather a missing source of aerosol above the surface layer.

CALIOP measurements provide additional evidence of aerosol production above the surface; however, our interpretation is limited by the lack of aerosol chemical speciation. *Ervens et al.* [2011] suggest that in regions where there are large biogenic VOC and anthropogenic emissions, high RH, and cloudiness, yields from aqueous formation of SOA can be significant. *Carlton et al.* [2008] show that including SOA formation through cloud processing modestly improves model simulations of airborne OA observations in the Northeastern U.S. *Sorooshian et al.* [2007] also observe elevated organic aerosol layers above clouds during field campaigns over Texas and California due to the formation of organic acids from aqueous phase reactions which subsequently undergo droplet evaporation. Cloudy conditions, as well as enhanced oxidant concentrations in summertime, could also augment sulfate production, which we show is moderately underestimated at the surface in summertime in our simulation. It has also been suggested that including reactions of stabilized Criegee intermediates with sulfur dioxide in models can produce a 10-25% increase in sulfuric acid (H_2SO_4) annually, and a 100% increase in July [*Pierce et al.*, 2013].

GEOS-5 meteorology does show higher cloud fractions over the SEUS in the summer compared to the winter, which is corroborated by observations from MODIS Aqua and Terra (Figure 2.7). If clouds are indeed serving as a medium for chemical production, this could

explain the increased aerosol loading above the surface layer that is seen in the cloud-free CALIOP profiles shown here. However, observations from the CloudSat Cloud Profiling Radar suggest that mid to low troposphere cloud cover is highest in the region in winter (Figure 2.7).

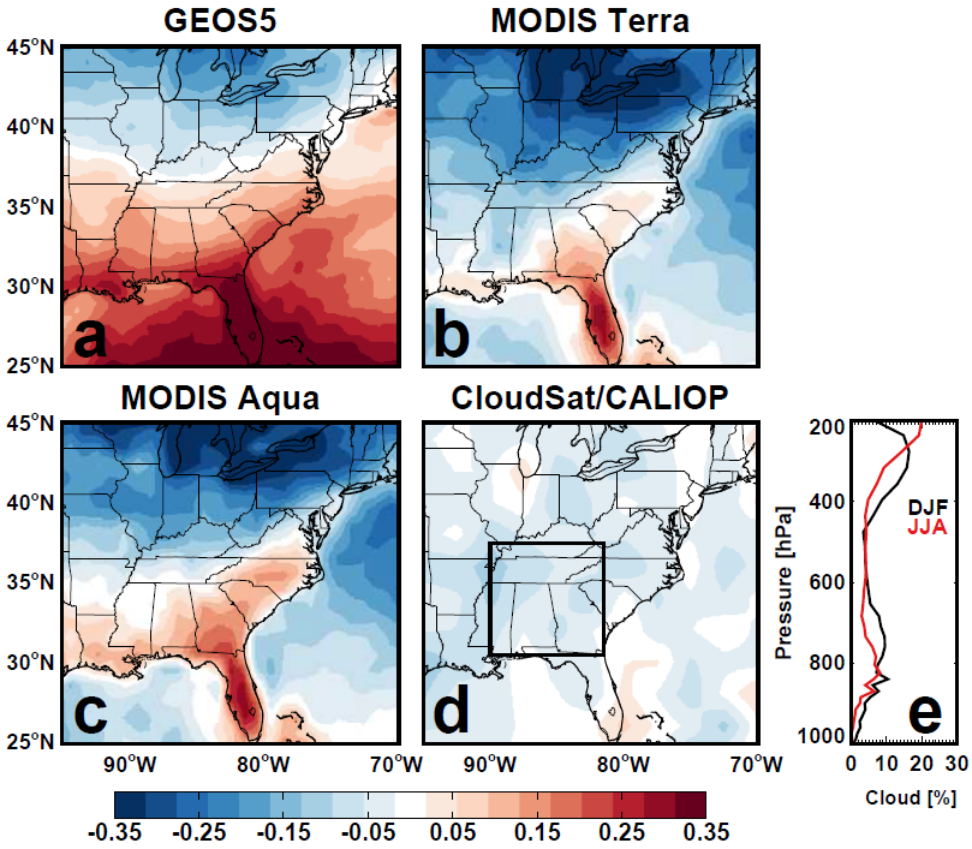

Figure 2.7 Differences (summer-winter) between average cloud fraction for 2006-2009 from (a) GEOS-5 sampled to afternoon overpass, (b) observed by MODIS onboard Terra and (c) Aqua, and (d) observed by CloudSat/CALIOP. Black outline shows region for profiles of average cloud fraction over the SEUS for winter (black) and summer (red) observed by CloudSat/CALIOP in 2009 shown in panel (e).

Furthermore, we see no evidence of a correlation in daily cloud optical depth and aerosol extinction reported by CloudSat and CALIOP over the region. These conflicting characterizations of cloud seasonality, and a lack of correspondence between cloud cover and aerosol loading, suggest that cloud liquid water may not be the limiting factor in summertime

aerosol production, but rather that the oxidation of biogenic VOCs whose emission peaks in summertime, is required to explain the observed aerosol enhancement above the surface layer. It is vital, therefore, to have in situ vertical measurements of aerosol composition in order to fully investigate this hypothesis of increased aqueous aerosol production aloft and determine whether it is organic or inorganic in nature. The upcoming SAS campaign aircraft data will be critical for resolving this issue and determining the aerosol characteristics in the SEUS, thus enabling a better prediction of how aerosol in this region is likely to evolve.

3 Exposure and Premature Mortality Estimates from Model and Satellite-based $PM_{2.5}$

3.1 Introduction

3.1.1 $PM_{2.5}$ and Health

Air pollution is the leading environmentally-related cause of premature mortality worldwide [*oecd.org*]. The World Health Organization (WHO) estimated that outdoor air pollution caused 3.7 million premature deaths in 2012 and indoor air pollution was responsible for about 4.3 million premature deaths. Many epidemiological studies have shown that chronic exposure to particulate matter is associated with an increase in the risk of mortality from respiratory diseases, lung cancer, and cardiovascular disease, with the underlying assumption that a causal relationship exists between PM and health outcomes [*Dockery et al.*, 1993; *Jerrett et al.*, 2005; *Krewski et al.*, 2000; *Pope et al.*, 1995; 2002; 2004; 2006]. This has been shown through single and multi-population time series analyses, long-term cohort studies, and meta-analyses, which combine data from several other studies. The range of the effects found in some of these studies is given in Figure 3.1 (not an exhaustive list), which shows the relative risk (RR) of mortality from different diseases due to a $10\mu g/m^3$ increase in $PM_{2.5}$. The RR is a ratio of the probability of mortality occurring in a population exposed to a certain level of pollution to the probability occurring in a population that is not exposed. Values greater than one suggest an increased risk, while a value of one would suggest no change in risk.

While it seems reasonable that long-term exposure to high pollutant concentrations could have negative impacts on human health and lung and heart function, it may seem implausible that short-term exposure (on the scale of 1-2 days) would lead to a statistically significant increase in a population's mortality. One might instead expect that the physiological effects

would be limited or temporary, such as asthma attacks [*Nastos et al.*, 2010] and noted increases in same day hospital admissions [*Schwartz*, 1999; *Wellenius et al.*, 2005; *Dominici et al.*, 2006; *Bell et al.*, 2008, 2009]. However, studies have also shown increased risk of mortality from cardiovascular disease [*COMEAP*, 2006; *Ostro et al.*, 2006], ischemic heart disease [*Schwartz et al.*, 1996; *Klemm and Mason*, 2003], and respiratory disease [*Klemm and Mason*, 2003; *Pope and Dockery*, 2006] associated with acute exposure; although the increased risk of mortality due to short-term exposure is usually much less than for long-term exposure (Figure 3.2, not exhaustive). Generally, the highest risks are for vulnerable populations, i.e. children, the elderly, people without access to proper healthcare, and people with pre-existing conditions [*Sacks et al.*, 2011].

To estimate the burden of mortality associated with exposure to particulate matter requires information on atmospheric concentrations and health impacts on both daily and annual timescales. It is beyond the scope of this study to assess the short-term and long-term health impacts of air pollution on a person-level, using specific hospital and mortality information and personal air quality monitoring. Instead, here we use "satellite-based" $PM_{2.5}$ for attributing health outcomes in the U.S. and China. These satellite-based estimates offer much needed observational constraints for population-level exposure estimates in regions where surface air quality monitoring is limited [*Crouse et al.*, 2012; *Hystad et al.*, 2012; *Evans et al.*, 2013]. We also compare our calculations of the burden of premature mortality using relative risk ratios derived from short-term events and long-term average concentrations.

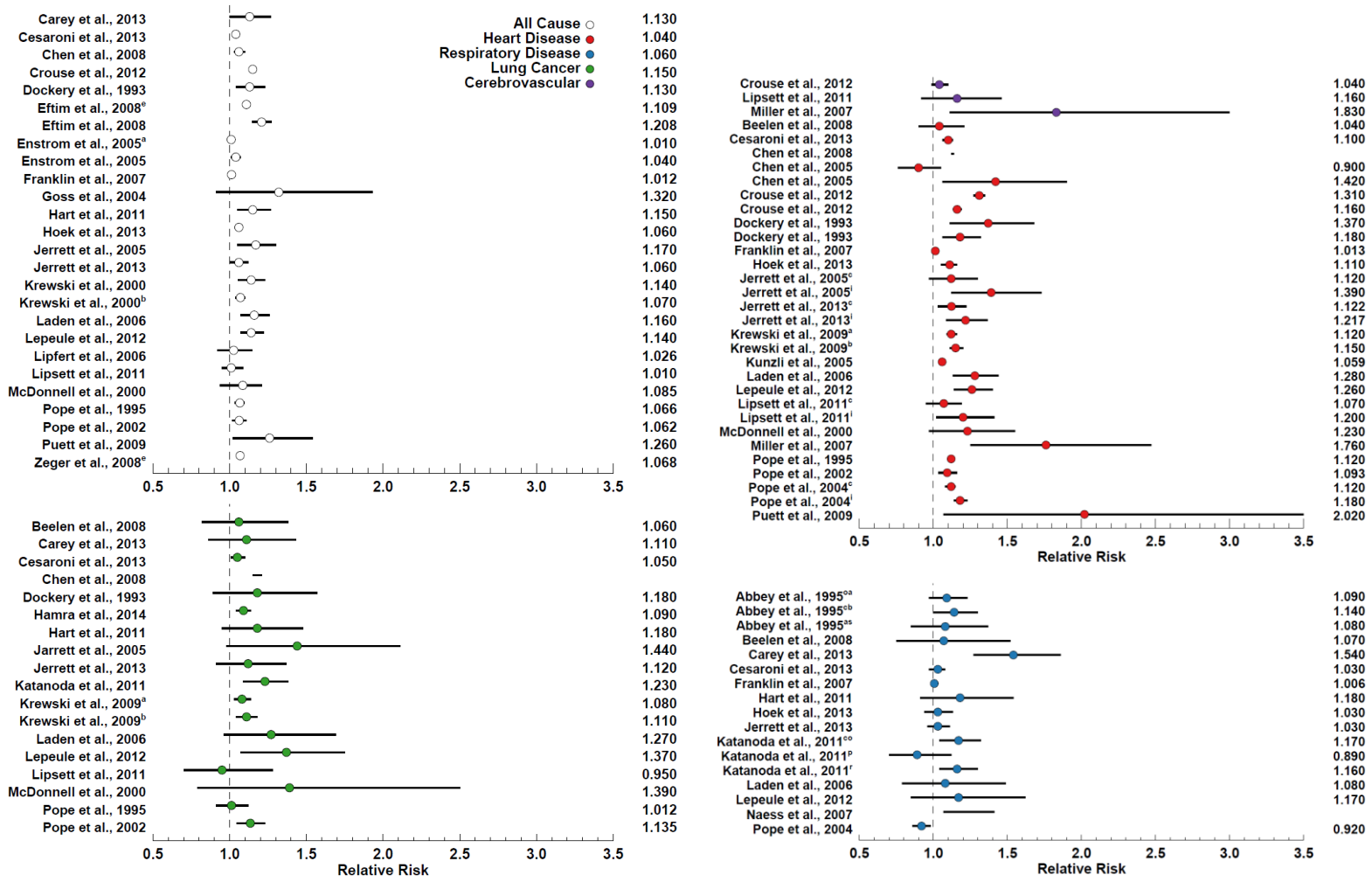

Figure 3.1 Relative risk ratios from select previous studies for mortality due to chronic exposure to $PM_{2.5}$ (given as per $10\mu gm^{-3}$ increase) colored by cause of death.

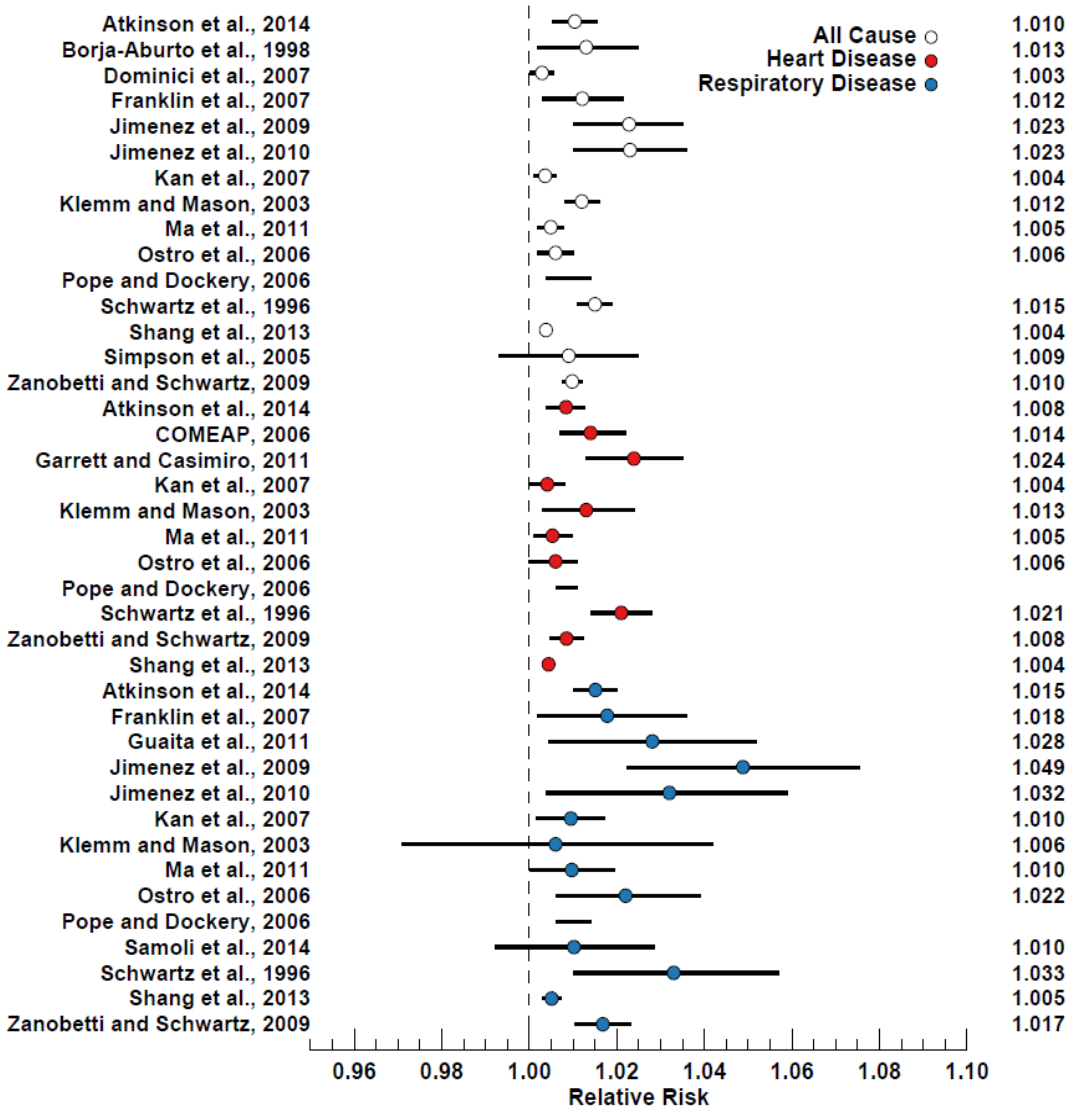

Figure 3.2 Relative risk ratios from select previous studies for mortality due to acute exposure to $PM_{2.5}$ (given as per $10\mu gm^{-3}$ increase) colored by cause of death.

3.1.2 Scarcity of daily long-term surface measurements

One of the main obstacles in attributing health impacts of $PM_{2.5}$ is determining personal exposure. *Jerrett et al.* [2005] suggest personal monitors would be the optimal method, but point out that the financial costs and time-intensiveness limit widespread use. Many studies have instead relied on fixed-site monitors within a certain radius to estimate exposure. However, these

monitoring networks are generally located in urban regions (Figure 3.3) and provide no information on concentration gradients between sites. Fixed-site monitoring networks can also be costly to operate and maintain, and the sampling time period for many of these monitors is often only every third or sixth day. Due to the high spatial and temporal variability in aerosol concentrations, this makes it difficult to determine exposure and widespread health impacts. In an environmental justice study by *Miranda et al.* [2011], the authors highlight that less than 20% of counties in the U.S. currently have daily $PM_{2.5}$ measurements, and data availability is often skewed away from populations that are potentially at greater risks for negative health effects due to particulate exposure (elderly and people without sufficient health care). Worldwide, monitoring networks are even scarcer with many developing countries lacking any long-term measurement.

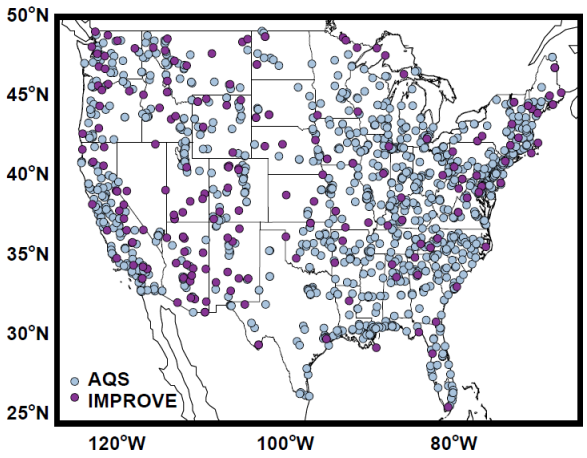

Figure 3.3 IMPROVE (purple) and AQS (blue) $PM_{2.5}$ monitoring sites in the U.S. active during 2004-2011.

3.1.3 Potential to Use Remote Sensing of Aerosols

Satellites can provide a consistent measure of atmospheric constituents with near-global coverage and provide information on gradients between fixed-site monitors. However, these instruments measure the entire atmospheric column and therefore require additional information

in order to be useful for diagnosing surface air quality. "Satellite observations of aerosols" refers to remote measurements of aerosol optical depth (AOD), or the integrated vertical extinction. The relationship between extinction and mass depends on more than simply the amount of aerosols in the atmospheric column.

AOD (τ) can be calculated for a specific wavelength using the extinction efficiency (Q_{ext}), the column mass loading (M), effective radius (r_{eff}), and particle mass density (ρ) in the following equation [*Tegen and Lacis*, 1996]:

$$\tau = \frac{3Q_{ext}M}{4r_{eff}\rho} = M\alpha \qquad (3.1)$$

These variables depend on the aerosol type, aerosol size, and water uptake, and are typically calculated for spherical particles using Mie scattering code. Yet for air quality and exposure assessments, we are most interested in the amount of aerosol mass at the surface. Therefore, in order to use AOD as a measure for changes in the mass concentration at the surface, we need to properly account for the vertical distribution and aerosol optical and physical properties which can be highly variable.

3.1.3.1 Empirical relationships of AOD and $PM_{2.5}$

There have been many previous attempts to translate satellite aerosol optical depth measurements into surface concentrations. The earliest of these used regression coefficients and simple correlations between measurements of aerosol optical depth and available ground observations [e.g. *Wang and Christopher*, 2003; *Engel-Cox et al.*, 2004, 2006; *Schaap et al.*, 2009]. However, as these studies found, there is substantial temporal and spatial variability in the relationship between AOD and surface $PM_{2.5}$, and the relationship is location-specific. The relationship between the column AOD and surface concentration is influenced by the vertical

distribution of aerosols, the aerosol optical properties, meteorology, and other factors which make it difficult to apply any locally-derived relationship to a different region (or time period) where (when) ground-based measurements are lacking.

Therefore rather than attempting to determine unique time- and space-dependent AOD-$PM_{2.5}$ relationships from measurements, several recent studies have instead relied on models to derive surface concentrations from satellite measurements of AOD [*Liu et al.*, 2004, 2007; *van Donkelaar et al.*, 2006, 2010, 2011]. This method uses a chemical transport model to simulate the spatially and temporally varying relationship between AOD and surface $PM_{2.5}$ by accounting for all the aerosol properties and variables and then constraining these results by "real" (i.e. satellite) measurements of AOD. The following sections describe this method and its potential usefulness for exposure assessments.

3.1.3.2 Introduction of Eta (η)

The method which has been previously used to derive surface $PM_{2.5}$ from satellite observations of AOD relies on the following relationship [*van Donkelaar et al.*, 2006, 2010]:

$$PM_{2.5,surface} = \eta \times AOD_{satellite} \quad (3.2)$$

Where the satellite-derived $PM_{2.5}$ is estimated by multiplying the satellite observed AOD by the value η, which is the ratio of model simulated surface $PM_{2.5}$ to simulated AOD at the time of the satellite overpass. For a polar orbiting satellite, this reflects sampling at approximately the same local time world-wide. The η value encompasses the aerosol size, type, extinction efficiency, water uptake and vertical distribution. Figure 3.4 shows an example of this process using the GEOS-Chem model (geos-chem.org).

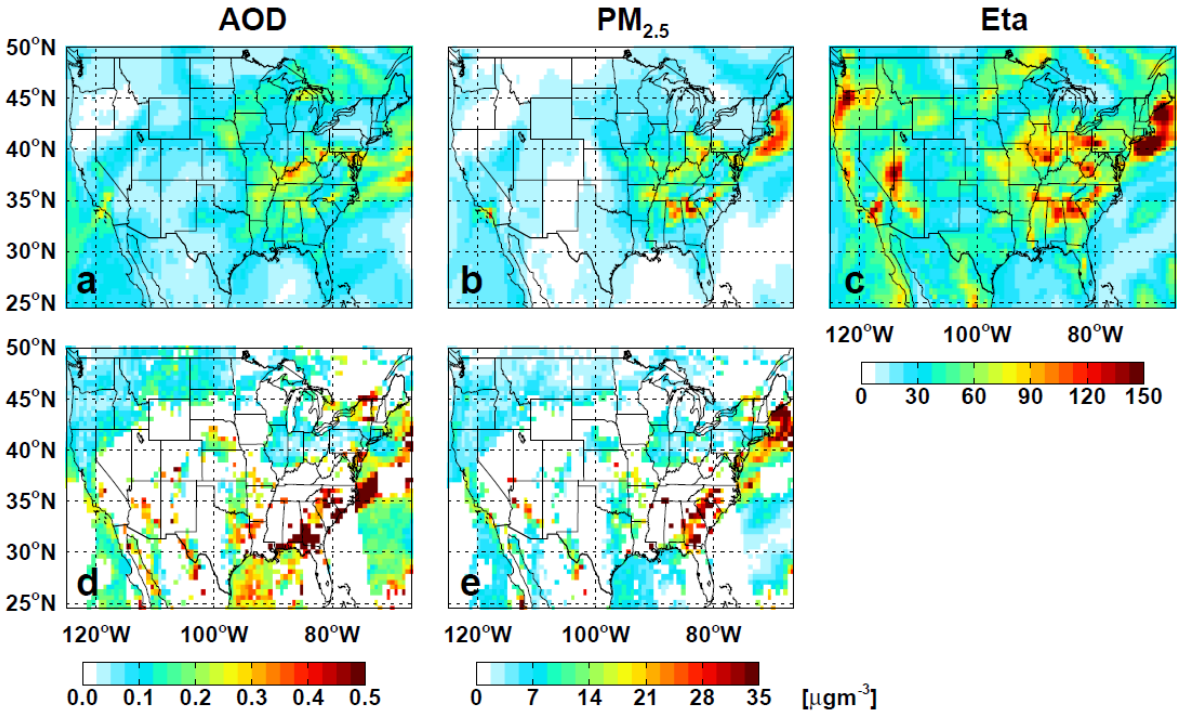

Figure 3.4 GEOS-Chem aerosol simulation of (a) AOD, (b) simulated PM2.5, (c) resulting simulated η for July 3, 2011. Also shown are (d) MODIS satellite AOD, (e) and the resulting satellite-based $PM_{2.5}$ for the same day.

3.1.3.3 Previous Studies methods and results

This method of using GEOS-Chem to "scale" satellite AOD to a surface concentration was first introduced by *Liu et al.* [2004] and relied on MISR observations over the U.S. Their resulting annual MISR $PM_{2.5}$ concentrations are strongly correlated (r=0.79, slope of 0.91) with $PM_{2.5}$ concentrations measured at EPA sites. The method was further developed by *van Donkelaar et al.* [2006] to include observations from MODIS and then extended to produce a global map of satellite-based $PM_{2.5}$ [*van Donkelaar et al.*, 2010]. Several iterations of this product have been created, for example to: examine extreme events with relaxed cloud filtering in MODIS retrievals, include climatological bias corrections to improve daily estimates [*van Donkelaar et al.*, 2012], reduce biases by including vertical profiles from the CALIOP

instrument in an optimal estimation framework [*van Donkelaar et al.*, 2013], and include SeaWIFS observations to determine long-term trends in global $PM_{2.5}$ [*van Donkelaar et al.*, 2014].

The combined conclusions from these studies suggest that while we could rely solely on model estimates of $PM_{2.5}$ for exposure assessment, using the satellite to constrain the results is extremely useful in regions where emissions inventories and model processes are potentially less certain. However, although the satellite-based approach relieves some of the reliance on the model, the derived $PM_{2.5}$ is still dependent on the assumed aerosol distributions and properties. In Chapter 4, we discuss how some of these assumptions can impact the resulting satellite-based $PM_{2.5}$ concentrations and exposure estimates and the sensitivity of the AOD-$PM_{2.5}$ relationship.

This dataset of satellite-derived $PM_{2.5}$ [*van Donkelaar et al.*, 2010] has been made publicly available and been used in a variety of epidemiology and health-related studies. *Fleischer et al.* [2014] use it in conjunction with hospital birth records to show that outdoor $PM_{2.5}$ concentrations are associated with low birth weights. In *Weichenthal et al.* [2014], satellite-derived $PM_{2.5}$ is used along with mortality data from the Agricultural Health Study Cohort to show positive associations with cardiovascular mortality among men (but not women) while *Evans et al.* [2013] use it to estimate global mortality attributable to exposure. Other studies that use this dataset as a proxy for long-term exposure, show that $PM_{2.5}$ enhances the risk of developing diabetes [*Chen et al.*, 2013a] and is associated with hypertension [*Chen et al.*, 2013b], but is not associated with childhood asthma (at the community-level) [*Anderson et al.*, 2012].

3.2 Descriptions of Model and Datasets

3.2.1 GEOS-Chem Model

To determine η values for this analysis, we use simulated AOD and surface concentrations from the global GEOS-Chem model (geos-chem.org). We use v9.01.03 of the chemical transport model, driven by GEOS-5 meteorology, in the nested grid configuration over North America and Asia (0.5°x0.667° horizontal resolution) for 2004-2011. The GEOS-Chem aerosol simulation includes sulfate, nitrate, ammonium [*Park et al.*, 2004], primary carbonaceous aerosols [*Park et al.*, 2003], dust [*Fairlie et al.*, 2007; *Ridley et al.*, 2012], sea salt [*Alexander et al.*, 2005], and secondary organic aerosols (SOA) [*Henze et al.*, 2008]. Most anthropogenic and biofuel emissions over the USA are from the EPA NEI05 inventory [*Hudman et al.*, 2007, 2008] and from *Streets et al.* [2006] over Asia [*Streets et al.*, 2003, 2006; *Zhang et al.*, 2007, 2009b]; although anthropogenic emissions of black and organic carbon over North America follow *Cooke et al.* [1999] with the seasonality from *Park et al.* [2003]. Biogenic VOC emissions are calculated interactively following MEGAN [*Guenther et al.*, 2006], while year-specific biomass burning is specified according to the GFED2 inventory [*van der Werf et al.*, 2006]. Surface dry $PM_{2.5}$ is calculated by combining sulfate, nitrate, ammonium, elemental carbon, organic matter, fine dust, and accumulation mode sea salt concentrations in the lowest grid box.

3.2.2 Satellite Observations

We compare results of derived $PM_{2.5}$ using AOD observations from different space-borne instruments. The Moderate Resolution Imaging Spectroradiometer (MODIS) measures radiances at 36 wavelengths and flies on both the EOS-Terra and EOS-Aqua platforms. MODIS-Terra has

a morning overpass of ~10:30 LT while MODIS-Aqua flies as part of the A-Train constellation with an equator crossing time of ~13:30 LT. There are two different algorithms to separately retrieve AOD over the ocean and land in order to account for differences in surface properties [*Remer et al.*, 2002b, 2005; *Levy et al.*, 2013]. Because surface characteristics are determined from a look up table and not directly from measurements, specific regions (such as bright deserts, mountainous regions, highly urbanized areas, and snow fields) have known biases. For this work we use Level 2, Collection 6, Atmosphere Products for Aqua as well as Level 2, Collection 5 for Terra and Aqua, as has been used in previous studies. These Level 2 products are calculated along the orbit of the MODIS swath with a 10 km x 10 km resolution. We filter these data for observations of high quality and low cloud fraction and remove observations with high AOD (>2.0), as cloud contamination causes known biases in the AOD.

Also onboard EOS-Terra is the Multi-angle Imaging SpectroRadiometer (MISR) instrument [*Martonchik et al.*, 2002, 2009; *Diner et al.*, 2005]. MISR employs nine different cameras to make multi-angle radiance measurements in four spectral bands (visible to near infrared). Retrievals have a nominal spatial resolution of 17.6 km and provide global coverage in approximately 4-7 days. Additionally, the MISR algorithm can provide information on particle size, shape (spherical vs. non-spherical), and type.

3.2.3 Surface Observations

For surface-based observations of AOD, we use measurements from the global AErosol RObotic NETwork (AERONET) of sun photometers. AOD and aerosol properties are recorded at eight wavelengths in the visible and near-infrared (0.34-1.64μm) and are often used to validate satellite measurements [e.g. *Remer et al.*, 2005]. AERONET AOD has an uncertainty of 0.01-

0.015 [*Holben et al.*, 1998]. For this work, we use hourly Version 2 Level 2 measurements sampled to two hour windows around the times of the satellite overpasses. We also perform a least-square polynomial fit in order to interpolate measurements to 550 nm.

We also use surface measurements of $PM_{2.5}$ mass from both the Interagency Monitoring of Protected Visual Environments (IMPROVE) and the EPA Air Quality System (AQS) database. The IMPROVE network measures $PM_{2.5}$ over a 24-hour period every third day and these measurements are then analyzed for concentrations of fine, total, and speciated particle mass [*Malm et al.*, 1994]. We use the reconstructed fine mass (RCFM) values (unlike in Chapter 2), which are the sum of ammonium sulfate, ammonium nitrate, soil, sea salt, elemental carbon and organic matter. While water can contribute significantly to aerosol mass in the atmosphere, filters are equilibrated at 35% relative humidity prior to weighing so that the RCFM is generally considered a measure of the dry particle mass. However, aerosols can still contain water even at this humidity, with water accounting for up to 20% of the particle mass [*Khylstov et al.*, 2005]. This could explain some of the discrepancy between the model-simulated concentrations and surface measurements.

3.2.4 Method and Datasets for Determining Attributable Premature Mortality

To estimate the burden of premature mortality due to $PM_{2.5}$ exposure, we rely on equations (3.3) and (3.4), where the change in relative risk (RR, given as per 10 $\mu g/m^3$) linearly depends on the concentration (C, in $\mu g/m^3$). Here we assume that this concentration represents dry aerosol mass. The attributable fraction of mortality due to $PM_{2.5}$ exposure then depends on this relative risk value and the baseline mortality (M_b). In this equation C_0 is the background or "threshold" concentration, or the concentration below which there is no effect. As studies have

shown that there is no concentration level below which there is no adverse health effect for PM (e.g. *Pope et al.*, 2002), in our initial analysis, we assume this value is zero. However, other studies often set C_0 to the value of the lowest measured concentration observed in the epidemiology study from which the RRs are derived [e.g. *Evans et al.*, 2013]. We will discuss the impact of using a threshold on our results.

$$\Delta RR = (RR-1) \times (C-C_0)/10 \quad (3.3)$$

$$\Delta M = M_b \times \Delta RR / (\Delta RR+1) \quad (3.4)$$

To determine baseline mortalities in the U.S. for cardiovascular disease, lung cancer, and respiratory disease, we use crude death rates for all ages from the Center for Disease Control (cdc.gov) for each year and each state which we then multiply by the population to obtain the total baseline mortality. An example of M_b for the year 2004 is given in Appendix 1, with noted variability in death rates between states.

Mortality values are not as readily available for China, so instead of using province-level baseline mortality, we rely on country-wide values. Therefore, variations in M_b (Appendix 1) are due to variations in population and not death rates. Because studies have shown variability in death rates in different locations, this method of using country-wide averages could also be a source of error in our estimates. In order to account for some variability, we use a population threshold in order to distinguish between urban and rural region for lung cancer mortality [*Chen et al.*, 2013c].

For population data, we use the Gridded Population of the World, Version 3 (GPWv3), created by Center for International Earth Science Information Network (CIESIN) and available from the Socioeconomic Data and Applications Center (SEDAC). This gridded dataset has a

native resolution of 2.5 arc-minutes (~5km at the equator) and provides population estimates for 1990, 1995, and 2000, and projections (made in 2004) for 2005, 2010, and 2015.

We temporally interpolate between available years to get population estimates for years not provided. Population density for China and the United States for the year 2000 are shown in Figure 3.5 along with the projected change in population density by the year 2015, illustrating continued growth of urbanized areas (at the expense of rural regions in China). From these population estimates, we are then able to estimate the number of people exposed to certain pollution levels by comparing to concentration data that we grid to the same resolution as the population data.

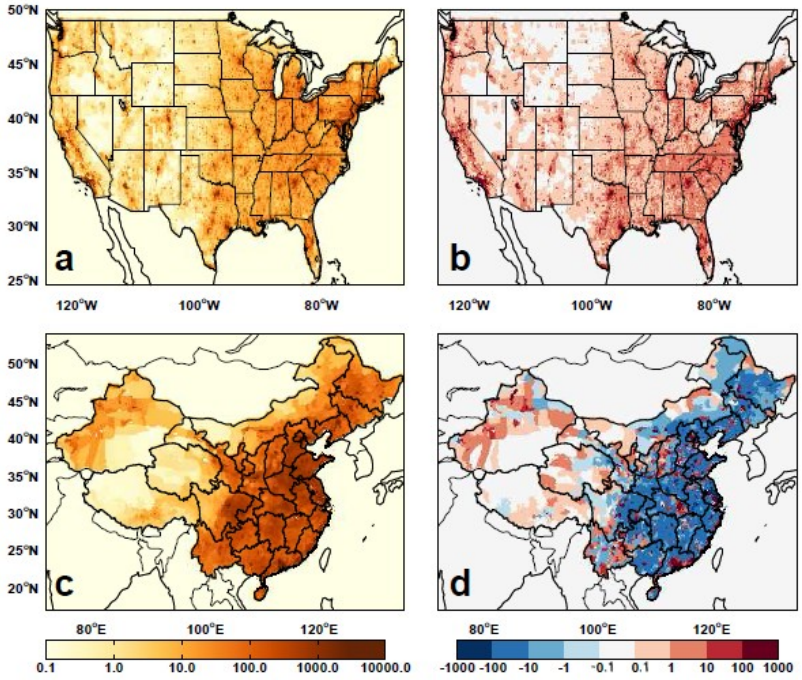

Figure 3.5 Population density [per km^2] for the year 2000 from the GPWv3 data for (a) the continental U.S. and (c) China. The projection for increase in population density by the year 2015 for (b) the continental U.S. and (d) China.

As we are using population and concentration data from the past decade, the results we will show here are the estimates of deaths incurred as opposed to deaths avoided if concentrations are reduced to some future value.

3.3 Unconstrained model simulations

We begin our analysis by comparing the unconstrained model simulations to observations. This provides examples of where and when satellite observations could potentially help improve model representations of surface air quality. Additionally, because η depends on the model simulation of AOD and $PM_{2.5}$, it is important to determine if the model skills are similar for both.

3.3.1 Comparisons with surface $PM_{2.5}$ at AQS and IMPROVE sites

Previous studies have generally shown good agreement between measurements and GEOS-Chem simulations of $PM_{2.5}$ [e.g. *Park et al.*, 2003, 2004]. In Figure 3.6, we show the long-term average of $PM_{2.5}$ at AQS and IMPROVE sites in the U.S. overlaid on simulated values. We focus our analysis here on the United States, where long-term observations are more readily available to test our simulation. In general, GEOS-Chem agrees better with measurements at IMPROVE sites, possibly because these are located in rural regions where simulated values will not be as impacted by the coarse model resolution. There are noted discrepancies in California (as also noted by *Schiferl et al.*, 2014) and the Appalachia/Ohio River Valley region where the model is biased low. These discrepancies in the long-term average can impact chronic exposure estimates. However, because we are also interested in short-term exposure, it is important to note that there can be larger discrepancies at daily timescales. We also show the 95^{th} percentile $PM_{2.5}$ concentrations in Figure 3.6, which show that the highest observations are often much higher than the model, suggesting that the model does not capture

extreme events which are potentially associated with elevated health risks. The greatest discrepancies in the 95th percentile comparisons are generally in urban regions, but the model also fails to capture extremes at some sites in the northwest likely due to fires. In general, although the model often simulates slightly higher mean concentrations (Figure 3.6), the distribution of daily concentrations in the observations has longer tails, in part due to the model resolution's inability to capture very fine scale features.

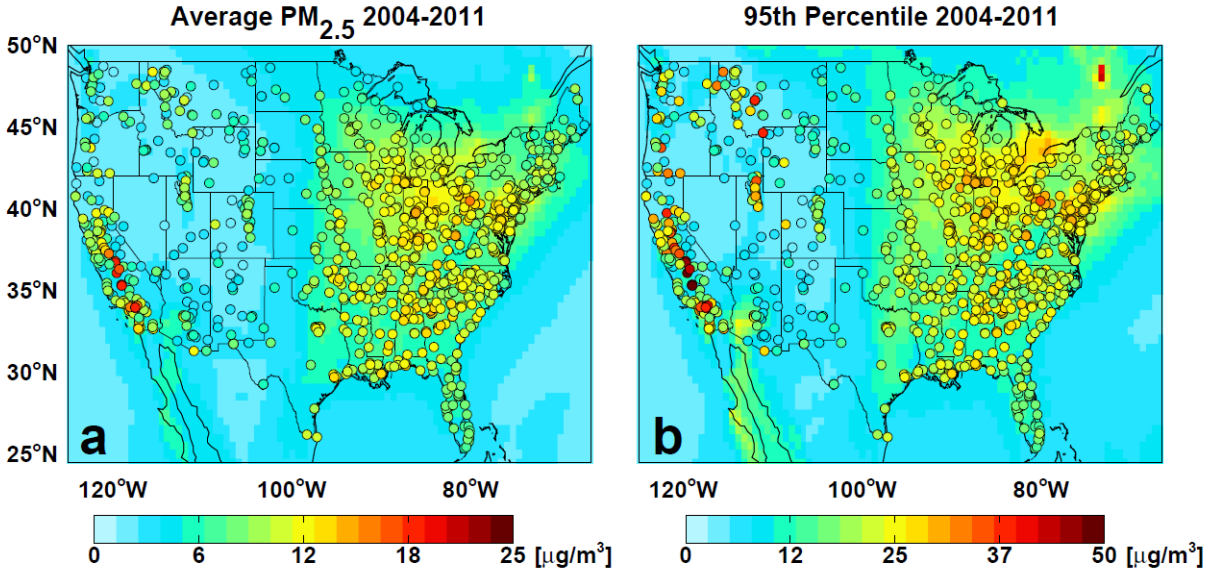

Figure 3.6 GEOS-Chem simulated (a) average and (b) 95th percentile in surface $PM_{2.5}$ mass for years 2004-2011 overlaid with measurements at IMPROVE and AQS sites.

We also compare the composition of aerosol mass simulated in the model and measured at different network sites in the United States. In Figure 3.7, we see that GEOS-Chem model captures much of the variability in different dominant aerosol types across regions. It does underestimate organic matter [*Heald et al.*, 2011; *Pye et al.*, 2010; *Chen et al.*, 2014], but tends to overestimate nitrate, especially in the Midwest [*Heald et al.*, 2012]. The discrepancies in the specific components often compensate for each other, so that the mean difference in the total

mass is closer to zero. Using satellite observations to constrain the model could potentially reduce the bias in total $PM_{2.5}$ mass; however, if the model simulation of a specific species is incorrect and it has a different vertical profile, this could bias the η used to derive the satellite-based $PM_{2.5}$.

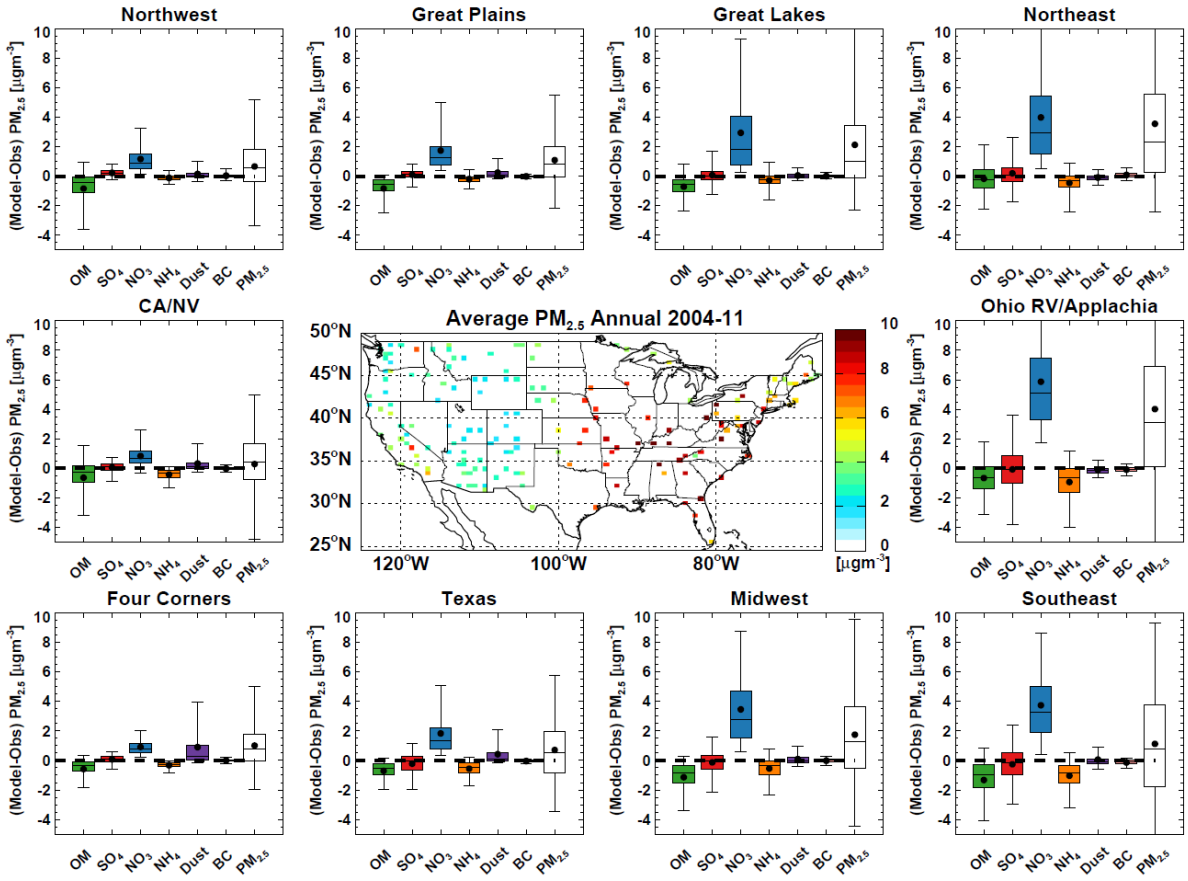

Figure 3.7 Box and whisker plots of the difference between simulated and measured $PM_{2.5}$ mass for different regions of the U.S. Whiskers denote the 5th and 95th percentiles, box outlines 25th and 75th percentiles, middle line denotes the median and the dot represents the mean.The location of sites and the long-term average $PM_{2.5}$ at the sites is shown in the middle map.

Robust long-term measurements of $PM_{2.5}$ speciation are not routinely (publicly) available for China in order to repeat these comparisons for that region [*Chan and Yao*, 2008]. However, previous studies have shown that in general the model is able to capture the spatial distributions

and seasonality in sulfate, nitrate, ammonium, and black carbon; although there is a slight low bias in annual sulfate concentrations, a low bias in ammonia, and a high bias in nitrate, especially in East Asia [*Yang et al.,* 2013a; 2013b, 2014]. These discrepancies have been attributed to model underestimates in deposition (specifically with regards to large precipitation events), poor representation of pollution plumes, and uncertainties in NO_x, SO_2, and NH_3 emissions [*Wang et al.,* 2013a; 2013b, 2014]. Additionally, *Wang et al.* [2013a; 2014] show that comparisons are much worse on shorter timescales, and the model has significant difficulties simulating high concentrations on days with extreme pollution.

3.3.2 Comparisons of AOD

For AOD, we compare the simulation to that of observations from the MISR and MODIS spaceborne instruments as shown in Figure 3.8 (using spatially interpolated values, see Appendix II for original). Here we show the normalized mean bias (Equation 3.5) in daily values, where *M* is the model value and *O* is the observation (in grid box *i*).

$$\text{NMB} = \left(\frac{\sum_1^N (M_i - O_i)}{\sum_1^N O_i}\right) \times 100\% \quad (3.5)$$

In the U.S., the model often over predicts AOD over the Ohio River Valley, and under predicts AOD in the west (although the satellite observations could also be biased high in the western U.S. due to high surface reflectivities). In China, results are similar, in that the model often under predicts AOD over the desert regions of the west and over predicts AOD in the urbanized central, North China Plain, and east (Sichuan Basin). While the magnitude of AOD differs between instruments, the spatial pattern of AOD and the differences compared to GEOS-Chem are similar. These differences provide us with insight into where the satellite-based $PM_{2.5}$ will

also be different, although we point out that these are the normalized mean biases of daily values and are different than comparisons of annual means.

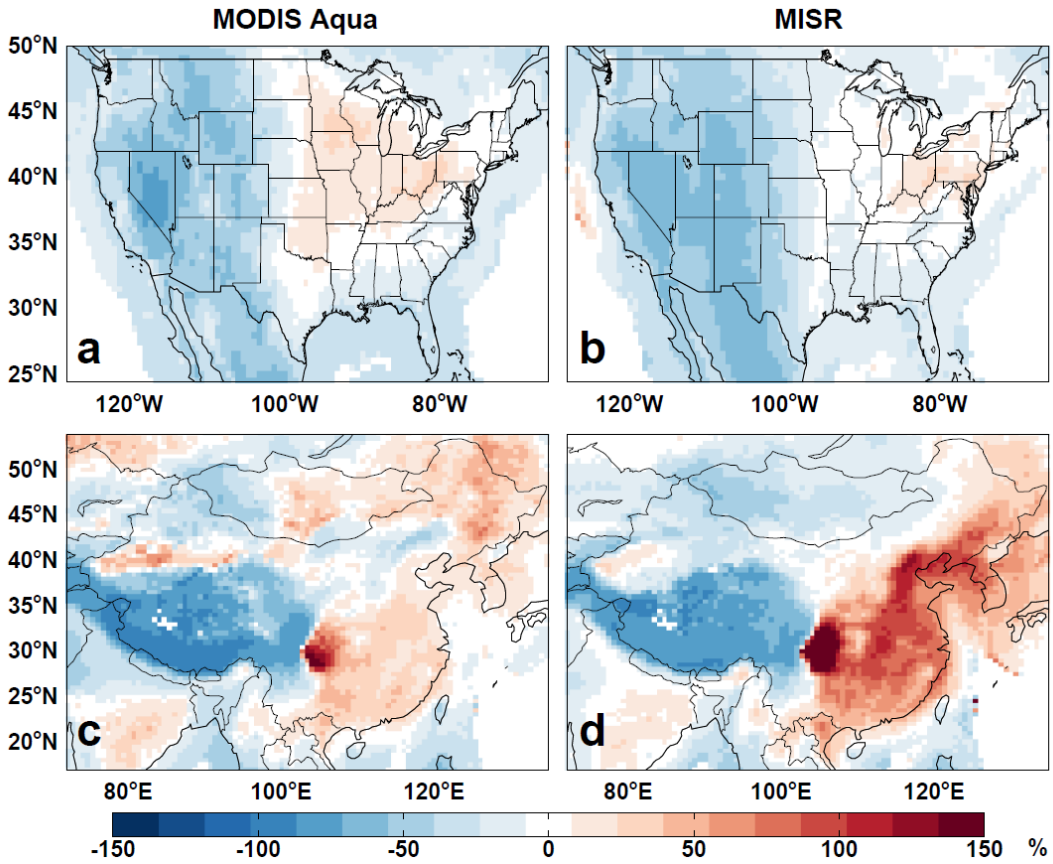

Figure 3.8 Normalized mean bias [%] between daily GEOS-Chem simulated AOD and (a, c) MODIS Collection 6 Aqua and (b, d) MISR for 2004-2011 over the U.S. and China, respectively.

We also compare the trend in AOD over the time period of 2004-2011 simulated in the model and compared to the satellite observations (Figure 3.9). GEOS-Chem simulates a significant decreasing trend in AOD over the eastern U.S. which is also noticeable in the satellite observations. This decreasing trend can be attributed to declining SO_2 emissions in the U.S. as noted in *Leibensperger et al.* [2012]. The AOD from MODIS Aqua tends to suggest some increasing trends in the western U.S. (as with IMPROVE concentrations, not shown); however,

AOD has greater uncertainty in this region (larger biases compared to AERONET) and can be heavily influenced by large fires in the time series. Additionally, due to the limited time period, the only significant trend in the satellite observations is in the northeastern U.S.

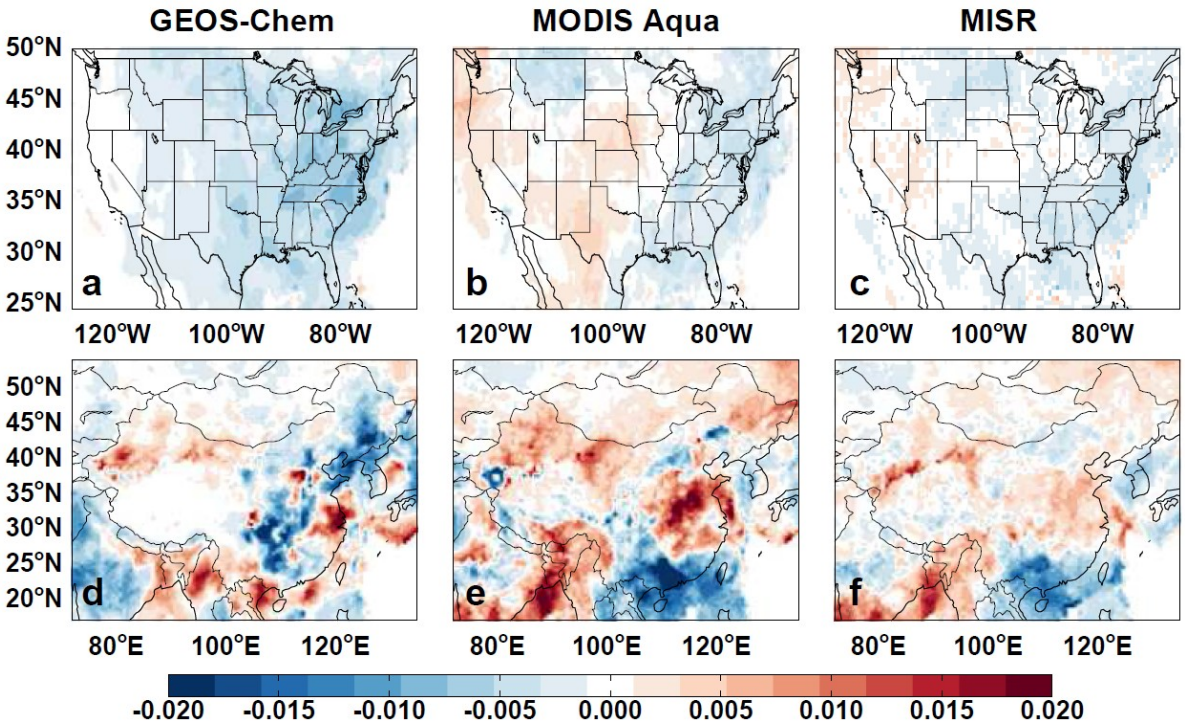

Figure 3.9 Trends in AOD [AOD/year] for 2004-2011 (a,c) as simulated by GEOS-Chem and as observed by (b,e) MODIS Aqua and (c,f) MISR over the U.S. and China, respectively.

In China, GEOS-Chem suggests AOD is decreasing over the northeast (Jilin, Liaoning) and in central China (Hunan and Hubei) while increasing over the northwest (Xinjiang and Inner Mongolia) and the region surrounding Shanghai. The satellites agree with the increasing AOD simulated in the model over the northwest (due to dust), but suggest a stronger increase in AOD over the East (Shandong, Henan, etc.) and a strong decrease over southeast China (Pearl River Delta Region). However, again because this only covers a time period of eight years, most of these trends over China are not significant with the exception of the satellite observed trend over

the southeast. *Luo et al.* [2014] also examine trends in AOD from MODIS, but for the time period of 2000-2010 and note the increase in AOD over eastern and central China, however, they find a strong positive trend in AOD over the Pearl Delta Region. This discrepancy is likely due to the different time periods as *Zhong et al.* [2013] found increases in AOD from 2000 to 2007, followed by a decrease through 2011 over the Pearl River Delta Region, corresponding to similar trends in regional SO_2 emissions. If only 2004-2008 is used to determine a trend, the satellites suggest widespread positive trends over China. Additionally, emissions in China have been variable over this period in general [*Lu et al.*, 2011; *Zhao et al.*, 2013] and AOD is strongly variable between years, with some regions experiencing interannual variability in AOD of greater than 20% [*Luo et al.*, 2014]. It is difficult to represent the rapidly changing emissions in China in the model, which could suggest that it is an ideal place to employ satellite constraints.

3.3.3 Reducing Sampling biases

Estimating daily satellite-derived $PM_{2.5}$ for short term exposure estimates is hindered by the lack of available satellite AOD observations. MODIS provides 1-2 day coverage in the absence of clouds while MISR takes approximately 4-7 days for global coverage. Filtering for quality assurance can also reduce available data. Therefore, it has been previously suggested that interpolation schemes can help fill in data gaps. We assess here the impact of these different methods on both the column AOD and surface $PM_{2.5}$.

For this analysis, we rely on the model simulation. We first sample the model to the time and location of satellite observations, then apply the interpolation scheme and finally compare these new values to the original model simulation in order to determine the method that best represents the "real" values. We test 5 different methods. The first and second methods are

simply grid area means, where a missing value is replaced by the average of available "observations" in the (1) 8 surrounding (3x3 region) grid boxes or (2) 24 surrounding (5x5 region) grid boxes. We also test using an inverse distance weighted mean (IDW, data points closer to the grid points have more effect than those which are further away) [*Shepard*, 1968] with a search cutoff of a four grid box radius, a kriging method (where both data points and the spatial variance are used to determine trends) [*Isaaks and Srivastava*, 1989], and a triangulation method in which each interpolant is a linear combination of the three vertices of its enclosing Delaunay triangle and their vertices [*Lee and Schachter*, 1980]. We also compare using the full interpolated map generated from these three methods against only using interpolated values when there are no valid observations.

Our goal is to find a method which reduces the sampling bias and increases data volume. Simply sampling to available MODIS observation days leads to an underestimate of the annual average AOD over much of the U.S., due to the cloud filtering and the removal of observations with high AOD values from MODIS. The interpolation methods reduce the overall sampling bias in the yearly AOD as shown in Table 3.1. Regionally, there is still a low bias in the Great Plains and Great Lakes regions and a high bias in the Western U.S. is introduced (due to the limited number of observations and the available observations occurring over regions of high AOD, ie. California). For calculating the annual average AOD, the kriging, IDW, and triangulation methods produce very similar results. Because interpolation methods can smooth out finer scale variability, we find that using the full interpolated maps produces larger biases than using interpolated values only when there are no valid observations. We find the lowest average bias when using simple area means. However, we note that this method increases the amount of available days by much less than the other methods.

Table 3.1 Normalized Mean Biases (NMB) in annual average AOD and $PM_{2.5}$ for different interpolation/data filling methods over the GEOS-Chem North America nested grid region and the resulting number of available days for 2009.

Method	Yearly NMB [%]		# Observation Days (Average)
	AOD	$PM_{2.5}$	
Sampled	-4	2.2	91
3x3 Simple Mean	0	0.3	132
5x5 Simple Mean	0	0.6	148
Inverse Distance Weighted Mean	4.0	15.6	261
Kriging	4.5	14.4	249
Triangulation	5.2	18.7	327
Original values + IDW	3.5	14.7	262
Original values + Kriging	3.8	13.5	251
Original values + Triangulation	4.8	14.2	327

We also compare interpolation schemes for $PM_{2.5}$. From a methodology standpoint, this allows us to assess whether it is better to interpolate AOD before deriving $PM_{2.5}$ values or interpolate derived $PM_{2.5}$ to provide greater coverage. In general, there is greater temporal and spatial variability in $PM_{2.5}$ compared to AOD, thus, interpolating $PM_{2.5}$ concentrations creates greater biases than interpolating AOD. From these comparisons, we determine that using an inverse distance weighted mean is a good balance between reducing the bias and increasing the amount of available days.

We determine the impact this data filling method has on satellite-derived $PM_{2.5}$, by using the new (interpolated) AOD values from the model along with the original model η value to compute a derived $PM_{2.5}$ value which we then compare to the original $PM_{2.5}$ value. The resulting uncertainty in derived $PM_{2.5}$ due to both sampling and interpolating the AOD over the U.S. is shown in Figure 3.10 along with the original uncertainty just due to sampling. We see that interpolating the AOD generally reduces the bias associated with satellite sampling, particularly

in the Northern U.S. When translated to mortality estimates from chronic exposure, this interpolation can lead to a ~1% difference in our estimates.

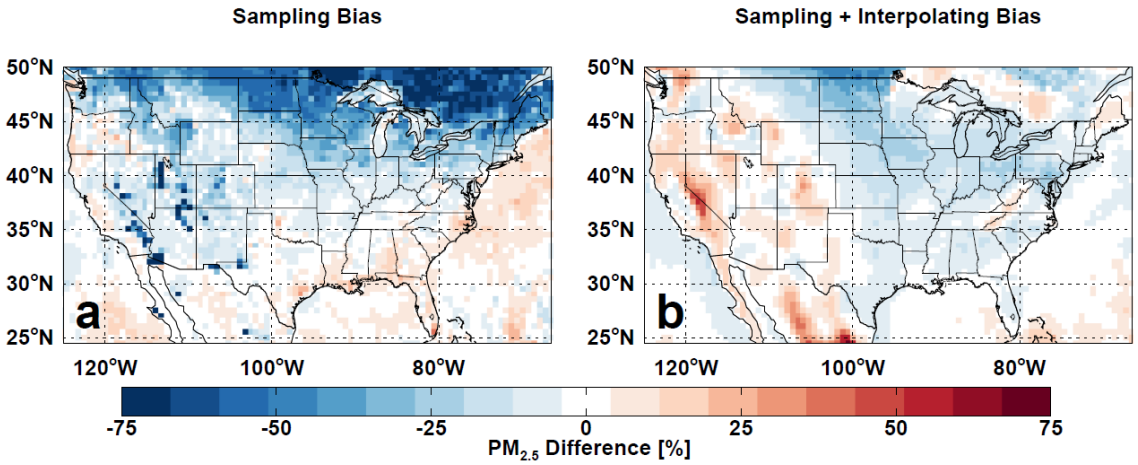

Figure 3.10 Bias in annual average $PM_{2.5}$ from sampling to satellite observation days (a) and the bias from sampling AOD to satellite observation days, interpolating AOD, and then deriving $PM_{2.5}$ (b).

3.4 Satellite-derived $PM_{2.5}$ estimates

In Figure 3.11, we show the long-term average of satellite-derived $PM_{2.5}$ for the U.S. and China using MODIS Aqua Collection 6. In the following sections, most of our results will be determined using Collection 6; but reference and comparisons will be made to other instruments as a measure of uncertainty. We sample GEOS-Chem to days and grid boxes with valid satellite observations to calculate the η used to translate the AOD to surface $PM_{2.5}$. Compared to the unconstrained model, the satellite-based estimate suggests that concentrations should be higher over much of the western U.S., particularly over California, Nevada, and Arizona. In China, the satellite-derived $PM_{2.5}$ is higher in Eastern China, around Beijing, Tianjin, and Shanghai, but less in many of the central provinces. *Yang et al.* [2013b] noted that the model underestimates $PM_{2.5}$ in the Sichuan basin, suggesting that the satellite-based estimate could reduce the bias in this province.

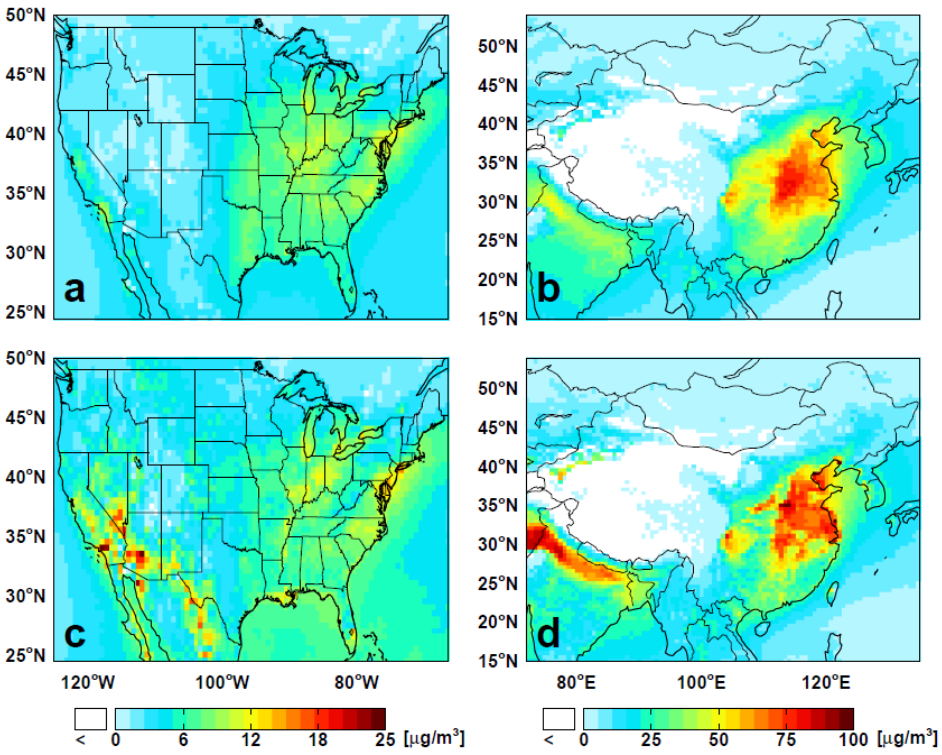

Figure 3.11 Long-term average (2004-2011) unconstrained model simulation of $PM_{2.5}$ for the (a) continental U.S. and (b) China, along with the satellite-derived $PM_{2.5}$ for the (c) continental U.S. and (d) China.

3.4.1 Estimates of Chronic Exposure to $PM_{2.5}$ and associated Premature Mortality

We compare $PM_{2.5}$ exposure estimates using both model and satellite-based approaches and in the U.S. and China. Figure 3.12 shows the distribution of population exposed to different long-term average concentrations. Not only are concentrations higher in China, but a greater fraction of the population is exposed to harmful concentrations. This figure suggests that 12% (with model concentrations, 18% with MODIS-Aqua, 14% with MISR) of the U.S. population has experienced a long-term average concentrations above 12 μgm^{-3} (national air quality standard), but 0% has experienced average concentrations of 35μgm^{-3} or greater. In China, 75% (with model concentrations, 78% with MODIS-Aqua, 50% with MISR) of the population was exposed to average concentrations of 35 μgm^{-3} (proposed national standard for China) or greater,

which suggests that should these standards go into effect in 2016 as anticipated, many regions will be out of compliance. However, this also emphasizes the greater need for significant measures to reduce poor air quality as it affects such a large percentage of the population. The lower exposure estimates for China obtained with MISR are due to a low bias (compared to MODIS) in AOD over the northeast (around Beijing) where the population is high [*Cheng et al.,* 2012].

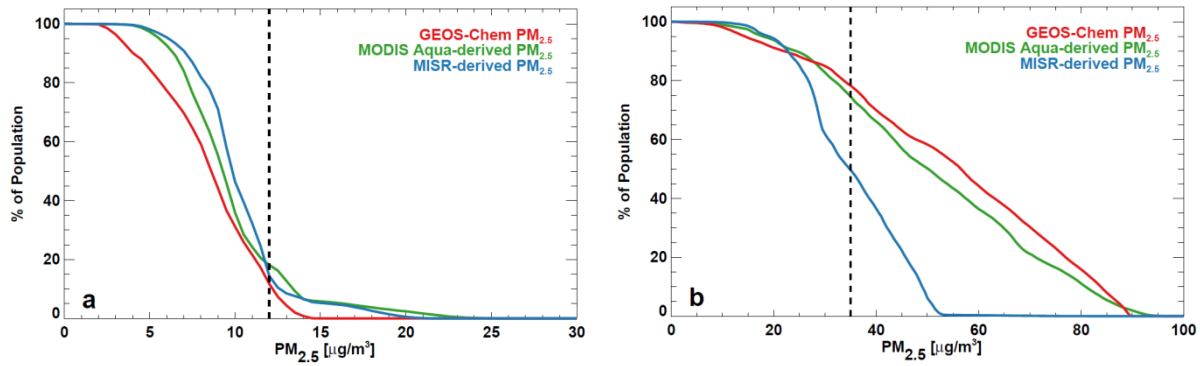

Figure 3.12 Estimates of the percent of the population of the (a) U.S. and (b) China that are exposed to different annual $PM_{2.5}$ concentrations. Colors denote method for calculating concentrations and dashed lines denote national air quality standards.

We also calculate the fraction of mortality associated with heart disease, lung cancer, and respiratory disease attributable to chronic exposure determined using annual average concentrations. Results are shown in Figure 3.13 with error bars denoting the range determined from the confidence intervals of the risk ratios. For our initial comparisons, we use risk ratios for cardiovascular and lung cancer premature mortality due to chronic exposure determined in *Krewski et al.* [2009], which is an extended analysis of the American Cancer Society study, and for respiratory disease, from *Laden et al*. [2006] which is an updated and extended analysis of the Harvard Six Cities study. We use the *Krewski et al.* [2009] risk ratios as they are widely used in similar studies due to the large study population with national coverage, 18 year time span,

and extensive analysis of confounding variables (ecological covariates, gaseous pollutants, weather, medical history, age, smoking, etc.). Using these same risk ratios makes our results more directly comparable to these other studies. Chronic exposure to $PM_{2.5}$ is most strongly associated with premature mortality due to heart disease. Additionally, baseline mortality rates are also higher for heart disease, resulting in most of the premature mortalities attributed here to exposure being linked to heart disease in both the U.S. and China.

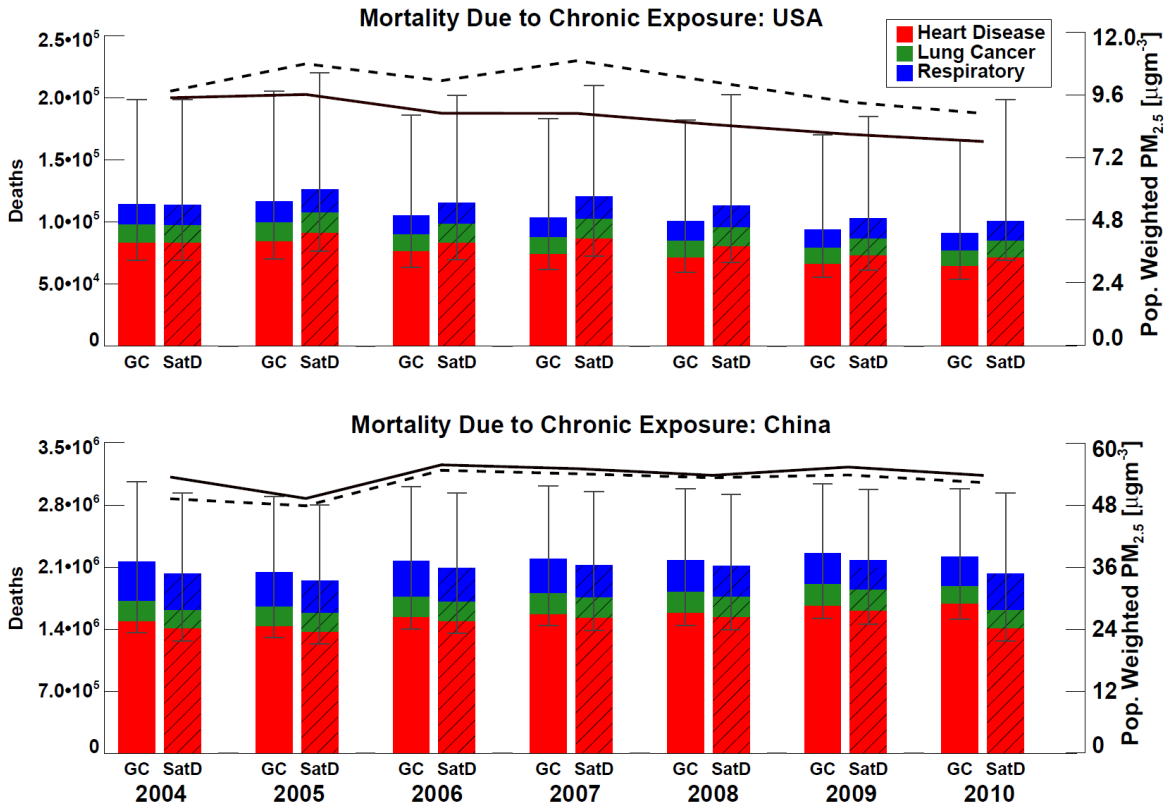

Figure 3.13 Time series for (a) the U.S. and (b) China using the unconstrained model estimates and the MODIS Collection 6 satellite-derived $PM_{2.5}$ to determine the annual mortality due to chronic exposure. Colors denote cause of death: cardiovascular (red), respiratory (blue), and lung cancer (green). Overlaid black lines are the population-weighted mean concentration from the unconstrained model (solid) and satellite based (dashed). Whiskers denote the confidence intervals associated with the risk ratios.

In Figure 3.13, we overlaid the country-wide population-weighted $PM_{2.5}$ concentrations. In the U.S., the model suggests that concentrations have generally decreased over time following declines in anthropogenic emissions, and consequently, the premature mortality rates have also decreased. This is less obvious with the satellite-based $PM_{2.5}$ concentrations, where high exposures in 2007 associated with large dust storms in the spring and a long stagnation event in the winter interrupt this downward trend.

In China, the trend in the population-weighted $PM_{2.5}$ is more complicated. Emissions of SO_2 in China steadily increased until about 2010, when they began to decline. However, emissions of black and organic carbon have continued to increase [*Lu et al.*, 2011] along with nitrogen oxides [*Zhao et al.*, 2013], and potentially ammonia, although ammonia emissions trends in China are fairly uncertain [*Wang et al.*, 2013b]. Even though sulfate contributes a significant portion of the aerosol mass in China, the decline in SO_2 emissions has not been reflected as strongly in surface $PM_{2.5}$, likely due to increases in the sources of these other pollutants. *Wang et al.* [2013b] suggest that in many places in China the decrease in sulfate aerosols has simply been met with an increase in nitrate aerosols. Additionally, the population has continued to grow (by 5% from 2004 to 2011), especially in urban regions with high $PM_{2.5}$ concentrations, which can also impact the population-weighted mean.

Maps of the average annual mortalities attributable to chronic $PM_{2.5}$ exposure are shown in Figure 3.14 (total number of deaths in a grid box). In conjunction with population density shown in Figure 3.5 and concentrations shown in Figure 3.11, these mortality estimates highlight the significant risk facing growing urban areas. The difference between mortality estimates using model or satellite concentrations is approximately 9% for the U.S. and 4% for China on a nationwide basis, although this figure shows that regionally the difference can be much greater.

These values suggest that chronic $PM_{2.5}$ could be responsible for 2-3% of the mortalities in the U.S. and 15-20% in China. We test the sensitivity of these estimates in Chapter 4.

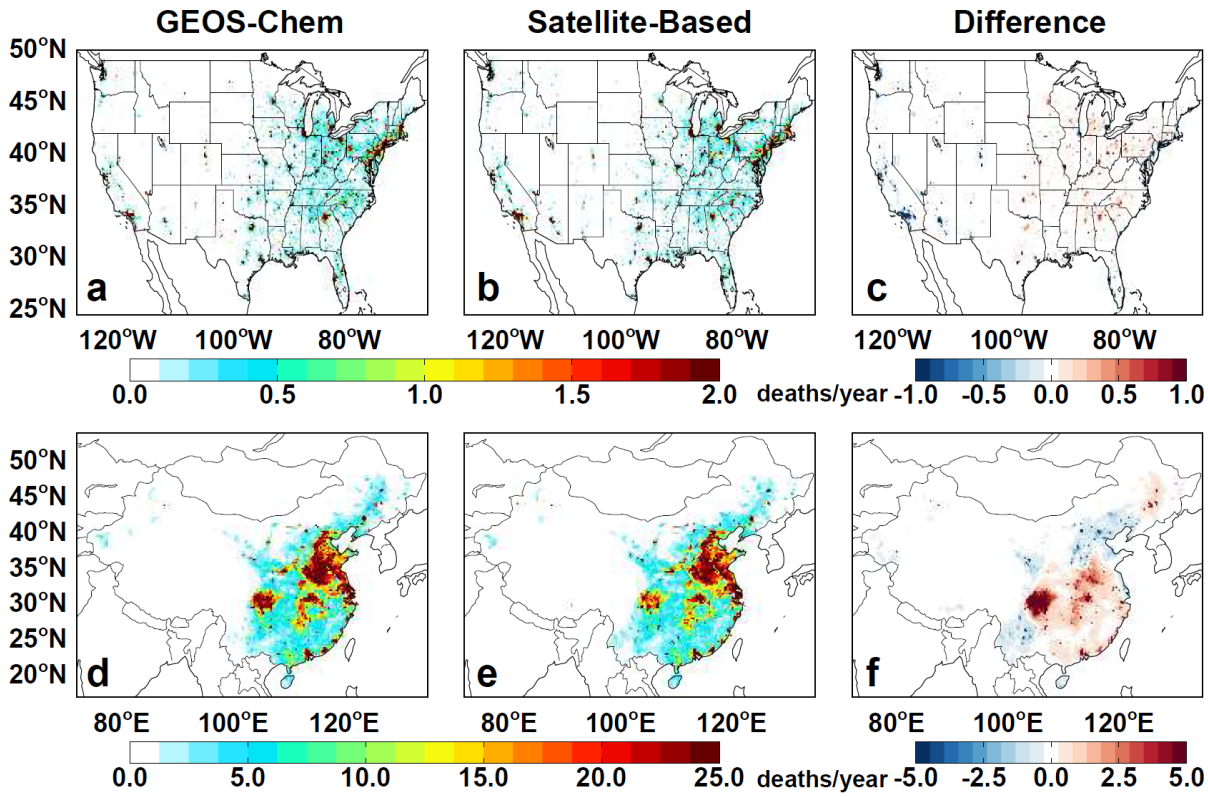

Figure 3.14 Maps of premature mortality (total deaths per grid box) due to $PM_{2.5}$ exposure using (a, d) GEOS-Chem and (b ,e) MODIS Aqua satellite-based concentrations and the difference (model-satellite, c, f) for the U.S. and China, respectively.

3.4.2 Premature Mortality due to Acute Exposure

We also calculate the fraction of the mortality burden attributable to acute exposure from daily concentrations. For acute exposure risk ratios, values are taken from *Ostro et al*. [2006], which uses data from the CALFINE study. These values are shown in Figure 3.2 and are mid-range estimates among the studies shown. We show results as the total number of deaths that occur throughout the year in Figure 3.15 with the error bars denoting the confidence intervals of

the risk ratios. For acute exposure, the risk ratio for mortality from respiratory disease is higher than for heart disease.

To determine the mortality burden attributable to acute exposure we use the spatial interpolated satellite observations. Additionally, we scale the final results by the ratio of the results calculated using the model sampled to the satellite observations to the unsampled model concentrations to account for missing daily observations.

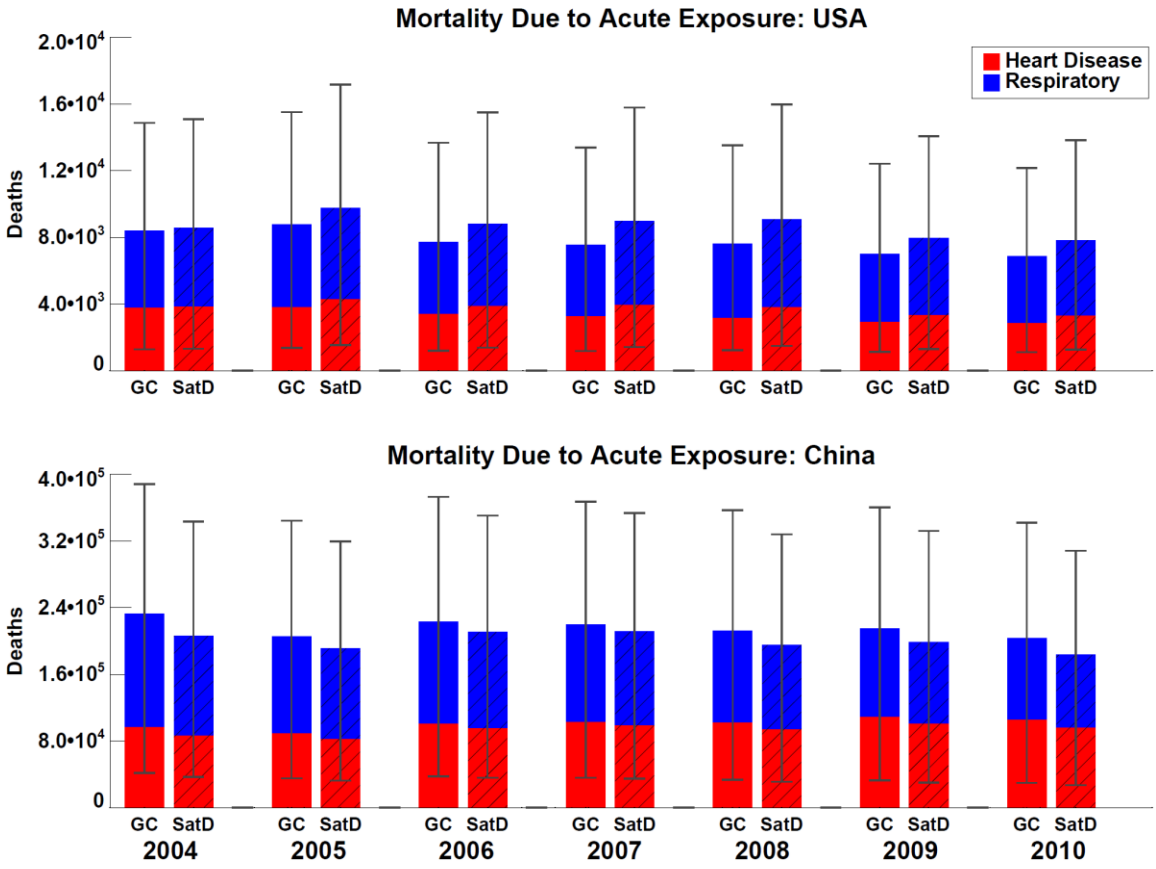

Figure 3.15 Time series for (a) the U.S. and (b) China using the unconstrained model estimates and the satellite-derived (hatched bars) $PM_{2.5}$ to determine the annual mortality due to acute exposure. Colors denote cause of death and whiskers denote the confidence intervals associated with the risk ratios.

In China, the fraction of premature mortality rates attributable to acute exposure is 10-12% of the estimates for chronic mortality, whereas in the U.S., acute exposure accounts for ~5% of total deaths from $PM_{2.5}$ exposure. Because the mortality estimates for chronic exposure have been decreasing over time in the U.S., and the mortality estimates due to acute exposure do not, this percentage increases over time. This suggests that while improvements in air quality in the U.S. have led to overall health improvements and should continue to do so; extreme events (defined as days with 24-hour concentrations over the daily standard of 35 μgm^{-3} in the U.S. or 75 μgm^{-3} in China, shown in Appendix 3), which are potentially harder to control, can still cause a significant number of premature deaths. Additionally, if these extreme events are removed from the annual average, the number of deaths from chronic exposure decreases by 3% (9%) in the U.S. using model (satellite) concentrations and 22% (19%) in China (although in China this removes a significant amount of data due to several regions in China which have annual average concentrations greater than the 75 μgm^{-3} proposed national daily standard.

Comparing the satellite-based estimates to the model-based estimates suggest that the model might produce mortality estimates that are too low by 2-16% in the U.S. and too high by 4-13% in China. However, while these datasets can provide potentially useful mortality estimates, it is important to understand the limitations of this method and properly constrain the potential uncertainties before suggesting that the satellite-based estimates are superior.

4 Uncertainties in estimates of the mortality burden using satellite-based and model $PM_{2.5}$ concentrations

Previous studies that have relied on satellite-based $PM_{2.5}$ datasets acknowledge that there is uncertainty in the estimated health impacts, partly due to the uncertainty in satellite-derived $PM_{2.5}$. However, the uncertainty is assumed to be rather small, such as in *Evans et al.* [2013] which suggests the uncertainty in their $PM_{2.5}$ contributes 5-28% of the uncertainty in their mortality estimates based on the error analysis performed in *van Donkelaar et al.* [2006]. However, these are also the error estimates in the long-term mean which can potentially average out greater daily errors. Here our goal is to conduct a more thorough analysis of the potential errors in using satellite-based $PM_{2.5}$ and estimate the range in premature mortality that can be attributed to acute and chronic exposure as in the previous chapter. In addition, this work will help to explain some differences in mortality burden estimates from previous studies. In the following sections, we explore the range of mortality burden estimates that can be determined using satellite-based $PM_{2.5}$ concentrations due to potential uncertainties in the model representation of the AOD-$PM_{2.5}$ relationship and the satellite AOD, along with the choices used for the mortality response function and choice of risk ratio.

4.1 Uncertainties in Model η

It behooves us to understand the limitations and uncertainties associated with a model-based η when applying it to the calculation of $PM_{2.5}$ exposure. Therefore, we perform multiple sensitivity tests in order to determine the impact that different aerosol properties, grid-size resolution and time scales will have on η and, ultimately, on the resulting satellite-based $PM_{2.5}$. These sensitivity tests are performed solely with model output, which can provide a complete

spatial and temporal record, and results from the modified simulations are compared to the standard model simulation. Thus these can be considered in the framework of an Observing System Simulation Experiment (OSSE), comparing results with a baseline "truth". Our analysis focuses on regions of the United States and China.

Because our goal is to better estimate surface $PM_{2.5}$, our error analyses in this section are calculated against the standard model daily simulated $PM_{2.5}$ concentrations. However, we calculate these values using a modified η to better understand how this η relationship impacts the derived $PM_{2.5}$. We note that these are errors with respect to our baseline simulation; we do not characterize how each sensitivity simulation may be "better" or "worse" compared to true concentrations of surface $PM_{2.5}$, but rather how different they are from the baseline, thus characterizing the uncertainty in derived $PM_{2.5}$ resulting from the model estimates of η. The timescale of the estimated $PM_{2.5}$ influences the error metric we choose for this analysis. We use the NMB (Equation 3.5) for estimating error associated with annual $PM_{2.5}$ exposure. This allows for the possibility that day-to-day errors may compensate, resulting in a more generally unbiased annual mean value. However when estimating errors for daily $PM_{2.5}$ estimates, we rely on the mean normalized gross error (MNGE) metric which is calculated using Equation 4.1, where *O* is the original model simulated concentration (at a location *x* and time *t*) and *P* is the concentration derived using the modified simulation.

$$\mathrm{MNGE} = \frac{1}{\mathrm{N}} \sum_{1}^{N} \left(\frac{|\mathrm{P_{x,t}} - \mathrm{O_{x,t}}|}{\mathrm{O_{x,t}}} \right) \times 100\% \tag{4.1}$$

We note that this metric characterizes the absolute error, which we choose to emphasize that both positive and negative biases constitute daily errors and should not be averaged out (such as in the NMB metric for example) when estimating the error associated with daily $PM_{2.5}$. These gross error comparisons are summarized in the figures on an annual basis (averaging daily values), but

we also discuss results on a seasonal basis to stress that influencing factors also vary with season. In the figures, we show the box and whisker plots of the normalized gross error but provide the MNGE in the text. In Section 4.1.1 we show only the "daily" gross errors, but we apply mean and gross errors to annual and daily $PM_{2.5}$ exposure respectively in Section 4.4.

4.1.1 Model Sensitivity Studies

In Figure 4.1and Figure 4.2 we show the MNGE and range of normalized gross errors in derived $PM_{2.5}$ for each of the sensitivity tests for 10 different regions over the U.S. and nine regions in China for 2006. The figures also include maps of the annual mean surface $PM_{2.5}$ to highlight geographical variability.

Our first sensitivity tests relate specifically to the methodology. To derive a satellite-based $PM_{2.5}$ with this method requires that model simulations are run for every day and that there are valid satellite observations. Running a model can be labor intensive and there are specific regions and time periods with poor satellite coverage. Therefore, it might be beneficial to be able to use a climatological η, climatological satellite AOD, or climatological satellite-based $PM_{2.5}$. In order to assess the usefulness of these methods, the first 3 tests simply use seasonally averaged $PM_{2.5}$, AOD and η model values instead of daily varying values. For using a seasonally averaged $PM_{2.5}$, we compute the error in derived $PM_{2.5}$ that results if we assume that the model $PM_{2.5}$ concentration is consistent throughout the season when calculating η. In some regions this is not a poor approximation. In the Southeast and Ohio RV/Appalachia during winter and the 4 Corners in summer and fall, surface concentrations are fairly consistent and the MNGE is ~35-50%. Conversely, the Great Lakes region, the Northeast in summertime, and the Great Plains in winter and spring have greater variability, and therefore potential errors are greater with MNGE values often greater than 100%. In the Great Plains (note that concentrations are low and therefore while

the percentage change is large, the absolute error is smaller), this variability can often be attributed to transported sources, while in the Northeast this is mainly due to local emission sources and meteorology (PBL, wind, etc.). In the Great Lakes region, variability can be driven by local sources and transport. Therefore, using seasonal averages of $PM_{2.5}$ can produce large MNGE (90-130%). However, as evident from the figures, this mean error is often weighted strongly by the outliers. While on most days, the error may be only 2-5 μgm^{-3}, there is a larger range of errors (sometimes >35 μgm^{-3}) in the estimated $PM_{2.5}$ that could lead to very erroneous short-term exposure estimates.

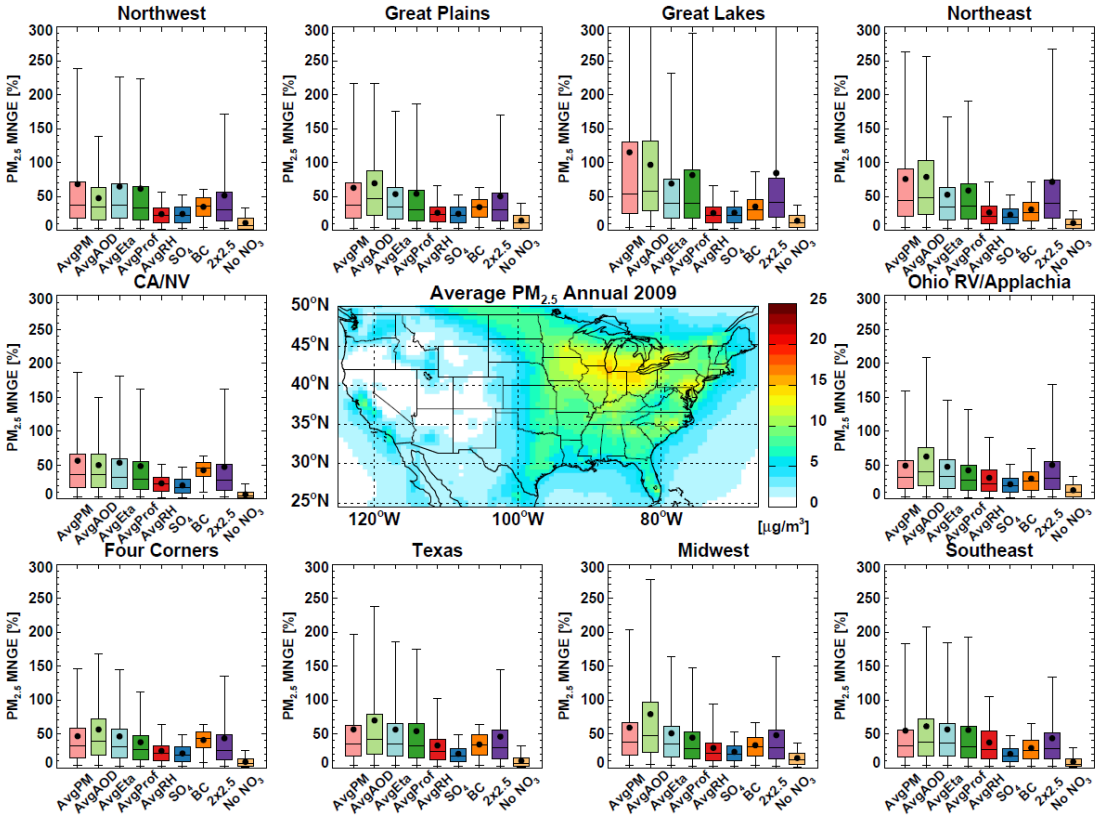

Figure 4.1 Box and whisker plots showing the mean and range of differences in derived daily $PM_{2.5}$ associated with the different sensitivity tests described in the text for different regions in the U.S. for the year 2009. Boxes outline the 25th and 75th percentiles, while whiskers show the 5th and 95th percentiles, dots show the mean, and the middle line shows the median. Absolute differences are given as a

percent different from the base case. Map shows annual average concentrations for the base case.

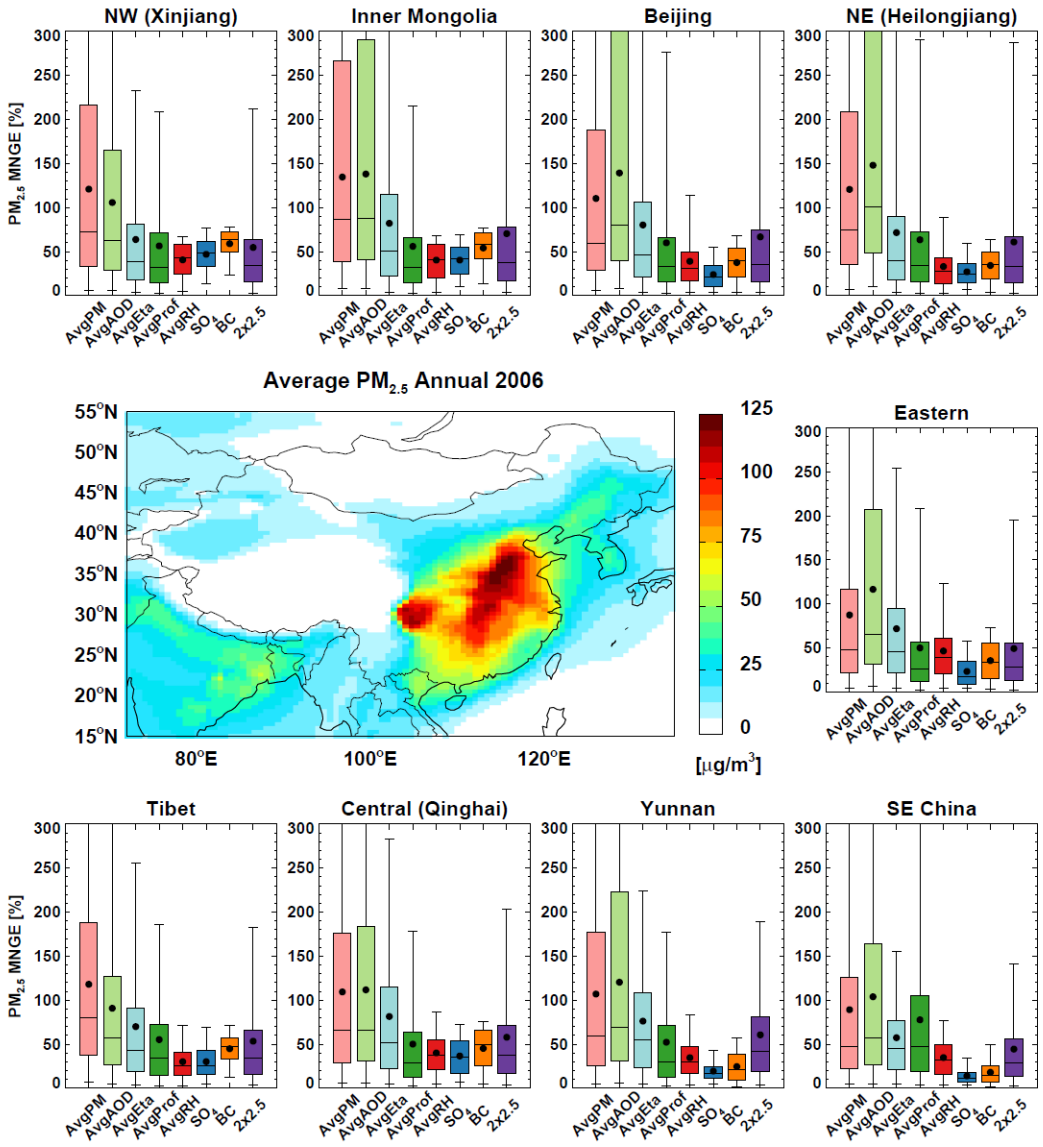

Figure 4.2 Similar to Figure 4.1 but for regions in China in 2006.

Because mass concentrations in China are generally much higher, the absolute value of potential errors can also be much greater. Additionally, the normalized error from using a mean $PM_{2.5}$ has a wider distribution as seen by the length of the whiskers on the box plots (Figure 4.2).

The largest potential errors are generally in the eastern urbanized region, especially around Beijing and the Heibei province, where there are many anthropogenic sources of aerosols. Absolute errors are greatest in the winter, when China often experiences strong stagnation events that lead to several days of very high surface concentrations. However, the MNGE is generally greater in the west where transport (specifically of dust) can cause significant magnitudes of variability in surface concentrations.

To test the importance of daily variability in AOD, we compute daily η values and then solve for daily surface $PM_{2.5}$ values using a seasonally averaged model simulated AOD. This mimics the error introduced by using seasonally averaged satellite observations, an attractive proposition to overcome limitations in coverage. This approximation often produces the greatest error. It produces smaller errors in regions with less variability in AOD (spring and fall in the Northwest and spring and summer in the Southeast U.S.). However, in regions where AOD varies more dramatically and specifically where transported layers aloft can significantly increase AOD (Four Corners, Great Plains, Great Lakes), using a seasonally averaged AOD, can lead to much larger errors. With few exceptions, these errors are often greater than those determined by using an average $PM_{2.5}$ concentration.

For China, the mean absolute errors using a seasonally averaged AOD are greater than using an average $PM_{2.5}$ value in all regions and seasons except in summer. This suggests strong caution in using AOD to constrain $PM_{2.5}$ estimates, because using incorrect AOD values can lead to extremely large daily errors in China (mean errors >60 μgm^{-3}). Similar to using an average $PM_{2.5}$, largest errors occur for winter in eastern China. These tests suggest the importance of high-time resolution satellite observations for constraining surface $PM_{2.5}$.

For the seasonally averaged η test, we estimate daily $PM_{2.5}$ values from the seasonally averaged η and daily AOD values. This test evaluates the necessity of using daily model output to define the η relationship. Regional η relationships can be more consistent over time than $PM_{2.5}$ or AOD. Therefore, using a seasonally averaged η generally produces less error than using a seasonally averaged AOD, and often less than a seasonally averaged $PM_{2.5}$ as well. This is especially true in summer where the mean and range of errors are smaller. Because the model η also inherently prescribes a vertical distribution of aerosol, this suggests that there is less variability in the vertical distribution in summer, either from lofted plumes (increase AOD but not $PM_{2.5}$) or boundary layer trapping (increase $PM_{2.5}$ and little impact on AOD). Using a seasonally averaged η appears to lead to the greatest errors in the winter, notably in the Southeast, Great Lakes, and Northwest U.S. (MNGE 60-80%). Much like the U.S., using a seasonally averaged η for regions in China generally produces a lower mean normalized error (50-100%) than using a seasonally averaged AOD or $PM_{2.5}$ and the absolute error is generally also greater in winter.

We further test the importance of the vertical distribution in the η relationship for predicting surface $PM_{2.5}$ concentrations by comparing values from the standard simulation against using an η resulting from a seasonally averaged vertical distribution. For this comparison, we allow the column mass loading to vary day-to-day, but we assume that the profile shape does not change (i.e. we re-distribute the simulated mass to the same seasonally averaged vertical profile). We note that this is not the same as assuming a constant η, as relative humidity and relative aerosol composition vary day-to-day. From Figure 4.1, we see that using a seasonally averaged vertical distribution often leads to large errors in surface concentrations in a similar manner to the sensitivity to η comparisons. However, because we do allow the mass (and

composition) to change, errors are generally not as great. This suggests that even if we lack information on the aerosol composition, information on how the pollutants are distributed is extremely important. While changes in AOD can be driven by changes in surface mass loading; elevated AOD can also be produced by layers of lofted aerosols that result from production aloft or transport. These can be infrequent so as not to strongly impact the seasonally averaged vertical profile, but when they are not accounted for can lead to significant overestimations in daily surface $PM_{2.5}$ concentrations as shown by the length of the whiskers. We see this is of greater concern across the northern U.S., in areas that are occasionally impacted by transported elevated biomass burning plumes. Additionally, we see that there are larger errors across most regions in the winter, which is likely again associated with trapping of pollutants near the surface compared to seasons where pollutants are more likely to be mixed out of the boundary layer.

Much like the normalized errors associated with a seasonally averaged η, the distribution of errors across seasons associated with using a seasonally averaged vertical profile is very similar in the U.S. and China, although the absolute errors vary across seasons due to seasonal variations in mass loading. In Beijing, Eastern and Southeastern China, the largest errors are found in winter. In the NW and Inner Mongolia, the lowest errors are found in winter and highest in spring. This is due to the fact that these regions are often influenced by transported dust from the Taklamakan and Gobi Deserts, from which emissions are generally greatest in spring [*Wang et al*., 2008].

We also test the sensitivity of derived $PM_{2.5}$ related to aerosol water uptake. This is done by recalculating η using a seasonally averaged relative humidity (RH) profile. This generally reduces the seasonally averaged AOD (less water uptake) in every season (because hygroscopic growth of aerosols is non-linear with RH). This leads to an overestimate of η that when applied

to the AOD values from the standard simulation, generally overestimates surface $PM_{2.5}$. This is because, for the same AOD, a higher η value would suggest more mass at the surface in order to compensate for optically smaller particles aloft. We see from Figure 4.1, this is particularly an issue in regions with potentially higher RH (Southeast and Ohio RV/Appalachia) and more hygroscopic aerosols (Northeast). In China, the largest errors are again in urban regions dominated by inorganics and in the summertime when the average RH is high (Figure 4.2, RH in Appendix 4). In wintertime, the tails on the error distributions are much greater in these regions, suggesting that there are several days where the RH deviates enough from the mean to significantly impact the assumed water uptake and AOD. However, the MNGE for all regions and seasons is generally 20-30% for the U.S. and 25-50% for China, less than errors associated with seasonally averaged AOD, $PM_{2.5}$, η, and the vertical profile.

We also compute two more sensitivity analyses related to assumptions about the aerosol type and water uptake. For these, we take the simulated mass concentrations and compute the AOD assuming that the entire aerosol mass is sulfate or, alternatively, hydrophobic black carbon. This test mimics a mischaracterization of the aerosol composition in the model. Black carbon has a high mass extinction efficiency, which is constant with RH given its hydrophobic nature; while sulfate is very hygroscopic, resulting in much higher extinction efficiencies at higher relative humidity values. Overall, assuming that all the mass is sulfate does not lead to large errors in the U.S. (MNGE 15-35%) as the AOD in many regions in the U.S. is dominated by inorganics. Errors are largest in regions and seasons with larger contributions of less hygroscopic aerosols (organic carbon and dust) and/or high relative humidity. In China, the normalized errors are also only 10-50% across seasons, but the mean absolute errors are > 20 μgm^{-3} in the NW, Inner

Mongolia, Beijing, and Eastern China in spring, due to the influence of dust transport. Absolute errors are much smaller in winter and only high for NW China during summer and fall.

Assuming the entire aerosol mass is black carbon leads to greater and more frequent (wider distributions) errors than sulfate (MNGE 22-47%), mainly because of the larger mass extinction at lower relative humidity values and the fact that hydrophobic black carbon generally makes up a small fraction of the mass loading in all regions in the U.S. When RH is low, this assumption increases the AOD, which leads to an under prediction in the derived $PM_{2.5}$. When RH is high, this decreases the AOD and leads to an over prediction in derived $PM_{2.5}$. The largest percentage changes occur in NV and the 4 Corners regions due to the low relative humidity, low mass concentrations, and large contribution of dust, while the largest absolute changes are in regions in the Eastern US. In China, much like the sensitivity test with sulfate, the absolute errors are generally small, but the absolute errors can be rather large. We also see similar behavior with regards to regions impacted by dust having the greatest absolute errors. The length of the whiskers on the box plots suggests that this would be a larger issue on specific days with high dust transport.

It is assumed that using a finer resolution model, although more computationally expensive, will be more representative of small scale variability and better suited for estimating surface air quality. *Punger and West* [2013] find that coarse resolution models often drastically underestimate exposure in urban areas. We therefore investigate the grid-size dependence of our simulated η to test the necessity of running the model at a finer resolution to determine η. For this, we determine the η values from a simulation running at 2°x 2.5° grid resolution (with the same emission inputs and time period), re-grid these values to the nested grid resolution (0.5°x0.666°) and solve for the derived $PM_{2.5}$ concentrations using the AOD values from the

nested simulation. From Figure 4.1 and Figure 4.2, we see larger discrepancies in regions which are dominated by more spatially variable emissions (California, the Northeast, the Great Lakes, around Beijing and eastern China) rather than areas with broad regional sources (Southeast U.S., 4 Corners, Tibet, Yunan). However, compared to the other sensitivity tests, using the coarser grid leads to mean errors of only 30-50% (or 1-5 μgm^{-3}) in the U.S. and 30-80% in China, which suggests that spatially averaged η are potentially more useful than temporally averaged η for constraining $PM_{2.5}$ values. This is in line with the conclusions from *Thompson and Selin* [2012] who show that very coarse grids can over predict pollutant concentrations and consequently health impacts (associated with exposure to ozone), but using very fine grids does not significantly decrease the error in simulated concentrations compared to observations.

As mentioned previously, GEOS-Chem has a known high bias in nitrate aerosol over the U.S. [*Heald et al*., 2012]. In our final sensitivity, we test how this nitrate bias could impact η and the derived $PM_{2.5}$. This should only be an issue if nitrate has a drastically different vertical profile (or η) from the other species. In some regions, such as the Southeastern U.S., the simulated nitrate profile is very similar to the profile of sulfate, but in the Midwest and Great Lakes regions, the nitrate profile is significantly different than that of the other species (Figure 4.3). We therefore compute η without nitrate aerosol, and then derive $PM_{2.5}$ using the standard AOD. From the figure, this appears to not be a large source of potential error (<15%), and only slightly larger errors occur in winter and in regions where nitrate is biased high (Great Lakes, Midwest, and Appalachia). Furthermore, these errors are less than the bias between the model and surface observations of nitrate (1-2 μgm^{-3} compared to 2-7 μgm^{-3}), suggesting that even though there is a known bias in the model, using satellite observations may largely correct for this by constraining the total AOD when estimating satellite-derived $PM_{2.5}$. We did not do this

comparison for China as measured nitrate concentrations are not widely available for evaluation in this region. However, *Yang et al.* [2014] do suggest that model nitrate is also too high in China.

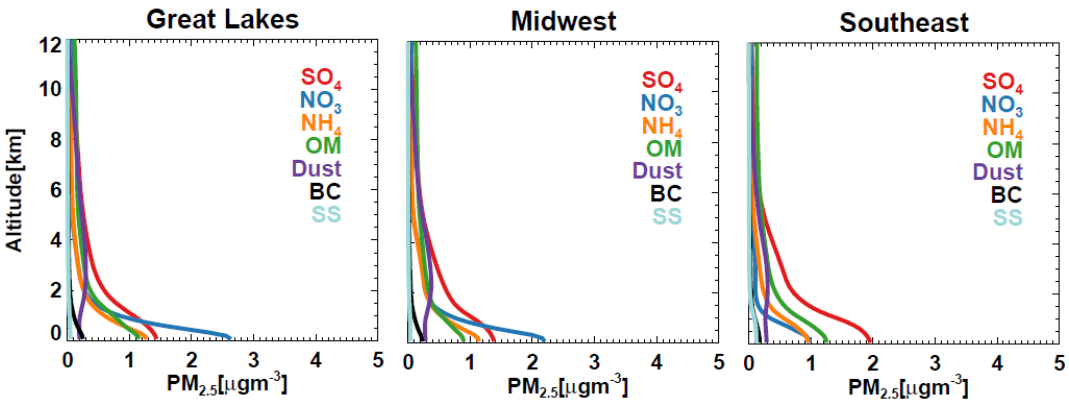

Figure 4.3 Simulated annual average vertical profiles of atmospheric aerosol species in several regions of the U.S. for 2009.

Compared to the MNGE values determined from daily concentrations, the normalized mean biases in annual average concentrations (shown in Appendix 5) are generally much less (range of ±20% in U.S. and ±50% in China). This is because the potential random errors in daily values can cancel out in the long-term mean, as is the case for our sensitivity tests regarding the vertical profile and relative humidity (for most places). However, because our method to test the sensitivity to aerosol type required assuming all aerosol mass was black carbon or sulfate, we introduced a systematic bias that is not significantly reduced in the annual NMB. For China, NMB values are also greater than for the U.S. reflecting the high daily variability that is more strongly impacted by using seasonal averages.

This analysis was performed for one year, however, there is also inter-annual variability and previously well-noted recent trends in $PM_{2.5}$ which may impact our uncertainty estimates. *Leibensperger et al.* [2012] and *Chin et al.* [2014] show that concentrations of anthropogenic aerosols in the U.S. have been decreasing over the last 20 years due to decreasing anthropogenic

emissions. The changes in AOD, as discussed previously, are less obvious because they also encompass transboundary pollution. However, in general, AOD over the U.S. has also declined, while AOD over China has increased in many regions following increases in fossil fuel and dust emissions as shown by *Chin et al.* [2014] and shown in Figure 3.9. This suggests that year-specific η values are also necessary. While 2009 appears to be a rather typical year in terms of concentrations and in general there is not a significant variability in the meteorology fields over this time period (Appendix VI**. Interannual Variability**

), additional work is still required to verify whether these uncertainty estimates might vary significantly year-to-year.

4.1.2 Context of Sensitivity Analysis

While we use the term "error" in our sensitivity analysis for comparing the original simulation of $PM_{2.5}$ to $PM_{2.5}$ derived using modified η values, we do not assume that these are actual model errors, but potential errors or uncertainties. In general, the model simulates $PM_{2.5}$ well (as shown in Figure 3.6) and represents important processes. Satellite AOD can help to constrain these estimates to better represent measured concentrations [*van Donkelaar et al.*, 2006], but in specific regions or periods of time, the potential errors in η could lead to discrepancies between satellite-derived and actual surface mass. These sensitivity tests can then provide a range of uncertainty in satellite-based $PM_{2.5}$ associated with model uncertainties. Our estimates in Section 4.1.1 convey the uncertainty in daily satellite-derived $PM_{2.5}$ associated with model assumptions.

However, there are also potentially more uncertainties in the model that we did not quantify here. We used seasonally averaged values for most of our sensitivity tests because we assume that on average the model does a good job simulating the relationship between AOD and $PM_{2.5}$, but it might be incorrect on specific days. In reality, the relationship could be further outside the range then is simulated either in the seasonal average or on specific days. We also did not test the sensitivity to particle density or size distribution, but these were previously found to have little impact on derived values [*van Donkelaar et al.*, 2006]. Additionally, derived values were shown to often be more sensitive to η in regions impacted by dust or biomass burning, which could suggest uncertainty linked to specific emission sources, which we did not test explicitly here.

From these sensitivity tests, we see that assumptions used to determine the η relationship can strongly impact the estimated daily surface $PM_{2.5}$ concentrations. Using a seasonally averaged value of AOD appears to be a poor option for determining short term exposure, often worse than using a seasonally averaged $PM_{2.5}$ value. This suggests that while constraining the model with satellite observations might improve $PM_{2.5}$ estimates, if, due to lack of coverage, the only option is to use a temporally averaged satellite AOD, the associated errors for daily estimates could outweigh the benefits. With regards to potential errors in the model calculation of η, the vertical distribution appears to be a larger source of potential error across locations and seasons than assumptions about the aerosol type and water uptake, although those assumptions can lead to the large errors on specific days. Additionally, while there is variability in the errors across regions, errors are often greatest in regions with high populations (Northeast, Great Lakes, California, Beijing and Eastern China), which implies greater uncertainty when determining exposure estimates.

Comparing MNGE for daily concentrations to NMB for annual average concentrations also highlights the differing potential impacts on exposure estimates due to systematic and random errors, an important distinction for determining the usefulness of this method. Systematic errors may not be as obvious on short timescales compared to random errors (related to meteorology and/or representation of plumes) that can lead to large biases in daily concentrations and acute exposure estimates. However, these random errors have less impact when we examine annual average concentrations and chronic exposure. Systematic errors, potentially related to sources or processes (as in Chapter 2), might be harder to counteract even on longer timescales even when the model is constrained by satellite observations.

We translate this potential uncertainty in η to potential uncertainty in mortality and exposure estimates determined from the satellite-based $PM_{2.5}$. Figure 4.4 is an updated version of Figure 3.12 with added range of potential uncertainty. This range is determined by using the normalized mean bias in annual $PM_{2.5}$ determined from the sensitivity tests for RH, the vertical profile, grid resolution, and aerosol composition. We determine the bias for each gridbox and then use these values to "bias correct" our satellite-based $PM_{2.5}$. This bias is then applied to the satellite-based $PM_{2.5}$ and used to re-calculate the percent of the population that is exposed to different annual $PM_{2.5}$ concentrations. We see that the potential uncertainty in η can substantially impact exposure estimates. In particular, these uncertainties are comparable to, or even larger than, the differences in exposure estimated using different satellite sensors.

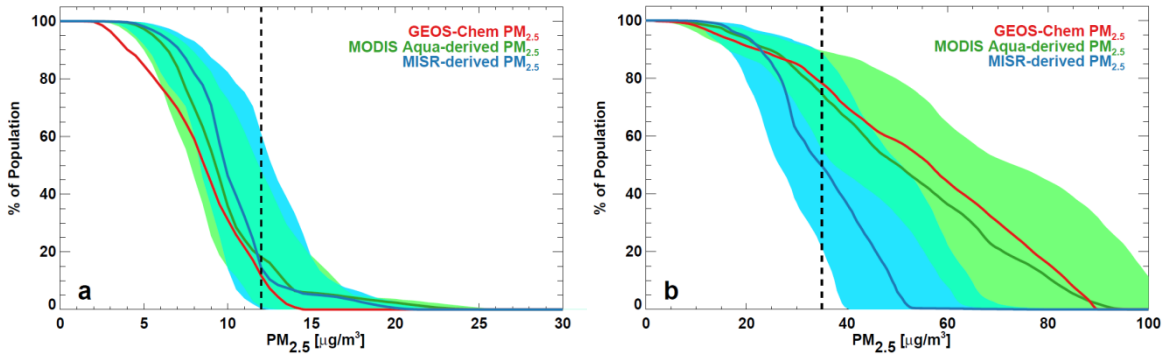

Figure 4.4 Similar to Figure 3.12 but with range of potential uncertainty in exposure estimates using satellite-based $PM_{2.5}$ denoted by shaded regions. Colors of shaded regions coordinate with line color.

4.2 Uncertainties in Satellite AOD

4.2.1 Method to determine uncertainty in satellite AOD

For "validation" of satellite products, studies have often relied on comparisons with AERONET ground sites [e.g. *Kahn et al*., 2005; *Levy et al*., 2010; *Remer et al*., 2005; 2008; *Zhang and Reid*, 2006]. Uncertainty in satellite AOD can be due to a variety of issues such as the presence of clouds, the choice of optical model used in the retrieval algorithm, and surface properties [*Zhang and Reid*, 2006; *Toth et al*., 2013]. The uncertainty in AOD over land from MODIS is estimated as 0.05 ± 15% [*Remer et al.*, 2005], while *Kahn et al.* [2005] suggested that 70 % of MISR AOD data are within 0.05 (or 20 % ×AOD) of AERONET AOD.

There are also discrepancies between AOD measured by the different instruments. Aqua has an afternoon overpass while Terra has a morning overpass. It might be expected that there would be some differences in retrieved AOD associated with diurnal variations in aerosol loading. However, the difference of 0.015 in the globally averaged AOD between MODIS onboard Terra and Aqua (Collection 5), although within the uncertainty range of the retrieval, is

attributed more to uncertainties and a drift in the calibration of the Terra instrument, noted in *Zhang and Reid* [2010] and *Levy et al.* [2010]. MODIS Collection 6 (as will be discussed further) is more consistent with MISR [*Levy et al.*, 2013]. MISR has a different measurement technique and swath width, such that the correlation between MISR AOD and MODIS AOD is only 0.7 over land (0.9 over ocean) [*Kahn et al.,* 2005].

There are also uncertainties due to the satellite sampling as discussed in Section 3.3.3 because of the potential that the concentrations on days with available observations are skewed. In order to assess the sampling bias, we use the model and compare the annual mean to the mean when sampled to days with valid observations (Figure 3.10). In general, sampling leads to an underestimation in AOD. This can partly be attributed to high aerosol loading mistakenly identified as cloud and possibly the removal of anomalously high AOD values (>2.0). This suggests that the average AOD values shown here can also be influenced by our filtering and data quality standards. Analysis of the impact of satellite data quality on the AOD to $PM_{2.5}$ relationship is discussed in *Toth et al.* [2014]. They find that using higher quality observations does tend to improve correlations between AOD and $PM_{2.5}$ across the U.S. (though in general correlations are low).

Not only are there discrepancies in AOD between instruments, there are also differences between product versions for the same instrument. The MODIS Collection 6 Level 2 AOD is substantially different from Collection 5.1 [*Levy et al.*, 2013]. In general, AOD decreases over land and increases over ocean with Collection 6. These changes are due to a variety of algorithm updates including better detection of thin cirrus clouds, a wind speed correction, a cloud mask that now allows heavy smoke retrievals, better assignments of aerosol types, and updates to the Rayleigh optical depths and gas absorption corrections [*Levy et al.*, 2013]. The normalized mean

bias in AOD estimated for Collection 5.1 against Collection 6 is shown in Figure 4.5. In general, Collection 5.1 is biased low over most regions, with the exception of more desert regions, where it is biased high. Results are similar over China and the United States. We do note that some of these differences are limited in part by the number of available observations; however, here we only show grid boxes that have at least 200 observations. These differences would also impact the derived $PM_{2.5}$ (and can explain some differences between our results and previous studies based on MODIS Collection 5.1). In particular, because Collection 6 suggests higher AOD over many of the urbanized regions, the derived $PM_{2.5}$ and resulting exposure estimates (all other variables constant) would be greater.

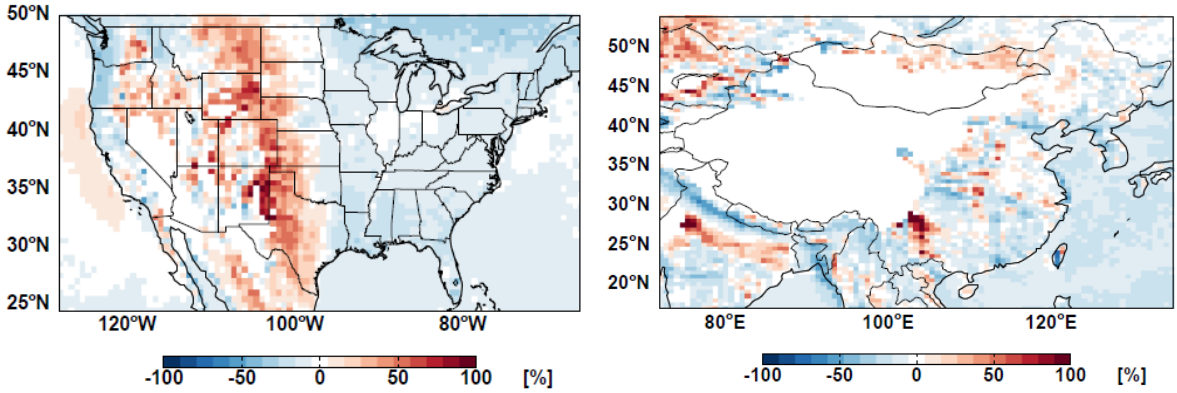

Figure 4.5 Percent Difference in AOD from using MODIS Aqua Collection 5 compared to AOD with Collection 6 for 2004-2011 over (a) the United States and (b) China (Collection 5-Collection 6/Collection 6).

We ascertain the potential uncertainty in satellite AOD by comparing satellite observations to AERONET. In Figure 4.6 and Figure 4.7, we show the mean normalized gross errors between AOD from each satellite instrument and AERONET sampled to all available days during the time period of 2004-2011 for the U.S. and China. Note that some of these sites were established for field studies and only provide a few months of data. As we use all available data,

we are not accounting for bias that might have some seasonality (i.e. changes in surface properties).

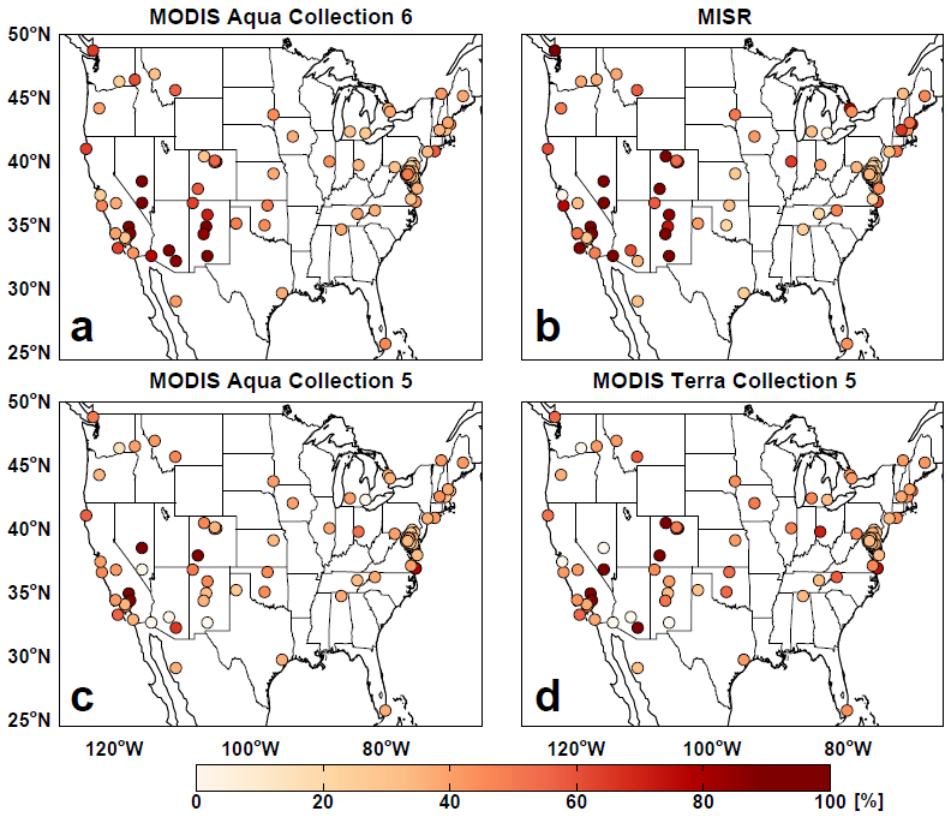

Figure 4.6 Mean Gross Error in AOD between satellites (a) MODIS Aqua Collection 6, (b) MISR, (c) MODIS Aqua Collection 5 and (d) MODIS Terra Collection 5 and AERONET AOD for 2004-2011.

From these comparisons, we see that the satellites generally agree with AERONET better in the eastern U.S. and eastern China and have larger biases in the west near deserts and at coastal regions where there could be issues distinguishing land and water in the retrieval algorithm. MNGEs at each AERONET site are generally similar among the instruments, with greater differences at these western sites. While Collection 6 does reduce the bias at several sites along the East Coast in the U.S., it is generally more biased at the sites in Arizona, Nevada, and

New Mexico. However, the biases between MODIS Collection 6 and AERONET are more similar to the biases in MISR than compared to Collection 5.

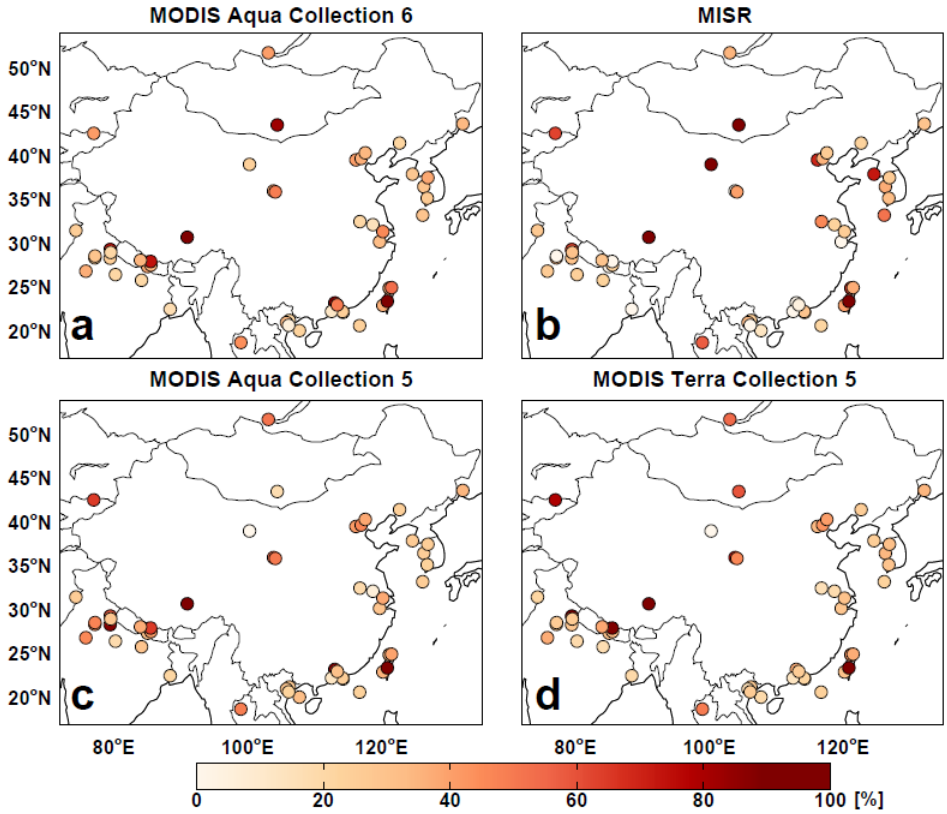

Figure 4.7 Same as Figure 4.6 but for China.

To calculate the uncertainty in premature mortality due to acute exposure from estimates made with the (MODIS Collection 6) satellite AOD, we use the MNGE (from daily observations) calculated from all sites in the west and all sites in the east for each country. To calculate the uncertainty for chronic exposure, we calculate a NMB (from annual averages). The NMB for each site is shown in Figure 4.8. There are fewer sites in Figure 4.8 because we require at least 50 (approximately one a week) valid observations in order to calculate an annual average. At many sites, the NMB is much less compared to MNGEs. This is expected, as annual values can average out larger daily fluctuations. However, sites where the NMB is on the same order as

the MNGE (i.e. the western U.S.), are suggestive of a systematic bias that cannot be averaged out over time. Table 4.1 summarizes the statistics for East and West U.S. and China. Compared to the standard MODIS AOD retrieval uncertainty, our overall NMB is less in the eastern U.S. (-1%) and western China (11%) and higher in the western U.S. (40%) and eastern China (18%). Our MNGEs are significantly greater than the standard uncertainty and our NMB values (Table 4.1).

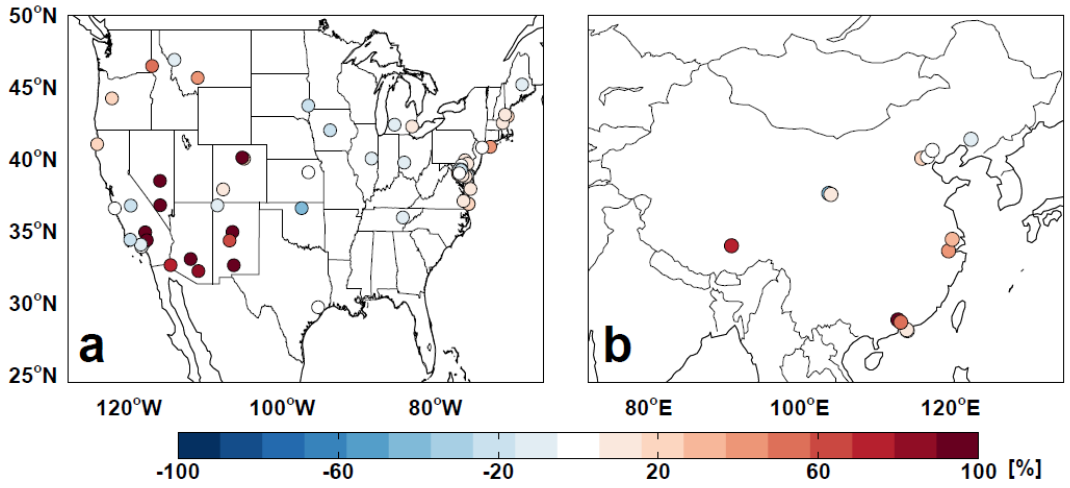

Figure 4.8 Normalized mean bias in annual average AOD from MODIS Collection 6 compared to AERONET sites in (a) the U.S. and (b) China.

Table 4.1 Regional satellite AOD MNGE and NMB values determined from comparison with AERONET AOD and used to determine uncertainty range in mortality estimates.

Region		NMB [%]	MNGE [%]
China	West	11	58
	East	18	36
U.S.	West	40	76
	East	-1	37

4.2.2 Methods to address biases associated with surface properties

There are potentially ways to reduce the biases in MODIS AOD as shown in previous

studies [*van Donkelaar et al.*, 2010; *Hyer et al.*, 2011]. AOD retrievals using the MODIS instruments assume surface reflectance properties that can lead to known (and unknown) biases in AOD, such as over the Western U.S. It was previously suggested by *Hyer et al.* [2011] that this bias can be minimized by using the MODIS Bidirectional Reflectance Distribution Function (BRDF) Albedo Product [*Schaff et al*., 2002] to derive an empirical relationship to correct AOD values. However, as shown in their study, minimal improvements were seen over the Western United States. Additionally, their empirical relationships were derived for Collection 5 and the strong filtering criterion that they employ greatly reduces our sample size, making it impractical for the daily estimates needed in this study.

Another method used by *van Donkelaar et al.* [2010] also uses the BRDF Albedo product, but instead of applying the empirical formula, uses it to establish land surface types (separated by their ratios of black sky to white sky albedo at different wavelengths). These regions were then grouped together for comparisons with AERONET stations and any outliers are removed. This was done for monthly averages where it is more likely that an outlier would suggest a systematic bias rather than simply a high AOD event. We therefore chose not to use this method. As our goal is not to produce a product but to delineate uncertainties, we instead use comparisons with AERONET AOD (Table 4.1) to determine the uncertainty in satellite AOD and determine the range of values that could be calculated if these regional biases are considered.

4.3 Uncertainties in Mortality Burden Estimates

We compare our results to premature mortality burden estimates from other studies in Table 4.2. In general, our estimates for the U.S. and China are higher than most previous global estimates, except for *Anenberg et al.* [2010] and *Fann et al*. [2012]. In particular, our estimates appear to be double those found by *Lelieveld et al.* [2013]. The spread among these studies can

be attributed to the data used (i.e. Collection 5 rather than Collection 6 and choice of baseline mortality rates) as well as the risk ratios and response functions. For example, *Evans et al.* [2013] also use satellite-based concentrations (using MISR/MODIS Collection 5 and GEOS-Chem), but their resulting mortality estimates are much lower. However, unlike this study, they use a threshold value and a different response function than we use for our initial results along with lower baseline mortalities.

Table 4.2 Premature mortality from $PM_{2.5}$ exposure by all-cause (All), cardiovascular disease (CVD) and lung cancer (LC) as estimated in other studies for the globe, U.S., and China. Values are for (x1000), and colors represent studies that give higher (red) or lower (blue) estimates.

Study	**U.S. (North America)**			**China (Asia)**			**Global**		
	All	Heart	LC	All	Heart	LC	All	Heart	LC
Evans et al., 2013							3313	1256	222
Fann et al., 2012	130								
Anenberg et al., 2010	141	124±37	17±7	2736	2584±618	152±53	3721	3499±864	222±80
Lelieveld et al., 2013	55	46	9.1	1006	898	108	2200	2000	186
Cohen et al., 2005				487			800	712	6.4
Lim et al., 2012							3200		
Fang et al., 2013		38	4.4		661	53		1532	95
Silva et al., 2013							2100	1950	150
This Study: Satellite	113	82	15	2084	1512	220	-	-	-
This Study: Model	104	75	14	2171	1570	231	-	-	-

For an initial metric of the uncertainty in the risk ratios that we used, we include whiskers on the mortality estimate bar plots (Figure 3.13 and Figure 3.15) to denote the range using the 95% confidence interval determined in the *Laden et al.* [2006] and *Krewski et al.* [2009] studies. Using just these confidence intervals as a measure of uncertainty suggests that there is still a large range of uncertainty in population-level health responses to exposure and caution should be exercised when attempting to transfer these values beyond the population from which they were determined in order to estimate national-level mortality burdens based on ambient concentrations. The choice of concentration response function and/or risk ratio will have a large

impact on our results. The difference between the satellite-based and model estimates is within this range, as well as results from many of these previous studies.

While we chose to use risk ratios from the *Laden et al.* [2006] and *Krewski et al.* [2009] studies, it is evident (Figure 3.1 and Figure 3.2) that there are a range of values from different exposure studies. While these studies attempt to account for differences in populations, lifestyles, pre-existing conditions, and co-varying pollutants, relative risk ratios determined from each study still differ. This is likely due to variables not taken into consideration, errors in exposure estimates ("exposure misclassification") [*Sheppard et al.*, 2012], and because, although the long-term effects of exposure to atmospheric pollutants have been well-documented, the pathophysiological mechanisms linking exposure to mortality risk are still unclear [*Chen and Goldberg*, 2009; *Sun et al.*, 2010; *Pope and Dockery*, 2013] making it difficult to determine how transferable results are from the context in which they were generated.

We also chose to use the same relative risk values for every location. However, studies have found that different populations have varied responses to exposure (potential for "effect modification") [*Dominici et al*., 2003]. One of the main uncertainties in our method is relying on risk ratios determined from epidemiology studies conducted in the United States, which may not represent the actual risks for populations in China. Long-term epidemiology studies examining exposure to $PM_{2.5}$ across broad regions of China are scarce, but studies using PM_{10} have suggested lower exposure-response coefficients than determined by studies conducted in the U.S. and Europe [*Aunan and Pan*, 2004].

Additionally, there is difficulty in separating the health impacts of acute and chronic exposure. Because short-term exposure disproportionally impacts sensitive populations, it has been suggested that time series analyses which find correlations between acute exposure and

mortality are merely noting the "harvesting" or "displacement" effect, in which there is a short-term displacement in time of death of these vulnerable populations and no impact on the absolute mortality [*McMichael et al*., 1998]. Calculating mortality from both acute and chronic exposure could then lead to an overestimate in annual death rates. If this were the case, it would be expected that there would be an increase in risk of mortality immediately following an event and then a period of decreased risk [*Schwartz*, 2001]. However, several studies that have regressed the relationship using lag correlations have shown that mortality risks do increase slightly 1-2 days following exposure but that the risk of mortality at longer lags (5, 20, 40 days) is even greater than the risk immediately following exposure [*Pope et al*., 1992; *Schwartz*, 2000; *Zanobetti et al*., 2002; 2003]. This suggests that the noted increase in mortality is not simply due to a harvesting effect, but that there can be larger long-term effects even without long-term exposure [*Pope*, 2007]. Additionally, studies have suggested that the biological response to short and long-term exposure could be different [*Pope et al.*, 2004; *Sun et al.*, 2010].

The distinction between effects from short-term and long-term exposure can be further confounded by considering that chronic exposure, which in studies generally refers to annual average concentrations, could be considered cumulative short-term exposure. Using annual averages can de-emphasize a time period, such as a fire or stagnation event, where concentrations are an order of magnitude larger than the average, but only for a short period of time. *Pope* [2007] suggests that we need a better understanding of the impacts of both the length and degree of exposure, as the responses could be very different. Additionally, it would be useful to have more studies that examine both acute and chronic exposure simultaneously to assess whether long-term exposure can lead to a vulnerable population that is more susceptible to acute exposure in the future [*Schwartz*, 2000]. Furthermore, it would be useful for determining health

benefits of improved air quality. *Correia et al.* [2012] found an increase in overall life expectancy with decreasing $PM_{2.5}$ concentrations. Comparing different time periods, *Laden et al*. [2006] found reducing atmospheric concentrations was associated with reductions in cardiovascular and respiratory mortality risk but little reduction in mortality risk due to lung cancer, suggesting that some individual effects might be irreversible.

Differences in mortality burden estimates between studies are attributable not only to the $PM_{2.5}$ concentrations determined, but also the methodologies and risk ratios that are used. As we show here, the response function can have a significant impact on our results.

As shown in Equation 3.3, we chose to use a linear response function for our initial comparisons. Some studies have suggested that linear response functions can greatly overestimate responses at high concentrations, where it would instead be expected that responses would start to level off [e.g. *Pope et al.*, 2011]. We therefore compare mortality estimates also using a log-linear function (equation 4.2 from *Krewski et al.*, 2009), a power law function for cardiovascular disease (equation 4.3, where I is the inhalation rate of $18 m^3 day^{-1}$, $\beta = 0.2730$, $\alpha = 0.2685$ from *Pope et al.*, 2011 and as used in *Marlier et al.*, 2013), and a ceiling value where we assume that the response is the same for any value over 30 $\mu g/m^3$. We also show results using a 5.8 $\mu g/m^3$ threshold (often used with the *Krewski et al.* [2009] RRs as it was the lowest concentration observed) and the *Laden et al.* [2006] RRs for cardiovascular and lung cancer.

$$RR = \exp [\beta (C - C_0)] \quad (4.2)$$

$$RR = 1 + \alpha (I \times C)^{\beta} \quad (4.3)$$

Table 4.3 gives the list of input for our comparisons of results using different CR functions and the resulting percent change compared to our initial conditions. These values are also shown in Figure 4.9 for chronic exposure. The largest difference is between using the risk

ratios from *Laden et al.* [2006] vs. using *Krewski et al.* [2009]. In the U.S., using a threshold value of 5.8 μgm^{-3} has a significant impact, reducing estimates by more than half (which consequently would put our estimates into agreement with *Lelieveld et al.* [2013]). The impact of a threshold is much less for estimates for China because annual concentrations are high enough that subtracting a threshold makes little difference. Conversely, using a ceiling value of 30 μgm^{-3} produces no difference in the U.S. (0% of the population experiences annual concentration values greater than 30 μgm^{-3}), while strongly reducing estimates in China.

Most experts in health impacts of ambient air quality agree that there is no population-level threshold (although there may be individual-level thresholds) [*Roman et al., 2008*]. Other studies have not used a threshold value, but a "policy-relevant" background concentration (generally >1μgm^{3}) to calculate mortality attributable to concentrations due to national anthropogenic emissions (not transport or natural sources, ie. concentrations that policies might be able to reduce).

We also see that the shape of the CR function produces different results between the U.S. and China. In the U.S., a log-linear CR function is almost equivalent to a linear response because of the low concentrations, while a power law function increases estimates as opposed to in China, where a log-linear relationship increase estimates and a power law function decreases estimates. This is because log-linear functions are very similar to linear function at low concentrations, higher at mid-range concentrations (~10-30 μgm^{-3}), and then lower at high concentrations. Power law functions are greater than linear and log-linear at lower concentrations, but tend to level off at lower concentrations than log-linear functions.

Table 4.3 Input for premature mortality estimate sensitivity tests and the resulting percent change in mortality due to chronic exposure as shown in Figure 4.9. Parentheses are for values determined from satellite-based concentrations.

	RR source	Threshold	CR shape	% Change USA	% Change China	Study using method
Krewski	*Krewski et al.*, 2009	No	Linear	base	base	*Evans et al.*, 2013; *Fan et al.*, 2012
Krewski-T	*Krewski et al.*, 2009	Yes	Linear	-62 (-56)	-8 (-8)	*Evans et al.*, 2013; *Annenburg et al.*, 2010
Laden	*Laden et al.*, 2006	No	Linear	89 (87)	45 (47)	*Fan et al.*, 2012
Laden-T	*Laden et al.*, 2006	Yes	Linear	-24 (-14)	36 (38)	
Krewski-30	*Krewski et al.*, 2009	No	Linear to 30 ugm^{-3}	0 (0)	-34 (-31)	*Anenburg et al.*, 2010
Krewski-LL	*Krewski et al.*, 2009	No	Log-Linear	5 (6)	21 (20)	*Evans et al.*, 2013; *Fann and Risley*, 2013
Krewski-PL	*Krewski et al.*, 2009	No	Power Law	52 (44)	-18 (-17)	*Marlier et al.*, 2013

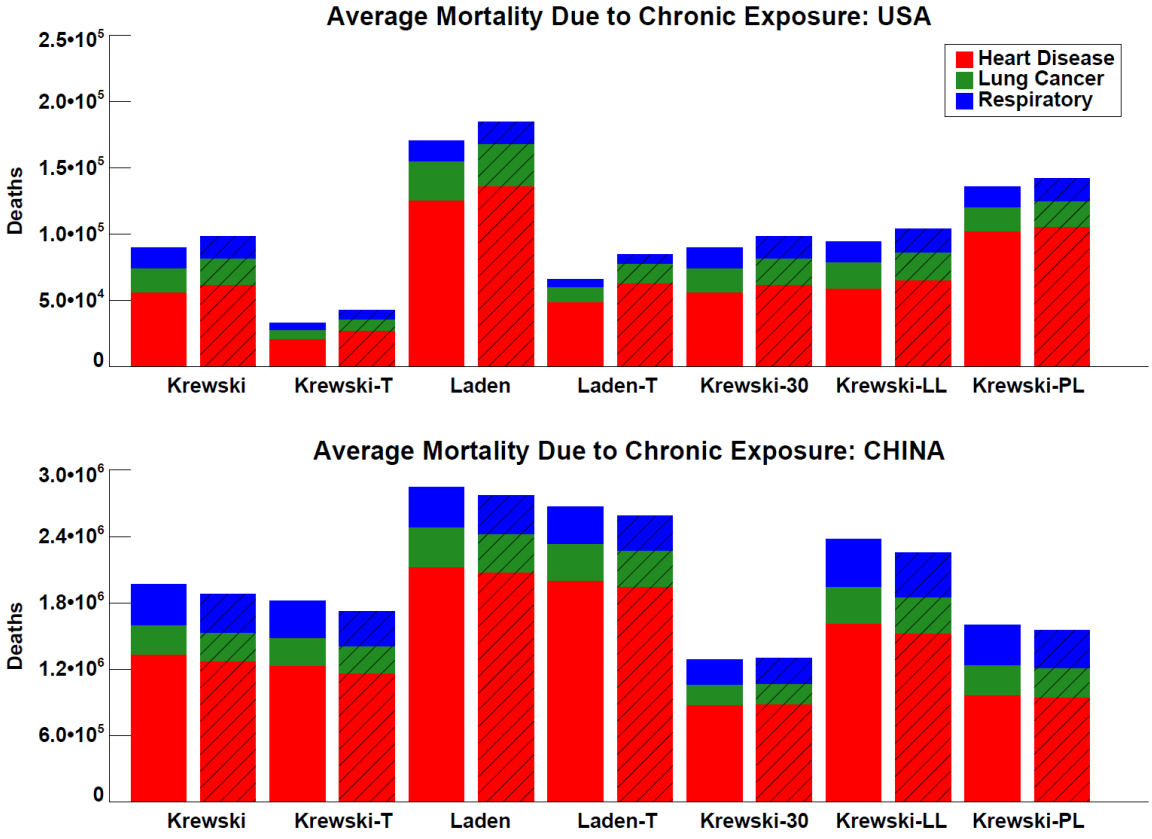

Figure 4.9 Average annual premature mortality due to $PM_{2.5}$ exposure in the U.S. and China as a function of the risk ratio and shape of C-R function. Colors indicate cause of death and satellite-based $PM_{2.5}$ estimates are hatched, model-based estimates are solid. Explanation of input for each estimate is given in Table 4.3.

We conduct similar tests for premature mortality due to acute exposure. In order to not confound these comparisons with sampling issues, we show results using the model for a single year (Figure 4.10). We again test the impact of using a threshold, ceiling and log linear function (Table 4.4). Results are similar to the sensitivity tests for chronic exposure in that the threshold has a significant impact on estimates for the U.S. and much less on estimates for China. Because most mortality associated with acute exposure occurs due to days with extremely high concentrations, using a ceiling value reduces estimates in China by about half. Using a log-linear response function makes the least difference, but does slightly increase estimates.

Table 4.4 Input for premature mortality estimate sensitivity tests for acute exposure and resulting percent change in mortality shown in Figure 4.10.

	Threshold	CR shape	% Change USA	% Change China
Ostro	No	Linear	base	Base
Ostro-T	Yes	Linear	-56%	-10%
Ostro-30	No	Linear to 30 ugm^{-3}	-6%	-51%
Ostro-LL	No	Log-Linear	+1%	+5%

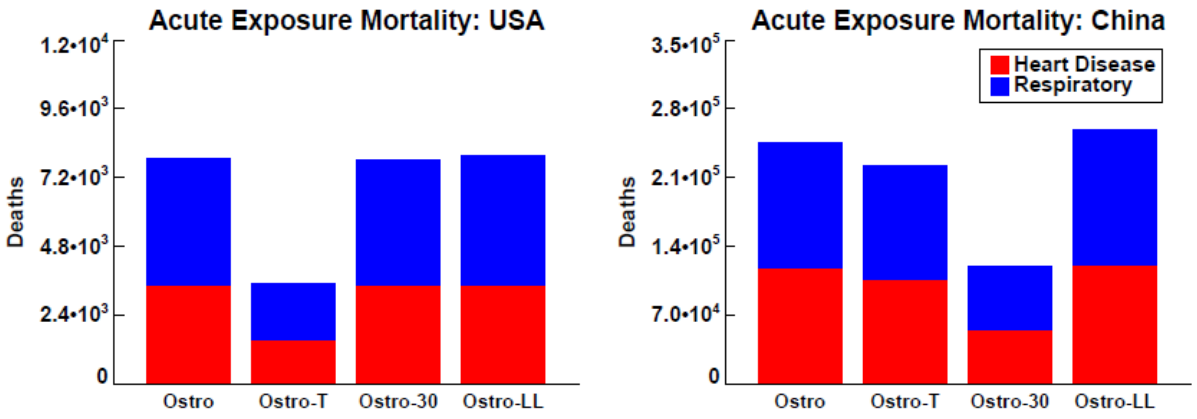

Figure 4.10 Average annual premature mortality due to $PM_{2.5}$ exposure (using model concentrations) in the U.S. and China as a function of the risk ratios and shape of C-R function with colors indicating cause. Explanation of input for each estimate is given in Table 4.4.

While we explore a variety of factors that could influence our results and those of other studies here, there are also many other factors that we have not considered explicitly. For

example, in this study we used bulk mass concentrations of $PM_{2.5}$ and did not take into account any differences in aerosol composition or size as there remains uncertainty about which size and chemical components of PM are most harmful to human health. It is generally suggested that fine particles (or $PM_{2.5}$) are more toxic than coarse particles [*Schwartz et al.*, 1996; *Burnett et al.*, 2000; *Cifuentes, Luis et al.*, 2001; *Franklin et al.*, 2008; *Zanobetti and Schwartz*, 2009]; however, *Ostro et al.* [2006] found stronger effects on mortality with coarser particles. This discrepancy could also be due to difference in composition as *Zanobetti and Schwartz* [2009] show stronger seasonal and regional variability in the coarse mode particle effects than in the fine, suggesting the toxicity could be more related to composition than size. *Krall et al.* [2013] suggest that organic and elemental carbon are more toxic than inorganic species. *Cao et al.* [2012] also find stronger respiratory mortality responses to organics; but for all-cause mortality, the response is stronger for ammonium nitrate. *Franklin et al.* [2008] suggest that the association between non-accidental mortality and $PM_{2.5}$ exposure is greatest when there are greater concentrations of aluminum, arsenic, sulfate, silicon, and nickel, while *Burnett et al.* [2004] suggest mortality is more associated with sulfate ion, iron, nickel, and zinc. *Ostro et al.* [2010] also find that mortality associations with exposure to $PM_{2.5}$ differ by season and composition, suggesting that risk ratios based only on $PM_{2.5}$ mass may underestimate the effects with regards to specific components. This range of conclusions highlights the need for further studies which focus on effects related to composition of particles as understanding which emission sources produce the most toxic particles could be very useful for regulation purposes [*Dominici et al.*, 2014].

4.4 Discussion

In the previous sections, we have discussed a variety of uncertainties and methodological choices made that could impact mortality burden estimates using model and satellite-based $PM_{2.5}$ concentrations. The impact this uncertainty has on our burden estimates is summarized in Figure 4.11 and Figure 4.12. Here we show the average mortality determined using the model or satellite-based concentrations, along with error bars for the confidence intervals on the risk ratios, the uncertainty due to satellite AOD, the potential uncertainty in η, and a range of estimates related to the shape of the C-R function and risk ratios. We show these results separately, but their effects can also be cumulative. From these figures, it is evident that the discrepancy between the satellite-based and model mortality estimates is within the range of uncertainty associated with the satellite AOD. The uncertainty in satellite AOD and the model η also creates a larger range of premature mortality estimates with regards to acute exposure rather than chronic exposure. For chronic exposure, we see that the choice of C-R function and RR clearly spans the largest range of mortality burden estimates. This is not as evident in our figure for acute exposure because we only show results from the limited sensitivity tests shown in the previous section. However, if we compare our results to the range denoted by the confidence intervals, the uncertainty in RR still overshadows the other uncertainty in our mortality burden estimates. If we included all of these uncertainties, we would cover the range of premature

mortality estimates given in Table 4.2.

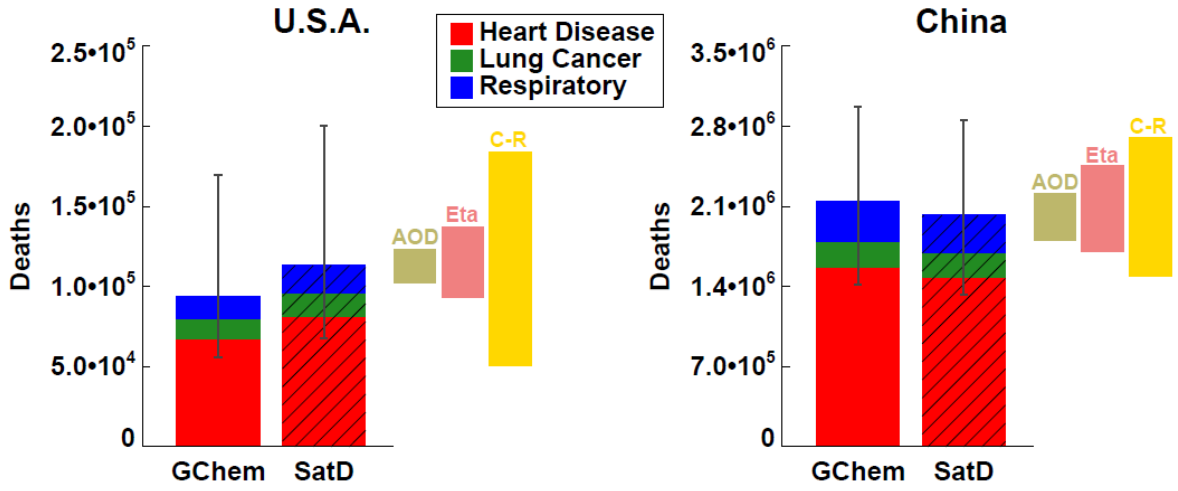

Figure 4.11 Premature mortality burden (colored by cause) due to chronic exposure and uncertainty estimates for the U.S. and China. Hatched bars indicate estimates using satellite-based concentrations; gray lines denote confidence interval for risk ratio. Boxes denoted range of uncertainty in satellite AOD, η, and choice of the concentration-response function and risk ratio for sensitivity tests in Figure 4.9.

It is also important to note that we are calculating national mortality burdens attributable to PM exposure, in which potentially large errors for specific locations can be canceled out in the aggregate. Therefore, regionally we might expect a larger range of estimates.

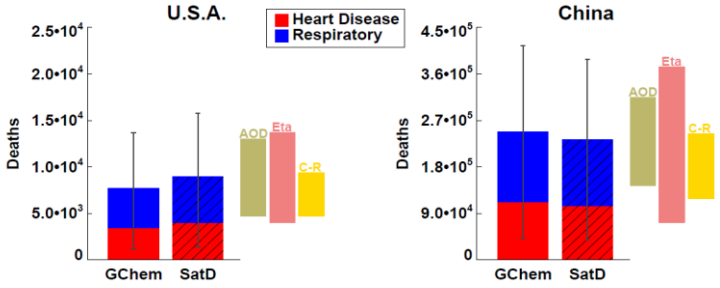

Figure 4.12 Premature mortality burden (colored by cause) due to acute exposure and uncertainty estimates for the U.S. and China. Hatched bars indicate estimates using satellite-based concentrations; gray lines denote confidence interval for risk ratio. Boxes denoted range of uncertainty in satellite AOD, η, and the choice of concentration-response function and risk ratio for sensitivity tests in Figure 4.10.

4.5 Conclusions

While exposure to $PM_{2.5}$ is associated with premature mortality, the magnitude of this response is still unclear and we can conclude that determining mortality estimates at the population level is a tentative process. When put in the perspective of uncertainties in AOD, choices regarding the response functions, and (potentially) the AOD-to-surface-$PM_{2.5}$ relationship; satellite-based or strictly model-based concentrations do not predict significantly different results at the national scale, even though they may produce different surface concentrations. Therefore, while using satellite observations to constrain model $PM_{2.5}$ concentrations can be very useful, there remain many areas linked to this methodology where improvements could reduce the uncertainty in mortality burden and exposure estimates, especially concerning acute exposure.

As we have shown that the potential model biases are not globally consistent and the η relationship is highly variable based on local sources, transported pollution and dust, and meteorology; further regional studies need to be conducted in order to assess and improve the model's ability to simulate AOD, $PM_{2.5}$, and the relationship between the two.

Our sensitivity analysis highlights the strong variability in $PM_{2.5}$ and AOD, suggesting time-averaged (running mean) values or intermittent sampling may not correctly represent concentrations to which people are exposed. Therefore, there should be a stronger push for more daily monitoring and again, while the satellite can help constrain the model in regions with less certainty in model emission inventories, this uncertainty will still have a strong impact on the resulting satellite-based concentrations. Improving spatial and temporal coverage of observations is therefore extremely important.

In regards to mortality burden estimates, a large range of values could be determined bases on our choice of the risk ratio or shape of concentration response function. While the epidemiology studies attempt to statistically account for differences in populations and confounding variables, there is still a large spread in determined risk ratios. Further studies will be necessary to better account for exposure misclassification and reduce uncertainties in potential effect modification and determine how transferable CR functions and RRs are for this kind of study.

5 Summary and Future Work

Exposure to atmospheric aerosols is currently one of the leading environmental concerns pertaining to human health, and these particles are also important for our climate system. Since 1999, the NASA constellation of earth observations has been providing insight into the transport, processing, and budget of atmospheric particles. These remote observations also offer a new set of constraints for improving and investigating model simulations of atmospheric aerosols.

The focus of this dissertation was to use satellite observations of aerosols in conjunction with a global chemical transport model to improve our understanding of surface air quality. The motivation for this work was two-fold: estimating the burden on health and potential impacts on climate. While health impacts were more explicitly discussed, implications from Chapter 2 are also important for climate.

Chapter 2 explored aerosol loading and composition in the Southeastern United States, a region of great interest for climate and for the potential interaction of biogenic and anthropogenic sources. This work revealed the disconnect between the observed seasonality in surface mass concentrations and satellite AOD, which was investigated in the context of a model simulation of aerosol loading in the region. The main conclusions of this section were as follows:

1.) The model underestimates summertime AOD over the SEUS as observed by the suite of NASA satellites and the magnitude of seasonal variation between summer and winter.
2.) The model also slightly underestimates the magnitude and seasonality in surface mass concentrations, but the seasonality in AOD is not being driven by changes in surface mass concentrations, as the magnitude of seasonality in AOD is 100-300% and surface observations only show ~10-40%.
3.) New measurements of vertically-resolved aerosol extinction from the CALIOP instrument suggest that the seasonality in AOD that is not captured by the model is due to an underestimation in extinction above the surface.

4.) Model sensitivity simulations suggest that this above-surface enhancement is not due to a bias in the mixing height or a current surface source but must be due to another missing source.

This study was the most comprehensive investigation of aerosol loading undertaken in the Southeastern United States to date. Despite this, we lacked the observations to pinpoint the source of aerosols in the lower free troposphere. This has however motivated follow-on work. Field measurements taken during the SOAS and SEAC^{4}RS campaigns that occurred in summer 2013 are currently being analyzed which may provide some useful insights (although 2013 had above average rainfall and fire activity). However, in order to determine how representative these field campaigns are to other years in terms of aerosol sources, loading, and meteorology; future work should also explore interannual and interseasonal variability in the region.

The Southeastern U.S. is of interest due to the observed regional cooling and the suggestion that aerosols could be driving temperature trends [*Portmann et al.*, 2009]. This hypothesis was not explored in this work, but our results could be useful for such an endeavor. Modeling studies that have attempted to simulate the "warming hole" have found it centered over the Midwest rather than the Southeastern U.S. [*Leibensperger et al.*, 2012]. A missing source of aerosol over the Southeastern U.S. could explain some of this discrepancy. Additionally, in order to link aerosol trends to temperature trends, knowledge of the vertical profile is important, especially for estimating radiative effects.

In a broader context, this work provides an example of how to combine information from a variety of observations and a model in order to determine what we do not understand (or do not simulate in the model) well about atmospheric processes or aerosol sources and then identifying the observations that would be useful to improve our understanding.

This example links to the other chapters in which we further explored how limited observations, potential model uncertainties, and or details we do not understand about health effects might impact national premature mortality burden estimates and could lead to a range of results. Multi-year average satellite-based $PM_{2.5}$ datasets have been previously used for calculating chronic exposure and mortality burdens. However, long-term mean concentrations are often insensitive to hidden peaks in daily

concentrations, so we explored whether satellite observations could also be used to determine acute exposure. This process is difficult due to the limitations of coverage for polar-orbiting satellites, but some spatial interpolation techniques were analyzed and found to improve coverage without causing drastic biases. We also found that because of strong daily variability, temporal averages are not a good substitute for missing data. While comparisons with surface observations in the U.S. showed that the model underrepresented days with the highest concentrations, the satellite-based estimates did provide more and higher daily concentrations and greater exposure, suggesting potentially better representations of extreme events (especially with regards to fires). Surface $PM_{2.5}$ measurements in China were not available for the time period studied here, but are now being taken more routinely. Future work comparing model simulated values to these observations will help us better understand model performance in China. The main conclusion from Chapter 3 was that acute exposure could account for 2-3% of the premature mortality due to $PM_{2.5}$ exposure in the U.S. and 15-20% in China. We provided these as a national-level burden because (much like concentrations); using nationally-representative response functions can cause large errors for specific locations, but some of these errors should cancel out in aggregate.

However, not all errors in mortality burden calculations cancel out on longer timescales or in national totals. Therefore, in Chapter 4, we explored the uncertainty in our mortality burden estimates which also provides a context for comparing the range of results from previous studies (Table 4.2) that used modeled, spatially interpolated surface observations, and satellite-based concentrations. In this section, we detail how different assumptions could account for some of the discrepancies in these estimates and where most of the uncertainty in these endeavors arises. In particular, we looked at potential model uncertainties, satellite AOD uncertainty, and different functions for determining mortality burden estimates. To our knowledge, this is the first time that the role of these different factors has been compared.

In terms of model uncertainty, our goal was to determine the variables that are the most important for the model to simulate correctly in each region and the range of potential model uncertainty. The sensitivity tests using the seasonally-averaged values stress the importance of potential systematic biases

and the need for daily-resolved values. We saw that the vertical profile was generally the most important variable, but grid resolution and relative humidity can also be important on specific days and in specific locations. These sensitivity analyses suggest that more regional studies (such as our example over the Southeastern U.S.) should be conducted in the future to better understand and better simulate the AOD-$PM_{2.5}$ relationship (η). The developing SPARTAN network will hopefully provide more continuous measurements of $PM_{2.5}$ collocated with AERONET observations in order to have measurements against which to compare the model η [*Snider et al.*, 2014].While satellite observations may be useful for constraining model concentration estimates, ultimately the goal for model development should be to improve the model simulation with updated emission inventories and better representation of atmospheric processes. This is specifically important in China as potential model uncertainty is much larger as shown not only by our sensitivity analysis, but also because of the lack of long-term surface measurements to validate against. As concerns over air quality continue to rise, more measurements of surface $PM_{2.5}$ are being made available and can provide necessary model validation.

We also compared the potential model uncertainty to uncertainty in the satellite retrieval of AOD, which, in an aggregate sense, was less. However, comparisons with AERONET sites showed greater errors at specific locations (desert regions) and on shorter timescales. These uncertainties in satellite AOD and model simulated AOD-$PM_{2.5}$ relationships more than accounts for the difference between the model estimates and the satellite-based estimates of mortality.

Reducing uncertainty in satellite-retrieved AOD is a matter of both improving coverage and better accounting for regional biases. Future geostationary satellites (such as TEMPO) will be useful for both of these endeavors and provide extremely valuable information for daily monitoring and tracking of air quality. It is difficult to globally correct MODIS for differing regional biases and sampling issues, but increased temporal coverage over a fixed location should aid in reducing potential biases. Additionally, the ability to track plumes would improve our understanding of local and regional influences and provide beneficial information for regulating and reducing exposure.

Finally, we examined the range of premature mortality burden estimates determined using different health-risk assumptions. In our work, we determine concentration not exposure. Therefore, to take atmospheric concentrations and then estimate the burden of mortality, several different assumptions have to be made:

(1) Is there a causal relationship?

(2) What is the shape of the concentration-response function (and is this response transferable to different regions and people groups)?

(3) What is the magnitude of this relationship (as different studies determine different risk ratios) and does it vary (can we use one value for all locations, at all times and under all circumstances)?

Our mortality estimates differed from estimates in other studies, not only because the $PM_{2.5}$ concentrations differ (associated with our use of the latest MODIS Collection 6 product, as well as the specific version of the GEOS-Chem model), but also because of our method to calculate this burden. We show that the range of results that can be obtained when applying different concentration-response functions was much larger than the range of uncertainty from the model or satellite AOD. This demonstrates the tentative nature of calculating national-level mortality burdens despite the strides made in using satellite to improving exposure estimates. This work emphasizes that there is a need for further bridges to be built between atmospheric chemistry research and epidemiology to improve our understanding of the health burden of atmospheric particulate matter.

This study has focused on the potential of using satellite-based PM2.5 as a measure of ambient concentrations. Compared to using fixed-site monitors, this could potentially reduce exposure misclassification. However, we are still using results on a coarse grid resolution to estimate population exposures. Future work is required to improve how we relate ambient concentrations to population- and personal-level exposure. Additionally, as there exists a large range of relative risk ratios in the literature, additional guidance from the epidemiological community on how transferable results from each study are beyond the context in which they were determined, both in terms of the population characteristics and the particulars of the PM mass (size, composition, etc.), is necessary.

In this dissertation, I explored how atmospheric observations and model simulations can be synthesized in order to investigate aerosol loadings and estimate health impacts. The complexity of atmospheric PM suggests that new observations of aerosol composition and properties, both in situ and remotely sensed, will be critical for further model development and application for health studies.

6 References

Alexander, B., R. J. Park, D. J. Jacob, Q. B. Li, R. M. Yantosca, J. Savarino, C. C. W. Lee, and M. H. Thiemens (2005), Sulfate formation in sea-salt aerosols: Constraints from oxygen isotopes, *J. Geophys. Res.*, *110*(D10), D10307, doi:10.1029/2004JD005659.

Alston, E. J., I. N. Sokolik, and O. V. Kalashnikova (2012), Characterization of atmospheric aerosol in the US Southeast from ground- and space-based measurements over the past decade, *Atmospheric Meas. Tech.*, *5*(7), 1667–1682, doi:10.5194/amt-5-1667-2012.

Anderson, H. R. et al. (2012), Satellite-based Estimates of Ambient Air Pollution and Global Variations in Childhood Asthma Prevalence, *Environ. Health Perspect.*, *120*(9), 1333–1339, doi:10.1289/ehp.1104724.

Anenberg, S. C., L. W. Horowitz, D. Q. Tong, and J. J. West (2010), An Estimate of the Global Burden of Anthropogenic Ozone and Fine Particulate Matter on Premature Human Mortality Using Atmospheric Modeling, *Environ. Health Perspect.*, *118*(9), 1189–1195, doi:10.1289/ehp.0901220.

Aunan, K., and X.-C. Pan (2004), Exposure-response functions for health effects of ambient air pollution applicable for China -- a meta-analysis, *Sci. Total Environ.*, *329*(1-3), 3–16, doi:10.1016/j.scitotenv.2004.03.008.

Bell, M. L., K. Ebisu, R. D. Peng, J. Walker, J. M. Samet, S. L. Zeger, and F. Dominici (2008), Seasonal and Regional Short-term Effects of Fine Particles on Hospital Admissions in 202 US Counties, 1999–2005, *Am. J. Epidemiol.*, *168*(11), 1301–1310, doi:10.1093/aje/kwn252.

Bell, M. L., K. Ebisu, R. D. Peng, J. M. Samet, and F. Dominici (2009), Hospital Admissions and Chemical Composition of Fine Particle Air Pollution, *Am. J. Respir. Crit. Care Med.*, *179*(12), 1115–1120, doi:10.1164/rccm.200808-1240OC.

Berman, S., J.-Y. Ku, J. Zhang, and S. T. Rao (1997), Uncertainties in estimating the mixing depth-comparing three mixing-depth models with profiler measurements, *Atmos. Environ.*, *31*(18), 3023–3039, doi:10.1016/S1352-2310(97)00118-0.

Burnett, R. T., J. Brook, T. Dann, C. Delocla, O. Philips, S. Cakmak, R. Vincent, M. S. Goldberg, and D. Krewski (2000), Association between particulate- and gas-phase components of urban air pollution and daily mortality in eight Canadian cities, *Inhal. Toxicol.*, *12 Suppl 4*, 15–39, doi:10.1080/08958370050164851.

Cao, J., H. Xu, Q. Xu, B. Chen, and H. Kan (2012), Fine Particulate Matter Constituents and Cardiopulmonary Mortality in a Heavily Polluted Chinese City, *Environ. Health Perspect.*, *120*(3), 373–378, doi:10.1289/ehp.1103671.

Carlton, A. G., B. J. Turpin, K. E. Altieri, S. P. Seitzinger, R. Mathur, S. J. Roselle, and R. J. Weber (2008), CMAQ Model Performance Enhanced When In-Cloud Secondary Organic

Aerosol is Included: Comparisons of Organic Carbon Predictions with Measurements, *Environ. Sci. Technol.*, *42*(23), 8798–8802, doi:10.1021/es801192n.

Chen, H., and M. S. Goldberg (2009), The effects of outdoor air pollution on chronic illnesses, *McGill J. Med. MJM*, *12*(1), 58–64.

Chen, H. et al. (2013a), Risk of Incident Diabetes in Relation to Long-term Exposure to Fine Particulate Matter in Ontario, Canada, *Environ. Health Perspect.*, *121*(7), 804–810, doi:10.1289/ehp.1205958.

Chen, H. et al. (2013b), Spatial Association between Ambient Fine Particulate Matter and Incident Hypertension, *Circulation*, CIRCULATIONAHA.113.003532, doi:10.1161/CIRCULATIONAHA.113.003532.

Chen, W., R. Zheng, S. Zhang, P. Zhao, G. Li, L. Wu, and J. He (2013c), The incidences and mortalities of major cancers in China, 2009, *Chin. J. Cancer*, *32*(3), 106–112, doi:10.5732/cjc.013.10018.

Chin, M., D. L. Savoie, B. J. Huebert, A. R. Bandy, D. C. Thornton, T. S. Bates, P. K. Quinn, E. S. Saltzman, and W. J. De Bruyn (2000), Atmospheric sulfur cycle simulated in the global model GOCART: Comparison with field observations and regional budgets, *J. Geophys. Res. Atmospheres*, *105*(D20), 24689–24712, doi:10.1029/2000JD900385.

Chin, M. et al. (2014), Multi-decadal aerosol variations from 1980 to 2009: a perspective from observations and a global model, *Atmos Chem Phys*, *14*(7), 3657–3690, doi:10.5194/acp-14-3657-2014.

Cifuentes, Luis, Borja-Aburto, V. H., Gouveia, Nelson, Thurston, George, and Davis, Devra Lee (2001), Assessing the health benefits of urban air pollution reductions associated with climate change mitigation (2000-2020): Santiago, Sao Paulo, Mexico City, and New York City, *Environ Health Perspect*, *109*(Suppl 3), 419.

Cohen, A. J. et al. (2005), The global burden of disease due to outdoor air pollution, *J. Toxicol. Environ. Health A*, *68*(13-14), 1301–1307, doi:10.1080/15287390590936166.

COMEAP (2006), *Cardiovascular disease and air pollution: a report by the Committee On the Medical Effects of Air Pollutants*, COMEAP, UK.

Cooke, W. F., C. Liousse, H. Cachier, and J. Feichter (1999), Construction of a 1° × 1° fossil fuel emission data set for carbonaceous aerosol and implementation and radiative impact in the ECHAM4 model, *J. Geophys. Res.*, *104*(D18), 22137–22,162, doi:10.1029/1999JD900187.

Crouse, D. L. et al. (2012), Risk of Nonaccidental and Cardiovascular Mortality in Relation to Long-term Exposure to Low Concentrations of Fine Particulate Matter: A Canadian National-Level Cohort Study, *Environ. Health Perspect.*, *120*(5), 708–714, doi:10.1289/ehp.1104049.

Diner, D. J. et al. (2005), The value of multiangle measurements for retrieving structurally and radiatively consistent properties of clouds, aerosols, and surfaces, *Remote Sens. Environ.*, *97*(4), 495–518, doi:10.1016/j.rse.2005.06.006.

Dockery, D. W., C. A. Pope, X. Xu, J. D. Spengler, J. H. Ware, M. E. Fay, B. G. Ferris, and F. E. Speizer (1993), An association between air pollution and mortality in six U.S. cities, *N. Engl. J. Med.*, *329*(24), 1753–1759, doi:10.1056/NEJM199312093292401.

Dominici, F., R. D. Peng, M. L. Bell, L. Pham, A. McDermott, S. L. Zeger, and J. M. Samet (2006), Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases, *JAMA J. Am. Med. Assoc.*, *295*(10), 1127–1134, doi:10.1001/jama.295.10.1127.

Dominici, F., M. Greenstone, and C. R. Sunstein (2014), Particulate Matter Matters, *Science*, *344*(6181), 257–259, doi:10.1126/science.1247348.

Van Donkelaar, A., R. V. Martin, and R. J. Park (2006), Estimating ground-level PM2.5 using aerosol optical depth determined from satellite remote sensing, *J. Geophys. Res. Atmospheres*, *111*(D21), n/a–n/a, doi:10.1029/2005JD006996.

Van Donkelaar, A., R. V. Martin, M. Brauer, R. Kahn, R. Levy, C. Verduzco, and P. J. Villeneuve (2010), Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application, *Environ. Health Perspect.*, *118*(6), 847–855, doi:10.1289/ehp.0901623.

Van Donkelaar, A., R. V. Martin, R. C. Levy, A. M. da Silva, M. Krzyzanowski, N. E. Chubarova, E. Semutnikova, and A. J. Cohen (2011), Satellite-based estimates of ground-level fine particulate matter during extreme events: A case study of the Moscow fires in 2010, *Atmos. Environ.*, *45*(34), 6225–6232, doi:10.1016/j.atmosenv.2011.07.068.

Van Donkelaar, A., R. V. Martin, A. N. Pasch, J. J. Szykman, L. Zhang, Y. X. Wang, and D. Chen (2012), Improving the accuracy of daily satellite-derived ground-level fine aerosol concentration estimates for North America, *Environ. Sci. Technol.*, *46*(21), 11971–11978, doi:10.1021/es3025319.

Van Donkelaar, A., R. V. Martin, R. J. D. Spurr, E. Drury, L. A. Remer, R. C. Levy, and J. Wang (2013), Optimal estimation for global ground-level fine particulate matter concentrations, *J. Geophys. Res. Atmospheres*, *118*(11), 5621–5636, doi:10.1002/jgrd.50479.

Drury, E., D. J. Jacob, R. J. D. Spurr, J. Wang, Y. Shinozuka, B. E. Anderson, A. D. Clarke, J. Dibb, C. McNaughton, and R. Weber (2010), Synthesis of satellite (MODIS), aircraft (ICARTT), and surface (IMPROVE, EPA-AQS, AERONET) aerosol observations over eastern North America to improve MODIS aerosol retrievals and constrain surface aerosol concentrations and sources, *J. Geophys. Res.*, *115*(D14), D14204, doi:10.1029/2009JD012629.

Duncan, B. N., R. V. Martin, A. C. Staudt, R. Yevich, and J. A. Logan (2003), Interannual and seasonal variability of biomass burning emissions constrained by satellite observations, *J. Geophys. Res. Atmospheres*, *108*(D2), ACH 1–1–ACH 1–22, doi:10.1029/2002JD002378.

Edgerton, E. S., B. E. Hartsell, R. D. Saylor, J. J. Jansen, D. A. Hansen, and G. M. Hidy (2005), The Southeastern Aerosol Research and Characterization Study: Part II. Filter-based measurements of fine and coarse particulate matter mass and composition, *J. Air Waste Manag. Assoc. 1995*, *55*(10), 1527–1542.

Edgerton, E. S., B. E. Hartsell, R. D. Saylor, J. J. Jansen, D. A. Hansen, and G. M. Hidy (2006), The Southeastern Aerosol Research and Characterization Study, part 3: continuous measurements of fine particulate matter mass and composition, *J. Air Waste Manag. Assoc. 1995*, *56*(9), 1325–1341.

Engel-Cox, J. A., R. M. Hoff, and A. D. J. Haymet (2004), Recommendations on the use of satellite remote-sensing data for urban air quality, *J. Air Waste Manag. Assoc. 1995*, *54*(11), 1360–1371.

Engel-Cox, J. A., R. M. Hoff, R. Rogers, F. Dimmick, A. C. Rush, J. J. Szykman, J. Al-Saadi, D. A. Chu, and E. R. Zell (2006), Integrating lidar and satellite optical depth with ambient monitoring for 3-dimensional particulate characterization, *Atmos. Environ.*, *40*(40), 8056–8067, doi:10.1016/j.atmosenv.2006.02.039.

Ervens, B., B. J. Turpin, and R. J. Weber (2011), Secondary organic aerosol formation in cloud droplets and aqueous particles (aqSOA): a review of laboratory, field and model studies, *Atmospheric Chem. Phys. Discuss.*, *11*(8), 22301–22383, doi:10.5194/acpd-11-22301-2011.

Evans, J., A. van Donkelaar, R. V. Martin, R. Burnett, D. G. Rainham, N. J. Birkett, and D. Krewski (2013), Estimates of global mortality attributable to particulate air pollution using satellite imagery, *Environ. Res.*, *120*, 33–42, doi:10.1016/j.envres.2012.08.005.

Fairlie, D. T., D. J. Jacob, and R. J. Park (2007), The impact of transpacific transport of mineral dust in the United States, *Atmos. Environ.*, *41*(6), 1251–1266, doi:10.1016/j.atmosenv.2006.09.048.

Fang, Y., V. Naik, L. W. Horowitz, and D. L. Mauzerall (2013), Air pollution and associated human mortality: the role of air pollutant emissions, climate change and methane concentration increases from the preindustrial period to present, *Atmos Chem Phys*, *13*(3), 1377–1394, doi:10.5194/acp-13-1377-2013.

Fleischer, N. L., M. Merialdi, A. van Donkelaar, F. Vadillo-Ortega, R. V. Martin, A. P. Betran, J. P. Souza, and M. S. O'Neill (2014), Outdoor Air Pollution, Preterm Birth, and Low Birth Weight: Analysis of the World Health Organization Global Survey on Maternal and Perinatal Health, *Environ. Health Perspect.*, doi:10.1289/ehp.1306837.

Ford, B., and C. L. Heald (2012), An A-train and model perspective on the vertical distribution of aerosols and CO in the Northern Hemisphere, *J. Geophys. Res.*, *117*(D6), D06211, doi:10.1029/2011JD016977.

Franklin, M., P. Koutrakis, and P. Schwartz (2008), The role of particle composition on the association between PM2.5 and mortality, *Epidemiol. Camb. Mass*, *19*(5), 680–689.

Goldstein, A. H., C. D. Koven, C. L. Heald, and I. Y. Fung (2009), Biogenic carbon and anthropogenic pollutants combine to form a cooling haze over the southeastern United States, *Proc. Natl. Acad. Sci.*, *106*(22), 8835–8840, doi:10.1073/pnas.0904128106.

Guenther, A., T. Karl, P. Harley, C. Wiedinmyer, P. I. Palmer, and C. Geron (2006), Estimates of global terrestrial isoprene emissions using MEGAN (Model of Emissions of Gases and Aerosols from Nature), *Atmos Chem Phys*, *6*(11), 3181–3210, doi:10.5194/acp-6-3181-2006.

Hallquist, å. M., M. Jerksjö, H. Fallgren, J. Westerlund, and å. Sjödin (2012), Particle and gaseous emissions from individual diesel and CNG buses, *Atmospheric Chem. Phys. Discuss.*, *12*(10), 27737–27773, doi:10.5194/acpd-12-27737-2012.

Hand, J. L., B. A. Schichtel, M. Pitchford, W. C. Malm, and N. H. Frank (2012), Seasonal composition of remote and urban fine particulate matter in the United States, *J. Geophys. Res.*, *117*(D5), D05209, doi:10.1029/2011JD017122.

Hansen, D. A., E. S. Edgerton, B. E. Hartsell, J. J. Jansen, N. Kandasamy, G. M. Hidy, and C. L. Blanchard (2003), The Southeastern Aerosol Research and Characterization Study: part 1--Overview, *J. Air Waste Manag. Assoc. 1995*, *53*(12), 1460–1471.

Heald, C. L., M. J. Wilkinson, R. K. Monson, C. A. Alo, G. Wang, and A. Guenther (2009), Response of isoprene emission to ambient CO2 changes and implications for global budgets, *Glob. Change Biol.*, *15*(5), 1127–1140, doi:10.1111/j.1365-2486.2008.01802.x.

Heald, C. L. et al. (2011), Exploring the vertical profile of atmospheric organic aerosol: comparing 17 aircraft field campaigns with a global model, *Atmos Chem Phys*, *11*(24), 12673–12696, doi:10.5194/acp-11-12673-2011.

Heald, C. L. et al. (2012), Atmospheric ammonia and particulate inorganic nitrogen over the United States, *Atmos Chem Phys*, *12*(21), 10295–10312, doi:10.5194/acp-12-10295-2012.

Henze, D. K., J. H. Seinfeld, N. L. Ng, J. H. Kroll, T.-M. Fu, D. J. Jacob, and C. L. Heald (2008), Global modeling of secondary organic aerosol formation from aromatic hydrocarbons: high- vs. low-yield pathways, *Atmos Chem Phys*, *8*(9), 2405–2420, doi:10.5194/acp-8-2405-2008.

Holben, B. N. et al. (1998), AERONET—A Federated Instrument Network and Data Archive for Aerosol Characterization, *Remote Sens. Environ.*, *66*(1), 1–16, doi:10.1016/S0034-4257(98)00031-5.

Hoyle, C. R. et al. (2011), A review of the anthropogenic influence on biogenic secondary organic aerosol, *Atmos Chem Phys*, *11*(1), 321–343, doi:10.5194/acp-11-321-2011.

Hudman, R. C. et al. (2007), Surface and lightning sources of nitrogen oxides over the United States: Magnitudes, chemical evolution, and outflow, *J. Geophys. Res.*, *112*(D12), D12S05, doi:10.1029/2006JD007912.

Hudman, R. C., L. T. Murray, D. J. Jacob, D. B. Millet, S. Turquety, S. Wu, D. R. Blake, A. H. Goldstein, J. Holloway, and G. W. Sachse (2008), Biogenic versus anthropogenic sources of CO in the United States, *Geophys. Res. Lett.*, *35*(4), L04801, doi:10.1029/2007GL032393.

Hunt, W. H., D. M. Winker, M. A. Vaughan, K. A. Powell, P. L. Lucker, and C. Weimer (2009), CALIPSO Lidar Description and Performance Assessment, *J. Atmospheric Ocean. Technol.*, *26*(7), 1214–1228, doi:10.1175/2009JTECHA1223.1.

Hyer, E. J., J. S. Reid, and J. Zhang (2011), An over-land aerosol optical depth data set for data assimilation by filtering, correction, and aggregation of MODIS Collection 5 optical depth retrievals, *Atmos Meas Tech*, *4*(3), 379–408, doi:10.5194/amt-4-379-2011.

Hystad, P., P. A. Demers, K. C. Johnson, J. Brook, A. van Donkelaar, L. Lamsal, R. Martin, and M. Brauer (2012), Spatiotemporal air pollution exposure assessment for a Canadian population-based lung cancer case-control study, *Environ. Health*, *11*(1), 22, doi:10.1186/1476-069X-11-22.

Isaaks, E. H., and R. M. Srivastava (1989), *An Introduction to Applied Geostatistics*, Oxford University Press, Oxford.

Jaeglé, L., P. K. Quinn, T. S. Bates, B. Alexander, and J.-T. Lin (2011), Global distribution of sea salt aerosols: new constraints from in situ and remote sensing observations, *Atmos Chem Phys*, *11*(7), 3137–3157, doi:10.5194/acp-11-3137-2011.

Jerrett, M., A. Arain, P. Kanaroglou, B. Beckerman, D. Potoglou, T. Sahsuvaroglu, J. Morrison, and C. Giovis (2005), A review and evaluation of intraurban air pollution exposure models, *J. Expo. Anal. Environ. Epidemiol.*, *15*(2), 185–204, doi:10.1038/sj.jea.7500388.

Khylstov, A., C. O. Stanier, S. Takahama, and S. N. Pandis (2005), Water Content of ambient aerosol during the Pittsburgh air quality Study, *J. Geophys. Res.*, *110*(D07s10), doi:10.1029/2004JD004651.

Klemm, R. J., and R. Mason (2003), *Revised analyses of time-series studies of air pollution and health, PartII*, Health Effects Institute, Boston, MA.

Koepke, P., M. Hess, I. Schult, and E. P. Shettle (1997), *Global Aerosol Data Set*, Max-Planck-Institut für Meteorologie, Hamburg.

Krall, J. R., G. B. Anderson, F. Dominici, M. L. Bell, and R. D. Peng (2013), Short-term exposure to particulate matter constituents and mortality in a national study of U.S. urban communities, *Environ. Health Perspect.*, *121*(10), 1148–1153, doi:10.1289/ehp.1206185.

Krewski, D., R. T. Burnett, M. S. Goldberg, K. Hoover, J. Siemiatyckis, M. Jerrett, M. Abrahamowicz, and W. H. White (2000), *Reanalysis of the Harvard Six Cities Study and the American Cancer Society Study of Particulate Air Pollution and Mortality*, A Special Report of the Institute's Particle Epidemiology Reanalysis Project, Health Effect's Institute, Cambride, MA.

Krewski, D. et al. (2009), Extended follow-up and spatial analysis of the American Cancer Society study linking particulate air pollution and mortality, *Res. Rep. Health Eff. Inst.*, (140), 5–114; discussion 115–136.

Laden, F., J. Schwartz, F. E. Speizer, and D. W. Dockery (2006), Reduction in Fine Particulate Air Pollution and Mortality, *Am. J. Respir. Crit. Care Med.*, *173*(6), 667–672, doi:10.1164/rccm.200503-443OC.

Lee, D. T., and B. J. Schachter (1980), Two algorithms for constructing a Delaunay triangulation, *Int. J. Comput. Inf. Sci.*, *9*(3), 219–242, doi:10.1007/BF00977785.

Leibensperger, E. M., L. J. Mickley, D. J. Jacob, W.-T. Chen, J. H. Seinfeld, A. Nenes, P. J. Adams, D. G. Streets, N. Kumar, and D. Rind (2012), Climatic effects of 1950–2050 changes in US anthropogenic aerosols – Part 1: Aerosol trends and radiative forcing, *Atmos Chem Phys*, *12*(7), 3333–3348, doi:10.5194/acp-12-3333-2012.

Lelieveld, J., C. Barlas, D. Giannadaki, and A. Pozzer (2013), Model calculated global, regional and megacity premature mortality due to air pollution, *Atmos Chem Phys*, *13*(14), 7023–7037, doi:10.5194/acp-13-7023-2013.

Levin, E. J. T., S. M. Kreidenweis, G. R. McMeeking, C. M. Carrico, J. L. Collett Jr., and W. C. Malm (2009), Aerosol physical, chemical and optical properties during the Rocky Mountain Airborne Nitrogen and Sulfur study, *Atmos. Environ.*, *43*(11), 1932–1939, doi:10.1016/j.atmosenv.2008.12.042.

Levin, E. J. T. et al. (2010), Biomass burning smoke aerosol properties measured during Fire Laboratory at Missoula Experiments (FLAME), *J. Geophys. Res. Atmospheres*, *115*(D18), D18210, doi:10.1029/2009JD013601.

Levy, R. C., S. Mattoo, L. A. Munchak, L. A. Remer, A. M. Sayer, F. Patadia, and N. C. Hsu (2013), The Collection 6 MODIS aerosol products over land and ocean, *Atmos Meas Tech*, *6*(11), 2989–3034, doi:10.5194/amt-6-2989-2013.

Lim, H.-J., and B. J. Turpin (2002), Origins of Primary and Secondary Organic Aerosol in Atlanta: Results of Time-Resolved Measurements during the Atlanta Supersite Experiment, *Environ. Sci. Technol.*, *36*(21), 4489–4496, doi:10.1021/es0206487.

Lim, S. S. et al. (2012), A comparative risk assessment of burden of disease and injury attributable to 67 risk factors and risk factor clusters in 21 regions, 1990-2010: a systematic analysis for the Global Burden of Disease Study 2010, *Lancet*, *380*(9859), 2224–2260, doi:10.1016/S0140-6736(12)61766-8.

Liu, Y., R. J. Park, D. J. Jacob, Q. Li, V. Kilaru, and J. A. Sarnat (2004), Mapping annual mean ground-level PM2.5 concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States, *J. Geophys. Res. Atmospheres*, *109*(D22), D22206, doi:10.1029/2004JD005025.

Liu, Y., P. Koutrakis, and R. Kahn (2007), Estimating fine particulate matter component concentrations and size distributions using satellite-retrieved fractional aerosol optical depth: part 1--method development, *J. Air Waste Manag. Assoc. 1995*, *57*(11), 1351–1359.

Luo, Y., X. Zheng, T. Zhao, and J. Chen (2014), A climatology of aerosol optical depth over China from recent 10 years of MODIS remote sensing data, *Int. J. Climatol.*, *34*(3), 863–870, doi:10.1002/joc.3728.

Malm, W. C., and J. L. Hand (2007), An examination of the physical and optical properties of aerosols collected in the IMPROVE program, *Atmos. Environ.*, *41*(16), 3407–3427, doi:10.1016/j.atmosenv.2006.12.012.

Malm, W. C., J. F. Sisler, D. Huffman, R. A. Eldred, and T. A. Cahill (1994), Spatial and seasonal trends in particle concentration and optical extinction in the United States, *J. Geophys. Res. Atmospheres*, *99*(D1), 1347–1370, doi:10.1029/93JD02916.

Mamouri, R. E., V. Amiridis, A. Papayannis, E. Giannakaki, G. Tsaknakis, and D. S. Balis (2009), Validation of CALIPSO space-borne-derived attenuated backscatter coefficient profiles using a ground-based lidar in Athens, Greece, *Atmos Meas Tech*, *2*(2), 513–522, doi:10.5194/amt-2-513-2009.

Marlier, M. E., R. S. DeFries, A. Voulgarakis, P. L. Kinney, J. T. Randerson, D. T. Shindell, Y. Chen, and G. Faluvegi (2013), El Nino and health risks from landscape fire emissions in southeast Asia, *Nat. Clim. Change*, *3*(2), 131–136, doi:10.1038/nclimate1658.

Marsik, F. J., K. W. Fischer, T. D. McDonald, and P. J. Samson (1995), Comparison of Methods for Estimating Mixing Height Used during the 1992 Atlanta Field Intensive, *J. Appl. Meteorol.*, *34*(8), 1802–1814, doi:10.1175/1520-0450(1995)034<1802:COMFEM>2.0.CO;2.

Martin, R. V., D. J. Jacob, R. M. Yantosca, M. Chin, and P. Ginoux (2003), Global and regional decreases in tropospheric oxidants from photochemical effects of aerosols, *J. Geophys. Res.*, *108*(D3), 4097, doi:10.1029/2002JD002622.

Martonchik, J. V., D. J. Diner, K. A. Crean, and M. A. Bull (2002), Regional aerosol retrieval results from MISR, *IEEE Trans. Geosci. Remote Sens.*, *40*(7), 1520–1531, doi:10.1109/TGRS.2002.801142.

Martonchik, J. V., R. A. Kahn, and D.J. Diner (2009), Retrieval of aerosol properties over land using MISR observations, in *Satellite Aerosol Remote Sensing over Land*, pp. 267–293, Springer, Berlin.

Miranda, M. L., S. E. Edwards, M. H. Keating, and C. J. Paul (2011), Making the Environmental Justice Grade: The Relative Burden of Air Pollution Exposure in the United States, *Int. J. Environ. Res. Public. Health*, *8*(6), 1755–1771, doi:10.3390/ijerph8061755.

Mona, L., G. Pappalardo, A. Amodeo, G. D'Amico, F. Madonna, A. Boselli, A. Giunta, F. Russo, and V. Cuomo (2009), One year of CNR-IMAA multi-wavelength Raman lidar measurements in coincidence with CALIPSO overpasses: Level 1 products comparison, *Atmos Chem Phys*, *9*(18), 7213–7228, doi:10.5194/acp-9-7213-2009.

Nastos, P. T., A. G. Paliatsos, M. B. Anthracopoulos, E. S. Roma, and K. N. Priftis (2010), Outdoor particulate matter and childhood asthma admissions in Athens, Greece: a time-series study, *Environ. Health*, *9*(1), 45, doi:10.1186/1476-069X-9-45.

Omar, A. H., J.-G. Won, D. M. Winker, S.-C. Yoon, O. Dubovik, and M. P. McCormick (2005), Development of global aerosol models using cluster analysis of Aerosol Robotic Network (AERONET) measurements, *J. Geophys. Res. Atmospheres*, *110*(D10), D10S14, doi:10.1029/2004JD004874.

Omar, A. H. et al. (2009), The CALIPSO Automated Aerosol Classification and Lidar Ratio Selection Algorithm, *J. Atmospheric Ocean. Technol.*, *26*(10), 1994–2014, doi:10.1175/2009JTECHA1231.1.

Ostro, B., R. Broadwin, S. Green, W.-Y. Feng, and M. Lipsett (2006), Fine particulate air pollution and mortality in nine California counties: results from CALFINE, *Environ. Health Perspect.*, *114*(1), 29–33.

Ostro, B., M. Lipsett, P. Reynolds, D. Goldberg, A. Hertz, C. Garcia, K. D. Henderson, and L. Bernstein (2010), Long-term exposure to constituents of fine particulate air pollution and mortality: results from the California Teachers Study, *Environ. Health Perspect.*, *118*(3), 363–369, doi:10.1289/ehp.0901181.

Paciorek, C. J., and Y. Liu (2009), Limitations of remotely sensed aerosol as a spatial proxy for fine particulate matter, *Environ. Health Perspect.*, *117*(6), 904–909, doi:10.1289/ehp.0800360.

Park, R. J., D. J. Jacob, M. Chin, and R. V. Martin (2003), Sources of carbonaceous aerosols over the United States and implications for natural visibility, *J GEOPHYS RES*, *108*, 2002–3190.

Park, R. J., D. J. Jacob, B. D. Field, R. M. Yantosca, and M. Chin (2004), Natural and transboundary pollution influences on sulfate-nitrate-ammonium aerosols in the United States: Implications for policy, *J. Geophys. Res. Atmospheres*, *109*(D15), n/a–n/a, doi:10.1029/2003JD004473.

Peltier, R. E., A. P. Sullivan, R. J. Weber, A. G. Wollny, J. S. Holloway, C. A. Brock, J. A. de Gouw, and E. L. Atlas (2007), No evidence for acid-catalyzed secondary organic aerosol formation in power plant plumes over metropolitan Atlanta, Georgia, *Geophys. Res. Lett.*, *34*(6), n/a–n/a, doi:10.1029/2006GL028780.

Petters, M. D., and S. M. Kreidenweis (2007), A single parameter representation of hygroscopic growth and cloud condensation nucleus activity, *Atmos Chem Phys*, *7*(8), 1961–1971, doi:10.5194/acp-7-1961-2007.

Pierce, J. R., M. J. Evans, C. E. Scott, S. D. D'Andrea, D. K. Farmer, E. Swietlicki, and D. V. Spracklen (2013), Weak global sensitivity of cloud condensation nuclei and the aerosol indirect effect to Criegee + SO2 chemistry, *Atmos Chem Phys*, *13*(6), 3163–3176, doi:10.5194/acp-13-3163-2013.

Pope, C. A. (2007), Mortality Effects of Longer Term Exposures to Fine Particulate Air Pollution: Review of Recent Epidemiological Evidence, *Inhal. Toxicol.*, *19*(s1), 33–38, doi:10.1080/08958370701492961.

Pope, C. A., and D. W. Dockery (2013), Air pollution and life expectancy in China and beyond, *Proc. Natl. Acad. Sci.*, *110*(32), 12861–12862, doi:10.1073/pnas.1310925110.

Pope, C. A., D. V. Bates, and M. E. Raizenne (1995), Health effects of particulate air pollution: time for reassessment?, *Environ. Health Perspect.*, *103*(5), 472–480.

Pope, C. A., J. B. Muhlestein, H. T. May, D. G. Renlund, J. L. Anderson, and B. D. Horne (2006), Ischemic Heart Disease Events Triggered by Short-Term Exposure to Fine Particulate Air Pollution, *Circulation*, *114*(23), 2443–2448, doi:10.1161/CIRCULATIONAHA.106.636977.

Pope, C. A., R. T. Burnett, M. C. Turner, A. Cohen, D. Krewski, M. Jerrett, S. M. Gapstur, and M. J. Thun (2011), Lung Cancer and Cardiovascular Disease Mortality Associated with Ambient Air Pollution and Cigarette Smoke: Shape of the Exposure-Response Relationships, *Environ. Health Perspect.*, *119*(11), 1616–1621, doi:10.1289/ehp.1103639.

Pope, C. A., 3rd, and D. W. Dockery (2006), Health effects of fine particulate air pollution: lines that connect, *J. Air Waste Manag. Assoc. 1995*, *56*(6), 709–742.

Pope, C. A., 3rd, R. T. Burnett, M. J. Thun, E. E. Calle, D. Krewski, K. Ito, and G. D. Thurston (2002), Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution, *JAMA J. Am. Med. Assoc.*, *287*(9), 1132–1141.

Pope, C. A., 3rd, R. T. Burnett, G. D. Thurston, M. J. Thun, E. E. Calle, D. Krewski, and J. J. Godleski (2004), Cardiovascular mortality and long-term exposure to particulate air pollution: epidemiological evidence of general pathophysiological pathways of disease, *Circulation*, *109*(1), 71–77, doi:10.1161/01.CIR.0000108927.80044.7F.

Remer, L. A., Y. J. Kaufman, Z. Levin, and S. Ghan (2002a), Model Assessment of the Ability of MODIS to Measure Top-of-Atmosphere Direct Radiative Forcing from Smoke Aerosols, *J. Atmospheric Sci.*, *59*(3), 657–667, doi:10.1175/1520-0469(2002)059<0657:MAOTAO>2.0.CO;2.

Remer, L. A. et al. (2002b), Validation of MODIS aerosol retrieval over ocean, *Geophys. Res. Lett.*, *29*(12), MOD3–1–MOD3–4, doi:10.1029/2001GL013204.

Remer, L. A. et al. (2005), The MODIS Aerosol Algorithm, Products, and Validation, *J. Atmospheric Sci.*, *62*(4), 947–973, doi:10.1175/JAS3385.1.

Ridley, D. A., C. L. Heald, and B. Ford (2012), North African dust export and deposition: A satellite and model perspective, *J. Geophys. Res.*, *117*(D2), D02202, doi:10.1029/2011JD016794.

Roelofs, G.-J. (2013), A steady-state analysis of the temperature response of water vapor and aerosol lifetimes, *Atmos Chem Phys, 13*, 8245-8254, doi: 10.5194/acp-13-8245-2013.

Rogers, R. R. et al. (2011), Assessment of the CALIPSO Lidar 532 nm attenuated backscatter calibration using the NASA LaRC airborne High Spectral Resolution Lidar, *Atmos Chem Phys*, *11*(3), 1295–1311, doi:10.5194/acp-11-1295-2011.

Roy, B., R. Mathur, A. B. Gilliland, and S. C. Howard (2007), A comparison of CMAQ-based aerosol properties with IMPROVE, MODIS, and AERONET data, *J. Geophys. Res.*, *112*(D14), D14301, doi:10.1029/2006JD008085.

Al-Saadi, J. et al. (2005), Improving National Air Quality Forecasts with Satellite Aerosol Observations, *Bull. Am. Meteorol. Soc.*, *86*(9), 1249–1261, doi:10.1175/BAMS-86-9-1249.

Sacks, J. D., L. W. Stanek, T. J. Luben, D. O. Johns, B. J. Buckley, J. S. Brown, and M. Ross (2011), Particulate matter-induced health effects: who is susceptible?, *Environ. Health Perspect.*, *119*(4), 446–454, doi:10.1289/ehp.1002255.

Schaap, M., A. Apituley, R. M. A. Timmermans, R. B. A. Koelemeijer, and G. de Leeuw (2009), Exploring the relation between aerosol optical depth and PM2.5 at Cabauw, the Netherlands, *Atmos Chem Phys*, *9*(3), 909–925, doi:10.5194/acp-9-909-2009.

Schiferl, L. D., C. L. Heald, J. B. Nowak, J. S. Holloway, J. A. Neuman, R. Bahreini, I. B. Pollack, T. B. Ryerson, C. Wiedinmyer, and J. G. Murphy (2014), An investigation of ammonia and inorganic particulate matter in California during the CalNex campaign, *J. Geophys. Res. Atmospheres*, *119*(4), 2013JD020765, doi:10.1002/2013JD020765.

Schwartz, J. (1999), Air pollution and hospital admissions for heart disease in eight U.S. counties, *Epidemiol. Camb. Mass*, *10*(1), 17–22.

Schwartz, J. (2000), The distributed lag between air pollution and daily deaths, *Epidemiol. Camb. Mass*, *11*(3), 320–326.

Schwartz, J., D. W. Dockery, and L. M. Neas (1996), Is daily mortality associated specifically with fine particles?, *J. Air Waste Manag. Assoc. 1995*, *46*(10), 927–939.

Shepard, D. (1968), A Two-dimensional Interpolation Function for Irregularly-spaced Data, in *Proceedings of the 1968 23rd ACM National Conference*, pp. 517–524, ACM, New York, NY, USA.

Sheppard, L., R. T. Burnett, A. A. Szpiro, S.-Y. Kim, M. Jerrett, C. A. Pope, and B. Brunekreef (2012), Confounding and exposure measurement error in air pollution epidemiology, *Air Qual. Atmosphere Health*, *5*(2), 203–216, doi:10.1007/s11869-011-0140-9.

Sheridan, P. J., E. Andrews, J. A. Ogren, J. L. Tackett, and D. M. Winker (2012), Vertical profiles of aerosol optical properties over central Illinois and comparison with surface and satellite measurements, *Atmos Chem Phys*, *12*(23), 11695–11721, doi:10.5194/acp-12-11695-2012.

Silva, R. A. et al. (2013), Global premature mortality due to anthropogenic outdoor air pollution and the contribution of past climate change, *Environ. Res. Lett.*, *8*(3), 034005, doi:10.1088/1748-9326/8/3/034005.

Simon, H., P. V. Bhave, J. L. Swall, N. H. Frank, and W. C. Malm (2011), Determining the spatial and seasonal variability in OM/OC ratios across the US using multiple regression, *Atmos Chem Phys*, *11*(6), 2933–2949, doi:10.5194/acp-11-2933-2011.

Snider, G. et al. (2014), SPARTAN: a global network to evaluate and enhance satellite-based estimates of ground-level particulate matter for global health applications, *Atmos Meas Tech Discuss*, *7*(7), 7569–7611, doi:10.5194/amtd-7-7569-2014.

Sorooshian, A., M.-L. Lu, F. J. Brechtel, H. Jonsson, G. Feingold, R. C. Flagan, and J. H. Seinfeld (2007), On the Source of Organic Acid Aerosol Layers above Clouds, *Environ. Sci. Technol.*, *41*(13), 4647–4654, doi:10.1021/es0630442.

Stanier, C. O., A. Y. Khlystov, and S. N. Pandis (2004), Ambient aerosol size distributions and number concentrations measured during the Pittsburgh Air Quality Study (PAQS), *Atmos. Environ.*, *38*(20), 3275–3284, doi:10.1016/j.atmosenv.2004.03.020.

Stephens, G. L. et al. (2002), THE CLOUDSAT MISSION AND THE A-TRAIN, *Bull. Am. Meteorol. Soc.*, *83*(12), 1771–1790, doi:10.1175/BAMS-83-12-1771.

Stephens, G. L. et al. (2008), CloudSat mission: Performance and early science after the first year of operation, *J. Geophys. Res.*, *113*(D8), D00A18, doi:10.1029/2008JD009982.

Streets, D. G. et al. (2003), An inventory of gaseous and primary aerosol emissions in Asia in the year 2000, *J. Geophys. Res. Atmospheres*, *108*(D21), 8809, doi:10.1029/2002JD003093.

Streets, D. G., Q. Zhang, L. Wang, K. He, J. Hao, Y. Wu, Y. Tang, and G. R. Carmichael (2006), Revisiting China's CO emissions after the Transport and Chemical Evolution over the Pacific (TRACE-P) mission: Synthesis of inventories, atmospheric modeling, and

observations, *J. Geophys. Res. Atmospheres*, *111*(D14), D14306, doi:10.1029/2006JD007118.

Sun, Q., X. Hong, and L. E. Wold (2010), Cardiovascular Effects of Ambient Particulate Air Pollution Exposure, *Circulation*, *121*(25), 2755–2765, doi:10.1161/CIRCULATIONAHA.109.893461.

Tasić, M., S. Rajšić, V. Novaković, Z. Mijić (2006), Atmospheric aerosols and their influence on air quality in urban areas. *Facta Univ Ser Phys Chem Techn*, *4*(1): 83 – 91.

Tegen, I., and A. A. Lacis (1996), Modeling of particle size distribution and its influence on the radiative properties of mineral dust aerosol, *J. Geophys. Res. Atmospheres*, *101*(D14), 19237–19244, doi:10.1029/95JD03610.

Volkamer, R., J. L. Jimenez, F. S. Martini, K. Dzepina, Q. Zhang, D. Salcedo, L. T. Molina, D. R. Worsnop, and M. J. Molina (2006), Secondary organic aerosol formation from anthropogenic air pollution: Rapid and higher than expected, *Geophys. Res. Lett.*, *33*(17), L17811, doi:10.1029/2006GL026899.

Wang, C., Q. Liu, N. Ying, X. Wang, and J. Ma (2013a), Air quality evaluation on an urban scale based on MODIS satellite images, *Atmospheric Res.*, *132–133*, 22–34, doi:10.1016/j.atmosres.2013.04.011.

Wang, J., and S. A. Christopher (2003), Intercomparison between satellite-derived aerosol optical thickness and PM2.5 mass: Implications for air quality studies, *Geophys. Res. Lett.*, *30*(21), 2095, doi:10.1029/2003GL018174.

Wang, X., Y. Wang, J. Hao, Y. Kondo, M. Irwin, J. W. Munger, and Y. Zhao (2013b), Top-down estimate of China's black carbon emissions using surface observations: Sensitivity to observation representativeness and transport model error, *J. Geophys. Res. Atmospheres*, *118*(11), 5781–5795, doi:10.1002/jgrd.50397.

Wang, Y., D. J. Jacob, and J. A. Logan (1998), Global simulation of tropospheric O3-NO x - hydrocarbon chemistry 3. Origin of tropospheric ozone and effects of nonmethane hydrocarbons, *J. Geophys. Res.*, *103*(D9), 10757–10,767, doi:10.1029/98JD00156.

Wang, Y., Q. Q. Zhang, K. He, Q. Zhang, and L. Chai (2013c), Sulfate-nitrate-ammonium aerosols over China: response to 2000–2015 emission changes of sulfur dioxide, nitrogen oxides, and ammonia, *Atmos Chem Phys*, *13*(5), 2635–2652, doi:10.5194/acp-13-2635-2013.

Wang, Y., Q. Zhang, J. Jiang, W. Zhou, B. Wang, K. He, F. Duan, Q. Zhang, S. Philip, and Y. Xie (2014a), Enhanced sulfate formation during China's severe winter haze episode in January 2013 missing from current models, *J. Geophys. Res. Atmospheres*, *119*(17), 2013JD021426, doi:10.1002/2013JD021426.

Wang, Z. et al. (2014b), Modeling study of regional severe hazes over mid-eastern China in January 2013 and its implications on pollution prevention and control, *Sci. China Earth Sci.*, *57*(1), 3–13, doi:10.1007/s11430-013-4793-0.

Weber, R. et al. (2003), Short-term temporal variation in PM2.5 mass and chemical composition during the Atlanta Supersite Experiment, 1999, *J. Air Waste Manag. Assoc. 1995*, *53*(1), 84–91.

Weber, R. J. et al. (2007), A study of secondary organic aerosol formation in the anthropogenic-influenced southeastern United States, *J. Geophys. Res.*, *112*(D13), D13302, doi:10.1029/2007JD008408.

Weichenthal, S., P. J. Villeneuve, R. T. Burnett, A. van Donkelaar, R. V. Martin, R. R. Jones, C. T. DellaValle, D. P. Sandler, M. H. Ward, and J. A. Hoppin (2014), Long-Term Exposure to Fine Particulate Matter: Association with Nonaccidental and Cardiovascular Mortality in the Agricultural Health Study Cohort, *Environ. Health Perspect.*, doi:10.1289/ehp.1307277.

Wellenius, G. A., J. Schwartz, and M. A. Mittleman (2005), Air pollution and hospital admissions for ischemic and hemorrhagic stroke among medicare beneficiaries, *Stroke J. Cereb. Circ.*, *36*(12), 2549–2553, doi:10.1161/01.STR.0000189687.78760.47.

Van der Werf, G. R., J. T. Randerson, L. Giglio, G. J. Collatz, P. S. Kasibhatla, and A. F. Arellano Jr. (2006), Interannual variability in global biomass burning emissions from 1997 to 2004, *Atmos Chem Phys*, *6*(11), 3423–3441, doi:10.5194/acp-6-3423-2006.

Wesely, M. L. (1989), Parameterization of surface resistances to gaseous dry deposition in regional-scale numerical models, *Atmospheric Environ. 1967*, *23*(6), 1293–1304, doi:10.1016/0004-6981(89)90153-4.

Winker, D. M. (2003), The CALIPSO mission: spaceborne lidar for observation of aerosols and clouds, vol. 4893, pp. 1–11, SPIE.

Yan, B., M. Zheng, Y. Hu, X. Ding, A. P. Sullivan, R. J. Weber, J. Baek, E. S. Edgerton, and A. G. Russell (2009), Roadside, urban, and rural comparison of primary and secondary organic molecular markers in ambient PM2.5, *Environ. Sci. Technol.*, *43*(12), 4287–4293.

Young, S. A., and M. A. Vaughan (2009), The Retrieval of Profiles of Particulate Extinction from Cloud-Aerosol Lidar Infrared Pathfinder Satellite Observations (CALIPSO) Data: Algorithm Description, *J. Atmospheric Ocean. Technol.*, *26*(6), 1105–1119, doi:10.1175/2008JTECHA1221.1.

Young, S. A., M. A. Vaughan, R. E. Kuehn, and D. M. Winker (2013), The Retrieval of Profiles of Particulate Extinction from Cloud–Aerosol Lidar and Infrared Pathfinder Satellite Observations (CALIPSO) Data: Uncertainty and Error Sensitivity Analyses, *J. Atmospheric Ocean. Technol.*, *30*(3), 395–428, doi:10.1175/JTECH-D-12-00046.1.

Zanobetti, A., and J. Schwartz (2009), The Effect of Fine and Coarse Particulate Air Pollution on Mortality: A National Analysis, *Environ. Health Perspect.*, *117*(6), 898–903, doi:10.1289/ehp.0800108.

Zeng, T., Y. Wang, Y. Yoshida, D. Tian, A. G. Russell, and W. R. Barnard (2008), Impacts of prescribed fires on air quality over the Southeastern United States in spring based on modeling and ground/satellite measurements, *Environ. Sci. Technol.*, *42*(22), 8401–8406.

Zhang, H., R. M. Hoff, and J. A. Engel-Cox (2009a), The relation between Moderate Resolution Imaging Spectroradiometer (MODIS) aerosol optical depth and PM2.5 over the United States: a geographical comparison by U.S. Environmental Protection Agency regions, *J. Air Waste Manag. Assoc. 1995*, *59*(11), 1358–1369.

Zhang, J., J. S. Reid, and B. N. Holben (2005), An analysis of potential cloud artifacts in MODIS over ocean aerosol optical thickness products, *Geophys. Res. Lett.*, *32*(15), L15803, doi:10.1029/2005GL023254.

Zhang, L., R. Vet, A. Wiebe, C. Mihele, B. Sukloff, E. Chan, M. D. Moran, and S. Iqbal (2008), Characterization of the size-segregated water-soluble inorganic ions at eight Canadian rural sites, *Atmos Chem Phys*, *8*(23), 7133–7151, doi:10.5194/acp-8-7133-2008.

Zhang, L., D. J. Jacob, E. M. Knipping, N. Kumar, J. W. Munger, C. C. Carouge, A. van Donkelaar, Y. X. Wang, and D. Chen (2012a), Nitrogen deposition to the United States: distribution, sources, and processes, *Atmos Chem Phys*, *12*(10), 4539–4554, doi:10.5194/acp-12-4539-2012.

Zhang, Q., D. G. Streets, K. He, and Z. Klimont (2007), Major components of China's anthropogenic primary particulate emissions, *Environ. Res. Lett.*, *2*(4), 045027, doi:10.1088/1748-9326/2/4/045027.

Zhang, Q. et al. (2009b), Asian emissions in 2006 for the NASA INTEX-B mission, *Atmos Chem Phys*, *9*(14), 5131–5153, doi:10.5194/acp-9-5131-2009.

Zhang, X., A. Hecobian, M. Zheng, N. H. Frank, and R. J. Weber (2010), Biomass burning impact on PM 2.5 over the southeastern US during 2007: integrating chemically speciated FRM filter measurements, MODIS fire counts and PMF analysis, *Atmos Chem Phys*, *10*(14), 6839–6853, doi:10.5194/acp-10-6839-2010.

Zhang, X., Z. Liu, A. Hecobian, M. Zheng, N. H. Frank, E. S. Edgerton, and R. J. Weber (2012b), Spatial and seasonal variations of fine particle water-soluble organic carbon (WSOC) over the southeastern United States: implications for secondary organic aerosol formation, *Atmos Chem Phys*, *12*(14), 6593–6607, doi:10.5194/acp-12-6593-2012.

Zhao, B., S. X. Wang, H. Liu, J. Y. Xu, K. Fu, Z. Klimont, J. M. Hao, K. B. He, J. Cofala, and M. Amann (2013), NOx emissions in China: historical trends and future perspectives, *Atmos Chem Phys*, *13*(19), 9869–9897, doi:10.5194/acp-13-9869-2013.

Zheng, M., G. R. Cass, J. J. Schauer, and E. S. Edgerton (2002), Source Apportionment of PM2.5 in the Southeastern United States Using Solvent-Extractable Organic Compounds as Tracers, *Environ. Sci. Technol.*, *36*(11), 2361–2371, doi:10.1021/es011275x.

Zhong, L., P. K. K. Louie, J. Zheng, K. M. Wai, J. W. K. Ho, Z. Yuan, A. K. H. Lau, D. Yue, and Y. Zhou (2013), The Pearl River Delta Regional Air Quality Monitoring Network – Regional Collaborative Efforts on Joint Air Quality Management, *Aerosol Air Qual. Res.*, doi:10.4209/aaqr.2012.10.0276.

Appendix I. Death Rates and Baseline Mortality

The following Figure shows the variability in death rates and baseline mortality for the U.S. Death rates for cardiovascular disease have declined from 2004 to 2010 in every state (except Alaska) while death rates for lung cancer and respiratory disease have been variable.

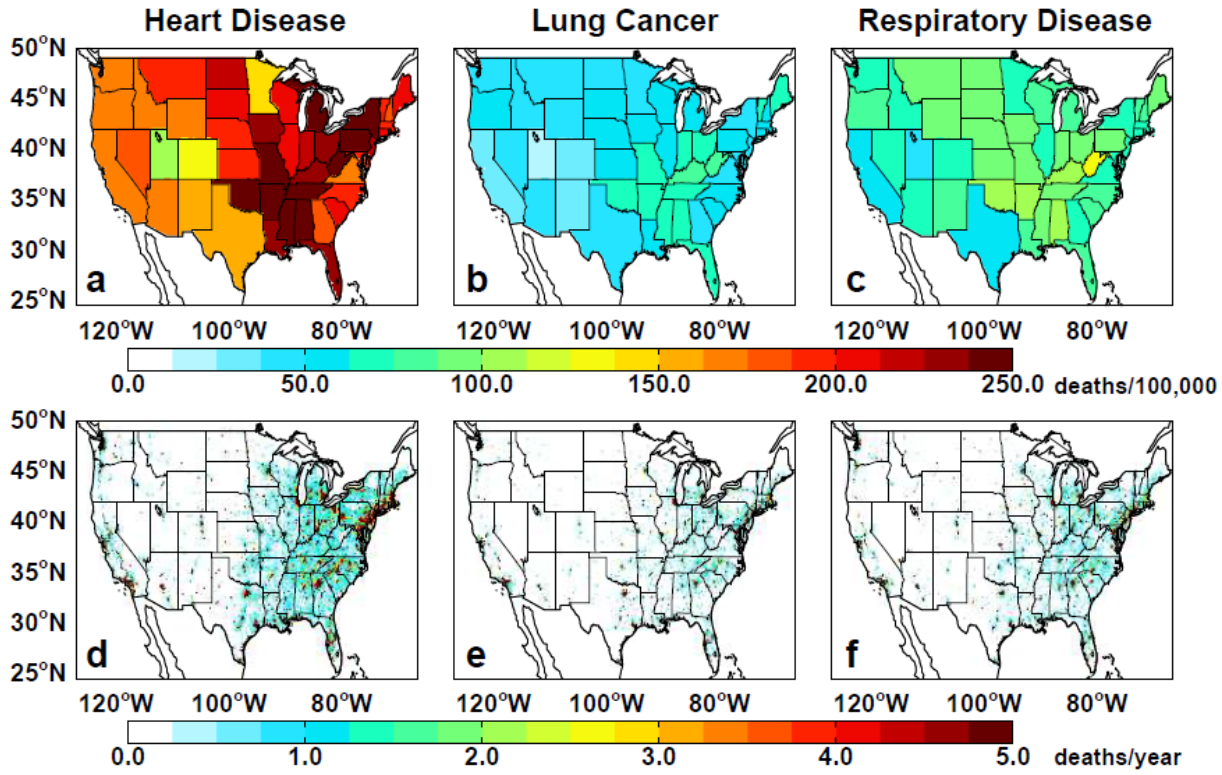

Figure A1 Crude death rates for (a) heart disease, (b) lung cancer and (c) respiratory disease (given as per 100,000 people) for the United States and (d,e,f) the respective baseline mortalities.

Appendix II. Normalized Mean Bias between Model and Satellite Observations

This following figure is similar to Figure 3.8; however, it uses the original satellite observations without spatial interpolation. General patterns are similar although not as evident due to lack of observations.

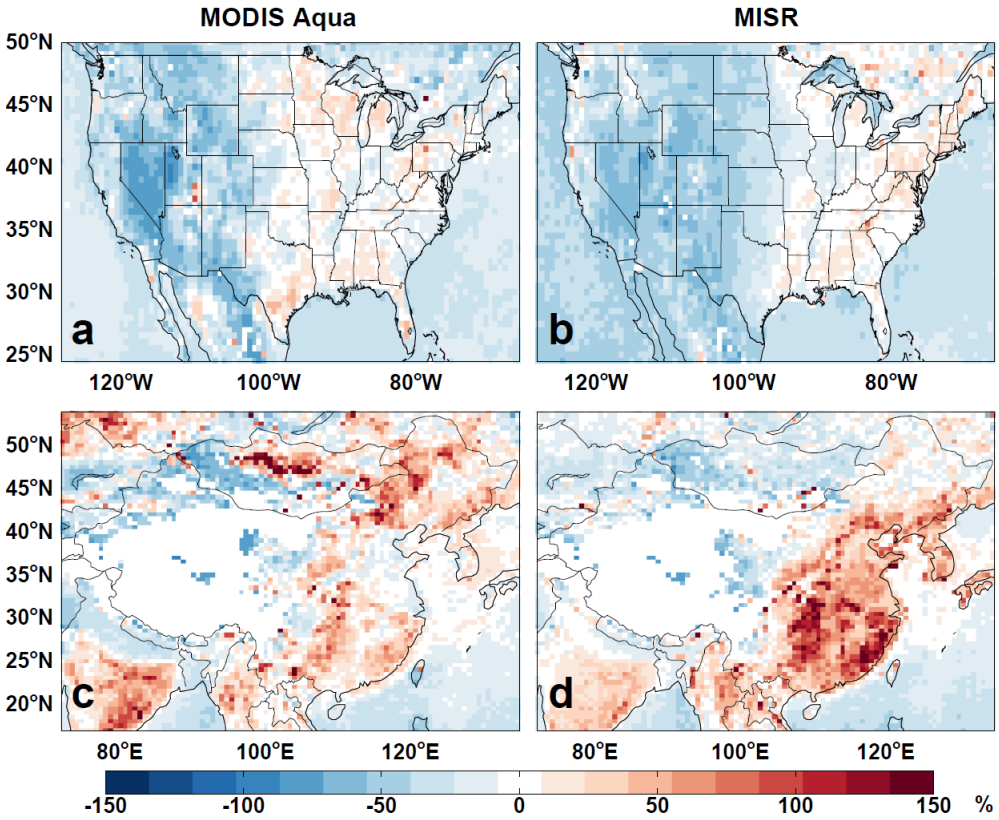

Figure A2 Same as Figure 3.8 except determined using observations which have not be spatially interpolated.

Appendix III. Number of days with concentrations above daily standard

The following figures show the number of days that the concentration in that grid box was higher than the daily (24 hour) standard of 35 μgm^{-3} for the U.S. and 75 μgm^{-3} for China. This is the proposed standard for China that is to take effect nationwide in 2016 although many key cities are already required to implement it. To note for the figure, these are sampled to days with available satellite observations and do not cover all days in the year. In the U.S., using the satellite-based concentrations suggests that there are significantly more days that are out of compliance, although we do see that the overall number has decreased over time. Additionally, from the maps, it is evident that these are often strongly related to fires such as the Bugaboo Scrub Fire in Georgia in 2007, the Murphy Complex Fire in Idaho in 2007, the multiple fires in California in 2008 (Montecito Tea, Sayre, Freeway Complex, etc.), which suggests that the model is potentially under predicting fire impacts (either in AOD or surface $PM_{2.5}$). Also to note is the region over the Southeastern United States where the satellite suggests a widespread region with more days out of compliance up to 2007. This is potentially related to the issue discussed in Chapter 2 where the model is not simulating the vertical distribution correctly.

Even with a higher standard, regions of China have substantially more days that would be out of compliance if these standards were already in place. The model suggests more days than the satellite-based. This could be linked to issues in emissions but could also be due to the satellites mischaracterization of heavy plumes.

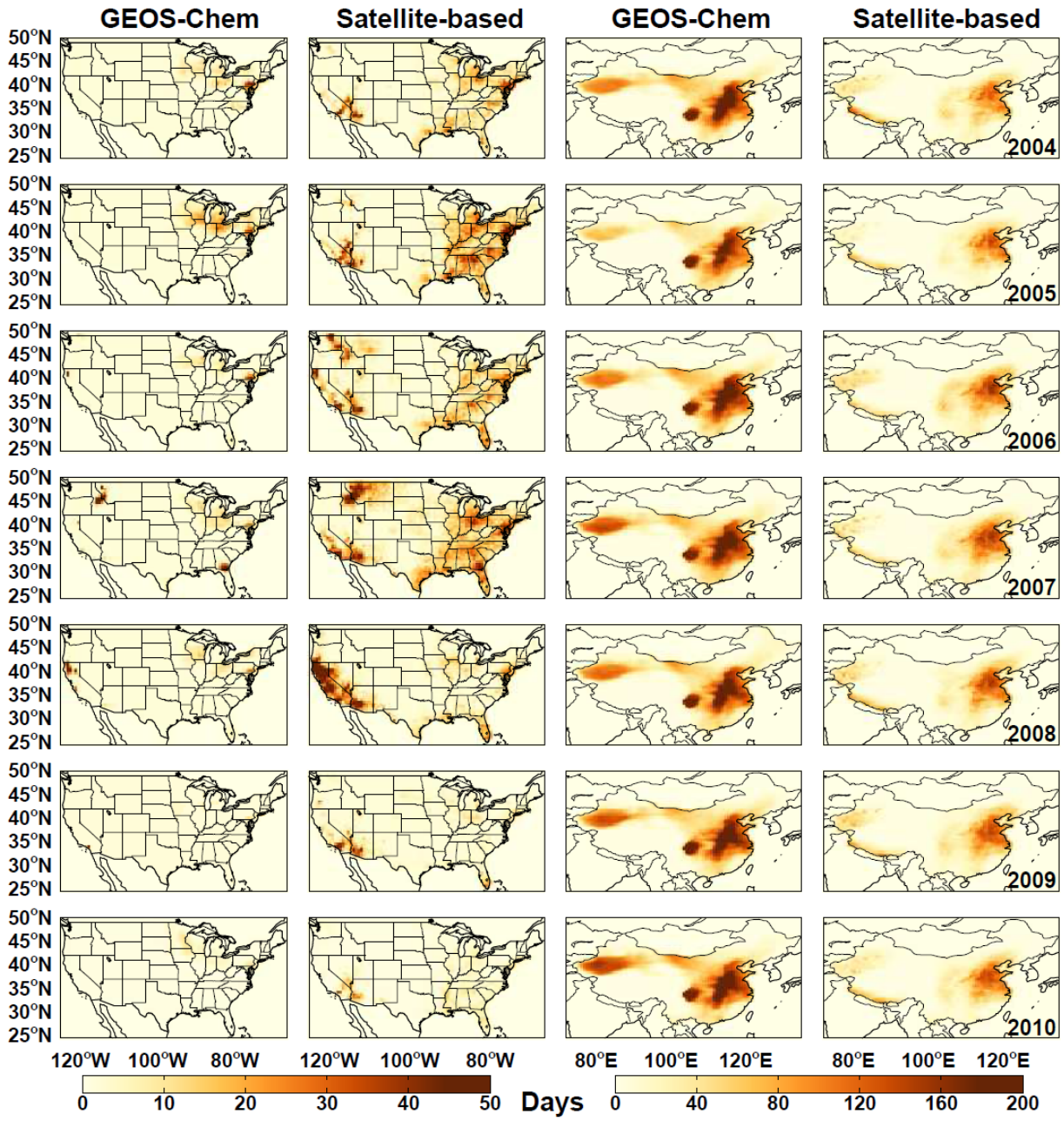

Figure A3 Number of days each year when the surface $PM_{2.5}$ is greater than the national standard of 35 μgm^{-3} for the U.S. and 75 μgm^{-3} for China using the model or satellite-based concentrations. These are sampled to days with satellite observations. Note the difference in color bars for the two regions.

Appendix IV. Seasonally averaged (2004-2011) surface relative humidity used in model simulations

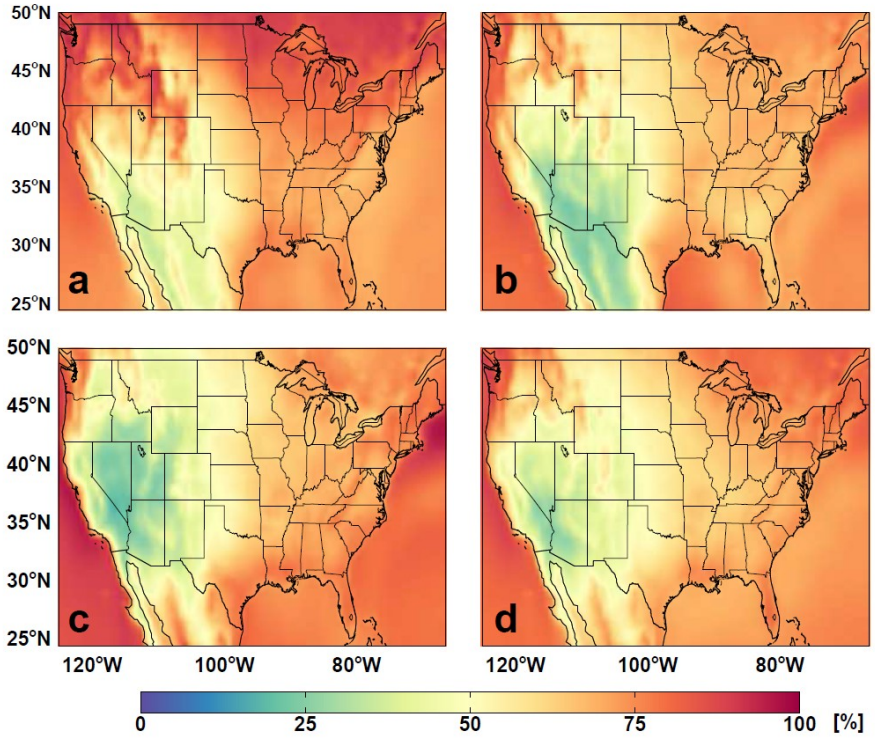

Figure A4 Average surface relative humidity during (a) winter (DJF), (b) spring (MAM), (c) summer (JJA), and (d) fall (SON) in the U.S. for 2004-2011.

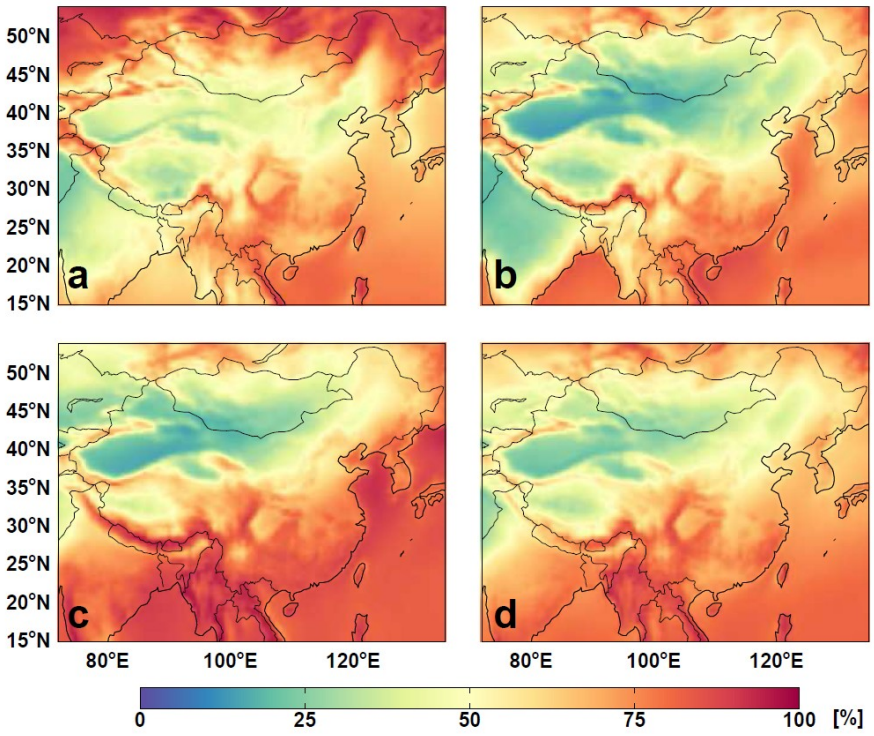

Figure A5 Same as Figure A4 but for China.

Appendix V. Normalized Mean Bias in Annual Average $PM_{2.5}$ from η Sensitivity Tests

These figures show the distribution of normalized mean biases (Equation 3.5) in annual average $PM_{2.5}$ determined from sensitivity tests and compared to the original model simulation. Note that because these rely on annual averages, the distributions shown represent spatial variations.

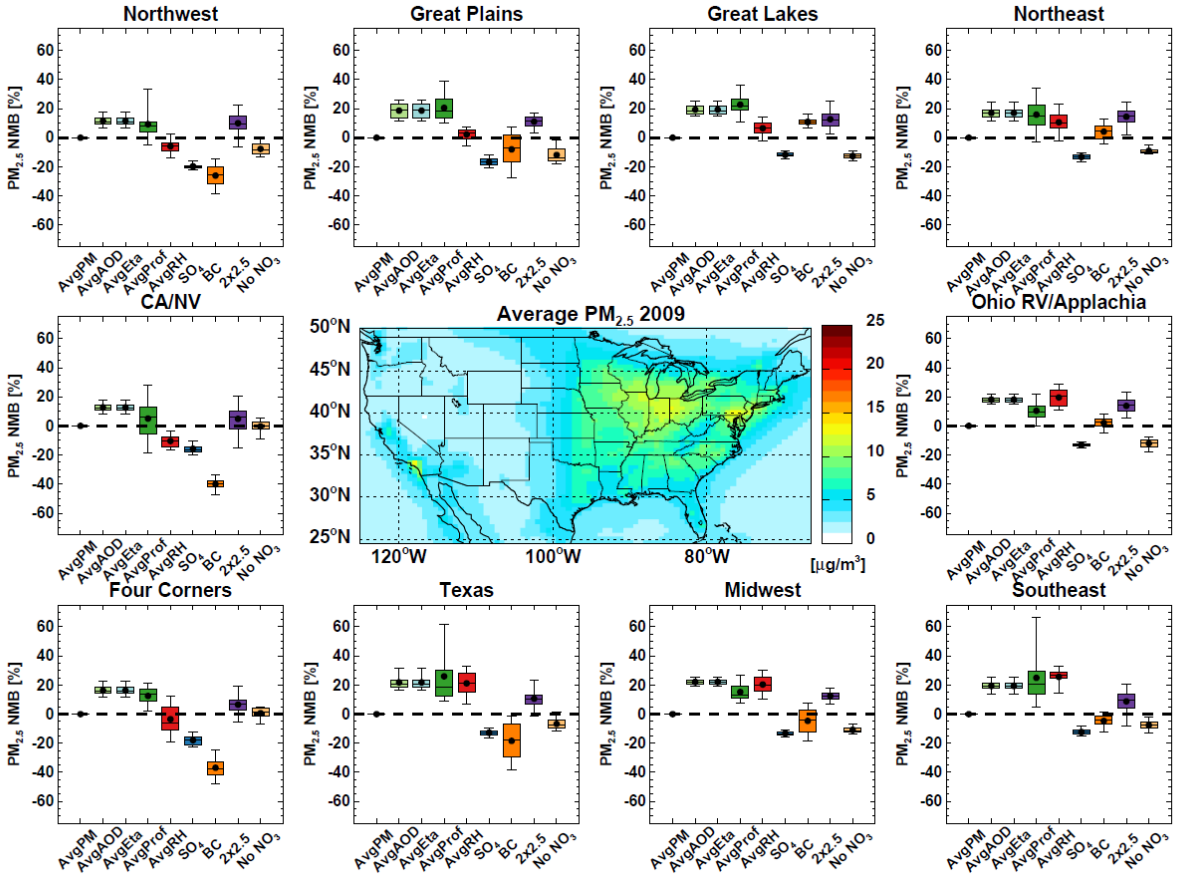

Figure A6 Distribution of normalized mean biases in annual average $PM_{2.5}$ for grid boxes in each region of the U.S. for 2009 determined from sensitivity tests discussed in Section 4.1.1. Map shows annual average simulated surface $PM_{2.5}$.

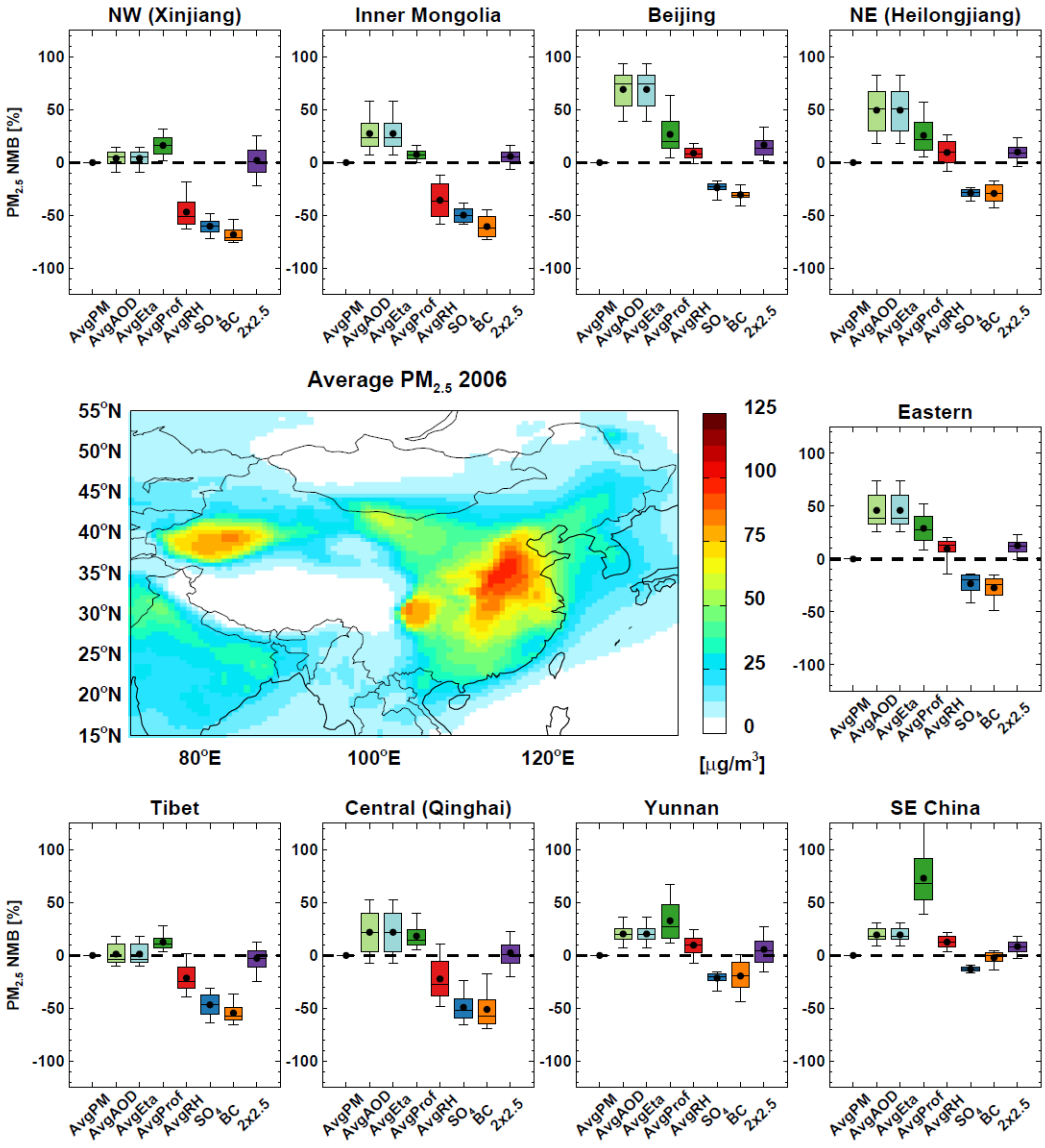

Figure A7 Similar to Figure A6 but for China 2006.

Appendix VI. Interannual Variability

In our sensitivity analyses, we used model simulations for a single year. As there will be some variability in weather and transport and $PM_{2.5}$ sources, sensitivities might also change between years. In order to determine if this is a substantial area of concern, we examine the interannual variability in AOD, η, $PM_{2.5}$, and RH in the following plots. From this series of figures, we can determine that, in general, we would expect interannual variability to impact our results by generally less than 5%.

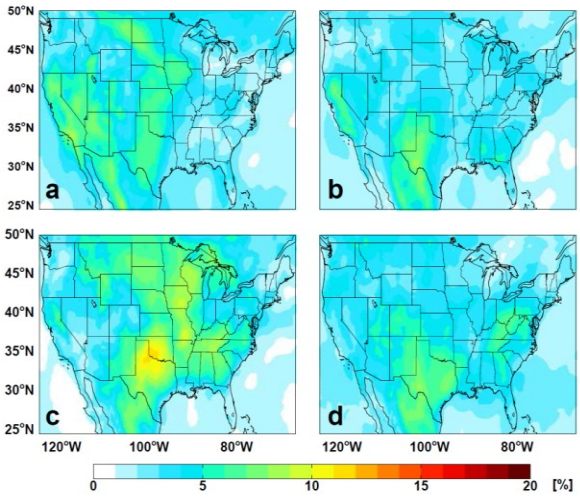

Figure A8 Standard deviation of relative humidity for each season (a) DJF, (b) MAM, (c) JJA, and (d) SON.

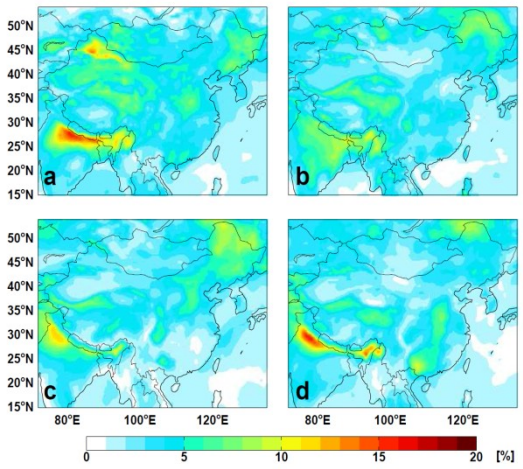

Figure A9 Same as Figure A8, but for China

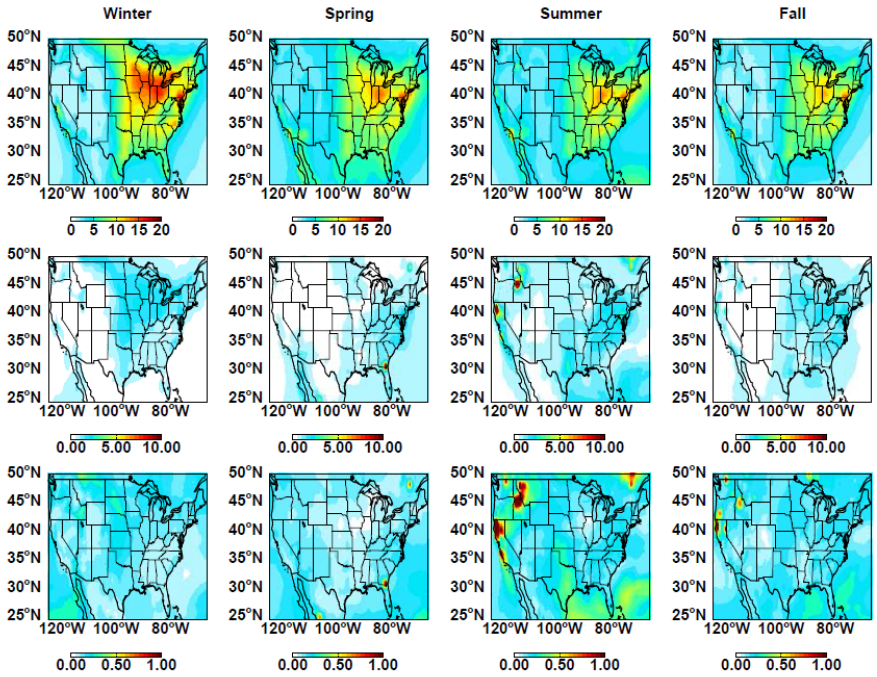

Figure A10 Multi-year average PM2.5 concentration (row 1), standard deviation in annual concentrations (row 2), and coefficient of variation (row 3) for the U.S.

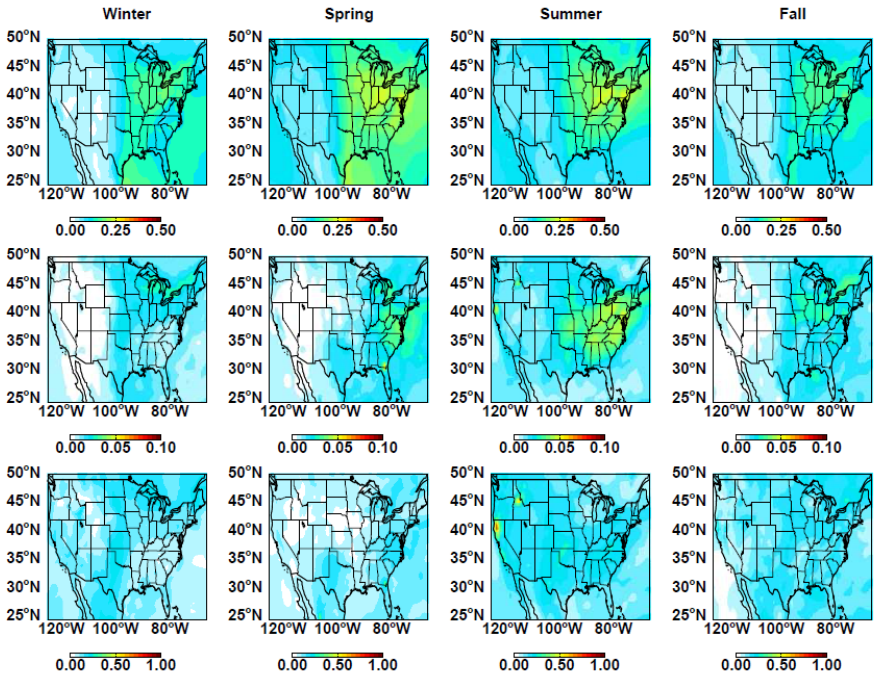

Figure A11 Similar to Figure A10 but for AOD.

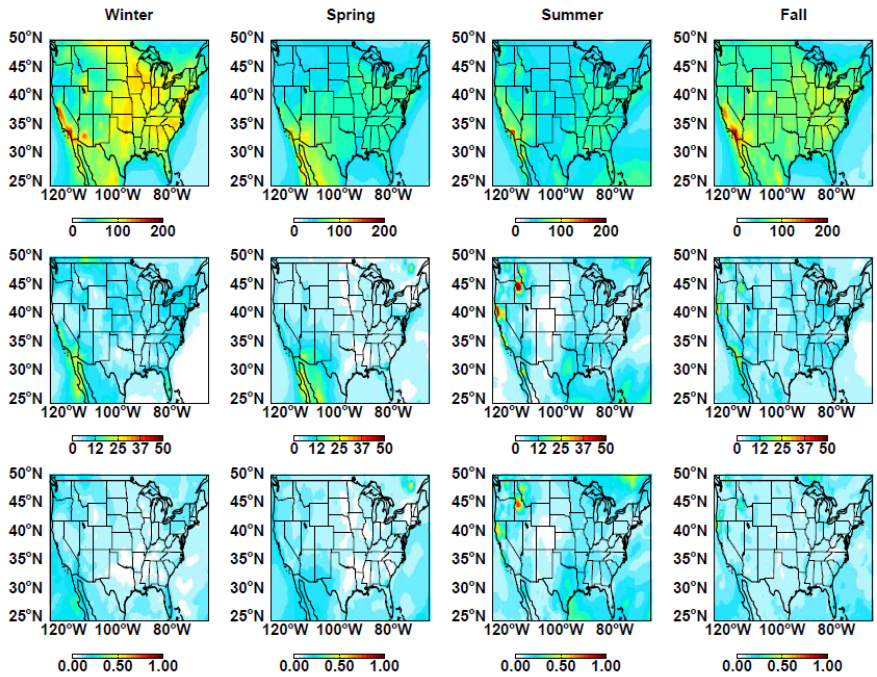

Figure A12 Similar to Figure A10 but for η.

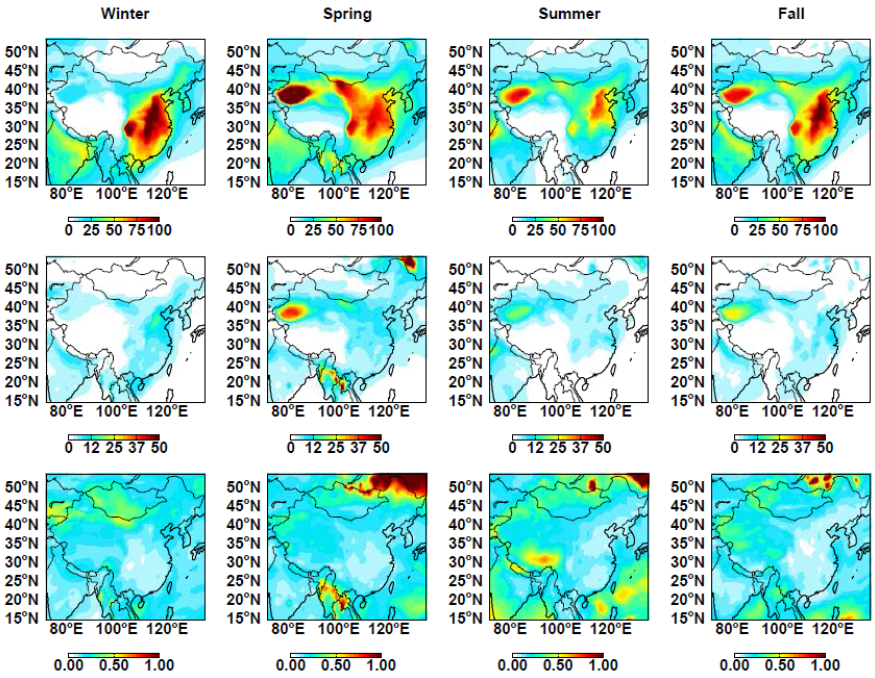

Figure A13 Similar to Figure A10 but for China.

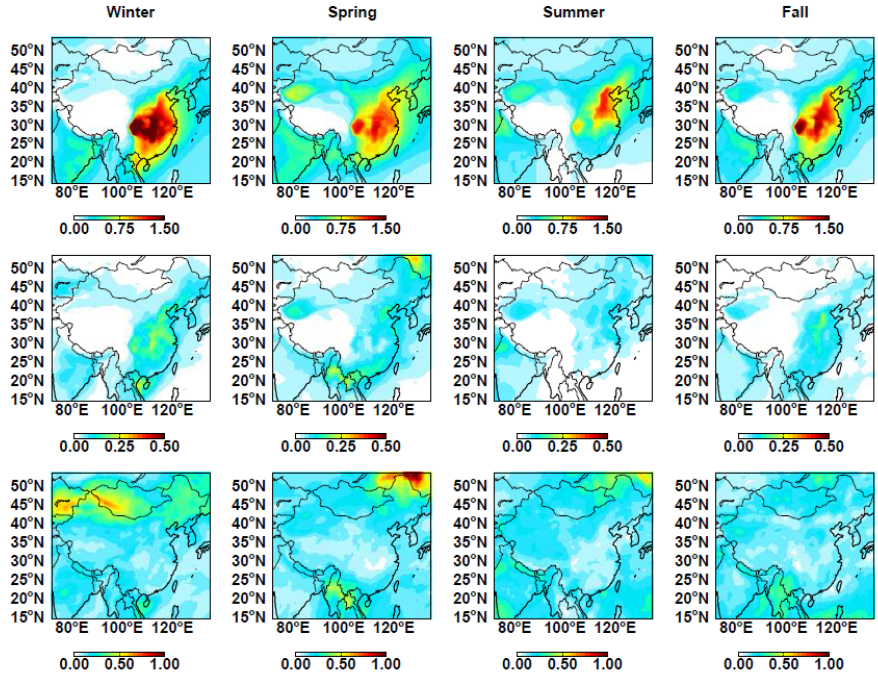

Figure A14 Similar to Figure A11 but for China.

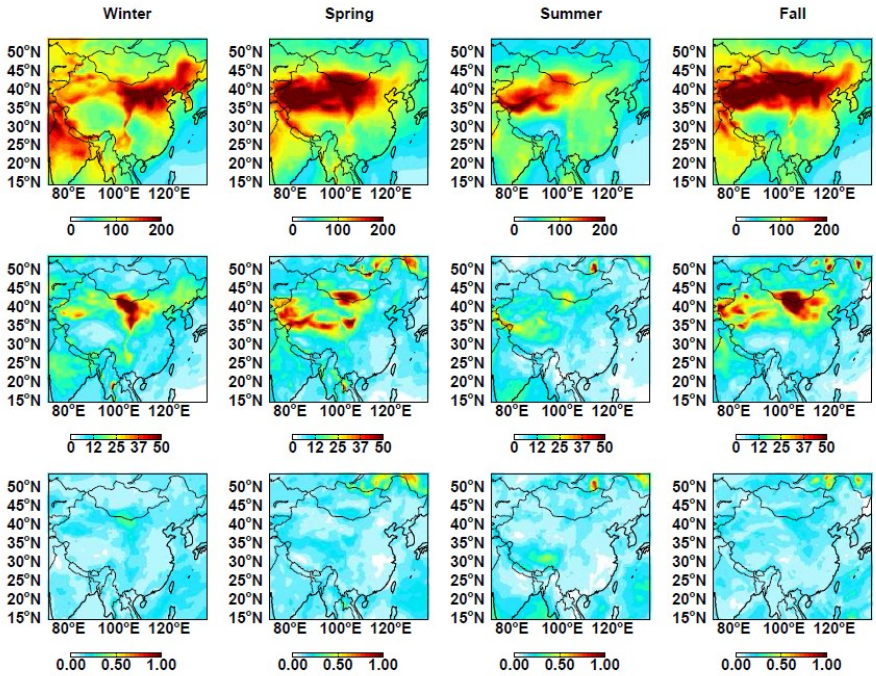

Figure A15 Similar to Figure A12 but for China.