



The Influence of Particulate Matter and Methane on Regional Air Quality and Climate in the United States and India

Citation

Cusworth, Daniel Harvey. 2018. The Influence of Particulate Matter and Methane on Regional Air Quality and Climate in the United States and India. Doctoral dissertation, Harvard University, Graduate School of Arts & Sciences.

Permanent link

<http://nrs.harvard.edu/urn-3:HUL.InstRepos:41121260>

Terms of Use

This article was downloaded from Harvard University's DASH repository, and is made available under the terms and conditions applicable to Other Posted Material, as set forth at <http://nrs.harvard.edu/urn-3:HUL.InstRepos:dash.current.terms-of-use#LAA>

Share Your Story

The Harvard community has made this article openly available.
Please share how this access benefits you. [Submit a story](#).

[Accessibility](#)

The influence of particulate matter and
methane on regional air quality and
climate in the United States and India

A DISSERTATION PRESENTED
BY
DANIEL CUSWORTH
TO
THE DEPARTMENT OF EARTH & PLANETARY SCIENCES
IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY
IN THE SUBJECT OF
EARTH & PLANETARY SCIENCES
HARVARD UNIVERSITY
CAMBRIDGE, MASSACHUSETTS
JUNE 2018

©2018 – DANIEL CUSWORTH
ALL RIGHTS RESERVED.

The influence of particulate matter and methane on regional air quality and climate in the United States and India

ABSTRACT

Atmospheric pollution is important for both air quality and climate considerations. In what follows, I explore three different atmospheric constituents – sulfate, particulate matter (PM_{2.5}) from fires, methane – and their respective influence on regional (United States, India) air quality and climate.

In situ surface observations show that downward surface solar radiation (SWdn) over the central and southeastern United States (U.S.) has increased by 0.58–1.0 Wm⁻²a⁻¹ over the 2000–2014 timeframe, simultaneously with reductions in U.S. aerosol optical depth (AOD) of 3.3 – 5.0 × 10⁻³a⁻¹. Establishing a link between these two trends, however, is challenging due to complex interactions between aerosols, clouds, and radiation. Here I investigate the clear-sky aerosol-radiation effects of decreasing U.S. aerosols on SWdn and other surface variables by applying a one-dimensional radiative transfer to 2000–2014 measurements of AOD at two Surface Radiation Budget Network (SURFRAD) sites in the central and southeastern United States. Observations characterized as “clear-sky” may in fact include the effects of thin cirrus clouds, and I consider these effects by imposing satellite data from the Clouds and

Earth's Radiant Energy System (CERES) into the radiative transfer model. The model predicts that 2000-2014 trends in aerosols may have driven clear-sky SWdn trends of $+1.35 \text{ Wm}^{-2}\text{a}^{-1}$ at Goodwin Creek, MS, and $+0.93 \text{ Wm}^{-2}\text{a}^{-1}$ at Bondville, IL. While these results are consistent in sign with observed trends, a cross-validated multivariate regression analysis shows that AOD reproduces 20-26% of the seasonal (June-September, JJAS) variability in clear-sky direct and diffuse SWdn at Bondville, IL, but none of the JJAS variability at Goodwin Creek, MS. Using *in situ* soil and surface flux measurements from the Ameriflux network and Illinois Climate Network (ICN) together with assimilated meteorology from the North American Land Data Assimilation System (NLDAS), I find that sunnier summers tend to coincide with increased surface air temperature and soil moisture deficits in the central U.S. The 1990-2015 trends in the NLDAS SWdn over the central U.S. are also of a similar magnitude as our modeled 2000-2014 clear-sky trends. Taken together, these results suggest that climate and regional hydrology in the central U.S. are sensitive to the recent reductions in aerosol concentrations. This work has implications for severely polluted regions outside the U.S., where improvements in air quality due to reductions in the aerosol burden could inadvertently pose an enhanced climate risk.

Since at least the 1980s, many farmers in northwest India have switched to mechanized combine harvesting to boost efficiency. This harvesting technique leaves abundant crop residue on the fields, which farmers typically burn to prepare their fields for subsequent planting. A key question is to what extent the large quantity of smoke

emitted by these fires contributes to the already severe pollution in Delhi and across other parts of the heavily populated Indo-Gangetic Plain located downwind of the fires. Using a combination of observed and modeled variables, including surface measurements of $\text{PM}_{2.5}$, I quantify the magnitude of the influence of agricultural fire emissions on surface air pollution in Delhi. With surface measurements, I first derive the signal of regional $\text{PM}_{2.5}$ enhancements (i.e., the pollution above an anthropogenic baseline) during each post-monsoon burning season for 2012-2016. I next use the Stochastic Time-Inverted Lagrangian Transport model (STILT) to simulate surface $\text{PM}_{2.5}$ using five fire emission inventories. I reproduce up to 25% of the weekly variability in total observed $\text{PM}_{2.5}$ using STILT. Depending on year and emission inventory, our method attributes 7.0–78% of the maximum observed $\text{PM}_{2.5}$ enhancements in Delhi to fires. The large range in these attribution estimates points to the uncertainties in fire emission parameterizations, especially in regions where thick smoke may interfere with hotspots of fire radiative power. Although our model can generally reproduce the largest $\text{PM}_{2.5}$ enhancements in Delhi air quality for 1-3 consecutive days each fire season, it fails to capture many smaller daily enhancements, which I attribute to the challenge of detecting small fires in the satellite retrieval. By quantifying the influence of upwind agricultural fire emissions on Delhi air pollution, our work underscores the potential health benefits of changes in farming practices to reduce fires.

Anthropogenic methane emissions originate from a large number of relatively small

point sources, often densely clustered and with a few anomalous sources contributing disproportionately to total emissions. Here I examine the potential of recently launched or planned satellites to promptly identify anomalous emitters among production sites in oil/gas fields through measurements of atmospheric methane, alone or supplemented by a surface observation network. I simulate atmospheric methane over a generic oil/gas field (20-500 production sites of different size categories in a $50 \times 50 \text{ km}^2$ domain) for a 1-week period using the WRF-STILT meteorological model with $1.3 \times 1.3 \text{ km}^2$ horizontal resolution. The simulations consider many random realizations for the occurrence and distribution of anomalous high-mode emitters in the field by sampling bimodal probability density functions (pdfs) of emissions from individual sites. The atmospheric methane field for each realization is then observed virtually with different satellite and surface observing configurations. Column methane enhancements observed from satellites are relatively small, even for high-mode emitters, so an inverse analysis is necessary. The inverse analysis can be regularized effectively using a L-1 norm to provide sparse solutions for a bimodally distributed variable. I find that the recently launched TROPOMI instrument (low Earth orbit, $7 \times 7 \text{ km}^2$ nadir pixels, daily return time) and the planned GeoCARB instrument (geostationary orbit, $2.7 \times 3.0 \text{ km}^2$ pixels, 2x or 4x/day return time) are successful at locating anomalous emitters for fields of 20 and sometimes 50 emitters within the $50 \times 50 \text{ km}^2$ domain, but unsuccessful for denser fields. GeoCARB does not benefit significantly from more frequent observations (4x/day vs. 2x/day). It performs better with a 5-km

error tolerance for localization, but a next-generation geostationary satellite instrument with $1.3 \times 1.3 \text{ km}^2$ pixels, hourly return time, and 1 ppb precision can successfully detect and locate the high-mode emitters for a dense field with up to 500 sites in the $50 \times 50 \text{ km}^2$ domain. The capabilities of TROPOMI and GeoCARB can be usefully augmented with a surface observation network of 10-20 monitors, and in turn these satellite instruments increase the detection capability that can be achieved from the surface monitors alone.

Contents

1	AEROSOL TRENDS AS A POTENTIAL DRIVER OF REGIONAL CLIMATE IN THE CENTRAL UNITED STATES: EVIDENCE FROM OBSERVATIONS	1
1.1	Introduction	1
1.2	Data and Methods	6
1.3	Long-term trends in surface SWdn	12
1.4	Short-term variability in SWdn	19
1.5	Meteorological impacts from enhanced SWdn in late summer	23
1.6	Discussion	32
2	QUANTIFYING THE INFLUENCE OF AGRICULTURAL FIRES IN NORTHWEST INDIA ON URBAN AIR POLLUTION IN DELHI, INDIA	39
2.1	Introduction	39
2.2	Data and Methods	44
2.3	Results	54
2.4	Discussion	76
3	DETECTING ANOMALOUS METHANE EMITTERS IN OIL/GAS FIELDS USING SATELLITE OBSERVATIONS	81
3.1	Introduction	81
3.2	Observing System Simulation Experiment	84
3.3	Results and discussion	100
3.4	Conclusions	106
	REFERENCES	110

THIS WORK IS DEDICATED TO MY WIFE, SYDNEY.

Acknowledgments

I would like to thank each of my coauthors for their help and guidance on each chapter of my thesis. For Chapter 1, I would like to thank Eric Leibensperger and Michael Iacono. For Chapter 2, I would like to thank Melissa Sulprizio, Tianjia (Tina) Liu, Miriam Marlier, Ruth DeFries, Sarath Guttikunda, and Pawan Gupta. For Chapter 3, I would like to thank Alex Turner, Jian-xiong Sheng, Laurent White, Cynthia Randles, and Jeremy Brandman.

I would like to acknowledge and thank Loretta Mickley and Daniel Jacob for their supervision and guidance over the course of my Ph.D. I would like to thank Loretta for helping me during the initial stages of my Ph.D to get excited about research and learn the ropes as quickly as possible. I would like to thank Daniel for helping me formulate and hone my last research project, and also for helping me refine the way I ask scientific questions. Without both advisors, none of this research would be possible. I would also like to thank the members of my research group for their guidance and friendship.

I would like to thank the Department of Earth & Planetary Sciences at Harvard University for creating a caring and intellectually stimulating community. I would like

to thank the Applied Mathematics and Institute for Applied Computation within the School of Engineering and Applied Sciences. Through my Master's Degree and Secondary in Computer Science coursework, I was exposed to and learned new technical skills which enhanced my research capability.

I would like to thank my parents and siblings for their constant support and encouragement. Most important, I would like to thank my wife Sydney. I would not have ever applied for graduate school at Harvard if she had not first taken the step of moving us out to the East Coast. I would not have survived the rigors of a Ph.D if she had not been there to always lend an ear and be my main source of encouragement. Thank you, I love you.

Chapter 1

Aerosol trends as a potential driver of regional climate in the central United States: Evidence from observations

1.1 INTRODUCTION

From 1930 to 2004, the eastern and central U.S. experienced significant cooling of as much as -0.12 K a^{-1} (Kumar et al., 2013). This phenomenon is known as the “warming hole”, as the temperature trend differs in sign from expected greenhouse gas warming (Pan et al., 2004). More recent observations from this region show average annual temperatures increased by 0.6–0.8 K between 1901–1960 and 1991–2012, signaling a reversal of the warming hole trend (Melillo et al., 2014). The causes of the warming hole and its subsequent reversal are uncertain. Previous studies have linked the U.S. warming hole to changing patterns in sea surface temperatures (SSTs) such as the Atlantic Multidecadal Oscillation (AMO) (Kumar et al., 2013; Zhang et al., 2013) or

to trends in anthropogenic aerosols, which may influence meteorology by interacting with solar radiation or clouds (Leibensperger et al., 2012b; Booth et al., 2012; Yu et al., 2014). Banerjee et al. (2017) concluded that while aerosols may have contributed to the warming hole, much of the observed cooling arose from unforced internal variability. Most of these studies relied on global or regional climate models, which are inherently uncertain. In this study, we use recent observations and simple models to better constrain the influence of aerosol-radiation interactions on U.S. regional meteorology. In response to tightening air quality regulations, emissions of aerosol sources are expected to decline worldwide over the 21st century (Westervelt et al., 2015), and so our results could have significance for regional climate elsewhere.

Some previous model studies have linked the emergence of a warming hole with changes in SST patterns or large-scale circulation. Pan et al. (2004) found that enhanced greenhouse gases produced a circulation response that increased the frequency of the southerly low-level jet over the southern Great Plains in late summer, which in turn replenished soil moisture and suppressed temperature extremes. They concluded this mechanism could induce a warming hole. Using a climate model forced with observed SSTs, Meehl et al. (2012) linked the warming hole cooling trend to the Interdecadal Pacific Oscillation (IPO). Kumar et al. (2013) found that among 22 climate models, those that had the best representation of the AMO also best reproduced the warming hole, although even these models show large discrepancies with the observed trend.

Alternative explanations for the warming hole involve the influence of aerosol trends on regional meteorology, which may include the impact of changing aerosols on SSTs. Over much of the United States, anthropogenic aerosols are dominated by light-colored, highly reflective species such as sulfate. Such aerosols have adverse health effects, and since the 1970s the U.S. EPA has worked to reduce their sources. Between 1990–2010, the Clean Air Act of 1970 and its amendments cut SO₂ emissions by 75 percent (USEPA, 2012), and this reduction may have affected regional climate. Focusing on June-July-August (JJA) during 2000–2011 across the United States, Yu et al. (2014) found positive correlations between monthly mean satellite observations of AOD and cloud optical depth (COD, $R = 0.76$) and between AOD and shortwave cloud forcing (SWCF, $R = 0.84$), as well as negative correlations between SWCF and maximum surface air temperatures ($R = -0.67$). They thus attribute the 20th century warming hole to aerosol-cloud interactions that lead to surface cooling. To quantify aerosol-radiation interactions, Gan et al. (2014) analyzed a sparse network of surface observations from the Surface Radiation Budget Network (SURFRAD) during 1995–2010, and found increasing trends in annual mean SWdn accompanying decreases in aerosol optical depth (AOD), especially among eastern U.S. sites. However, that study also detected an increase in clear-sky diffuse radiation, which is perplexing, given that declining aerosol would be expected to decrease such radiation. In contrast, Eshel (2016) diagnosed 1998–2014 surface observations at a site in upstate New York, and inferred that improved air quality has led to a strong increase in JJA SWdn there. The Eshel

(2016) result is similar to European studies that have tied aerosol reductions to enhanced SWdn (Philipona et al., 2009; Ruckstuhl et al., 2008). Tosca et al. (2017) compared the observed $+0.54 \pm 0.52$ K decade⁻¹ summertime temperature trend in the southeast U.S. to the -0.05 decade⁻¹ satellite retrieved AOD trend. The authors conclude that aerosols reductions enhanced surface temperature through aerosol-radiation interactions, but their study did not consider the covariability between observed clear-sky SWdn and AOD measurements, which we argue should exist for aerosol-radiation interactions.

In a modeling framework, Mickley et al. (2012) found that simple removal of U.S. aerosols exerted a top-of-atmosphere (TOA) radiative forcing of as much as $+4$ - 5 Wm⁻² over the central and eastern U.S. This forcing produced a positive feedback in which increases in surface shortwave radiation (SWdn) enhanced surface fluxes of sensible heat in late summer, drying out soils and reducing cloud cover, which further enhanced SWdn. To understand the climate response of a more realistic representation of U.S. aerosols, Leibensperger et al. (2012a) forced a global climate model using simulated historical aerosols, and found that high aerosol loading during 1970–1990s increased cloud cover and soil moisture by as much as 5% in the central and eastern U.S. The study also found that aerosol outflow to the Atlantic Ocean in this time frame may have cooled SSTs and increased mean JJA 850 hPa geopotential heights in the region of the Bermuda High (BH), a semi-permanent high-pressure system. Booth et al. (2012) found that stronger aerosol influence on surface forcing in a climate

model could better reproduce Atlantic sea surface temperatures (SSTs). Zhang et al. (2013) contested this result, pointing to mismatches between the Booth et al. (2012) model results and observations of North Atlantic upper ocean heat content and salinity. They proposed instead that variations in the Atlantic Multidecadal Overturning Circulation have driven recent changes in Atlantic SSTs. Mascioli et al. (2016) found competing effects on U.S. temperature extremes by changing aerosols and greenhouse gases over the 20th century, as expected, but temperature in the Southeast responded only weakly to aerosols in their simulation. Finally, using fifty-member ensembles, Banerjee et al. (2017) simulated aerosol-radiation interactions and the cloud albedo effect on U.S. climate, but not the cloud lifetime effect. They found that aerosol forcing could not entirely explain the 1951–1975 JJA decreasing trend in southeastern U.S. temperatures.

Nearly all these studies on the origin of the U.S. warming hole relied on climate or chemistry-climate models with their many uncertainties. For example, the response of soil moisture or low cloud cover to changing SWdn in such models may not be well captured (Soden and Held, 2006), and with few observations, aerosol concentrations in the early warming hole years are not well constrained. Aerosol composition is also not well known in the 1950s and 1960s, with black carbon emissions likely uncertain by at least a factor of two (Bond et al, 2007). The meteorological response to black carbon could be very different to that of sulfate (Koch et al., 2010; Bond et al., 2013), the most abundant anthropogenic aerosol in more recent decades.

In this paper, we turn to observational datasets to try to reconcile the apparently conflicting hypotheses of previous studies (e.g., Leibensperger et al., 2012b; Kumar et al., 2013; Gan et al., 2014; Yu et al., 2014). We extend previous analyses of SURFRAD trends (Long et al., 2009; Augustine and Dutton, 2013; Gan et al., 2014) by using more recent observations and by focusing on two central and eastern U.S. sites, where emission controls have had the largest influence on AOD. To gain knowledge of the potential influence of changing AOD on regional meteorology, we apply the observed AOD and cirrus cloud variables to a radiative transfer model and a simple statistical model. We further study regional meteorology during summers with enhanced SWdn to better understand how potential trends in SWdn could influence climate and soil hydrology, especially if the warming hole reversal continues, as suggested by some modeling studies (Leibensperger et al., 2012b). Our work has special relevance for developing countries that currently experience heavy aerosol loading but are planning emission reduction strategies (e.g., Lu et al., 2011).

1.2 DATA AND METHODS

We obtain surface SWdn observations from the SURFRAD Network, which consists of seven sites across the U.S. (Augustine et al., 2000). Although sparse, the network provides some of the longest *in situ* solar radiance measurements in the U.S., broken into diffuse and direct components. In this study we focus on 2000–2014 data from sites in Bondville, Illinois, and Goodwin Creek, Mississippi, as these sites are located

in the central and eastern U.S., and have experienced AOD reductions in the recent past (Gan et al., 2014). We exclude the SURFRAD Penn State site from this study, as the record is incomplete for much of 2009–2014. SURFRAD solar diffuse radiation is measured through a shaded Eppley Black and White Pyranometer, and direct solar radiation is measured with an Eppley Normal Incidence Pyrheliometer (NIP). Diffuse and direct measurements are summed to produce all-sky shortwave radiation fluxes. All broadband radiation measurements have a three-minute temporal resolution, taken as an average of one-second samples. There are uncertainties of 3% and 6% (4 and 20 Wm^{-2}) associated with the direct and diffuse measurements, respectively (Stoffel, 2005), where uncertainty is derived from the 95% confidence interval. SURFRAD stations also measure AOD in five spectral channels using a multifilter shadowband radiometer. The AOD data are also available as three minute averages, but only under cloud-free conditions. Both SURFRAD sites are located away from urban sources, so we expect them to be representative of the larger region. To verify consistency of SURFRAD trends against trends across the broader region, we compare the SURFRAD SWdn radiance data with pyranometer measurements from the U.S. Climate Reference Network (USCRN) (Diamond et al., 2013), but consider only those USCRN sites that have 10+ years of data, starting as early as 2003. We also compare SURFRAD SWdn to *in situ* pyranometer measurements from the Cary Institute of Ecosystem Studies (CIES; <http://www.caryinstitute.org>), located near Millbrook, New York, from 1990–2015.

SURFRAD also provides estimates of total clear-sky radiance using the all-sky observations, following the methods in Long and Ackerman (2000). Briefly, a power law model ($Y = A \cos(\theta)b$) is fit, where the initial guess for Y is all-sky SWdn, θ is the solar zenith angle, and A and b are the fitted coefficients. After eliminating cloudy measurements using various selection criteria, the power law model is refit following an iterative process until its coefficients converge, giving an estimate of clear-sky fluxes. A weakness of this fitting algorithm is that it may not remove the influence of thin cirrus clouds on the clear-sky flux (Long et al., 2009), and trends in cirrus cloud cover may potentially influence estimates of trends in clear-sky SWdn.

We use the column version of the Rapid Radiative Transfer Model for general circulation models (RRTMG_SW) to relate SURFRAD radiances to changes in AOD at the two sites (Iacono et al., 2008). RRTMG_SW relies on a correlated-k approach to approximate radiative fluxes and heating rates (Clough et al., 2005); multiple scattering is calculated through a two-stream approximation. We apply monthly mean profiles of atmospheric temperature, pressure, and ozone and water mixing ratios from the MERRA-2 Reanalysis (Rienecker et al., 2011) while keeping all other chemical profiles (e.g., CO₂ and N₂O) fixed to climatological means. The MERRA-2 ozone product is derived from a simple production and loss chemical scheme (Suarez et al., 2008), assimilated with measurements from the Ozone Monitoring Instrument and the Microwave Limb Sounder. MERRA-2 ozone fields below 260 hPa are not as reliable. We also apply MERRA-2 surface emissivities to the RRTMG_SW bands

in the infrared part of the spectrum (820–4000 cm^{-1}), and one minus the observed SURFRAD SW reflectance in the RRTMG_SW bands in the shortwave region (4000–50000 cm^{-1}).

We drive the model with observed monthly mean AOD from SURFRAD. For single scattering albedo and asymmetry parameters of the aerosol, we rely on measurements from two long-term AERONET sites located close to the SURFRAD sites: Bondville, IL, and Huntsville, AL (Dubovik and King, 2000). Gan et al. (2014) found close agreement between AOD measurements at the Bondville SURFRAD site and nearby AERONET sites, so we assume that AERONET aerosol properties represent those at SURFRAD. For information on thin cirrus clouds, we rely on cloud fraction, cloud water path, and ice and liquid radius data retrieved from the CERES instrument onboard the Terra and Aqua satellites (Minnis et al, 2011). CERES thin cirrus cloud optical depths have been shown to those retrieved by the Cloud-Aerosol Lidar and Infrared Pathfinder Satellite (launched in 2006) over land ($R = 0.65$). More detailed information about the cirrus retrieval uncertainty is currently being explored (Minnis, P; personal communication). Since we are interested in total column extinction and surface radiation values, we assume all aerosols are concentrated in the surface layer and fix cirrus fractional cloudiness at 300 hPa. In a sensitivity simulation, we find that whether we fix the aerosols at the surface layer or distribute the aerosol through the lower troposphere has little effect on modeled surface SWdn. Since the CERES Level 3 retrieval is available only since 2000, we perform all monthly radia-

tive transfer simulations over the 2000-2014 period. Our model setup is similar to the approach of Ruiz-Arias et al. (2013) who performed Weather Research and Forecasting (WRF) model simulations using RRTMG_SW driven by AERONET AOD and aerosol parameters during October 1–3, 2011. The authors found close agreement between 10-minute modeled and observed total, direct, and diffuse SWdn at SURFRAD and Atmospheric Radiation Measurement (ARM) sites, though they did not perform simulations for Goodwin Creek. We perform simulations with monthly mean data to avoid gaps in the AOD and CERES record. Missing daily observations at Goodwin Creek and Bondville range from 43-49%.

To assess the regional climate impacts of variations in SWdn at Bondville, we use tower data from the nearby Ameriflux site, also in Bondville, and Illinois Climate Network (ICN) sites (WARM, 2014). The Ameriflux site provides 9 years (1998-2007) of continuous observations of summertime radiation, temperature, heat flux, and soil moisture. The tower is located within an active corn/soybean agriculture field, but experiences little irrigation (Meyers, T; personal communication), which could influence the microclimate. The Bondville ICN tower sits in a non-irrigated grass field. Between 1983–2002, semimonthly soil moisture measurements were made at the tower, using a neutron probe instrument. Since 2003, the station has taken hourly soil moisture measurements using a hydraprobe sensor. The ICN site also provides shortwave global radiation, temperature, and soil temperature data from 1990–present.

We compare Ameriflux and ICN tower data with the estimates from the NASA

Land Data Assimilation System (NLDAS) over North America (Mitchell et al., 2004). The goal of the NLDAS project is to construct high-quality, consistent datasets for use in land surface models (LSMs). NLDAS utilizes a combination of gauge-based precipitation and meteorological data from the NCEP North American Regional Reanalysis (NARR) to drive an ensemble of LSMs, yielding estimates of soil moisture and surface energy fluxes. Version 2 of NLDAS also uses bias-corrected SWdn data from the University of Maryland Surface Radiation Budget dataset, which is based on GOES-8 satellite data (Pinker et al, 2003). Here we analyze the output from three LSMs in the NLDAS project: Mosaic, Variable Infiltration Capacity (VIC), and Noah (Xia et al, 2012; Koster and Suarez, 1994; Wood et al, 1997). Each LSM has a different treatment of land-atmosphere coupling, but all require that incoming solar radiation balance the sum of outgoing thermal radiation, latent and sensible heat losses, and diffusion of energy into the soil (Overgaard and Rosbjerg, 2011). Spatial resolution of these models is $1/8^\circ \times 1/8^\circ$.

We calculate annual trends in observations using monthly mean anomalies. For the SURFRAD dataset, we first find the mean diurnal profile for each month during 2000–2014. We then calculate monthly mean SWdn by averaging these diurnal profiles over daylight hours. We compute the monthly climatology over the 2000-2014 period and subtract that from each year’s monthly means to arrive at monthly SWdn anomalies. Trends for other data measured with hourly or daily frequency are computed using the same method. Radiative transfer simulations in RRTMG_SW are

performed with monthly average AOD and other environmental variables. As with the observations, we find monthly anomalies in RRTMG_SW by first calculating the 2000–2014 monthly climatology for each simulation and subtracting that from the corresponding time series of monthly means. We use least squares regression to estimate the slopes of the time series of both observed and modeled monthly anomalies. To test for statistical significance, we follow the method described by Weatherhead et al., (2008), a method also utilized by Gan et al., (2014). This method determines the significance of a least-squares trend based on variance of the noise (i.e., the residual from the straight line fit), the autocorrelation of the noise, and the number of data points were used to determine the trend. In this study, we set $p < 0.05$ as the threshold for statistical significance.

1.3 LONG-TERM TRENDS IN SURFACE SWDN

Observed 500 nm AOD decreases significantly at both Bondville (-0.047) and Goodwin Creek (-0.052) during the 2000–2014 time frame (Figure 1.1), providing evidence of the success of strengthening U.S. air quality regulations. In Figure 1.2 we show the corresponding trends in observed SWdn for all-sky and clear-sky conditions at the two sites. Both stations show significant increases in total (diffuse + direct) all-sky as well as clear-sky SWdn, as would be expected from the changes in AOD. However, diffuse SWdn is the dominant contributor to these clear-sky trends in both cases, a finding that is discussed further in this section and Section ??.

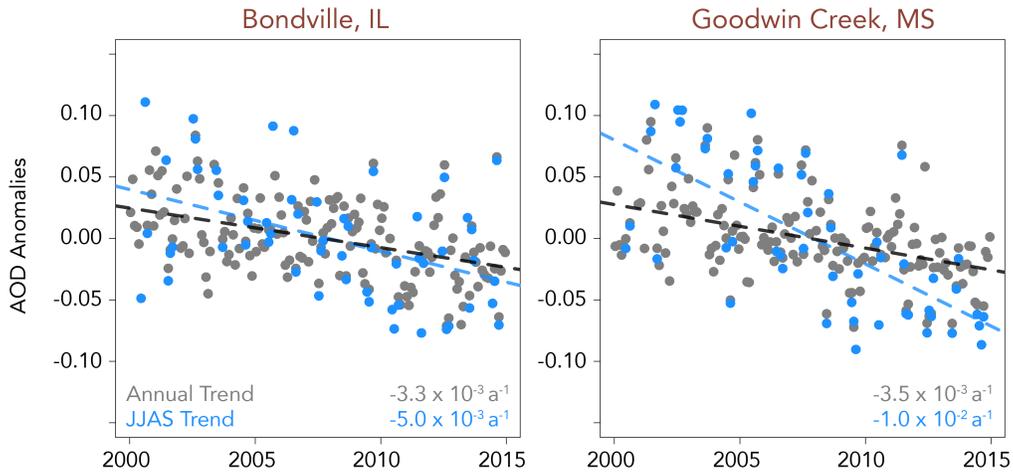


Figure 1.1: Trends in monthly anomalies of 500-nm aerosol optical depth (AOD) at Bondville, IL and Goodwin Creek, MS, during 2000–2014. All trend lines pass the significance threshold ($p < 0.05$).

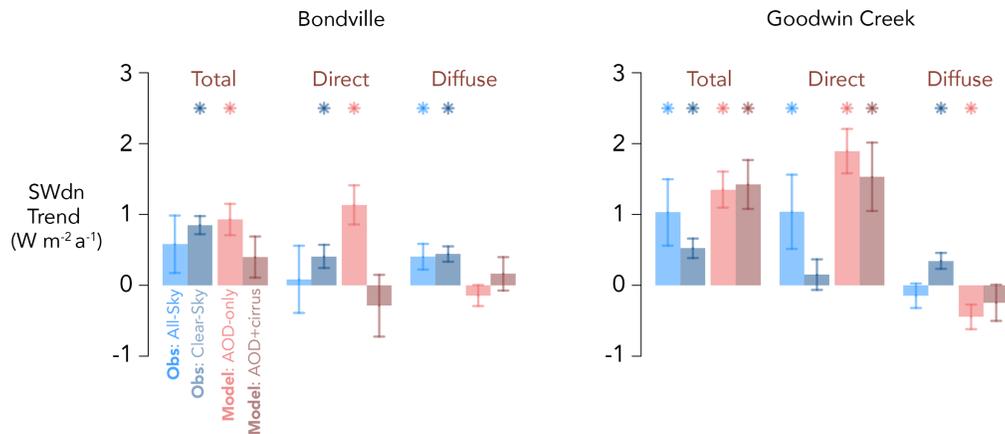


Figure 1.2: Annual trends in downward surface solar radiation (SWdn) during 2000–2014 at Bondville, IL (left), and Goodwin Creek, MS (right). Trends are determined from daytime data (10–23 UTC). Trends in observed all-sky SWdn from the SURFRAD network are shown in light blue; observed trends in clear-sky SWdn are in dark blue. Modeled trends in clear-sky SWdn are shown for an aerosol-only simulation (light red) and for a simulation with both aerosols and cirrus clouds included (dark red). Error bars are the standard error of the slope estimated from a linear regression fit. Asterisks above the bars indicate statistical significance ($p < 0.05$).

Figure 1.2 also shows the modeled trends in SWdn. At Bondville, the aerosol-only simulated trend in clear-sky SWdn agrees in magnitude and sign ($+0.93 \pm 0.22 \text{ Wm}^{-2}\text{a}^{-1}$) with that observed ($+0.85 \pm 0.13 \text{ Wm}^{-2}\text{a}^{-1}$). Breaking down the total SWdn into its direct and diffuse components, the aerosol-only simulation shows a result consistent with aerosol reductions, specifically large increases in direct SWdn accompanied by a decrease in diffuse SWdn. However, this result differs from clear-sky observations, which show both direct and diffuse SWdn increasing ($+0.41 \pm 0.16 \text{ Wm}^{-2}\text{a}^{-1}$ and $+0.44 \pm 0.11 \text{ Wm}^{-2}\text{a}^{-1}$).

As noted above, thin cirrus clouds may influence SWdn even under apparently clear-sky conditions. Figure 1.3 shows spatial trends in cirrus ice cloud fraction over 2000–2014 from CERES. Cirrus cloud fraction increases as much as $+0.5 \text{ \% a}^{-1}$ over parts of the eastern U.S., with a $+0.21 \text{ \% a}^{-1}$ increase over Bondville. Trends in two other cirrus cloud properties, cloud water path, and cloud particle radius are not as spatially coherent as those of cirrus cloud fraction, and we do not discuss these further. Incorporating cirrus cloud fraction and the other two cloud parameters into the aerosol-cirrus simulation still yields an increasing trend in total clear-sky SWdn as shown in Figure 1.2 ($+0.40 \pm 0.29 \text{ Wm}^{-2}\text{a}^{-1}$), but with a magnitude only about half that observed. Diffuse SWdn in this simulation is roughly a third of observed clear-sky trend, and this match comes at the expense of direct SWdn, which now shows a decreasing trend, in contradiction to the observations. Neither the direct nor diffuse SWdn trends in the aerosol-cirrus simulation are statistically significant.

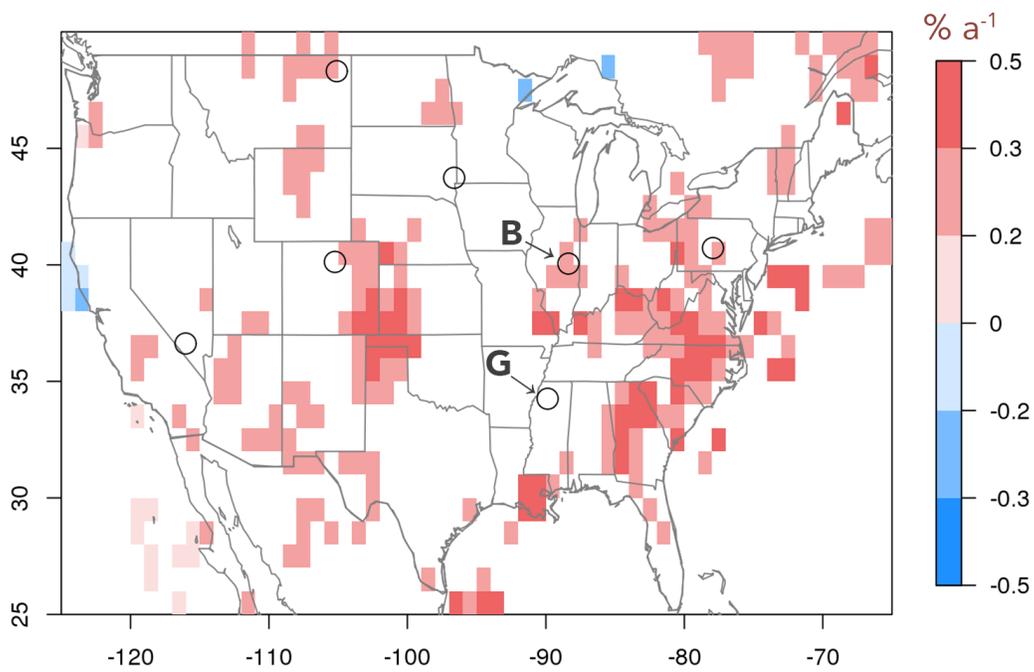


Figure 1.3: Annual trends in cirrus ice cloud fraction as retrieved from Clouds and Earth's Radiant Energy System (CERES) for 2000–2014. Black circles represent the locations of the SURFRAD stations (B = Bondville, IL; G = Goodwin Creek, MS). White indicates regions where trends are not statistically significant ($p > 0.05$).

At Goodwin Creek, the modeled aerosol-only simulation trend in clear-sky SWdn is $+1.35 \pm 0.25 \text{ Wm}^{-2}\text{a}^{-1}$, more than double the observed clear-sky trend ($+0.52 \pm 0.14 \text{ Wm}^{-2}\text{a}^{-1}$). As at Bondville, the diffuse component of the observed clear-sky SWdn at Goodwin Creek also exhibits an increasing trend ($+0.34 \pm 0.11 \text{ Wm}^{-2}\text{a}^{-1}$), even though cirrus cloud fraction shows no significant trend there (Figure 1.3). In fact, the diffuse SWdn trend in the aerosol-cirrus simulation at Goodwin Creek is still negative, though more positive than the aerosol-only simulation. This result contrasts with that in Bondville, where consideration of the large cirrus trend changed the sign and significance of direct and diffuse SWdn trends.

Reconciling the observed trends in diffuse, direct, and total SWdn at the two sites is challenging. In their analysis of SURFRAD data, Gan et al. (2014) also found increasing trends in clear-sky diffuse SWdn averaged over seven sites across the U.S. from 1995–2010. That study hypothesized that trends in this variable could be traced to increasing air traffic and enhanced thin cirrus cloud formation. The effect of aircraft contrails on total cirrus cloud fraction is uncertain. Analyzing trends in upper atmosphere humidity and cirrus cloud cover, Minnis et al. (2004) determined that the observed 1971–1995 $+1.0\%$ per decade ($+0.1\% \text{ a}^{-1}$) increase in cirrus cloud fraction over the U.S. was indeed caused by increased air traffic. The magnitude of this trend is similar to what we observe in Figure 1.3. Consistent with Minnis et al. (2004), Travis et al. (2004) found the diurnal temperature range (DTR) over the entire U.S. increased by 1.0 K compared to the 1971–2000 climatological mean after the 3-day

suspension of nearly all air traffic following the events on September 11, 2001, with especially large increases (1-2 standard deviations above the climatological mean) in regions such as Illinois that favor the formation of aircraft contrails at high altitudes. The conclusions of Minnis et al. (2004) finding, although controversial (Hong et al., 2008), suggest a strong influence of contrails on the surface energy budget. At the two SURFRAD sites, we do not find evidence in the diffuse SWdn record of a response to the abrupt halt to air traffic in September 2001 (Figure 1.4). At Bondville, a slight enhancement in diffuse radiation occurred one day after September 11, which then decayed to the 1995–2000 average after two weeks. Enhancements of similar magnitude occur previous to September 11, so excursions from the mean may be typical for diffuse SWdn at Bondville. Diffuse SWdn at Goodwin Creek site exhibits a jump on September 11 above the 1995–2000 climatology, but mostly stays within one standard deviation of the 1995–2000 average afterward. These jumps in diffuse SWdn around September 11th would seem to contradict an influence of contrails on surface SWdn observations. However, it is also possible that the sites are simply not representative of the larger domain during this short timeframe. However, since we do not see strong evidence of aircraft influencing the diffuse clear-sky SWdn at Bondville during Sept. 11-13, we are cautious in ascribing the increasing diffuse clear-sky SWdn trends simulated at Bondville (Figure 1.2) to aircraft contrails. The increasing trend in CERES cirrus cloud fraction (Figure 1.3) may be due to other factors that are outside the scope of this paper.

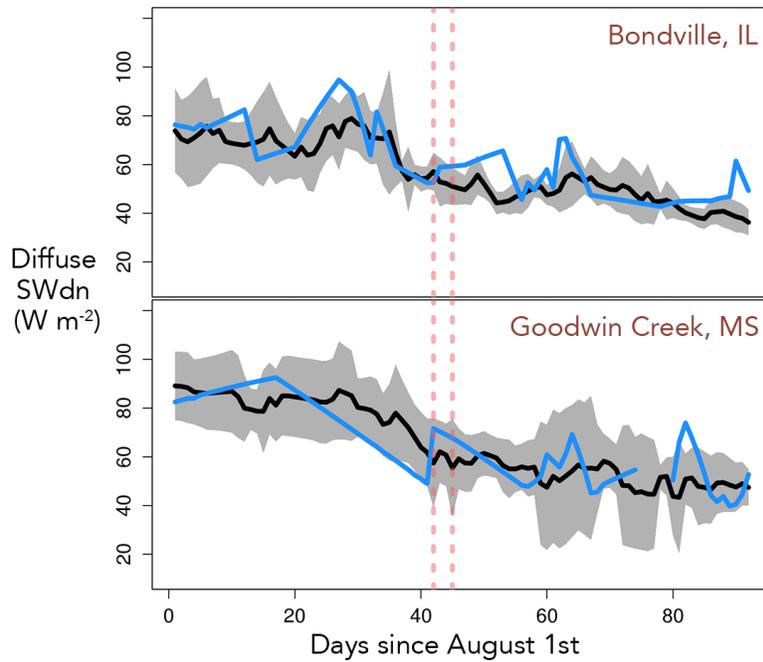


Figure 1.4: Observed diffuse downward surface solar radiation (SWdn) during August-October at Bondville, IL (top), and Goodwin Creek, MS (bottom). The blue line is the timeseries of daily mean diffuse SWdn at each site in 2001. The black line represents daily diffuse SWdn averaged over 1995–2000 at the sites, shaded by one standard deviation from the mean. The red vertical lines mark September 11th, when air traffic was temporarily suspended in 2001, and three days after, when air travel largely resumed.

1.4 SHORT-TERM VARIABILITY IN SWDN

??

The 2000–2014 trends in total observed and modeled clear-sky SWdn are positive at both Bondville and Goodwin Creek, implying a link between aerosols and SWdn. However, surface clear-sky SWdn would be expected to respond rapidly to changes in overhead aerosol, and here we check whether changes in aerosols and/or cirrus clouds can explain the monthly variability of clear-sky SWdn observations. Figure 1.5 shows the timeseries of observed and standardized monthly mean SWdn anomalies at Bondville and Goodwin Creek, together with SWdn results from the aerosol-only and aerosol-cirrus radiative transfer simulations. The standardized timeseries is constructed by differencing each month’s value with the long-term monthly mean and then dividing by the monthly standard deviation. The Bondville aerosol-only simulations show greater correlation with observations (e.g., $R = 0.49$ for total SWdn) than that for the aerosol-cirrus simulations ($R = 0.29$). At Goodwin Creek, results from neither monthly simulation are significantly correlated with monthly observations.

For AOD to be a controlling factor of clear-sky SWdn, we would expect covariability between variables. To check whether this connection is robust, we develop a statistical model to predict surface monthly mean clear-sky SWdn anomalies based on AOD and cirrus cloud properties using multivariable linear regression (MLR) with no lag. We perform MLR for June-July-August-September (JJAS). Aerosol load is

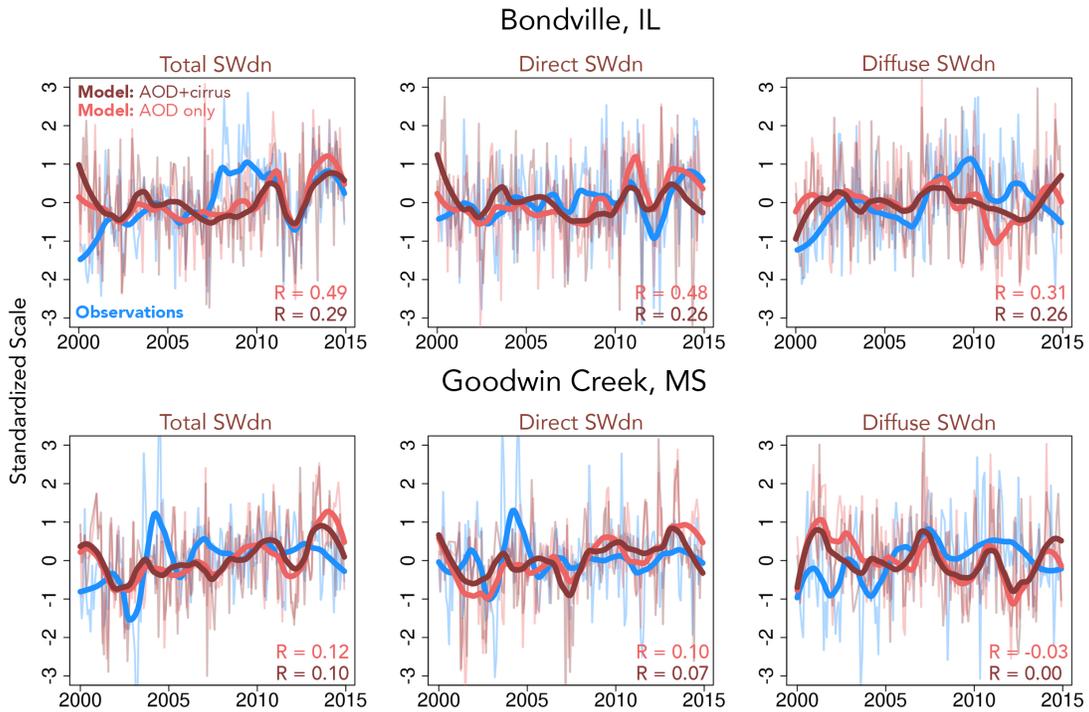


Figure 1.5: Time series of monthly mean anomalies (standardized by mean and standard deviation) in surface solar radiation (shortwave down or SWdn) at Bondville, IL, and Goodwin Creek MS, from 2000–2014. Blue curves denote observations, light red shows model results with observed aerosol optical depths (AOD) taken into account, and dark red shows results when both AOD and observed cirrus cloud fraction are included. Thick lines represent a 3-year locally-weighted scatterplot smoothing (lowess). The correlations R between the non-smoothed model simulations and observations are shown inset. All the correlations at Bondville are significant ($p < 0.05$), and no correlations at Goodwin Creek are significant.

generally highest in the U.S. during these months (Malm et al., 2004), the incoming solar flux is large, and feedbacks can extend the aerosol influence into late summer (Mickley et al., 2012). We find the optimal coefficients of the MLR by individually fitting independent models across all combinations of predictor variables. For predictor variables, we use AOD and the same cirrus cloud parameters used to drive RRTMG_SW aerosol-cirrus simulations cirrus cloud liquid water path (LWP), cirrus cloud ice water path (IWP), cirrus cloud liquid radius (R_L), cirrus cloud ice radius (R_I), and cirrus cloud fraction (C_f). We also include monthly average column ozone (pressure weighted) as a predictor. We optimize coefficients for each individual MLR fit using leave-one-out cross-validation. For each MLR fit, we calculate the Bayesian Information Criterion (BIC), which scores the MLR based on its goodness of fit, and penalizes based on the number of parameters included in the regression (Posada and Buckley, 2004). Thus we seek solutions that explain clear-sky SWdn using the fewest number of terms, so as to avoid over-fitting. Both clear-sky SWdn and the predictors are detrended, deseasonalized, and standardized before the MLR is fit.

Table 1.1 summarizes the coefficients of the optimal MLR fits to clear-sky SWdn. At Bondville, we find the optimal clear-sky total SWdn MLR model is driven by AOD and overhead ozone, explaining 15% of the variance. Direct SWdn is best explained (20%) by AOD alone. The MLR model with AOD and cirrus LWP best fits the observed variability in clear-sky diffuse SWdn, explaining 26% of the variance. The magnitude of the fitted AOD coefficients are of similar magnitude and opposite sign for

Table 1.1: Fitted regression coefficients to downward surface shortwave radiation (SWdn) at Bondville, IL, from 2000-2014. Fits represent the optimal multivariate regression model chosen using the Bayesian Information Criterion (BIC) and fit to column-averaged ozone, aerosol optical depth (AOD), cirrus cloud liquid water path (LWP), cirrus cloud ice water path (IWP), cirrus cloud liquid radius (R_L), cirrus cloud ice radius (R_I), and cirrus cloud fraction (C_f).

SWdn	Regression Fits*	R ²
Total	-0.27 AOD + 0.30 Ozone	15%
Direct	-0.45 AOD	20%
Diffuse	0.44 AOD + 0.29 LWP	26%

*Predictor variables were detrended, deseasonalized, and normalized by their mean and standard deviation before being fitted to SWdn anomalies.

direct and diffuse SWdn, consistent with the expectation for aerosol-radiation interactions. The coefficients of determination (R^2) in Table 1.1 are similar in magnitude to the correlations between observed SWdn fluxes and those calculated by the radiative transfer model (Figure 1.5), showing a consistency between the two methods to interpret the influence of AOD on SWdn.

At Goodwin Creek, neither observed AOD nor any property of cirrus clouds can explain the variability in direct or diffuse SWdn, casting doubt on these variables as influences on SWdn at this site. That two separate sites with similar 2000-2014 AOD reductions could have different MLR results underscores the possible multiplicity of drivers of SWdn and the uncertainty in resolving local radiation budgets. RRTMG_SW simulations driven by AOD have been shown to capture SWdn fluxes at SURFRAD sites well (Ruiz-Arias et al., 2013). However, Ruiz-Arias et al. (2016) found a 4% reduction in monthly mean AOD compared to daily AOD in fine aerosol regimes. (Ruiz-Arias et al. segregate aerosol regimes using the Angstrom exponent,

which is a direct function of the average size of the aerosol mixture). Among the Bondville and Goodwin Creek AOD timeseries and the CERES retrievals, nearly half the days are missing coincident observations, hence the need to bin the daily data into monthly observations.

1.5 METEOROLOGICAL IMPACTS FROM ENHANCED SWDN IN LATE SUMMER

Previous work has suggested that the land-atmosphere coupled response to increased SWdn involves a cascade of meteorological phenomena (Shindell et al., 2003; Budyko, 1969). Gu et al. (2012) saw evidence of coupling between net radiation, heat fluxes, and soil moisture using Ameriflux observations during the 2005 growing season in Missouri. The model study of Mickley et al. (2012) found that U.S. aerosol reductions lead to enhanced latent heat fluxes in early summer which transition to enhanced sensible heat fluxes by late summer/ early fall. We probe the observational record for evidence of these feedbacks by first analyzing 1998–2007 tower data from the Ameriflux site at Bondville. We classify the data for each year into either a sunny or cloudy regime depending on whether the JJAS mean SWdn for that year is above or below the climatological JJAS median. We chose JJAS as the timeframe of reference due to the large summer to early fall SWdn response from reduced aerosols seen in Mickley et al. (2012). We then compare the responses in monthly mean surface fluxes for these two regimes (Figure 1.6). We do not consider aerosols directly here, as the short time series of the Ameriflux data limits the ability to assess long-term trends, but

we can use our results to understand the regional sensitivity of the land-atmosphere system to SWdn changes driven by AOD trends. We also do not consider microclimate feedbacks from irrigation at the Ameriflux station. Although irrigation has been shown in other studies to suppress extreme temperatures in the Midwest (Mueller et al., 2016), we find that the monthly-average temperatures recorded at the Ameriflux Bondville site nearly match the temperature readings at the nearest airport in Champaign, IL (not shown). Also, since the Ameriflux temperatures and SWdn data correspond closely to these value in the 1/8 NLDAS dataset (Figure 1.8), we assume that the effect of micro-scale irrigation is small.

Figure 1.6 shows that the difference in all-sky daytime SWdn between these two regimes during JJAS is $+28.1 \text{ Wm}^{-2}$. Enhanced SWdn during sunny years leads to increased latent heat fluxes in May-June ($+5.3 \text{ Wm}^{-2}$), which transition to increased sensible heat fluxes in August-October ($+9.5 \text{ Wm}^{-2}$). Volumetric soil water content follows the latent and sensible heat fluxes, with greater soil moisture in June ($+2.7\%$) and drier conditions in August (-3.5%). Because of the increased sensible heating in the sunny regime, we would expect a corresponding change in temperature. However, the difference in JJAS maximum temperature at this site is slightly negative (-0.86 K), with cooler temperatures during sunny years. This negative change in temperature is corroborated by a negative change in upward surface longwave radiation (LWup) at the nearby SURFRAD site (-5.4 Wm^{-2}). Precipitation at Bondville does not differ significantly between sunny and cloudy regimes (not shown). Summer 2004,

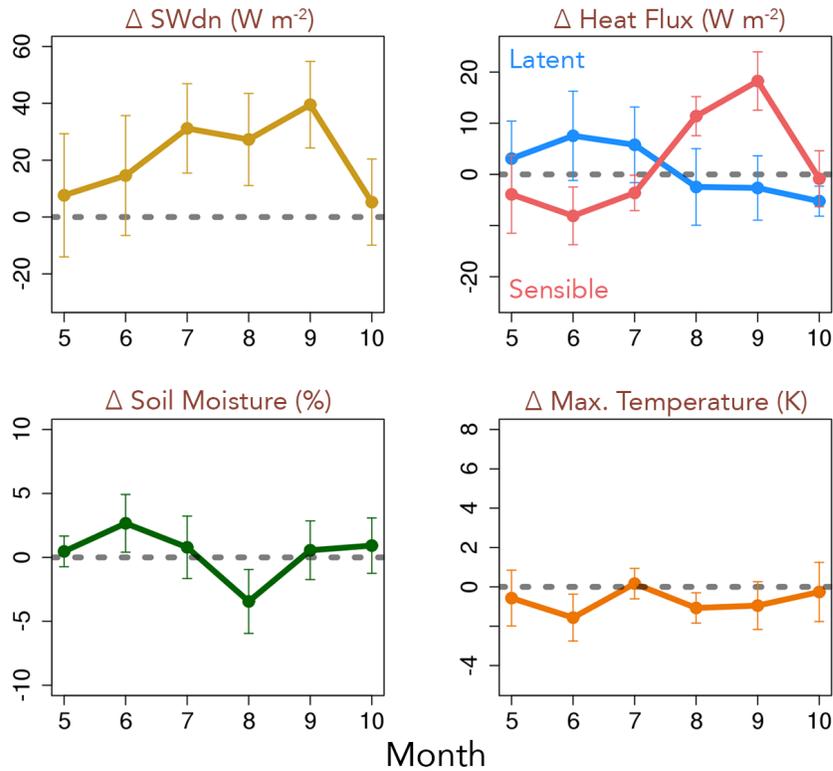


Figure 1.6: Differences in surface measurements at the Bondville Ameriflux site (1998–2007) between sunny and non-sunny summers, where summer is defined as June–July–August–September and a sunny summer is defined as one with average all-sky downward surface solar radiation (SWdn) greater than the median 1998–2007 all-sky summer SWdn. Error bars represent the 95% confidence interval of the difference.

classified as sunny, was paradoxically the coolest summer in the Ameriflux record, with a mean maximum temperature 1.7 K cooler than the 1998–2007 average at this site. Indeed much of the central U.S. experienced cool temperatures that summer (w2.weather.gov/dtx/2004annualclimatesummary). The surprising result of a cool but sunny summer in 2004 at the Ameriflux site points to the possible problem of relying on its short (10 year) record of observations.

We next examine ICN measurements of SWdn and surface variables, including soil moisture, using the same method of segregating sunny and cloudy summers. Since the method of measuring ICN soil moisture changed in 2003, we analyze differences in sunny vs. cloudy regimes within each measurement period, with PER1 corresponding to 1990–2002, and PER2 corresponding to 2003–2014. Though no sensible or latent heat measurements are available here, the ICN data include a record of 4 inch soil temperatures under sod together with soil moisture and maximum air temperatures (Figure 1.7). We see a similar pattern across the two instrument regime periods, specifically that sunny JJAS summers (PER1 = $+16.3 \text{ Wm}^{-2}$, PER2 = $+8.5 \text{ Wm}^{-2}$) favor hotter soils (PER1 = $+1.2 \text{ K}$, PER2 = $+2.9 \text{ K}$) and reduced soil volumetric water content (PER1 = -3.0% , PER2 = -3.4%). In both instrument periods, the maximum JJAS temperature increases during sunny summers (PER1 = $+2.0 \text{ K}$, PER2 = $+2.5 \text{ K}$), in contrast with the Ameriflux measurements. In the PER2 ICN data, summer 2004 is categorized as cloudy, hence its cooler temperature is expected. The LWup change at the nearby SURFRAD site during PER2 is consistent with the temperature increase ($+4.8 \text{ Wm}^{-2}$). Precipitation again does not differ significantly between sunny and cloudy regimes for each instrument period (not shown). The changes in SWdn for both ICN and Ameriflux stations are strongly positive, with the change during both PER1 and PER2 about half that in the Ameriflux record.

The record of observations at Ameriflux Bondville is relatively short, and the inconsistent soil instrument record at the ICN station complicates long-term analysis.

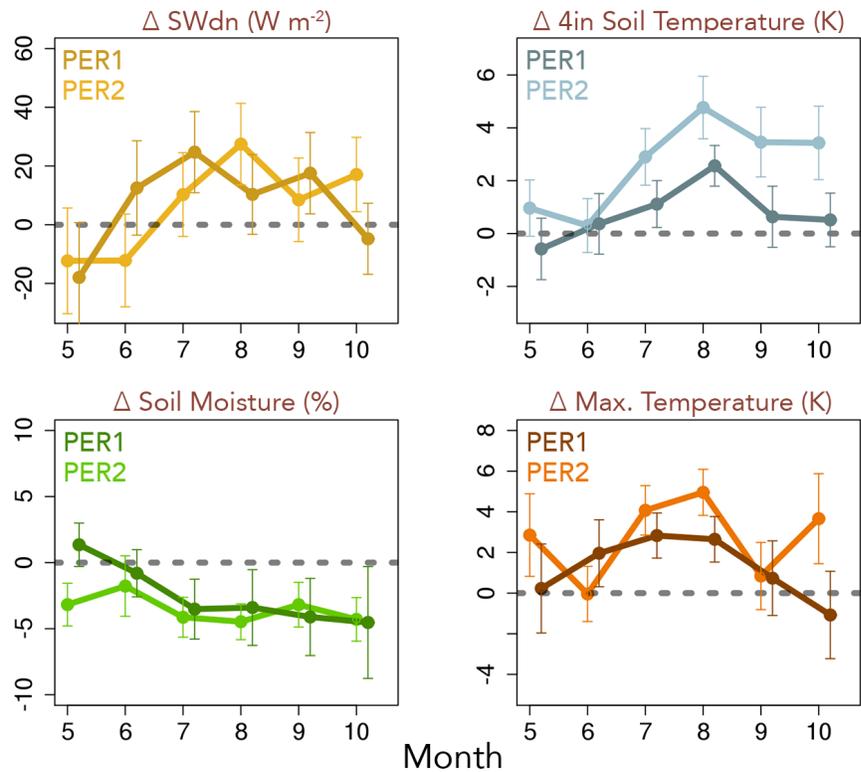


Figure 1.7: Differences in surface measurements at the Bondville ICN site between sunny and non-sunny summers. The technique for measuring soil moisture changed at this site in 2003, and the panels show differences separately for the periods before and after this change: period 1 (PER1, 1990–2002) and period 2 (PER2, 2003–2014). As in Figure 1.6, summer is defined as June–July–August–September and a sunny summer is defined as one with average all-sky downward surface solar radiation (SWdn) greater than the median all-sky summer SWdn of each respective period. Error bars represent the 95% confidence interval of the difference. For clarity, monthly means for PER1 are slightly offset with respect to the abscissa.

Ameriflux and ICN sites yield consistent SWdn and soil moisture results, but different surface temperature responses, perhaps due to the short-term measurement record. To investigate long-term land-atmosphere changes from 1990–2015, we use assimilated meteorology and other variables from the NLDAS dataset. To assure the quality of the NLDAS driving variables, we compare three of these variables monthly mean SWdn, precipitation, and air temperatures with observations from Ameriflux stations with at least 5 years of data (11 total). We also compare NLDAS SWdn with that from SURFRAD. Figure 1.8 shows statistically significant agreement among these datasets, with R ranging from 0.60 to 0.77, depending on the variable. We find much weaker, but still statistically significant, agreement between the LSM results and the Ameriflux observations, with R ranging from 0.24 to 0.47, depending this time on both the variable and model (Figure 1.8). Best matches between models and measurements are obtained for the sensible heat flux. The relatively weak correlations underline the difficulty in resolving land-atmosphere coupling at $1/8^\circ \times 1/8^\circ$ resolution.

Using the NLDAS dataset and mindful of the uncertainties in LSM results, we expand our focus to look at spatial trends in the relevant variables across the contiguous United States for JJAS over the 1990–2015 time period. Figure 1.9 shows that surface JJAS all-sky SWdn has increased significantly by $+0.78 \text{ Wm}^{-2}\text{a}^{-1}$ from 1990 to 2015 across the central U.S. (30–50N, 105–85W, denoted by the green box in Figure 1.9). Figure 1.9 also reveals a close correspondence between the all-sky NLDAS trend

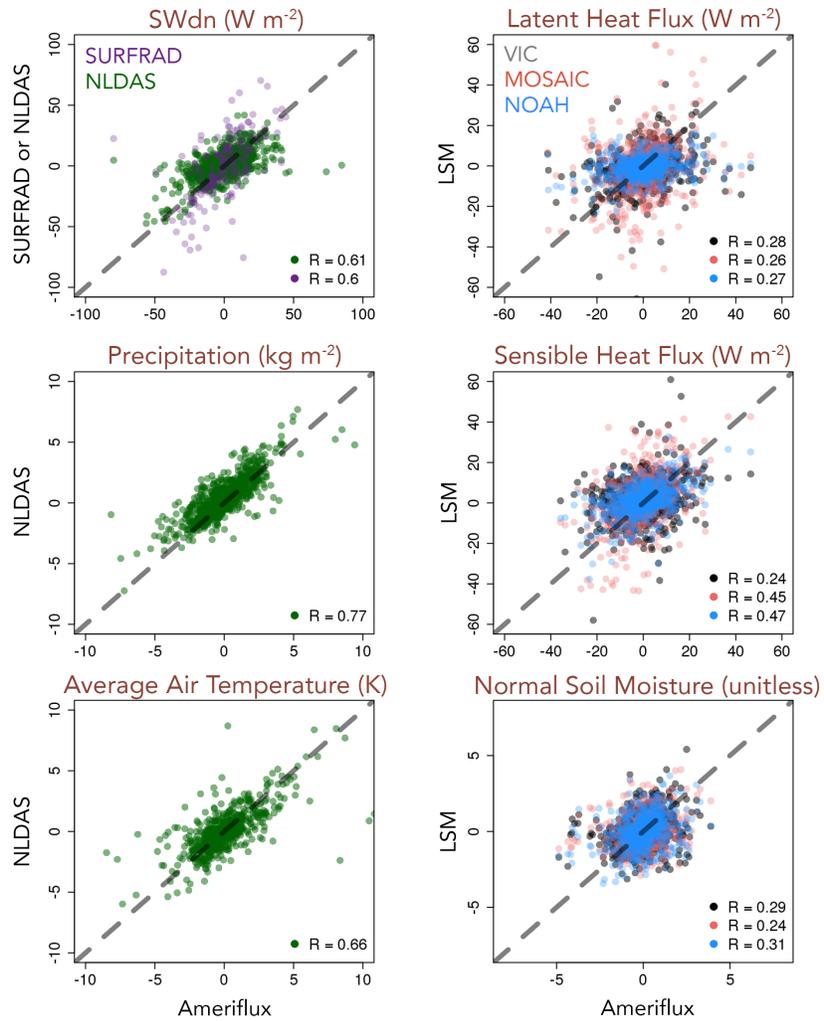


Figure 1.8: Comparison between Ameriflux/SURFRAD stations (11 total) with 5+ years of data and the corresponding North American Land Data Assimilation System (NLDAS) data. The left column compares observed forcing data to the NLDAS forcing data. The right column compares the land surface model (LSM) results to the Ameriflux observations. Each point represents the monthly anomaly calculated by removing the monthly climatology from the monthly mean of the data. Soil moisture has been standardized by its mean and standard deviation to remove biases in unit conversion.

and the all-sky trends derived from site measurements in the SURFRAD (1997–2014), USCRN (2003–2014), and CIES (1990–2015) networks, increasing confidence in the NLDAS dataset. Deviations between data products may be explained by the inconsistent time periods of comparison. Accompanying the change in SWdn is an increase in average JJAS air temperatures over much of the central and eastern U.S. (+0.07 K a⁻¹). Precipitation decreases slightly in the central U.S. (-0.19 kg m⁻² a⁻¹), mostly in a few isolated regions over the Great Lakes. The total JJAS enhancement in NLDAS all-sky SWdn over the central U.S. over the 1990–2015 time period is +20 Wm⁻², similar in magnitude to the increase observed during sunny years at the Ameriflux (+28.1 Wm⁻²) and ICN (PER1= +16.3 Wm⁻², PER2 = +8.5 Wm⁻²) sites. The 2000–2014 clear-sky SWdn enhancements that we simulate with RRTMG_SW at Bondville and Goodwin Creek (+13.9 and +6.2 Wm⁻²) are about half the NLDAS enhancements averaged over the central U.S. for the longer time period of 1990–2015.

Figure 1.9 also shows the soil moisture response to increasing SWdn and warmer temperatures over 1990–2015, as calculated by the three LSMs. We combine LSM results by first determining which of the model trends agree in sign in each grid cell and then taking the mean of just those models that agree. The combined trend reveals decreased soil moisture across the central U.S. between 1990 and 2015 (-0.85 kg m⁻²a⁻¹, averaged over the region defined in Figure 1.9), accompanied by an increase in sensible heating in the same region (+0.28 Wm⁻² a⁻¹). This decrease in soil moisture translates to a 1990–2015 decrease in volumetric soil water content of -2.2%,

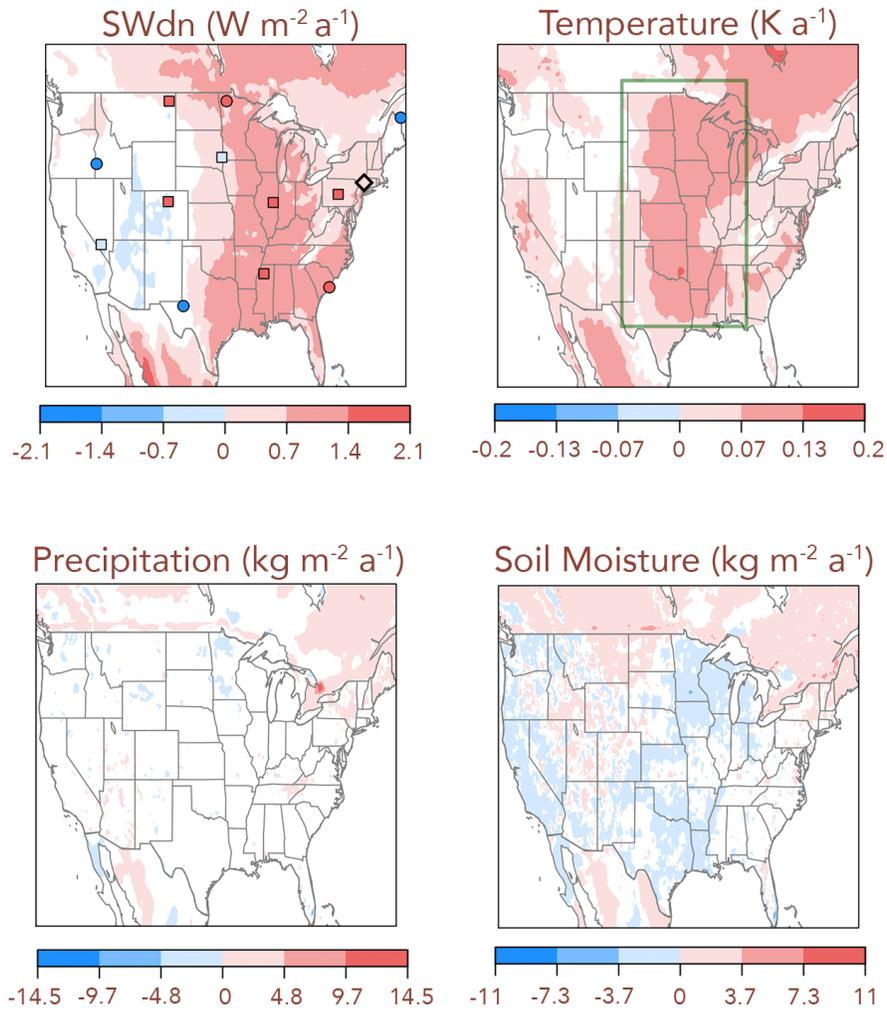


Figure 1.9: June-July-August-September (JJAS) trends in surface meteorological variables from 1990 to 2015. Surface downward solar radiation (SWdn), air temperature, and precipitation are assimilated from observations in the NASA Land Data Assimilation System (NLDAS). Overlaid on the SWdn plot are observed trends from the U.S. Climate Reference Network (USCRN: circles, 2003–2014), SURFRAD (squares, 1997–2015), and the Cary Institute for Ecosystem Studies (CIES: diamond, 1990–2015). The green box in the temperature panel represents the central U.S. (30–50N, 105–85W). In the soil moisture panel, we combine model results by first determining which of the LSM trends agree in sign in each grid cell and then showing the mean of just those models that agree. White indicates those regions that fail statistical significance (i.e., $p > 0.05$).

within the range of what is observed as the JJAS difference between sunny and cloudy years at the Ameriflux (-0.14%) and ICN (PER1 = -3.0%, PER2 = -3.4%) sites in Illinois. Since the precipitation patterns seem to be unchanging or small over much of the central U.S. during 1990-2015, the LSM soil moisture results provide evidence of a climate response to greater evapotranspiration in the presence of enhanced SWdn.

We emphasize that our main goal in examining the 1990–2015 NLDAS dataset is to probe the regional response of soil moisture and temperature to trends in all-sky SWdn. Coincidentally, the trends in all-sky SWdn trend across the central U.S. in the NLDAS dataset are similar in magnitude to the trends in clear-sky SWdn trends in the RRTMG_SW model at Bondville for a shorter time period. Given the similar magnitude of the observed all-sky and clear-sky SWdn trends and the covariability of AOD with clear-sky SWdn at Bondville, we infer the potential meteorological consequences of the observed AOD trends in the central U.S. with greater confidence. This inference is limited due to the availability of just one surface site that measures direct and diffuse components of all-sky SWdn, clear-sky SWdn, and AOD, but can be used as an example of assessing aerosol-climate interactions in future studies.

1.6 DISCUSSION

Here we assemble the evidence of changing surface climate in the U.S. and consider the possible role of aerosols in driving this change. During the 2000–2014 time period, observed AOD decreases significantly at both Bondville, IL (-0.047) and Good-

win Creek, MS (-0.052) Clear-sky total SWdn increases at these sites by 12.7 Wm^{-2} (Bondville) and 7.8 Wm^{-2} (Goodwin Creek) over the same time period, suggesting that the declining aerosols are at least partly responsible for these trends. All-sky total SWdn also increases at Goodwin Creek. However, the diffuse component of clear-sky SWdn increases at both Bondville ($+6.6 \text{ Wm}^{-2}$) and Goodwin Creek ($+5.2 \text{ Wm}^{-2}$), consistent with previous studies (e.g., Gan et al., 2014) and the cause of these increases remains an open question. We do not find evidence of aircraft contrails influencing observed diffuse SWdn at SURFRAD sites in the aftermath of Sep. 11th, 2001, and thus caution against attributing increasing clear-sky diffuse SWdn trends to increasing U.S. air traffic as has been done in previous studies (e.g., Long et al., 2009; Augustine and Dutton, 2013).

Using the RRTMG_SW radiative transfer model driven by observed AOD, we simulate increases in total and direct clear-sky SWdn at both sites that are consistent with observations and decreases in diffuse SWdn that are contrary to the observations. Previous studies invoked trends in aircraft contrails to explain the unexpected changes in diffuse SWdn at SURFRAD sites across the U.S. (e.g., Gan et al., 2014), but application of observed cirrus cloud fraction to RRTMG_SW does not resolve this issue. A cross-validated multivariate regression analysis further shows that observed monthly mean AOD accounts for 20% of the JJAS variability in clear-sky direct SWdn at Bondville, with cirrus cloud liquid water path and AOD reproducing 26% of the variability in clear-sky diffuse SWdn at this site. No combination of predic-

tors, however, explains the variability of clear-sky direct or diffuse SWdn at Goodwin Creek, casting doubt on the role of aerosol-radiation interactions on local meteorology at this site. Besides AOD and cirrus cloud cover, we are left with few other variables that could influence the direct/diffuse partitioning of clear-sky SWdn, and the cause of the observed increase in SWdn during the 2000-2014 timeframe is not clear. With more sites across the Southeast, we could diagnose these inconsistent results as either site-specific, or symptomatic of the larger region.

Our analysis of the Ameriflux data (1998–2007) and ICN data (1990–2014) suggests that soil moisture declines in response to enhanced solar radiation. In particular, we see possible evidence of a soil moisture feedback at the Bondville Ameriflux station, where the difference in JJAS SWdn between sunny and cloudy summers is nearly 30 Wm^{-2} , peaking in September. A sunny summer reduces soil moisture, especially in August (-3.5%), and enhances sensible heat fluxes by $+8.7 \text{ Wm}^{-2}$, with peak values in September. The SURFRAD data show an all-sky annual SWdn trend of $+0.58 \text{ Wm}^{-2} \text{ a}^{-1}$ at Bondville and a $+1.0 \text{ Wm}^{-2} \text{ a}^{-1}$ at Goodwin Creek for 2000–2014. This rate translates to changes in SWdn of $+8.7 \text{ Wm}^{-2}$ and $+15 \text{ Wm}^{-2}$ over this time period at these sites, or roughly one-fourth to one-half the change in SWdn between sunny and cloudy years at the Ameriflux station. Despite large spatial heterogeneity, the land surface models in Section 1.5 show a reduction of volumetric soil water content of as much as -2.2% over the central U.S. from 1990–2015. Given the observed trends in SWdn, this result is consistent in sign with the 0.14-3.4% decrease in soil moisture

content during sunnier summers at the Ameriflux and ICN sites in Illinois. The observed soil moisture and temperature responses between the sunny and cloudy regimes are of the same order of magnitude as those simulated by Mickley et al. (2012) for aerosol vs. no aerosol regimes over the eastern US, lending confidence to the conclusions of that model study. Our results are also consistent with Eshel (2016), who found that the observed SWdn increase from 1988–2014 at a rural site in the Northeast could be explained in a radiative transfer model only when considering trends in anthropogenic aerosols.

Our Ameriflux, ICN, and NLDAS results show the climate response to increasing surface SWdn. The RRTMG_SW results in Bondville show that changing aerosols influence SWdn trends. Taken together, the observations and modeled results suggest that aerosol-radiation interactions are significant at Bondville, a conclusion doubted in previous studies (Long et al., 2009; Augustine and Dutton, 2013). We find evidence that these interactions play a role in the observed climate trends at Bondville, similar in response to modelling studies that probed the influence of aerosols on the warming hole (Leibensperger et al, 2012b). Changes in overhead aerosol contribute about one-fourth of the interannual variability in direct and diffuse clear-sky SWdn. If the aerosol trends at Bondville are representative of the larger region, the recent decline in AOD may partly account for the 1.8 K increase in surface temperatures across the larger region for JJAS in the NLDAS dataset. In contrast, aerosol-radiation interactions do not appear to contribute to the interannual variability in SWdn at Goodwin

Creek, casting doubt on the role of the direct aerosol effect in the reversal of the U.S. warming hole in the Southeast.

Our results underscore the difficulty in attributing the warming hole to only aerosols or only aerosol-radiation interactions using observations. Even though we find evidence of aerosol-radiation interactions at Bondville, our result at Goodwin Creek contrasts with Tosca et al. (2017), who concluded that the observed increase in surface SWdn between 2007 and 2017 at this site was a result of aerosol-radiation interactions. Though we agree with these authors on the sign of the observed surface SWdn trend and find that RRTMG_SW can indeed reproduce a positive clear-sky SWdn trend when driven by aerosols (Figure 1.2), we do not find evidence of interannual covariability between AOD and clear-sky SWdn at Goodwin Creek. Hence we disagree that the SURFRAD observations point to aerosol-radiation interactions at this site, as we believe that evidence of covariability between AOD and SWdn on interannual timescales is a necessary condition in asserting aerosol-radiation interactions. Yu et al., (2014), however, found evidence of aerosol-cloud interactions in the southeastern U.S., so such interactions could potentially be important in that region.

Agriculture is a major industry in the central U.S., and decreases in soil moisture from increased SWdn may have made the region more vulnerable to drought, as suggested by previous model studies (Mickley et al., 2012; Leibensperger et al, 2012b). From our analysis of tower fluxes and the NLDAS assimilation, we find a consistent land-atmosphere response to SWdn as seen in these model studies. Specifically, soil

moisture responds to local enhancements in SWdn that further amplifies SWdn, especially in late summer. Previous studies have diagnosed the strong influence of tropical Pacific SSTs on drought occurrence in the U.S. (e.g., Schubert et al., 2004; Seager and Hoerling, 2014). However, drought models that rely on Pacific SSTs predict a prolonged drought during the 1970s, a period in reality characterized by increased precipitation, especially in the central and eastern United States (e.g., Seager and Hoerling, 2014). We speculate that high loading of anthropogenic aerosol during the 1970s may have led to more moist conditions in the central U.S., countering the SST influence and reducing drought risk. While the model results of Leibensperger et al. (2012a,b) are consistent with this hypothesis, more rigorous model studies with state-of-the-science hydrology and chemistry are needed to confirm it. We underscore the findings of other studies (e.g., Milly and Dunne, 2011) that caution modeling studies against projecting hydrological change in models without finding consistency with surface energy balance changes.

A drawback of this study is that it relies on relatively short-term records of aerosols and surface SWdn. There is also some uncertainty in the SURFRAD measurements, at least when compared to the derived trends per year of SWdn. In trends reported here, however, we find that the standard deviation of the residual noise is greater than the instrument uncertainty. Finally, our study relies on just a few measurement sites to infer relationships of AOD with other variables across a broad region. We find evidence of aerosol-radiation interactions at Bondville but not Goodwin Creek.

Though we use site-specific data to come to these conclusions, the good match between site measurements and assimilated NLDAS SWdn data, however, lends confidence that increased surface SWdn has indeed occurred over a broad region. As more *in situ* measurements of SWdn and AOD are recorded and as USCRN and other national networks are expanded, we expect the discrepancies between SURFRAD sites will be better explained.

This study provides observational evidence of the influence of AOD on SWdn and key variables such as soil moisture. By linking trends in AOD, SWdn, and soil moisture, our results point to the importance of considering atmospheric composition as an additional driver of drought. Currently, many regions of the developing world (e.g., China and India) have much higher aerosol loading than the U.S. and are planning strategies to reduce aerosol sources and improve air quality (Lu et al., 2011). These regions also depend on favorable meteorological conditions for agricultural production to feed growing populations. Our study suggests there may be inadvertent consequences of aerosol reduction on regional climate.

Chapter 2

Quantifying the influence of agricultural fires in northwest India on urban air pollution in Delhi, India

2.1 INTRODUCTION

Residents of the heavily populated Indo-Gangetic Plain (IGP) in India experience elevated health risks due to poor air quality. The National Capital Territory of Delhi (hereafter referred to as Delhi) sits within the IGP and has a population of ~ 16.5 million. The larger National Capital Region of Delhi which is centered on Delhi but also includes regions of Haryana, Uttar Pradesh, and Rajasthan is estimated to exceed a population of 46 million (Registrar General, India, 2011). Daily mean levels of surface particulate matter ($\text{PM}_{2.5}$) pollution in Delhi often exceed the World Health Organization threshold for unhealthy air (24-hour average of $25 \mu\text{g m}^{-3}$) as well as the daily mean threshold set by the Indian Central Pollution Control Board (CPCB, $60 \mu\text{g}$

m^{-3}). Delhi is often in exceedance of these standards during the post-monsoon season (Oct.-Nov.), and its ambient $\text{PM}_{2.5}$ concentrations are subject to large episodic spikes. For example, from Nov. 1st - Nov. 15th, 2016, the $\text{PM}_{2.5}$ concentration at the Mandir Marg CPCB site averaged $350 \mu\text{g m}^{-3}$, but reached as high as a daily-average of $692 \mu\text{g m}^{-3}$ on Nov. 5th (<http://www.cpcb.gov.in/CAAQM/>). One major uncertainty is the extent to which smoke emissions from post-monsoon agricultural fires in rural areas influence the already high concentrations of urban air pollution in the IGP. This study aims to quantify the magnitude of the contribution of these fire emissions to $\text{PM}_{2.5}$ pollution in Delhi during the post-monsoon burning season over the 2012-2016 time frame. The attribution of surface $\text{PM}_{2.5}$ due to fires versus other anthropogenic sources is critical in developing strategies to reduce overall pollution exposure.

Exceedances of $\text{PM}_{2.5}$ standards in Delhi occur year-round, with an annual mean $\text{PM}_{2.5}$ concentration of more than $100 \mu\text{g m}^{-3}$ (e.g., Twiari et al., 2013). Much of the pollution comes from coal-fired power plants, transportation, and domestic combustion sources (Guttikunda and Jawahar, 2014; Gurjar et al., 2016). $\text{PM}_{2.5}$ and its precursors (e.g., SO_2 , NO_x) have led to an estimated 30% of Delhi's population suffering from respiratory disorders (Kandlikar and Ramachadran, 2000). Nagpure et al. (2014) estimated a 60% increase in Delhi mortality due to the degradation of air quality between 2000 and 2010. Residents of Delhi have been found to suffer from diseases related to air pollution at a rate 12 times higher than the national average (Kandlikar and Ramachadran, 2000). More broadly, Dey et al. (2012) estimated that about

half the subcontinent’s population experienced pollution levels above the WHO PM_{2.5} annual mean standard of 35 $\mu\text{g m}^{-3}$ during 2000-2015, with the greatest pollution exposure in the IGP.

India’s agricultural “breadbasket” is located in the northwestern-most region of the country, mostly in the state of Punjab but also in the neighboring state of Haryana. Agriculture in these states is typically characterized by two growing seasons: a predominantly winter wheat crop, harvested in April-May, and a predominantly summer rice crop, harvested in October-November (Vadrevu et al., 2011). Increasing utilization of mechanized harvesters over the last 30 years has decreased costs and improved efficiency for farmers, and studies have found that more than 75% of rice is harvested using a combine harvester in Punjab (Kumar et al., 2015). However, this harvesting method leaves more crop residue on the fields than traditional methods using a sickle, and many farmers burn this residue to ready fields for the next growing season (Kaskaoutis et al., 2014). Smoke from these fires consists of black carbon and organic particulate matter. The post-monsoon rice harvest season coincides with post-monsoon conditions that favor stagnation and weak surface northwesterly winds in the IGP (Singh and Kaskaoutis, 2014). These conditions allow smoke to slowly permeate throughout the IGP, including Delhi, about 350 km downwind from Punjab.

Previous work has diagnosed co-variability between fire emissions in Punjab and observed urban pollution levels in the region and downwind. For example, using ground-based sensors in the Punjab city of Patiala, Mittal et al. (2009) reported PM_{2.5} en-

hancements as high as $547 \mu\text{g m}^{-3}$ during the 2007 burning season of October-November. Using remotely sensed data from the Moderate Resolution Imaging Spectroradiometer (MODIS), Mishra and Shibata (2012) found enhancements of 0.1-0.3 in 850-nm aerosol optical depth (AOD) during the 2009 post-monsoon burning season over the IGP. These authors used lidar to probe the vertical distribution of the aerosols over Punjab and Delhi and found the largest 532-nm backscatter coefficients ($>0.0035 \text{ km}^{-1} \text{ sr}^{-1}$) close to the surface (below 1.0 km altitude). Consistent with this study, Kaskaoutis et al. (2014) found daily maximum MODIS 550-nm AOD to often be in excess of 2.0 during the 2012 post-monsoon burning season. Observations from two Aerosol Robotic Network (AERONET) sites in the IGP show that aerosols tend towards larger volume and smaller particle size during the post-monsoon burning season (Kaskaoutis et al., 2014); such attributes are characteristic of fresh soot. Our previous work (Liu et al., 2018) used back trajectory analysis to define an airshed region upwind of Delhi during both pre-monsoon (Apr - May) and post-monsoon burning seasons. The study focused on relating available data on PM_{10} and other air quality measurements to fire radiative power (FRP) in the airshed for both burning seasons, accounting for meteorological conditions. We found that post-monsoon MODIS FRP within the airshed correlates with observed concentrations of surface PM_{10} , visibility, and AOD in Delhi, suggesting a coupling between upwind fires, meteorology, and urban pollution.

Missing from recent studies is an estimate of the magnitude of surface $\text{PM}_{2.5}$ in

Delhi that can be attributed to agricultural fire emissions. Building on the work of Liu et al. (2018) and other studies, this study aims to address this gap by combining analysis of surface $\text{PM}_{2.5}$ observations in Delhi with chemical transport modeling. We first use daily 2012-2016 surface observations to estimate the seasonal $\text{PM}_{2.5}$ enhancement due to regional (i.e., fire) sources. We compare these enhancements to model-derived $\text{PM}_{2.5}$ driven by a suite of fire emission inventories. We rely mostly on a Lagrangian-based modeling framework, which uses back trajectories to simulate chemical transport and map the sensitivity of $\text{PM}_{2.5}$ levels in Delhi to fire activity upwind. We find that our model can capture much of the weekly observed $\text{PM}_{2.5}$ variability in Delhi, as well as at least some of the extreme peaks in daily $\text{PM}_{2.5}$ during the post-monsoon burning season. We further fine-tune these simulated $\text{PM}_{2.5}$ estimates with a statistical model fit with local meteorology. Discrepancies between the model and observed $\text{PM}_{2.5}$ in Delhi point to the difficulty in detecting small fires from satellite, especially when clouds and/or smoke interfere with detection. Smoke from satellite-detected fires that are detected can contribute more than half the total observed $\text{PM}_{2.5}$ across Delhi during the post-monsoon burning season.

2.2 DATA AND METHODS

2.2.1 SURFACE AND SATELLITE OBSERVATIONS

The CPCB provides online hourly observations of a variety of pollutants including $\text{PM}_{2.5}$ at 12 sites within Delhi (<http://www.cpcb.gov.in/CAAQM>). We focus on observed $\text{PM}_{2.5}$ during the post-monsoon burning season (here defined as Oct. 17 - Nov. 30) during 2012-2016. We find that at least 90% of Oct. - Nov. FRP over the north-western IGP during 2012-2016 is detected during this time window. No CPCB site provides a complete record of $\text{PM}_{2.5}$ observations during the entire course of 2012-2016. The U.S. Embassy in Delhi (<https://in.usembassy.gov/embassy-consulates/new-delhi/air-quality-data/>) also provides daily $\text{PM}_{2.5}$ from 2013-2016, and is mostly complete during that time span. Finally, we rely on observations from a new monitoring network, #Breathe (<http://api.indiaspend.org/dashboard/>), launched in 2016 by IndiaSpend, a grassroots initiative to monitor air quality at 10 sites in Delhi and elsewhere in India. Figure 2.1 shows the spatial configuration of all surface sites where $\text{PM}_{2.5}$ was available sometime during 2012-2016. We aggregate and validate these surface observations with satellite AOD (described in Section 2.3.1) retrieved from the MODIS Level 3 Aqua Deep Blue algorithm (MYD08D3; Hsu et al., 2013). The Deep Blue algorithm is designed to provide AOD retrievals over bright surfaces, and was found to correlate well with the AERONET station in Kanpur, India ($0.70 \leq R \leq 0.86$; Sayer et al., 2013).

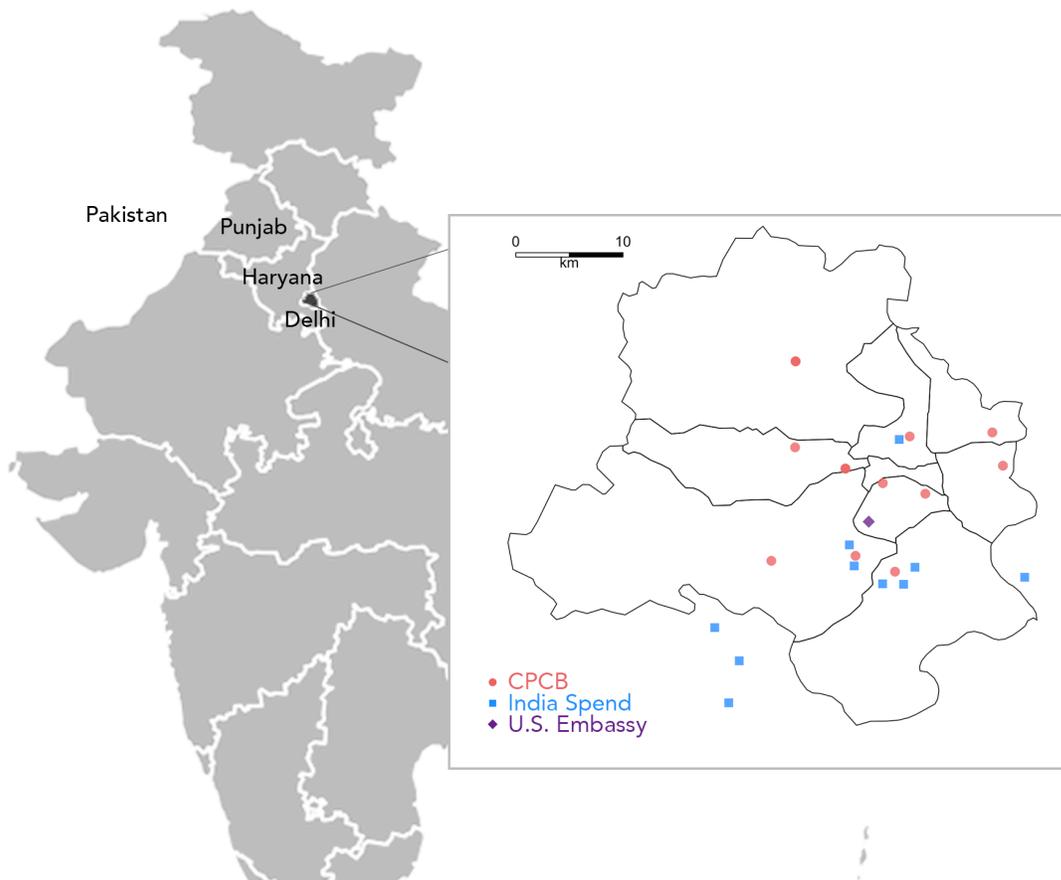


Figure 2.1: Distribution of Central Pollution Control Board (CPCB), India Spend, and U.S. Embassy $PM_{2.5}$ monitoring sites located in and around the Delhi. Solid lines represent districts within the National Capital Territory of Delhi.

2.2.2 FIRE EMISSION INVENTORIES

In situ information that can be used to quantify regional fire emissions on the daily scale in Punjab and Haryana is limited. Thus, we consider top-down fire emission inventories that are based on satellite information. As we shall see, many assumptions are needed to translate the satellite retrievals into emissions. The inventories considered in this study are the Fire Inventory from NCAR (FINN; Wiedinmyer et al., 2011), the Global Fire Emissions Database version 4 with small fires (GFED4.1s; van der Werf et al., 2017, Giglio et al., 2013; Randerson et al., 2012), the Global Fire Assimilation System (GFAS; Kaiser et al., 2012), and The Quick Fire Emissions Dataset (QFED; Darmenov and da Silva, 2015). Each of these fire emission inventories are based in part on thermal anomalies detected by MODIS (Giglio et al., 2006). However, they each differ in their treatment of emission factors and land cover that translate these thermal anomalies into emission estimates, and they also have different methods for treating gaps in the MODIS record. FINN aggregates 1-km MODIS active fire detections and produces a daily emission estimate given these detections and MODIS retrieved land cover. GFED4.1s relies primarily on monthly MODIS MCD64A1 500-m burned area maps, derived from observed changes in surface reflectance, to generate emission estimates. The dataset then adds 1-km active fire information to incorporate the influence of small fires that may have not been accounted for by the burned area product (Randerson et al., 2012). GFED4.1s derives a daily

emission estimate by applying the ratio of daily to total monthly active fire counts in each grid cell. GFAS also aggregates the 1-km MODIS FRP and uses emission factors similar to those in GFED4.1s. To account for FRP obscured by sub-grid clouds and other interferences, GFAS assumes that an obscured FRP pixel is equivalent in value to its adjacent non-obscured pixel, as long as they are not over a body of water (Kaiser et al., 2012). GFAS further uses a Kalman filtering method of data assimilation, in which the optimal estimate of FRP for a given day is a weighted average of the optimal FRP estimate from the previous day and the FRP estimate for the current day (Kaiser et al., 2012). Like GFAS, QFED uses information from adjacent pixels to estimate obscured thermal anomalies; QFED also relies on FRP estimates from the previous day. However, the QFED algorithm weights adjacent pixel information via the error covariance between pixels, and it allows the estimate of the previous day’s FRP to decay according to a characteristic timescale derived for each land type (Darmenov and da Silva, 2015). GFAS, QFED, and FINN are available in near real time, whereas GFED4.1s requires several months of processing before public release. We include another inventory, here called GFED + Agriculture. In this inventory, we apply the GFED4.1s dry matter emission estimates, but assume the land cover within each burning grid cell is 100% agricultural, thus only using emission factors associated with agricultural burning. We also increase the GFED4.1s emission factors associated with agricultural burning by a factor of three. The factor of three scaling is based on the laboratory findings of Oanh et al. (2010), who found that the particulate

matter emissions of rice straw approximately tripled when the straw was piled instead of spread evenly on the ground. In reality, we expect a mixture of partial and whole field burning (Vadrevu et al., 2008), so the GFED + Agriculture emissions represent a high derived upper bound on agricultural burning in the region.

2.2.3 PARTICLE DISPERSION, CHEMICAL TRANSPORT, AND STATISTICAL MODELING

We perform 2012-2016 simulations of daily surface $PM_{2.5}$ in Delhi using the Stochastic Time-Inverted Lagrangian Transport (STILT) model (Lin et al., 2003), driven by $0.5^\circ \times 0.5^\circ$ Global Data Assimilation meteorology (GDAS; Houser et al., 2001). STILT is a receptor-oriented Lagrangian particle dispersion model. This modeling framework has previously been used to assess the influence of wildfires on urban air pollution in Salt Lake City, Utah (Mallia et al., 2015). By tracing an ensemble of theoretical particles or air-mass trajectories from a receptor site backwards in time, STILT computes the sensitivity of $PM_{2.5}$ concentration at the receptor to emissions in the surrounding region. The resulting flux footprint reveals those regions where emissions likely influenced $PM_{2.5}$ at the receptor. We simulate daily footprints of the sensitivities of Delhi pollution to fire emissions upwind by sending 500 simultaneous air-mass trajectories backwards in time for 5 days. We choose 500 ensembles in order to account for random turbulence air-masses experience, especially in the boundary layer (Lin et al., 2013). We choose five days as this timeframe should allow an air mass to traverse

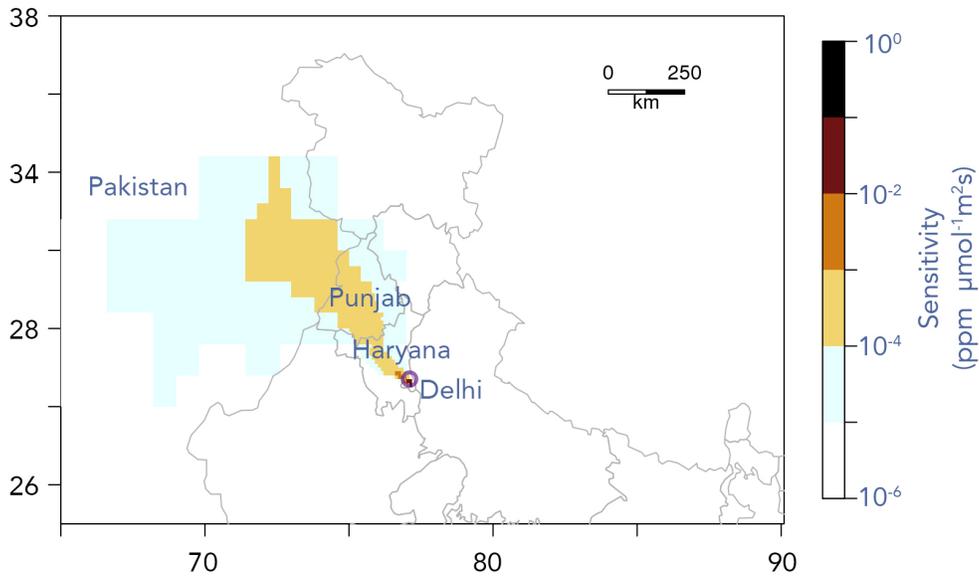


Figure 2.2: Median 2012-2016 STILT sensitivity of Delhi (28.62 °N, 77.21 °E, purple circle) $\text{PM}_{2.5}$ observations to surrounding fire emissions during the post-monsoon burning season (Oct. 17 - Nov 30). Sensitivities below $10^{-6} \text{ ppm } \mu\text{mol}^{-1} \text{ m}^2 \text{ s}$ are not shown.

the approximately 800 km between Delhi and the farthest upwind burning regions even under the weak wind conditions prevalent at this time of year, which is often less than 5 m s^{-1} according to GDAS. Figure 2.2 shows the spatial footprint of the median 2012-2016 sensitivities of a Delhi receptor (28.62°N, 77.21°E) to the surrounding emissions during the burning season. We see that Delhi is highly sensitive ($10^{-3} \text{ ppm } \mu\text{mol}^{-1} \text{ m}^2 \text{ s}$) to the upwind burning regions in Punjab. Similar to Koplitz et al. (2016), we assume that the $\text{PM}_{2.5}$ reaching Delhi from upwind fires is in its primary BC or OC form.

Liu et al. (2018) performed a back-trajectory analysis using the Hybrid Single Particle Lagrangian Integrated Trajectory Model (HYSPLIT; Stein et al., 2015) to create an airshed region upwind of Delhi. The boundaries of this airshed determined a region where agricultural fire emissions could potentially influence downwind air pollution in Delhi. Through the use of STILT, we make this relationship more explicit by quantifying explicitly how much those upwind emissions contribute to a particular downwind pollution observation. In other words, each STILT footprint can be coupled to an emissions inventory in order to simulate surface $\text{PM}_{2.5}$ concentrations. The footprint for the i th receptor location and time can be expressed as a vector $\mathbf{k}_i = (\partial y_{stilt,i} / \partial \mathbf{x})^T$, where \mathbf{x} is a vector of upwind emissions from the previous 5 days (units of $\mu\text{mol m}^{-2} \text{s}^{-1}$), and $y_{stilt,i}$ is the modeled $\text{PM}_{2.5}$ enhancement due to those emissions. If we couple this footprint to an emissions estimate (e.g., FINN, QFED, etc.) from the previous 5 days, we can simulate surface $\text{PM}_{2.5}$ enhancement from fires using the relation $y_{stilt,i} = \mathbf{k}_i \cdot \mathbf{x}$. These simulated surface concentrations can then be compared to the observed network average of $\text{PM}_{2.5}$ observations.

We simulate the urban fate of primary $\text{PM}_{2.5}$ from fires and assume no chemistry using STILT. To account for additional $\text{PM}_{2.5}$ production from other anthropogenic sources, we determine a background or baseline from observations (described further in Section 2.3.2). We compare this baseline to a simulated anthropogenic $\text{PM}_{2.5}$ from the 3-D global chemical transport model, GEOS-Chem (geos-chem.org). GEOS-Chem is here driven by assimilated meteorological data from the Goddard Earth Observ-

ing System (GEOS-5) at the NASA Global Modeling Assimilation Office (GMAO). The aerosol simulation in GEOS-Chem includes sulfate, nitrate, ammonium, dust, and black and organic carbon (Kim et al., 2015), and many previous studies have examined PM_{2.5} pollution in Asia using GEOS-Chem (e.g., Wang et al., 2013; Mu and Liao, 2014; Geng et al., 2015). Here we utilize the emission inventory for Model Inter-Comparison Study for Asia (MIX) for anthropogenic aerosol precursor emissions (Li et al., 2015). We follow Bond et al. (2007) for anthropogenic emissions of primary black and organic carbon. For this study, we perform nested grid simulations for the 2012 burning season at 0.50° x 0.67° resolution over most of eastern Asia, with lateral boundary conditions provided by a global simulation at 2.0° x 2.5° horizontal resolution.

2.2.4 STATISTICAL MODELING

We tune the STILT simulation of PM_{2.5} for a certain receptor ($y_{stilt,i}$) using a statistical model that relies on local variables that may not be well captured in the 0.5° re-analysis, e.g., local precipitation, mixing layer height, and wind speed. Our statistical prediction of surface PM_{2.5} from fires takes the form:

$$y_{stat,i} = \mathbf{h}_i \cdot \mathbf{w} \tag{2.1}$$

where \mathbf{h}_i is a $1 \times d$ vector consisting of meteorological parameters and the STILT-

driven $\text{PM}_{2.5}$ prediction ($y_{\text{still},i}$), and \mathbf{w} is a $d \times 1$ vector of coefficient weights that represent the relative importance of each predictor in hi to the prediction of $\text{PM}_{2.5}$. The optimal value of these weights is solved for empirically. For example, if we aggregate all daily observed network-averaged surface observations above the anthropogenic baseline (\mathbf{y}_{obs}), the traditional ordinary least square setup determines the optimal value of coefficient weights (\mathbf{w}) by the following relation:

$$\mathbf{w}^* = (\mathbf{H}^T \mathbf{H})^{-1} \mathbf{H}^T \mathbf{y}_{\text{obs}} \quad (2.2)$$

where \mathbf{H} is an $n \times d$ matrix and n is the number of observations. Each column of \mathbf{H} represents the time series of daily mean values of a particular predictor. To avoid overfitting in solving for \mathbf{w}^* , we follow the method of the least absolute shrinkage and selection operator (LASSO; Tibshirani, 1996), which reduces the magnitude of the coefficients of correlated predictors and those predictors offering little information. Here the optimal coefficients are determined through the following algorithm:

$$\mathbf{w}^* = \min_{\mathbf{w} \in \mathbb{R}^d} \{ \|\mathbf{H}\mathbf{w} - \mathbf{y}_{\text{obs}}\|_2^2 + \lambda \|\mathbf{w}\|_1 \} \quad (2.3)$$

In the above equation, the first term on the right-hand side of the equation penalizes mismatch between model and observations using the square loss function, hence the “2” subscript. The second term of the equation regularizes the fit (i.e., reduces overfitting) by penalizing the magnitude of \mathbf{w} via the absolute loss, also known as the

L1 norm, hence the “1” subscript. The algorithm is optimized over a grid of λ values (to control the degree of regularization), using three-fold cross validation. This method randomly separates the data into three sets and fits the statistical model on two of these sets, and then these fitted coefficients are applied to the remaining set (called the validation set), yielding the root mean squared error (RMSE) between the prediction (\mathbf{y}_{stat}) for that set and the observations (\mathbf{y}_{obs}). This process is repeated three times, and the value of λ that provides the best RMSE on the reserved validation sets is retained. In addition to $\mathbf{y}_{\text{stilt}}$, the array of local meteorological variables at Delhi used as predictors include wind speed and wind direction from the surface to the boundary layer and from the boundary layer to 500 hPa, as well as boundary layer height, precipitation, surface temperature, and surface pressure. All variables are taken from the Integrated Global Radiosonde Archive (Durre et al., 2006) and the Global Historical Climatology Network (Menne and Williams, 2009). IGRA estimates boundary layer heights over the Safdarjung airport (28.58°N, 77.2°E) using the parcel method, which locates the altitude where virtual potential temperature is equivalent to surface virtual potential temperature (Seibert et al., 2000).

2.3 RESULTS

2.3.1 CREATING A NETWORK-AVERAGE AND ANTHROPOGENIC BASELINE OF $\text{PM}_{2.5}$

Due to data inconsistencies among the CPCB sites, we employ data quality preprocessing before calculating a city-wide network average of urban $\text{PM}_{2.5}$ for Delhi. Figure 2.3 shows the number of daily averaged $\text{PM}_{2.5}$ observations available at each site during the burning season for each year. As previously noted, the U.S. embassy data are available only for 2013-2016, and IndiaSpend data only for 2016. Both data sources provide near complete measurement records over these time periods. Few CPCB sites have a record of observations of more than three years during 2012-2016. To represent mean pollution exposure across the city through the years, we implement a two-step data-cleaning procedure. First, we compare daily averaged CPCB $\text{PM}_{2.5}$ with corresponding MODIS AOD for each site during the burning season, using all available observations during 2012-2016. Then we select only those sites whose correlation with the AOD timeseries exceeds $R=0.5$ and is statistically significant ($p < 0.05$). The purpose of this step is to consider only those sites whose variability corresponds to regionally influenced pollution. Hence, if a surface site's daily averaged $\text{PM}_{2.5}$ correlates reasonably with the coarser 1° MODIS Deep Blue AOD retrieval, we assume that site to be sensitive to pollution from regional sources. Next, for each CPCB site that meets this correlation criterion, we calculate the mean absolute difference of daily-averaged $\text{PM}_{2.5}$ at that site compared to the network average of daily-

averaged $\text{PM}_{2.5}$ at the other sites in the network. We also compute the standard deviation of that difference across the five years. This step produces a metric revealing how much daily $\text{PM}_{2.5}$ at each site tends to deviate from $\text{PM}_{2.5}$ at the other sites on average. We find that these deviations are distributed normally, so that for any given day, if the absolute difference in $\text{PM}_{2.5}$ at a particular site deviates more than ± 2.5 standard deviations from the mean absolute difference associated with that site, we exclude that $\text{PM}_{2.5}$ observation from the network average for that day. Thus for any given day, we remove from consideration those sites that either experience instrument malfunction and/or appear to be heavily influenced by strong local sources. For each station, we could alternatively consider removing observations that deviate too much from that station's mean $\text{PM}_{2.5}$ concentration during the post-monsoon burning season. However, we did not follow this approach because surface $\text{PM}_{2.5}$ varies widely during large fire episodes. For example, the Mandir Marg CPCB site recorded daily-averaged surface $\text{PM}_{2.5}$ concentrations ranging between 120 - 692 $\mu\text{g m}^{-3}$ during Nov. 1-15, 2016. The higher $\text{PM}_{2.5}$ enhancements could erroneously be marked as outliers from local sources and/or instrument malfunction, when in fact all sites experienced large fluctuations in $\text{PM}_{2.5}$ during this time. Therefore, we wish to consider outliers as a function of the network average of monitors, as we expect all surface monitors to jointly respond to the regional signal of fire emissions and transport. In 2016, we have data from CPCB, U.S. Embassy, and India Spend $\text{PM}_{2.5}$ observations. We compare each data source (Figure 2.4) and find close correlation between datasets ($R =$

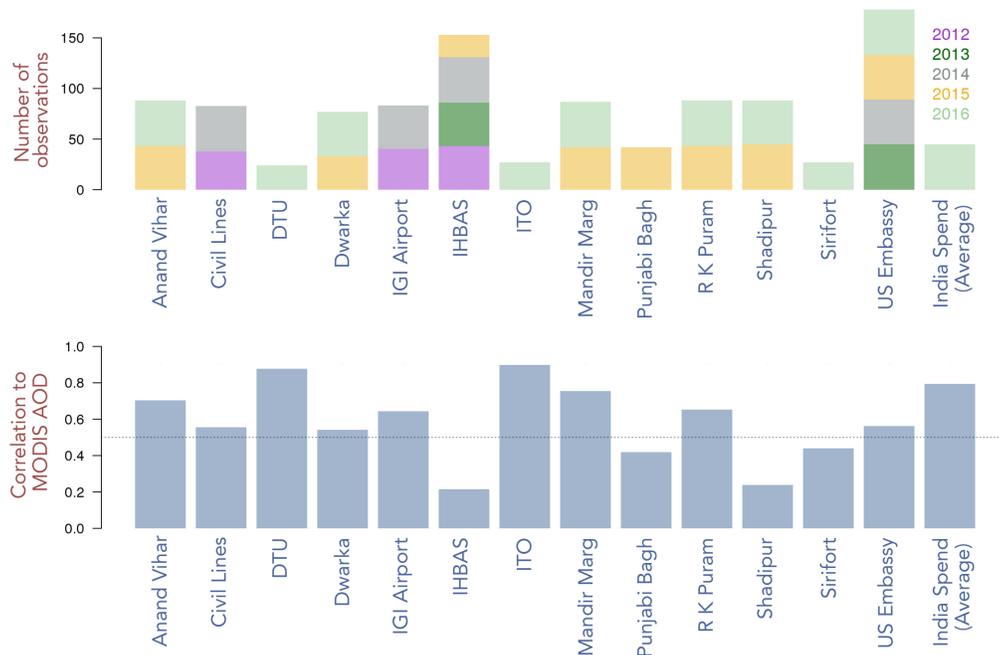


Figure 2.3: (Top) Number of daily-averaged $PM_{2.5}$ observations available at the Central Pollution Control Board (CPCB), U.S. Embassy, and India Spend sites during the post-monsoon burning season (Oct 17 - Nov 30) for each year during 2012-2016. (Bottom) Correlations R between observed $PM_{2.5}$ and satellite aerosol optical depth (AOD) over Delhi. The horizontal line at $R = 0.5$ corresponds to the threshold used to determine if a site is included in the $PM_{2.5}$ network average. All correlations above $R = 0.5$ are statistically significant ($p < 0.05$).

0.91-0.92).

2.3.2 BASELINE $PM_{2.5}$ DERIVATION DURING FIRE SEASON

We next determine a $PM_{2.5}$ baseline in Delhi to represent typical pollution levels in the absence of smoke from agricultural fires. This baseline represents the mean anthropogenic contribution to total $PM_{2.5}$ during the post-monsoon burning season.

Quantification of this baseline is important as we use it to derive a $PM_{2.5}$ enhance-

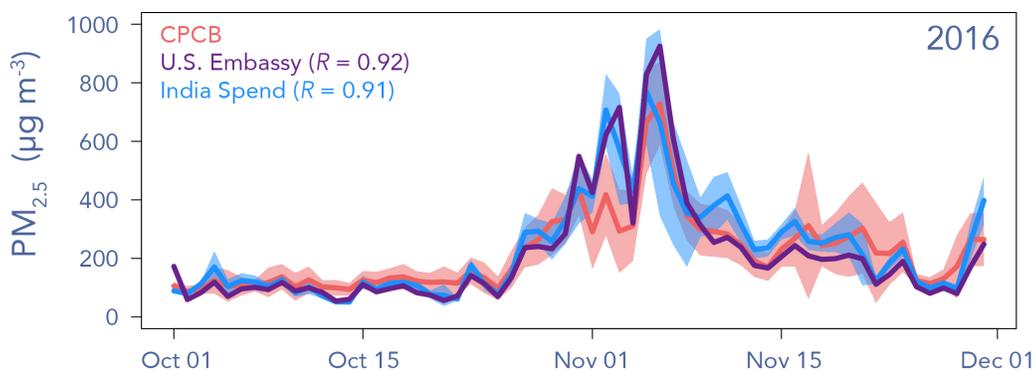


Figure 2.4: Daily-averaged $\text{PM}_{2.5}$ observations for CPCB (pink), U.S. Embassy (purple), and India Spend (blue) for 2016. Shading for the CPCB and India Spend curves represent one standard deviation from the network-averaged $\text{PM}_{2.5}$. The U.S. Embassy observations correlate well with the CPCB network average ($R = 0.92$), and the India Spend observations correlate well with the CPCB network ($R = 0.91$).

ment from observations ($\mathbf{y}_{\text{obs}} = \text{total observed } \text{PM}_{2.5} - \text{baseline}$). Baseline anthropogenic $\text{PM}_{2.5}$ in post-monsoon months consists of elemental carbon, organic matter, and secondary sulfate-nitrate-ammonium from gasoline exhaust, coal combustion, dust, and urban biomass combustion (Pant et al., 2015). For simplicity, we assume that baseline levels are constant during a given burning season. However, we anticipate that baseline $\text{PM}_{2.5}$ likely changes over the years due to changes in the surface monitoring network and local emission sources. For these reasons, we compute a unique baseline $\text{PM}_{2.5}$ for each year during 2012-2016. We apply three different methods with different assumptions in order to test the robustness of our baseline estimates.

- **Method 1:** This method relies on the daily variability of the GFAS fire emissions. We choose GFAS due to its assimilation properties which account for some missing or obscured fire pixels. For each fire season, we analyze the time series of these emissions summed over all grid cells in the burning regions upwind of Delhi. We specify low-fire days as those days when total fire emissions fall below a specified threshold at the low end of the frequency distribution for that season - e.g., below the 10th percentile. On days when fire emissions fall below that threshold, we assume that Punjab and Haryana are not burning significantly. If emissions remain below the threshold during the next N days, we tag the observation for that N th day as representative of the baseline. The baseline is then the average of all tagged days during the fire season. We vary N between 1-5 and the emission percentile threshold between 10-30% to check the robustness of our baseline estimate. We assume that N represents the transport time for smoke from fires to ventilate out of the IGP.
- **Method 2:** In this method, we take advantage of STILT sensitivity estimates. For each day of the fire season, STILT provides gridded sensitivities to upwind emissions for each observation in Delhi. If the map of sensitivity overlaps with cells containing fire emissions, the model predicts a pollution enhancement due to fire downwind in Delhi in the subsequent days. To compute the anthropogenic baseline, we count the number of fire emitting pixels for a particular day. We then count the number of those fire pixels that overlap with the STILT sensitivity map. If the ratio of overlapping pixels to total fire pixels is sufficiently low (e.g., less than a threshold of 0.1) on a given day, we assume that the urban pollution for that day has little influence from fires. We collect each of these non-fire days during each fire season and take their average as the baseline. We vary the ratio threshold between 0.1-0.7 to assess the sensitivity of this method to its underlying assumptions.
- **Method 3:** For this method, we compute the weekly block average of $\text{PM}_{2.5}$ within the city for each week (Sunday through Saturday) of the burning season. We then average the M lowest weekly averages to determine the baseline. We vary M between 1-4 to check the sensitivity of this method to this parameter.

Figure 2.5 shows the interannual variability in baseline estimates of urban pollu-

tion in Delhi for 2012-2016. Depending on the year and method chosen, the baseline can vary from 130 - 290 $\mu\text{g m}^{-3}$. The Method 3 baseline is consistently lower than the other baselines, however each baseline estimate is at least twice the CPCB daily air quality standard of 60 $\mu\text{g m}^{-3}$. Method 3 shows the greatest interannual stability, and predicts an average baseline across 2012-2016 of about 150 $\mu\text{g m}^{-3}$, which is within the annual average range of $122.3 \pm 90.7 \mu\text{g m}^{-3}$ total $\text{PM}_{2.5}$ reported by Twiari et al. (2013) for Delhi in 2011. The mean network averaged $\text{PM}_{2.5}$ during the month prior to the post-monsoon burning season (here Sep. 17 - Oct. 16) ranges from 90-150 $\mu\text{g m}^{-3}$ during 2012-2016, which is slightly lower but near the Method 3 baseline estimate. We compare these baseline estimates of Delhi $\text{PM}_{2.5}$ to that provided by GEOS-Chem. For this comparison, we perform the GEOS-Chem simulation with emissions from agricultural fires turned off. Figure 2.6 shows the resulting distribution of daily average urban $\text{PM}_{2.5}$ during the burning season of 2012. The distribution is centered on a mean of 99 $\mu\text{g m}^{-3}$, but is slightly skewed towards larger $\text{PM}_{2.5}$ values, with a maximum at 200 $\mu\text{g m}^{-3}$. Our observation-driven method for determining the 2012 $\text{PM}_{2.5}$ baseline yields values ranging from $147 \pm 47.9 \mu\text{g m}^{-3}$ to $287 \pm 21.9 \mu\text{g m}^{-3}$ (Figure 2.5), or about 1.5-3 times the mean GEOS-Chem simulated baseline.

2.3.3 VARIABILITY OF SURFACE $\text{PM}_{2.5}$

We first probe how well the STILT modeling framework reproduces the variability of $\text{PM}_{2.5}$ in Delhi during the burning season. Our approach is to couple daily STILT

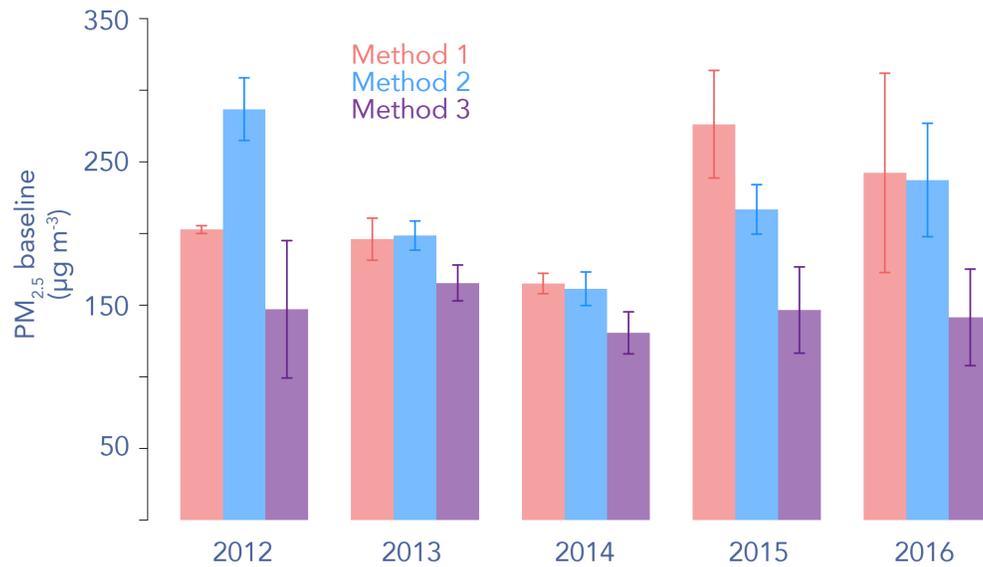


Figure 2.5: Estimates of the anthropogenic PM_{2.5} background in Delhi during the burning season (Oct. 17 - Nov. 30). Method 1 determines the baseline by averaging all observations on the last day of N days of no fires in the Punjab. Method 2 compares overlapping fire and STILT sensitivity grid cells, and determines a baseline if little or no overlap is detected. Method 3 averages the lowest M weekly average PM_{2.5} observations. Error bars represent 1 standard deviation when baseline parameters (e.g., N , M) are varied, as described in the text.

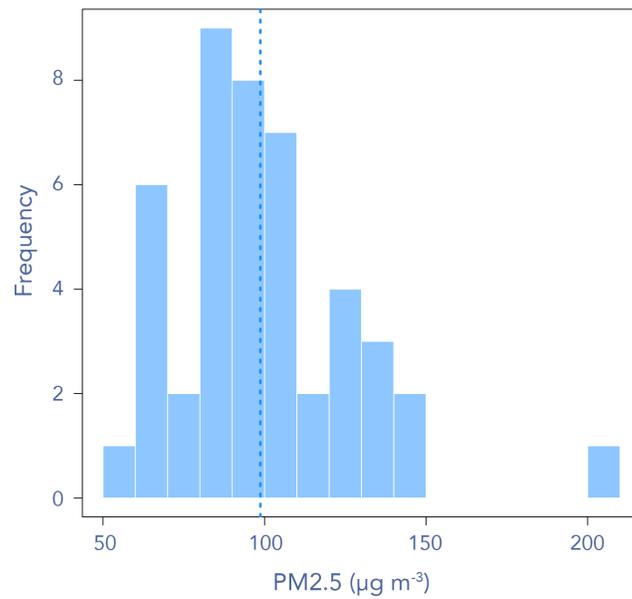


Figure 2.6: Frequency of PM_{2.5} daily observations in Delhi derived from a GEOS-Chem simulation performed at $0.5^\circ \times 0.667^\circ$ horizontal resolution during the burning season of 2012 (Oct 17. - Nov. 30). The dashed vertical line represents the mean of the distribution.

Table 2.1: Correlation and root mean squared error (RMSE) between modeled and observed $\text{PM}_{2.5}$ enhancements in Delhi for 2012-2016. Ranges are determined by the method (1-3) used to determine the anthropogenic baseline (see Section 2.3.2).

Model	STILT ¹		STILT+LASSO ²	
	Correlation	RMSE	Correlation	RMSE
GFED	0.43-0.50	80-109	0.72-0.78	53-62
QFED	0.41-0.46	79-101	0.69-0.72	59-65
FINN	0.29-0.45	80-98	0.70-0.73	59-64
GFAS	0.38-0.42	81-109	0.66-0.70	62-68

¹Correlation and RMSE between observed and modeled $\text{PM}_{2.5}$. The $\text{PM}_{2.5}$ enhancements are simulated using the Stochastic Time-Inverted Lagrangian Transport (STILT) model driven with several fire emission inventories.

²Correlation and RMSE between observed and modeled $\text{PM}_{2.5}$. Here the results from STILT are combined with local observed meteorology from sondes (precipitation, wind speed, wind direction, mixing height) and fit to the observed $\text{PM}_{2.5}$ enhancements using the least absolute shrinkage and selection operator (LASSO), a form of regularized linear regression.

sensitivity maps to each of the fire emission inventories described in Section 2.2.2 and compare the resulting $\text{PM}_{2.5}$ enhancements in Delhi to those observed when averaged across the network and with the derived $\text{PM}_{2.5}$ baseline subtracted. To reduce noise and variability arising from local emissions, we consider only weekly-averaged modeled and observed $\text{PM}_{2.5}$ enhancements. Results show that each of the emission inventories to some degree captures the variability in the surface observed surface $\text{PM}_{2.5}$ ($0.29 < R < 0.50$, Table 2.1), suggesting that smoke from fires upwind drives at least part of the weekly variability of Delhi $\text{PM}_{2.5}$. This modeling result agrees with previous studies that report significant correlations between urban AOD, PM_{10} , visibility, and $\text{PM}_{2.5}$ and MODIS FRP (Liu et al., 2018; Kaskaoutis et al., 2014).

As a measure of the mean bias of our predicted $\text{PM}_{2.5}$ compared to Delhi observa-

tions, we compute the RMSE (Table 2.1). We find that driving the model with STILT alone accounts for an RMSE between 79 - 109 $\mu\text{g m}^{-3}$, depending on the baseline method and emissions inventory, revealing that even though we can predict much of the observed surface $\text{PM}_{2.5}$ variability using STILT, we greatly underestimate the magnitude of the enhancements. A potential reason for this underestimate could be that the GDAS reanalysis used to drive STILT poorly characterizes the local meteorology. We add information from local meteorological sources and fit a statistical model to the observed $\text{PM}_{2.5}$ enhancements. Results of the statistical model are shown in Table 2.1. Adding local meteorological factors improves the correlation of predicted vs. observed $\text{PM}_{2.5}$ in each fire emission scenario ($0.66 < R < 0.78$). Figure 2.7 presents the normalized regression coefficient weights for just the GFED4.1s simulation. Regression coefficients for other statistical models fit with different emission inventories are shown in Figure 2.8. The STILT-GFED4.1s predictor is one of the most significant contributors, as expected by the presence of significant correlation ($0.43 < R < 0.50$) between observed and GFED4.1s STILT-derived $\text{PM}_{2.5}$ enhancements. The next two dominant predictors of observed $\text{PM}_{2.5}$ are wind speed below the boundary layer and precipitation. This result underscores the importance of local meteorology as drivers of urban $\text{PM}_{2.5}$ variability and suggests that the assimilated GDAS meteorology may not capture such meteorological effects at 0.5° resolution. The statistical model yields RMSE values ranging from 53 - 68 $\mu\text{g m}^{-3}$, substantially lower than those from the purely STILT-driven model, but still rather large. We hy-

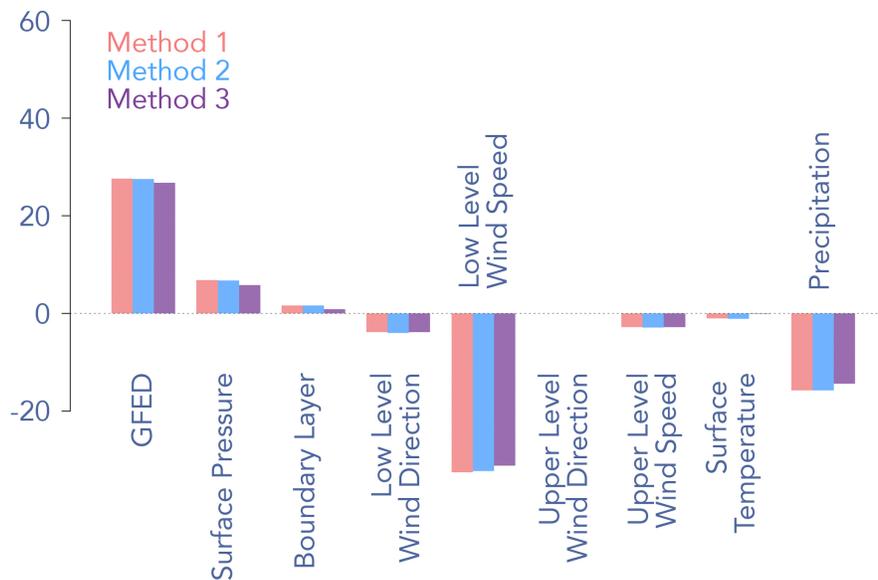
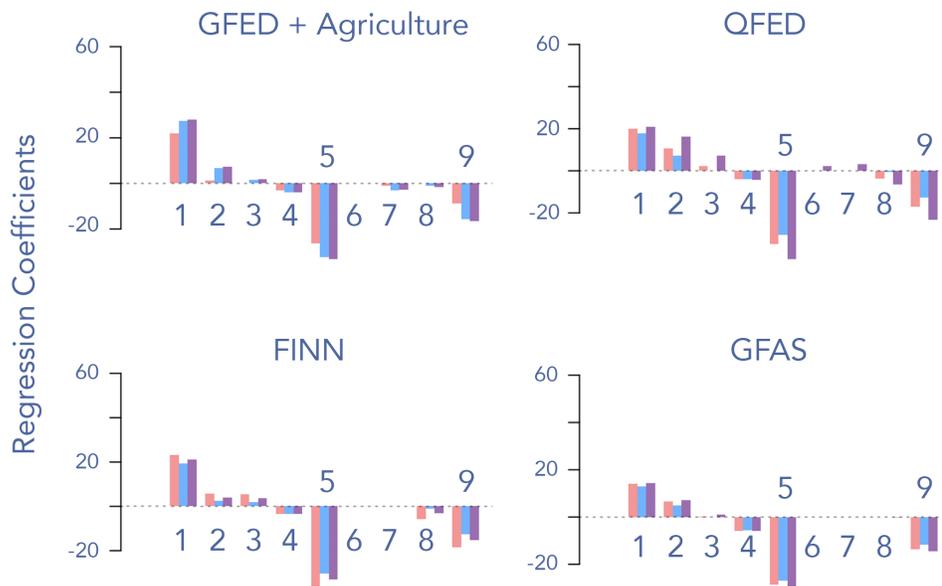


Figure 2.7: Standardized regression coefficients ($\mu\text{g m}^{-3}$ standard deviation⁻¹) fit to daily $\text{PM}_{2.5}$ enhancements, derived from three different baseline methods. See text for description of these methods. The GFED term is the $\text{PM}_{2.5}$ prediction based on driving STILT with GFED4.1s. The other predictors are derived from surface or sonde observed meteorology.

pothesize that other unaccounted factors (e.g., the smoke from small fires that escape satellite detection) could lead to model bias.

2.3.4 MAXIMUM DAILY ENHANCEMENT OF $\text{PM}_{2.5}$ DURING BURNING SEASON

While we capture the variability of $\text{PM}_{2.5}$ with both STILT and the statistical model, in both cases we find a high RMSE when compared to observations. Here we focus on



(1) STILT PM_{2.5} simulation (2) Surface pressure (3) Boundary Layer
 (4) Low level wind direction (5) Low level wind speed (6) Upper level wind direction
 (7) Upper level wind speed (8) Surface temperature (9) Precipitation

Figure 2.8: Standardized regression coefficients ($\mu\text{g m}^{-3}$ standard deviation⁻¹) fit to daily PM_{2.5} enhancements, derived from three different baseline methods. See text for description of these methods. The first term, labeled "1, STILT PM_{2.5} simulation," represents simulated PM_{2.5} using one of four fire emission inventories - GFED + Agriculture, which assumes 100% agricultural landcover and emission factors increased by a factor 3; QFED; FINN; and GFAS.

smoke extremes during each fire season to probe whether the model systematically underestimates surface $\text{PM}_{2.5}$. We also quantify the contribution of smoke $\text{PM}_{2.5}$ derived from observations or STILT (\mathbf{y}_{obs} or \mathbf{y}_{stilt} , respectively) to total $\text{PM}_{2.5}$ during these extreme events.

Figure 2.9 shows the model simulated maximum daily smoke enhancement in each burning season - i.e., the enhancement on that day each season characterized by the greatest simulated $\text{PM}_{2.5}$ value. For years when STILT simulations disagree on which day should produce maximal $\text{PM}_{2.5}$, we choose the day for which most models agree. The plot also shows the observed $\text{PM}_{2.5}$ enhancement and total observed $\text{PM}_{2.5}$ that correspond to the day where the STILT simulation predicted the maximal urban pollution enhancement. We compare these values in Figure 2.9 to the maximum observed $\text{PM}_{2.5}$ enhancement for each burning season, regardless of when the STILT simulation predicted a large enhancement. The largest observed $\text{PM}_{2.5}$ enhancements occur in 2012 and 2016 (492 and 648 $\mu\text{g m}^{-3}$ respectively, averaged across all baseline methods). The maximum observed enhancements are much lower during 2013-2015 (130-264 $\mu\text{g m}^{-3}$), which could be a result of lower fire activity or other local pollution-causing events. The magnitude and interannual variability in the maximum observed $\text{PM}_{2.5}$ enhancement differs from STILT, for which the largest simulated $\text{PM}_{2.5}$ enhancement occurs in 2013 (65-232 $\mu\text{g m}^{-3}$). The STILT simulated enhancements show roughly interannual consistency during 2012-2016 when averaged across all inventories (99-160 $\mu\text{g m}^{-3}$). However, several of the days over 2012-2016 where the

observations alone predict the largest seasonal enhancements are not consistent with the days STILT predicts. When we instead compare the maximum STILT enhancements to the same-day corresponding observed $\text{PM}_{2.5}$ enhancement ($108\text{-}299 \mu\text{g m}^{-3}$), we find closer agreement, though the observations still show more interannual variability than STILT. The FINN and GFED + Agriculture emission inventories often give the largest estimate of magnitude of the observed $\text{PM}_{2.5}$ enhancement in Delhi ($145\text{-}231 \mu\text{g m}^{-3}$ and $147\text{-}255 \mu\text{g m}^{-3}$, respectively). We find the largest mismatch between observed and modeled enhancements during 2012 and 2016 across all models, when we expect the largest enhancement of $\text{PM}_{2.5}$ due to fires. In these years, depending on emission inventory, the maximum STILT derived enhancements are $45\text{-}147$ and $37\text{-}255 \mu\text{g m}^{-3}$, respectively. We see large variability across emission inventories in their ability to capture the observed maximum daily enhancement of $\text{PM}_{2.5}$, and we find that the standard deviation across modeled enhancements for a given year during 2012-2016 ranges between $50\text{-}91 \mu\text{g m}^{-3}$.

Table 2.2 shows the percent contributions of smoke $\text{PM}_{2.5}$ to total $\text{PM}_{2.5}$ on extreme smoke days predicted by STILT - i.e., the day during the season where STILT predicts that the smoke enhancement is greatest. This provides a metric of the contribution of fires during the largest predicted episodes each season to total surface particulate pollution observed in Delhi. The observed $\text{PM}_{2.5}$ enhancement on days when STILT predicted a pollution maximum accounts for 21-72% of the total observed $\text{PM}_{2.5}$, depending on the year and baseline method used, implying that $\text{PM}_{2.5}$

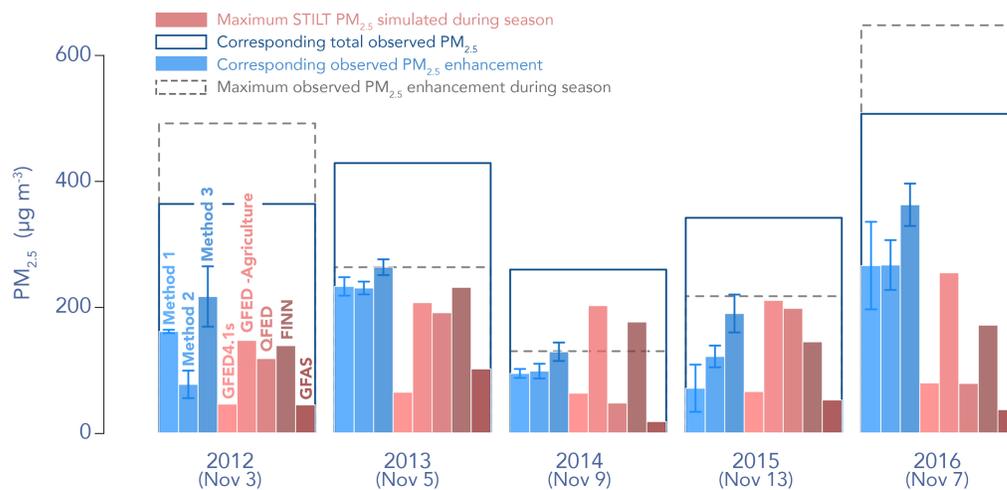


Figure 2.9: The maximum of daily simulated enhancements of $PM_{2.5}$ due to fires upwind fires during each post-monsoon burning season. In parentheses is the day in which the STILT simulation of $PM_{2.5}$ reached its maximum during each burning season from 2012-2016. In shades of red are the different model simulated $PM_{2.5}$ enhancements using different fire emission inventories that correspond to the date in parentheses. In shades of blue are the different network-averaged observed $PM_{2.5}$ enhancement estimates above the anthropogenic baseline for three different baseline methods that correspond to the date in parentheses. The outlined dark blue box represents the total observed $PM_{2.5}$ for the date in parentheses. The outlined grey box represents the maximum observed $PM_{2.5}$ enhancement regardless of the date when the STILT simulation predicted largest enhancement during the post-monsoon burning season.

from a regional source (here assumed to be fires) can constitute a large fraction of the total $\text{PM}_{2.5}$ concentration. For STILT $\text{PM}_{2.5}$, the GFED + Agriculture and FINN simulations provide large $\text{PM}_{2.5}$ estimates, and can account for as much as 78% and 68% percent of the total corresponding observed $\text{PM}_{2.5}$ in 2014, respectively. In other years, these two inventories can account for as much as and 40-62% and 28-54% of the total corresponding observed $\text{PM}_{2.5}$, respectively. This result means that on days when STILT predicts a large enhancement in Delhi from agricultural fires, the smoke from these fires constitutes a large portion of the total $\text{PM}_{2.5}$. On the lower end, the GFAS simulation accounts for just 7.0-24% of the corresponding total $\text{PM}_{2.5}$. Since all inventories use MODIS fire detections to constrain emissions, the variability in $\text{PM}_{2.5}$ estimates that arise from these inventories can be attributed to differing emission factors, allocation of additional fires from burned area maps, model assimilation, and MODIS gap-filling methods. Figure 2.9 and Tables 2.1-2.2 show the large sensitivity in our $\text{PM}_{2.5}$ estimates to the underlying assumptions used to translate satellite retrievals to actual emissions.

The results of Figure 2.9 and Table 2.2 show that STILT can reproduce much of the observed $\text{PM}_{2.5}$ enhancement in Delhi (depending on the emission inventory used), a result that appears at odds with the very high RMSE between observed and modeled enhancements in Table 2.1. To further investigate the reasons driving the discrepancies between observed and modeled $\text{PM}_{2.5}$ enhancements, we plot a sample time series of observed and simulated $\text{PM}_{2.5}$ enhancements for the 2013 post-monsoon

Table 2.2: The percentage of the maximum PM_{2.5} simulated STILT enhancements to corresponding total observed PM_{2.5} for each burning season in Delhi during 2012-2016. OBS refers to the range of PM_{2.5} enhancements derived using the three baseline methods (see Section 2.3.2). Each of the other columns reports simulated PM_{2.5} enhancements from STILT.

Year	Maximum enhancement					
	OBS ¹	GFED	GFED+AGRI ²	QFED	FINN	GFAS
2012	21-60%	13%	40%	33%	38%	12%
2013	54-61%	15%	48%	45%	54%	24%
2014	36-50%	24%	78%	18%	68%	7.0%
2015	21-56%	19%	62%	58%	42%	15%
2016	52-72%	16%	50%	16%	34%	7.3%

¹OBS corresponds to the network-averaged PM_{2.5} enhancement that was observed on same day that the maximum STILT-simulated PM_{2.5} enhancement occurred.

²GFED+AGRI is an emissions inventory based on GFED dry matter emissions, with 100% agriculture landcover assumed and emissions factors increased by a factor of three.

burning season (Figure 2.10). We show observed and simulated PM_{2.5} for 2012 and 2014-16 in Figure 2.11 and include the daily GEOS-Chem simulation of PM_{2.5} for 2012. For 2013, three versions of the STILT model - those driven by FINN, QFED, and GFED + Agriculture emissions - are able to match the PM_{2.5} enhancement on November 5th almost exactly. However, during the days before and after this large pollution enhancement, these models predict little or no PM_{2.5}.

There are several potential reasons for the mismatches between modeled and observed enhancements in smoke PM_{2.5}. On the days preceding the Nov. 5th maximum, MODIS may have been unable to detect many small agricultural fires upwind. Only when a sufficient number of these small fires become detectable is a pollution enhancement predicted by the STILT model. The challenge in detecting small fires

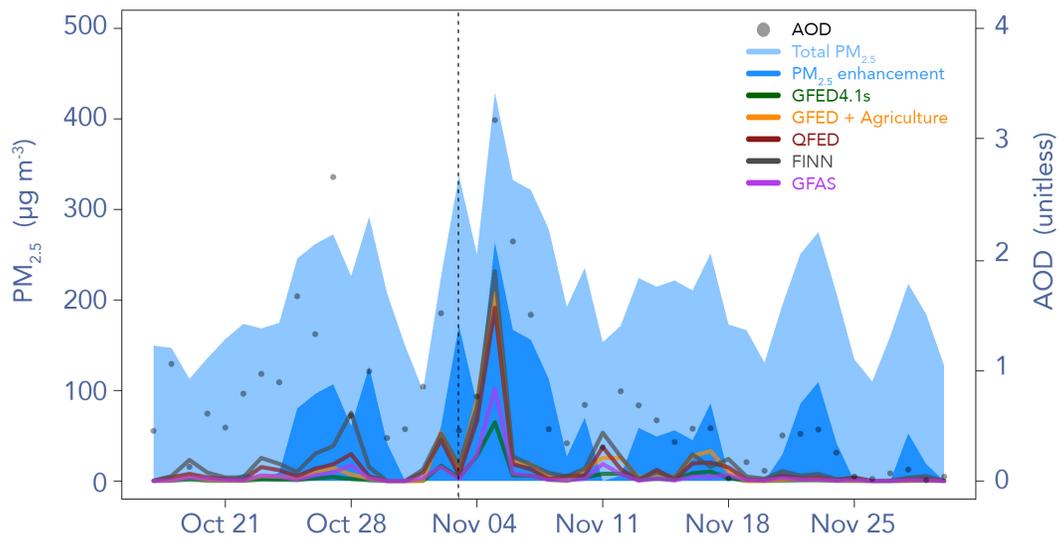


Figure 2.10: Time series of observed and modeled $PM_{2.5}$ during the 2013 burning season. The blue envelopes represent the observed total $PM_{2.5}$ and the $PM_{2.5}$ enhancement derived by subtracting the daily $PM_{2.5}$ by the mean $PM_{2.5}$ of the lowest week during the season. Each colored line represents a model simulation with a different fire emission inventory. The black dots are the MODIS AOD retrievals during the burning season. The dashed vertical line on represents the start of the Diwali festival for 2013 (Nov. 3rd).

from satellites is a well-known problem (Randerson et al., 2012). November 3rd was also the start of Diwali in 2013, a Hindu religious holiday celebrated with an abundance of firecrackers and sparklers. Though Diwali lasts a week, most firecrackers are lit on the first night of the festival (Singh et al., 2009). Without controlling for other factors, Singh et al. (2009) found PM_{10} concentrations to increase by a factor of 2-6 before and after Diwali in Delhi during 2002-2007, and found the effect to be strongest at night. In Figure 2.10, Nov. 3rd total $\text{PM}_{2.5}$ is observed to be especially high ($338 \mu\text{g m}^{-3}$); however, the STILT model simulations predict a small, near-zero $\text{PM}_{2.5}$ enhancement and the observed AOD is also relatively low. Thus an alternative explanation for the observed/modeled $\text{PM}_{2.5}$ mismatch on Nov. 3rd could instead be the effects of Diwali, which may not be captured in fire emission inventories and the coarser AOD product. Diwali generally occurs during the post-monsoon season, though not always during peak agricultural burning. In Figure 2.11, we show the post-monsoon time series of observed and modeled $\text{PM}_{2.5}$ for 2012 and 2014-2016. In 2012, Diwali occurred a week after peak burning and peak observed $\text{PM}_{2.5}$. In 2016, Diwali occurred a week before peak burning and peak observed $\text{PM}_{2.5}$. Though potentially a factor in 2013, the incongruous timing of post-monsoon burning and Diwali during the other seasons implies that observed $\text{PM}_{2.5}$ results from Figure 2.9 and Table 2.2 may sometimes be influenced, but are not driven principally by Diwali.

Returning to 2013 (Figure 2.10), for the days succeeding the Nov. 5th $\text{PM}_{2.5}$ enhancement, local meteorology may have deviated from the coarser 0.5° GDAS winds,

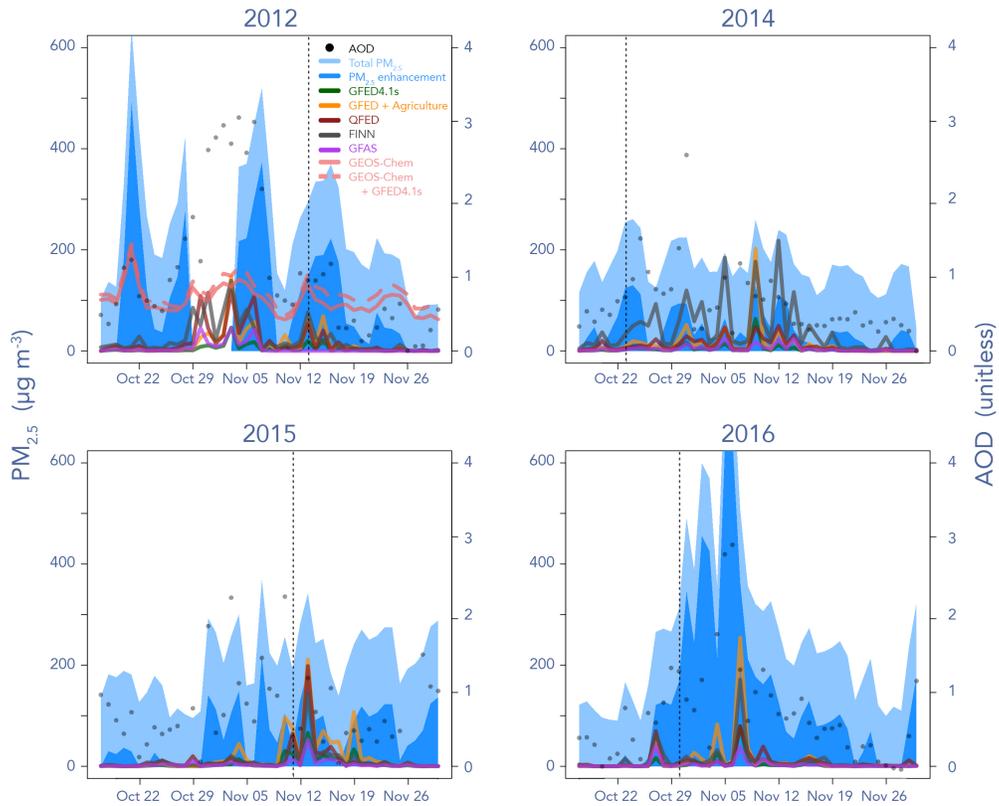


Figure 2.11: Time series of observed and modeled $PM_{2.5}$ during the 2012, 2014-2016 post-monsoon burning seasons. The blue envelopes represent the observed total $PM_{2.5}$ and the $PM_{2.5}$ enhancement derived by subtracting the daily $PM_{2.5}$ by the mean $PM_{2.5}$ of the lowest week during the season. Each colored line represents a model simulation with a different fire emission inventory. The black dots are the MODIS AOD retrievals during the burning season. The dashed vertical lines represent the first day of the Diwali festival. For 2012, we add two additional lines GEOS-Chem + GFED4.1s which represents GEOS-Chem $PM_{2.5}$ from both anthropogenic and fire sources and GEOS-Chem which represents $PM_{2.5}$ from anthropogenic sources alone.

favoring increased stagnation within the city and potentially amplifying surface $\text{PM}_{2.5}$ exposure. Stagnation could have been further amplified by boundary layer stabilization from enhanced $\text{PM}_{2.5}$ aloft, a feedback previously examined as an amplifier of pollution in China (e.g., Petaja et al., 2016; Wang et al., 2014; Ding et al., 2016).

We also hypothesize that dense smoke from fires may sometimes obscure the signal of fire activity at the earth's surface. Figure 2.12a shows True Color Terra reflectance imagery from MODIS as well as MODIS Aqua + Terra fire detections on a sample day over the IGP (November 6, 2016). Figure 2.12b shows the Visible Infrared Imaging Radiometer Suite (VIIRS) reflectance imagery with VIIRS fire detections. VIIRS detects many more fires on this day than does MODIS, perhaps because VIIRS has a finer resolution and different fire detection algorithm than MODIS (375 m compared to 1 km; Schroeder et al., 2014). The MODIS cloud product misidentifies the thick smoke plumes over the Punjab as clouds on this day. The Collection 6 MODIS fire product accounts for thick smoke from fires by relaxing the thresholds that determine whether a pixel is cloud-obscured (Giglio et al., 2016). In fact, on the day illustrated in Figure 2.12 (Nov. 6th, 2016), the MODIS fire product assumes that no pixels over Punjab and Haryana are obscured by clouds, even though the MODIS cloud product reports cloud cover (Figure 2.12c). Even so, fire detections still appear minimal in regions where the smoke is thickest. Thus we hypothesize that the large model underestimates of smoke $\text{PM}_{2.5}$ enhancements in 2016 may be due in large part to layers of dense smoke interfering with satellite detection of thermal anomalies.

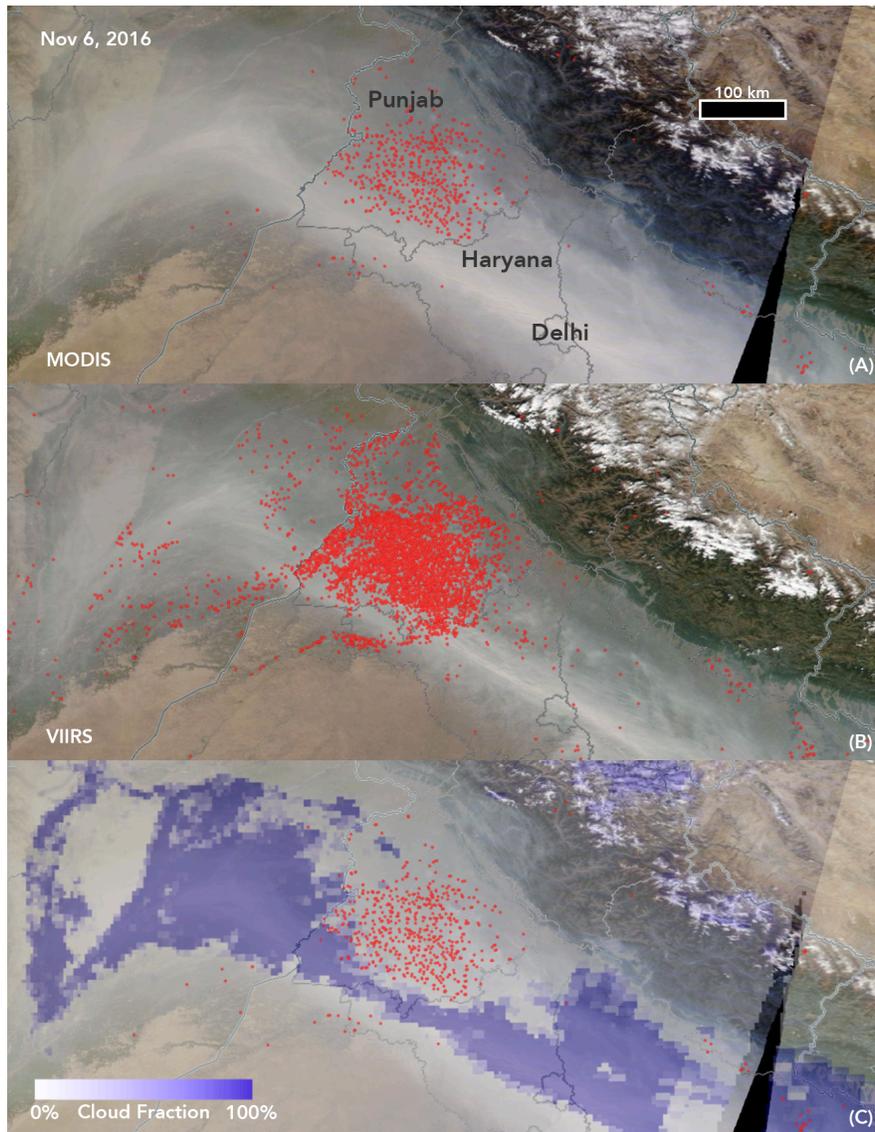


Figure 2.12: MODIS or VIIRS surface reflectance maps for November 6, 2016 overlaid with different fire and cloud detection algorithms. The top panel (A) shows the Terra and Aqua MODIS 1 km fire counts used in part to drive the fire emission inventories used in this paper. The middle panel (B) shows 375 m VIIRS day and night fire detections. The third panel (C) shows MODIS fire detections with MODIS Terra daytime cloud fraction overlaid. Comparison of the top and middle panels show that the resolution of the satellite sensor could influence the number of fires detected, meaning that many smaller fires may be undetected with current MODIS capabilities. Comparison with the bottom panel shows that thick smoke in the Indo-Gangetic Plain may be detected as clouds, which may interfere with surface thermal anomalies.

2.4 DISCUSSION

We estimate the contribution of smoke from upwind agricultural fire emissions to $\text{PM}_{2.5}$ exposure in Delhi during the burning season (Oct 17 - Nov 30). We apply two methods: (1) an observationally based method using CPCB and other surface observations, in which we determine daily enhancements above background levels, averaged over Delhi, and (2) application of the Lagrangian particle dispersion model STILT, in which we implement a suite of fire emission inventories. We find that the two approaches yield timeseries of weekly-averaged $\text{PM}_{2.5}$ that correlate significantly ($0.29 < R < 0.50$) with each other, implying that smoke from agricultural fires upwind accounts for much of the weekly variability of $\text{PM}_{2.5}$ in Delhi during the burning season. Addition of local meteorological factors (precipitation, wind speed, wind direction, temperature, and mixing heights) improves the correlation further ($0.66 < R < 0.78$). The maximum $\text{PM}_{2.5}$ smoke concentration calculated by the STILT model during each burning season is of similar magnitude as its corresponding observed $\text{PM}_{2.5}$ enhancement. For example, in 2013, the maximum simulated $\text{PM}_{2.5}$ enhancements (occurring on Nov. 5th) from GFED + Agriculture, QFED, and FINN are 48%, 45%, and 54% of the corresponding observed maximum $\text{PM}_{2.5}$, respectively, close to the 54-61% range derived from observations (Table 2.2). This result implies that smoke from agricultural fires contributes significantly to $\text{PM}_{2.5}$ pollution in Delhi during intense episodes. However, in general, the $\text{PM}_{2.5}$ simulations greatly underestimate the

enhancements implied by the observations over the entire burning season, with RMSE of 79 - 109 $\mu\text{g m}^{-3}$, indicating that further improvements to fire emission inventories are needed.

We find that although we can predict the magnitude of the maximum $\text{PM}_{2.5}$ enhancement during most seasons using STILT, we miss many smaller $\text{PM}_{2.5}$ enhancements. In the case of 2013, many smaller fires were likely undetected due to limitations in the resolution of the MODIS retrieval. Active fire detection using higher resolution (375m) VIIRS data may provide a promising new avenue to quantify the contribution from small fires. For other fire seasons, as in 2016, STILT underestimates the maximum $\text{PM}_{2.5}$ enhancement more severely, even though Delhi experienced much greater concentrations of $\text{PM}_{2.5}$ than compared to previous seasons. The fires in 2016 were especially strong, but analysis of visual MODIS imagery, fire counts, and cloud cover suggests that many fires were either missed due to the coarse resolution of MODIS detection or were not observed by satellites due to interference of thick smoke. Collection 6 of MODIS FRP accounts for thick smoke in its algorithm by relaxing the cloud-obscuring threshold, which means that regions of thick smoke where no thermal anomalies are detected may now be considered unburned instead of cloud-obscured. If there are missed fires due to the interaction of thick smoke with surface thermal anomalies, this could potential represent a large source of underestimation in assimilated fire emission inventories. As GFAS and QFED estimate FRP in cloud-obscured pixels by using information from adjacent non-obscured pixels, an omitted

or false-negative thermal anomaly under thick smoke would not be assimilated in the fire emission inventory. In Punjab and Haryana, where thick smoke is prevalent during the post-monsoon season due to agricultural fires and low boundary layers, this problem could particularly exacerbate low fire emission estimates.

Some uncertainty in this analysis can be traced to the methods of obtaining a seasonal $\text{PM}_{2.5}$ baseline. We incorporate three different methods to isolate the $\text{PM}_{2.5}$ enhancement due to fires. However, each of these methods shows considerable sensitivity to its various threshold parameters, and there is much variability between each of the methods (e.g., the baseline for 2016 ranges from 140 to 240 $\mu\text{g m}^{-3}$). As more monitors become available in Delhi, distinguishing a regional signal from local enhancement will become less challenging. Inversion methods to optimize emission factors or the spatial allocation of emissions could then be applied with more confidence, since these methods rely on the accuracy of the observed $\text{PM}_{2.5}$ enhancement. Instead of computing the baseline from the observations, one could instead simulate the $\text{PM}_{2.5}$ baseline using a chemistry model such as GEOS-Chem over the entire time domain. However, the result of such simulations would depend strongly on the quality of the emissions used to drive the model and on the extent to which we understand pollution chemistry in this region. In our 2012 GEOS-Chem simulation, we find that the model underestimates the $\text{PM}_{2.5}$ baseline by at least a factor of 1.5-3, compared to the baselines derived from observations.

Another shortcoming of this study is our designation of the whole of Delhi as a one-

point receptor in the STILT simulations. With a denser network of $\text{PM}_{2.5}$ monitors, we could further simulate the variability of $\text{PM}_{2.5}$ exposure due to fires within the city. This approach would also require higher resolution meteorology than the 0.5° resolution used here. Recent initiatives by grassroots efforts like India Spend show a promising step in making these types of analyses more feasible.

Many studies have assessed the human health impact of elevated particulate pollution in Delhi (Nagpure et al., 2014; Kandlikar and Ramachadran, 2000). Our work builds on these studies by quantifying the contribution of agricultural burning in the Punjab and Haryana to the degradation of Delhi air quality. Although officially banned nationally and enforced on the state level by the National Green Tribunal Act of 2010 (Nain Gill, 2010), the practice of agricultural burning is cheap and commonplace for farmers after harvest. India's population is expected to surpass China 2022, and reach 1.7 billion by 2050 (United Nations, 2015). Delhi is projected to grow to a population of 36 million by 2050 (Hoornweg and Pope, 2014). Thus the need for efficient and inexpensive agricultural production is paramount to feeding the increasing population. However, the adverse effects of fire emissions need to continue to be seriously considered and more accurately quantified as the populations of Delhi and the greater IGP continue to grow, leaving more people at risk. Building on the approaches in previous studies (e.g., Liu et al., 2018), the modeling approach presented in this paper can be used to infer not just the co-variability of urban pollution and upwind fires, but also the percent contribution of smoke to the already intense urban

PM_{2.5} in Delhi. As estimates of fire emissions improve and the distribution of air quality monitors in Delhi expands, such an approach will reduce uncertainty in the impacts of current agricultural practices that involve fire. This information can provide policymakers with a quantitative sense of the consequences of current agricultural burning practices in regions upwind of the city in order to inform decision-making.

Chapter 3

Detecting anomalous methane emitters in oil/gas fields using satellite observations

3.1 INTRODUCTION

Anthropogenic methane emissions originate from a large number of relatively small point sources including oil/gas production and processing devices, livestock operations, coal mine vents, landfills, and wastewater facilities (Maasakkers et al., 2016). These sources are often densely clustered, and a small number of sources with anomalously high emissions may contribute disproportionately to a regional emission total. In particular, a large fraction of methane emissions from oil/gas fields is typically due to a small fraction of malfunctioning devices (Zavala-Araiza et al., 2015). Identification of such devices through atmospheric monitoring of methane concentrations can enable corrective action and conducting this monitoring from space using public satellites

has advantages of coverage, cost, and convenience. Here we conduct observing system simulation experiments (OSSEs) to examine the potential of current and near-future satellites to detect malfunctioning production sites in dense oil/gas fields, alone or complemented by a surface observation network.

Oil/gas fields include a very large number of production sites (wells). For example, the Barnett Shale in Northeast Texas has over 20000 oil and gas well pads spread over a 300×300 km² domain (Lyon et al., 2015). Each production site is a point source of methane. The probability density function (pdf) of methane emission from a given site follows a bimodal distribution (Zavala-Araiza et al., 2015). The prevailing low-emission mode is associated with normal operations, while the anomalous high mode is associated with sporadic operations (e.g., venting) or with equipment malfunction. Intensive ground-based observations during the Barnett Shale Coordinated Campaign in 2013 found that 6.6% of wells contributed 50% of total well emissions (Rella et al., 2015). Lyon et al. (2015) found that 40% of the total oil/gas methane emission within the Barnett Shale originated from wells. Identifying high-emitting wells from space may thus be of significant benefit.

Satellites measure atmospheric columns of methane by backscattered solar radiation in the shortwave infrared (SWIR), with near-uniform sensitivity down to the surface under clear-sky conditions (Jacob et al., 2016). The satellite record for methane began with the SCIAMACHY instrument (2003-2012; Frankenberg et al., 2005), which provided coarse resolution measurements (30×60 km² in nadir). The currently oper-

ating GOSAT instrument (2009-; Kuze et al., 2016) has finer resolution (10-km diameter pixels) but sparse coverage (250 km apart). The TROPOMI instrument, launched in October 2017, will provide global daily measurements at 7×7 km² resolution once the data become operational (Hu et al., 2018). The geostationary GeoCARB instrument, to be launched in the early 2020s, is currently planned to provide 2.7×3 km² pixel resolution with a return time that may range from one to four times per day (Polonsky et al., 2014; O'Brien et al., 2016). Other geostationary methane satellite missions have been proposed with various combinations of more frequent coverage, finer pixel resolution, and higher instrument precision (Fishman et al., 2012; Butz et al., 2015; Xi et al., 2015).

A number of studies have examined the value of satellite observations for quantifying methane sources (Jacob et al., 2016). Inverse analyses of SCIAMACHY and GOSAT data have focused on quantifying emissions at 100-1000 km regional scales (Bergamaschi et al., 2013; Wecht et al., 2014a; Alexe et al., 2015; Turner et al., 2015). Previous OSSEs have shown the potential for TROPOMI and GeoCARB to effectively constrain emissions at the 25-100 km scale (Wecht et al., 2014b; Sheng et al., 2018). Other OSSEs have examined the potential for satellites to quantify large point sources from observations of the plumes (Buchwitz et al., 2013; Rayner et al., 2014; Varon et al., 2018). A recent study by Turner et al. (2018) evaluated the capability of TROPOMI and GeoCARB to quantify emissions in the Barnett Shale down to the kilometer scale. They found that GeoCARB should have some capability for constant

sources but not for transient sources.

Here we target a different problem. Given a dense population of production sites in an oil/gas field, can satellites detect inadvertent anomalous emissions sufficiently promptly to enable corrective action? In this problem, quantifying emissions is not as important as prompt detection and localization. The locations of the sources are known, but the mode of emission of any particular source (normal low-mode or anomalous high-mode) is unknown. Once a well starts emitting in the anomalous high mode, it continues doing so until corrective action is taken. The capability of a satellite to detect and localize these leaky wells may be limited by return frequency, cloud cover, pixel resolution, error in the atmospheric transport model used to relate the plume to the location of emission, and limitations in the inverse method for identifying sparse anomalous sources. Here we will evaluate the potential of different satellite observing configurations and inverse methods to address this problem with application to TROPOMI, GeoCARB, and finer-resolution geostationary data. We will also examine whether the information from satellites can be usefully complemented with a supporting network of surface observations.

3.2 OBSERVING SYSTEM SIMULATION EXPERIMENT

We consider a hypothetical oil/gas field of dimension $50 \times 50 \text{ km}^2$ with 20, 50, 100, or 500 randomly placed production sites, corresponding to site densities of 0.008 km^{-2} , 0.02 km^{-2} , 0.04 km^{-2} , and 0.2 km^{-2} , respectively. We create a large ensemble of

emission scenarios for this field where different random subsets of sites of different size categories are in the anomalous high-emission mode, and we simulate the resulting atmospheric methane concentrations with the WRF meteorological model at $1.3 \times 1.3 \text{ km}^2$ resolution. We then sample this pseudo-atmosphere with different satellite and surface observing configurations and apply different inverse methods to detect the high emitters. Detection success is evaluated for each observing configuration and inverse method using statistics for the ensemble of scenarios. We describe in this section the different elements of the OSSE.

3.2.1 CONSTRUCTING AN ENSEMBLE OF EMISSION FIELDS

Production sites within the $50 \times 50 \text{ km}^2$ domain are randomly placed on the $1.3 \times 1.3 \text{ km}^2$ WRF model grid, with at most one site per grid cell. Emission statistics for the sites are based on observations from the Barnett Shale Coordinated Campaign. For each scenario we randomly assign a size category to each production site (small, medium, large) with 23% of the sites as small (10-100 Mcf/d), 62% as medium (100-1000 Mcf/d), and 15% as large (1000+ Mcf/d), following the statistics of Rella et al. (2015). We then assign an emission rate for each site by randomly sampling the bimodal probability density functions (pdfs) defined for each category (Lan et al., 2015; Rella et al., 2015; Yacovitch et al., 2015). We assume no other sources in the domain.

Figure 3.1 shows the pdfs of methane emissions for each facility size category. Here

we flag production sites to be in the high-emission mode if they exceed an emission threshold of 40 kg h^{-1} (axis break in Figure 3.1), which corresponds on average to 5% of all the sites. The high-emission mode averages for large and medium sites are 150 and 100 kg h^{-1} , respectively, which is comparable to typical emissions from large processing plants (Lyon et al., 2015). Anomalous emissions from small facilities are much smaller, centered around 24 kg h^{-1} , and would be difficult to distinguish from the normal emission mode. Thus we do not attempt to detect them as high-mode emitters.

Figure 3.2 shows a sample realization of the oil/gas field with 24 small production sites, 67 medium sites, and 9 large sites (100 total) within the $50 \times 50 \text{ km}^2$ domain. In this realization there are five sites in the high-emission mode. We generate 500 emission scenarios in the same fashion as Figure 3.2 by randomly assigning size categories for each site (small, medium large) and randomly sampling the emission pdfs from Figure 3.1.

3.2.2 CONSTRUCTING PSEUDO-OBSERVATIONS OF ATMOSPHERIC METHANE

We use the meteorological environment previously generated by Turner et al. (2018) for a 1-week period (October 19-25, 2013) in the Barnett Shale, using the Weather Research and Forecasting Model (WRF; Skamarock et al., 2008) at 1.3 km horizontal resolution to drive the Stochastic Time-Inverted Lagrangian Transport (STILT) model (Nehrkorn et al., 2010). STILT is a receptor-oriented Lagrangian particle dis-

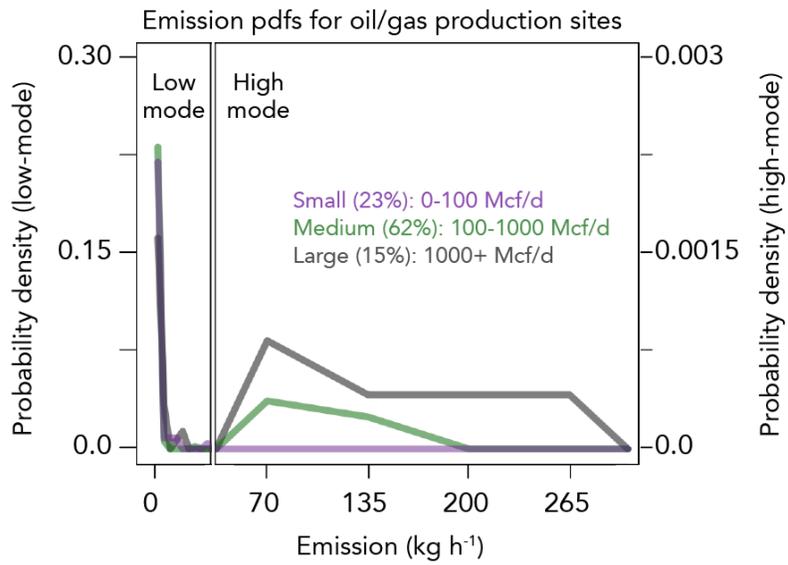


Figure 3.1: Probability density functions (pdfs) of emissions for oil/gas production sites of different size categories, taken from Barnett Shale observations (Lan et al., 2015; Rella et al., 2015; Yacovitch et al., 2015). Note the difference in scales between the left (low-mode) and right (high-mode) panels. The axis break represents the threshold for flagging an emitter as anomalously high.

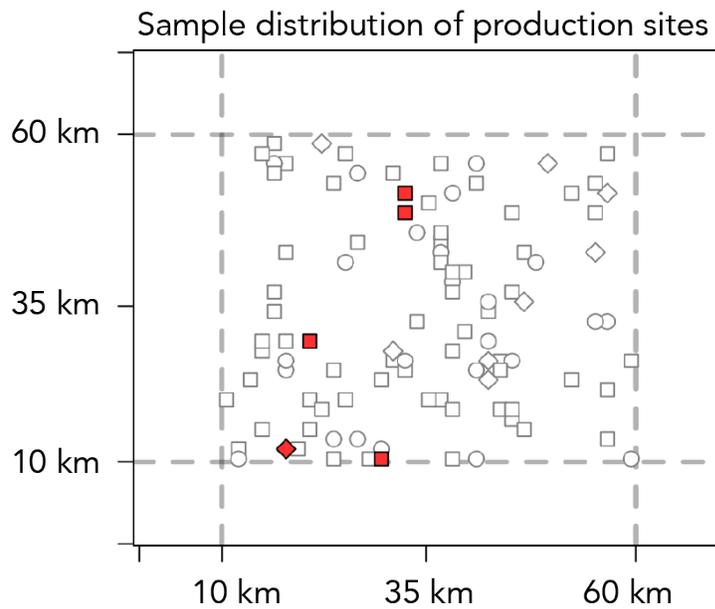


Figure 3.2: Sample realization of emissions from a hypothetical oil/gas production field with 100 production sites of different size categories (symbols) within a $50 \times 50 \text{ km}^2$ domain (dashed line). Different production size categories are shown with symbols. Red shading indicates high-mode emitters. Blue symbols mark the locations of five ground monitors placed according to the k -means algorithm.

persion model that defines the source footprints for individual atmospheric observations. Turner et al. (2018) applied it to generate $1.3 \times 1.3 \text{ km}^2$ hourly footprints for any daytime surface or atmospheric column observation in a $70 \times 70 \text{ km}^2$ domain. The column footprints were weighted with a typical near-uniform SWIR averaging kernel for satellite observations (Worden et al., 2015). We use their ensemble of footprints and add to it hourly footprints for surface observations at night. The $70 \times 70 \text{ km}^2$ observing domain encompasses our $50 \times 50 \text{ km}^2$ oil/gas field plus 10 km outside the boundaries (Figure 3.2) to account for plume transport.

We use this $70 \times 70 \text{ km}^2$ archive of WRF-STILT footprints in combination with any realization of the emission field (Section 3.2.1) to simulate the observed atmospheric methane concentrations that would be observed by a satellite or ground instrument at a given location and time i . Each location and time i has a specific footprint. Figure 3.3 gives an example. Column footprints are about an order of magnitude smaller than surface footprints because of the dilution effect of the column measurement. Taking the footprints to represent the true atmospheric transport relating emissions to atmospheric concentrations for that location and time, we can combine them with any realization of our emission field (Section 3.2.1) to generate the true methane concentrations to be sampled by the instruments.

Satellite observations of methane column concentrations are conventionally in unit of dry column mean mixing ratio (ppb), which is the ratio of the vertical column density of methane to the vertical column density of dry air (Jacob et al., 2016). The

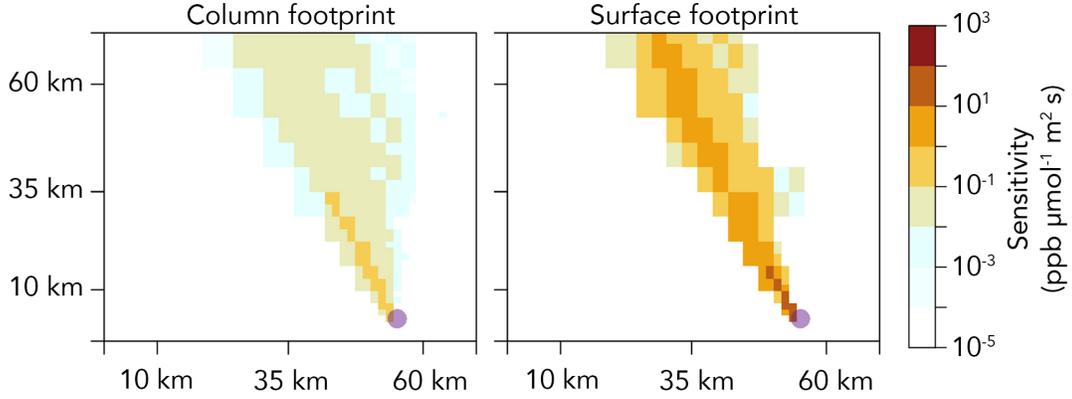


Figure 3.3: Sample sensitivities of observed atmospheric concentrations (column and surface) to surface emissions upwind, defining the emission footprint for that observation. Values are shown here for a particular observation point (purple dot) and time (October 19, 2013 at 09 LT). Concentrations are in mixing ratio units of ppb (dry column mean mixing ratio for the column) and emissions are in units of $\mu\text{mol m}^{-2} \text{s}^{-1}$.

footprint for location and time i is mathematically represented as $\mathbf{h}_i = (\partial y_i / \partial \mathbf{x})^T$ (units $\text{ppb } \mu\text{mol}^{-1} \text{m}^2 \text{s}^{-1}$) where y_i is the methane concentration (ppb) for that location and time, and \mathbf{x} ($\mu\text{mol m}^{-2} \text{s}^{-1}$) is a vector of dimension n describing the gridded emission field for the n emitters in the domain. From the archived footprints \mathbf{h}_i covering the complete set of observing locations and times, the true atmospheric concentration can be constructed for any emission field \mathbf{x} as $y_i = \mathbf{h}_i \cdot \mathbf{x} + b$, where \cdot denotes the scalar product and b is a background assumed here to be constant.

A given methane observing configuration makes m observations of the domain over the 1-week simulation period. The true methane concentrations for that observation ensemble can be assembled as an m -dimensional vector $\mathbf{y}_{true} = \mathbf{H}\mathbf{x}$ where

$\mathbf{H} = \partial \mathbf{y}_{true} / \partial \mathbf{x}$ is the $m \times n$ Jacobian matrix of footprints with rows \mathbf{h}_i^T . The pseudo-observations are then generated as $\mathbf{y} = \mathbf{y}_{true} + \sigma \boldsymbol{\varepsilon}$ where σ is the instrument precision (one standard deviation) and the vector $\boldsymbol{\varepsilon}$ is a random realization of Gaussian noise with mean value of zero and standard deviation of unity for each vector element.

3.2.3 SATELLITE AND SURFACE OBSERVING CONFIGURATIONS

Table 3.1 describes the different satellite observing configurations evaluated in this work including TROPOMI, GeoCARB with 2 or 4 return times per day, and an aspirational next-generation geostationary instrument with $1.3 \times 1.3 \text{ km}^2$ pixel resolution, 1 ppb precision, and hourly return frequency between 8 and 17 local time (LT). Clouds would interfere with satellite observations but here we assume clear-sky conditions to simplify the discussion.

We also wish to determine the benefit of a well-positioned surface air monitoring network for supplementing the satellite observations. Assume that we have M fixed surface monitors to deploy sampling methane concentrations in situ. We want to place them in a configuration that maximizes the information that they would provide, assuming an isotropic wind for generality. Given a known spatial distribution of emitters (the locations of the production sites), we use the k -means spatial clustering approach (Hartigan and Wong, 1979) that selects the monitoring site locations to minimize the distances to emitter locations. Figure 3.2 shows the selected locations for five surface monitors. We assume that the surface monitors report hourly data with 1

Table 3.1: Observing configurations considered in this work.

Instrument	Observation frequency	Pixel size (km ²)	Precision ^a (ppb)	Number of observations ^b
<i>Satellites</i>				
TROPOMI	Daily ^c	7.0 × 7.0	11 ^d	567
GeoCARB 2x/day	2x daily ^e	2.7 × 3.0	4.0 ^f	7700
GeoCARB 4x/day	4x daily ^g	2.7 × 3.0	4.0	15400
Next-generation ^h	Hourly ⁱ	1.3 × 1.3	1.0	164500
<i>Surface monitors^j</i>				
	Hourly ^k	point	1.0	840-3360 ^l

^aDry column mean mixing ratio for the satellite observations

^bOne week of clear-sky conditions in the 70 × 70 km² subdomain

^c13 local solar time (LT)

^dButz et al. (2012)

^e12 and 16 LT

^fO'Brien et al. (2016)

^g10, 12, 14, and 16 LT

^hAspirational instrument combining the characteristics of instruments currently at the proposal stage (Fishman et al., 2012; Butz et al., 2015; Xi et al., 2015)

ⁱBetween 8 and 17 LT

^jSurface monitors are *in situ* observations

^kDay and night

^lFor 5 to 20 surface monitors

ppb precision.

An important consideration for satellite measurements is that methane column enhancements from point sources are small relative to instrument precision, even in the high-emitting mode. This reflects the general problem of methane emissions originating from a large number of relatively small point sources (Jacob et al., 2016; Varon et al., 2018). Figure 3.4 shows the pixel-resolved distribution of atmospheric methane column enhancements above the background for a single pass of the different satellite instruments sampling the emission field of Figure 3.2. The enhancements are less than 1 ppb even for 1.3 km pixels and are diluted at coarser pixel resolution. This is less than the single-scene precision of the satellite instruments (Table 3.1). Successful detection of high-mode emitters thus requires the sampling of many pixels, either across the plume or through repeated sampling, to reduce the noise. This is less of an issue for surface air measurements, where methane enhancements are an order of magnitude higher (Figure 3.3). On the other hand, surface monitors are spatially sparse. For both satellite and surface air observations, a formal inverse analysis of the ensemble of atmospheric observations accounting for plume transport is required for detection of the high-mode emitters.

3.2.4 INVERSE METHODS

Given a set of observations y and Jacobian matrix \mathbf{H} , we need an inverse method to determine the best solution ($\hat{\mathbf{x}}$) of the emission field \mathbf{x} . The inversion should be able

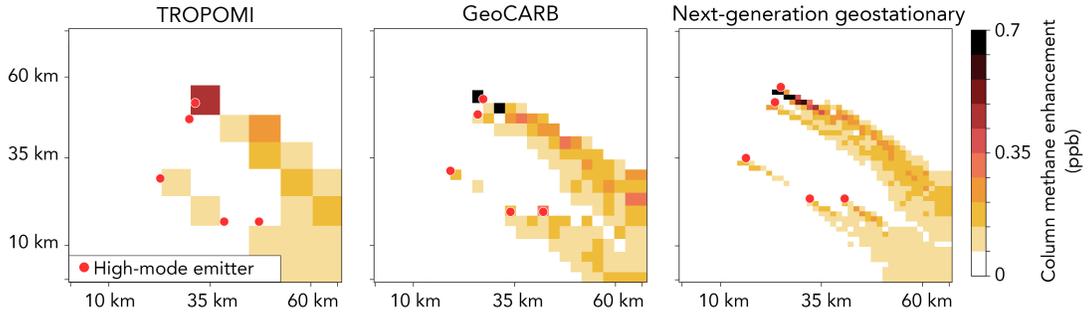


Figure 3.4: Simulated noiseless methane column enhancement for sampling by single overpasses of TROPOMI, GeoCARB, and a next-generation high-resolution geostationary satellite (Table 3.1). Emission field is that of Figure 3.2. The locations of the five high-mode emitters in that field are indicated. Values are for 22 October 2013 at 13 LT.

to detect the small fraction of sources in the high-emitting mode, with detection being more important than quantification. This is known as a sparse-solution problem, where most elements of the emission field \mathbf{x} are very small (for which an optimized value of zero would be acceptable), and a few of the elements are relatively large. We use a regularized least squares regression (Evgeniou et al., 2000), where the solution is found by minimizing the cost function $J(\mathbf{x})$,

$$J(\mathbf{x}) = (\mathbf{H}\mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x} - \mathbf{y}) + \lambda \|\mathbf{x}\|_L \quad (3.1)$$

Here the first term on the right hand side represents the ordinary least-squares loss function, such that the solution of \mathbf{x} would be found to minimize the residuals between the prediction $\mathbf{H}\mathbf{x}$ and the observations weighted by the observational error

covariance matrix \mathbf{R} . The second term represents a norm of \mathbf{x} , or a measure of the magnitude of the vector \mathbf{x} , with λ an adjustable parameter. Adding this second term in the cost function penalizes the total magnitude of \mathbf{x} in the solution, which reduces overfitting and regularizes the solution. When $L = 1$, this is known as L-1 regularization or the least absolute shrinkage and selection operator (LASSO; Tibshirani, 1996), and Equation 3.1 takes the following form:

$$J(\mathbf{x}) = (\mathbf{H}\mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x} - \mathbf{y}) + \lambda \sum_{k=1}^n |\mathbf{x}_k| \quad (3.2)$$

When $L = 2$, Equation 3.1 takes the form known as L-2 regularization or Ridge Regression (Evgeniou et al., 2000):

$$J(\mathbf{x}) = (\mathbf{H}\mathbf{x} - \mathbf{y})^T \mathbf{R}^{-1} (\mathbf{H}\mathbf{x} - \mathbf{y}) + \lambda \mathbf{x}^T \mathbf{x} \quad (3.3)$$

Equation 3.3 is equivalent to standard Bayesian optimization (Rodgers, 2000) assuming Gaussian distributions, a prior emission estimate of zero, and uniform prior error variance of λ^{-1} . The observational error covariance matrix $\mathbf{R} = (r_{ij})$ accounts for both instrument and model transport errors. The diagonal terms add the corresponding error variances in quadrature:

$$r_{ii} = \sigma_I^2 + \sigma_M^2 \quad (3.4)$$

where σ_I is the instrument error standard deviation as given by the precision in Table 3.1, and σ_M is the model transport error standard deviation previously estimated to be 4 ppb (Turner et al., 2018). We use the same model transport error for surface concentrations and satellite columns. Off-diagonal terms account for model transport error correlation between different observations. Following Turner et al. (2018), we assume a temporal error correlation length scale (τ) of 2 hours and a spatial error correlation length scale (ℓ) of 40 km:

$$r_{ij} = \sigma_M^2 \times \exp\left\{\frac{d}{\ell}\right\} \exp\left\{\frac{-t}{\tau}\right\} \quad (3.5)$$

where d and t are the distance and elapsed time between observations y_i and y_j .

Additional model transport error correlation applies when combining satellite and surface air observations in the inversion, since the footprints can be similar (Figure 3.3). To quantify this error correlation, we use the work of Sheng et al. (2018), who jointly compared column (TCCON) and surface air (NOAA) measurements of methane at Lamont, Oklahoma with GEOS-Chem transport model simulations. By correlating the coincident model-observation differences for the column (i) and surface air (j) observations we find a model transport error correlation coefficient $\text{cor}(i, j) = 0.65$ that we apply to the corresponding off-diagonal terms:

$$r_{ij} = \text{cor}(i, j) \times \sigma_M^2 \times \exp\left\{\frac{d}{\ell}\right\} \exp\left\{\frac{-t}{\tau}\right\} \quad (3.6)$$

Inverse solutions derived using L-1 regularization produce sparser solutions than the L-2 counterpart (Tibshirani, 1996), which is desirable for our application. Here we will perform both L-1 and L-2 inversions and compare the results. Minimization of $J(\mathbf{x})$ in Equations 2 and 3 to obtain the solution $\hat{\mathbf{x}}$ corresponding to $dJ/d\mathbf{x} = 0$ is done numerically using coordinate gradient descent (Friedman et al., 2009). The regularization parameter is chosen so that the mismatch between model and observations is small, but not so small that the solution $\hat{\mathbf{x}}$ is overfit to random noise, which would occur when $\lambda = 0$. We use the process of 5-fold cross-validation to select an optimal λ value (Arlot and Celisse, 2010). This process randomly samples \mathbf{H} and \mathbf{y} into a training and validation set. Minimization of J is done on the training set using an array of λ values. The process is repeated five times, and the value of λ that on average minimizes the residual error on the validation set is retained.

3.2.5 DETECTION OF HIGH-EMISSION MODES

Success in the detection of high-mode emitters from the distribution of $\hat{\mathbf{x}}$ can be determined by comparison to the actual occurrence and location of these emitters as defined in Section 3.2.1 and illustrated in Figure 3.2. In a real-world application we would not know the actual pdfs of emissions (Figure 3.1), so we need to diagnose the occurrence of high-mode emitters on the basis of anomalies in the distribution of $\hat{\mathbf{x}}$. We define high-mode elements as being more than S standard deviations from the mean of the $\hat{\mathbf{x}}$ distribution, where S is varied in the 1.65-2.5 range to examine the as-

sociated sensitivity.

The detection of high-mode emitters is graded into four categories: 1) true positives (TP), or the inversion correctly identifying the location of the high-mode emitters, 2) true negatives (TN), or the inversion correctly identifying the location of the low-mode emitters, 3) false positives (FP), or the inversion signaling a high-mode emitter when in reality the emitter is in the low-mode, and 4) false negatives (FN), or the inversion signaling a low-mode emitter when in reality the emitter is in the high mode. We compile these grades into three overall performance metrics (Brasseur and Jacob, 2017). The probability of detection (POD) is defined as the ratio of true positives to true positives plus false negatives:

$$POD = \frac{\Sigma TP}{\Sigma TP + \Sigma FN} \quad (3.7)$$

This metric measures the ability to detect high-mode emitters. The false alarm ratio (FAR) is defined as the ratio of false positives to false positives plus true positives:

$$POD = \frac{\Sigma FP}{\Sigma TP + \Sigma FP} \quad (3.8)$$

This metric measures the reliability of high-mode emission occurrences detected by the inversion. A perfect observing system would have a POD of one and a FAR of zero. Here we define a successful observing system as achieving a POD of 0.8 (80%)

and a FAR of 0.2 (20%).

We combine the POD and FAR metrics into one overall performance metric called the Equitable Threat Score (ETS; Wang, 2014):

$$ETS = \frac{\Sigma TP - \alpha}{\Sigma TP + \Sigma FP + \Sigma FN - \alpha} \quad (3.9)$$

where α is the number of TP predictions that are expected by chance:

$$\begin{aligned} \alpha &= \frac{(\Sigma TP + \Sigma FP)(\Sigma TP + \Sigma FN)}{\Sigma TP + \Sigma FP + \Sigma FN + \Sigma TN} \\ &= \frac{1}{N} \times \frac{\Sigma FP}{FAR} \times \frac{\Sigma TP}{POD} \end{aligned} \quad (3.10)$$

Where $N = \Sigma TP + \Sigma FP + \Sigma FN + \Sigma TN$. The ETS measures how well the high-mode emitters detected by the observing system correspond to the actual occurrences, beyond what could be achieved by chance. A perfect observing system has an ETS of one, and a system performing worse than chance would have a negative ETS. A successful observing system is here defined as having an ETS of 0.65, which corresponds to POD of 0.8 and FAR of 0.2 for a field where approximately 5% of emitters are in the high mode.

3.3 RESULTS AND DISCUSSION

3.3.1 PERFORMANCE OF DIFFERENT SATELLITE AND SURFACE OBSERVING SYSTEMS

We test the ability of each satellite configuration of Table 3.1 to detect high-mode emitters from a field of 20-500 randomly scattered production sites within the 50×50 km² domain as shown in Figure 3.2. We conduct each test for 500 different realizations of the emission field assigning each production site to a size category (small, medium, large) and sampling randomly the pdfs of Figure 3.1. Figure 3.5 shows the POD, FAR, and ETS results for a field of 100 emitters for each observing configuration and compares the results of L-1 and L-2 regularizations. The values represent the mean results for the ensemble of 500 realizations of the emission field, and the error bars represent the range of results when the high-mode detection threshold S is varied from 1.65 to 2.5. We find that L-1 regularization provides better predictions across all performance metrics. This is especially the case for the next-generation satellite, where L-1 regularization produces a POD of 0.85 with a near-perfect FAR of 0.04. L-2 regularization is more conducive to spreading emissions across a broader array of state vector elements. The better performance of L-1 regularization is also observed for other site densities (not shown). We will use L-1 regularization in what follows.

Figure 3.5 also compares the performance of the satellite observing systems to that

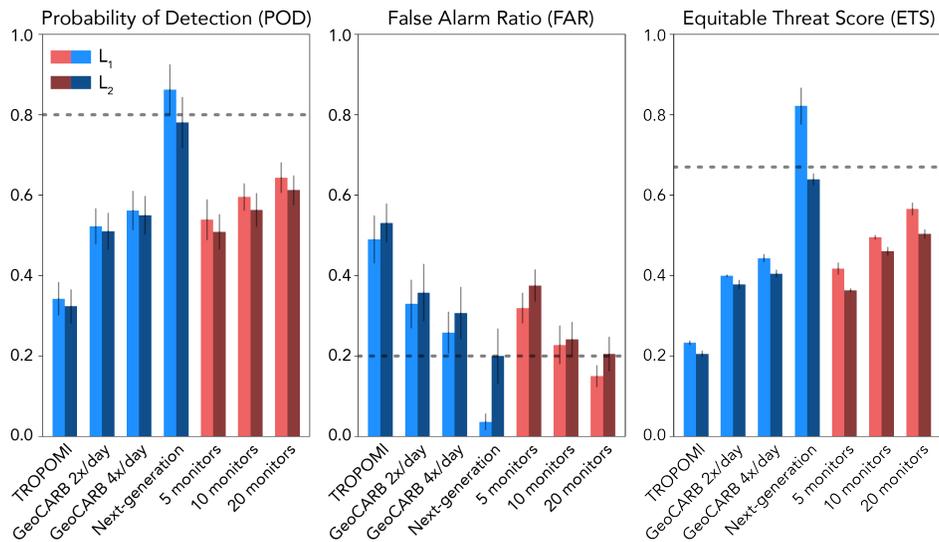


Figure 3.5: Probability of detection (POD), false alarm ratio (FAR), and Equitable Threat Score (ETS) of high-mode emitters for each satellite and surface observing configuration. Each bar represents the mean of 500 observing system simulation experiments (OSSEs), where 100 production sites in a $50 \times 50 \text{ km}^2$ domain were used to construct 500 random realizations of an emission field including different subsets of high-mode emitters. For each observing configuration, the left bar (lighter color) shows results for the inversion with L-1 regularization, and the right bar (darker color) is for the L-2 regularization. The dashed lines represent the POD, FAR, and ETS criteria for successful observing systems.

of an ensemble of 5-20 optimally placed (k -means) surface monitors. We find that the surface monitors perform comparably to GeoCARB. We explore combining satellite and surface observations into a single prediction in Section 3.3.3.

The results from Figure 3.5 show that TROPOMI and GeoCARB are unsuccessful in locating high-mode emitters for a 100 site field. We examined the sensitivity of this result to site density. Figure 3.6 compares the detection results for fields of 20, 50, 100, and 500 production sites within the $50 \times 50 \text{ km}^2$ domain. For a field of only 20 emitters, TROPOMI is successful and GeoCARB produces near perfect results. For a field of 50 emitters, TROPOMI is no longer successful, but GeoCARB 2-4x/day are both nearly successful. Across all site densities, GeoCARB 2-4x/day produce similar results, with GeoCARB 4x/day showing only slight improvement over twice a day sampling. This is due to the temporal error correlation between successive GeoCARB observations. Accounting for cloud cover would show more benefit from 4x/day observations, since a higher frequency of observations allows for cloud clearing, although the benefit depends on the cloud persistence time scale (Sheng et al., 2018).

The ability of a satellite observing configuration to localize high-mode emitters depends not only on repeat time, resolution, and precision, but also on the density of emitters within a field. For the high-density fields of 100 and 500 emitters we find that only the next-generation satellite instrument is successful. The density with 500 sites (0.2 km^{-2}) is comparable to the average density in the Barnett Shale (20000 sites in a $300 \times 300 \text{ km}^2$ domain). Our ability to explore higher site densities with ran-

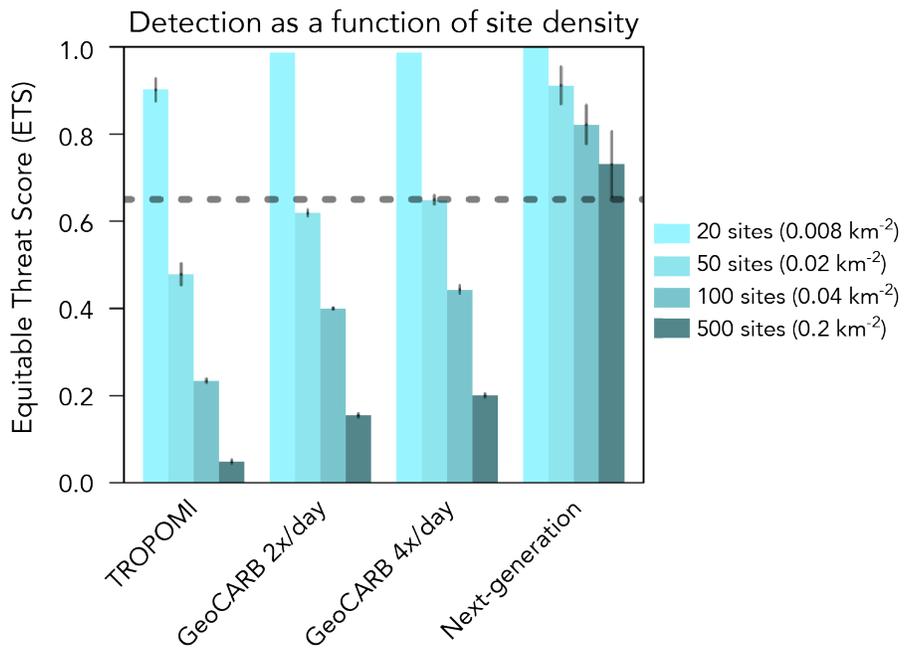


Figure 3.6: Equitable Threat Score (ETS) for each satellite observing configuration, varying the density of production sites (20-500 sites in 50×50 km² domain). Results are from the L-1 inversion. The dashed line represents the ETS criterion for successful observation.

dom placement is limited by the $1.3 \times 1.3 \text{ km}^2$ grid resolution of our WRF simulation.

3.3.2 SPATIAL TOLERANCE IN DETECTION OF HIGH-MODE EMITTERS

The results from Figure 3.6 are somewhat pessimistic regarding the ability of near-future satellite observations (TROPOMI and GeoCARB) to detect the locations of high-mode emitters in dense emission fields. But the localization criterion can be relaxed. If the observing system detects a false positive that is sufficiently close to the actual location of a high-mode emitter, then the detection still has some value. In our OSSE setup, localization is effectively limited by the 1.3 km grid resolution of the WRF simulation. To examine the sensitivity to localization, we repeated the analysis allowing for 3-5 km tolerance of false predictions. Figure 3.7 shows the results. We find that spatial tolerance significantly improves the performance of GeoCARB, as the FAR decreases below 0.2 for 3 km tolerance and below 0.1 for 5 km tolerance. The POD improves to 0.7 which is still below the 0.8 criterion for success, and thus the ETS is still below the 0.65 success threshold. Allowing for spatial tolerance does not fully solve the detection problem for GeoCARB.

3.3.3 COMBINING SATELLITE AND SURFACE OBSERVATIONS

We saw in Section 3.3.1 that only the next-generation satellite instrument can successfully detect anomalous high-mode emitters when the site density is high. Here we examine if a combination of satellite and surface observations could improve de-

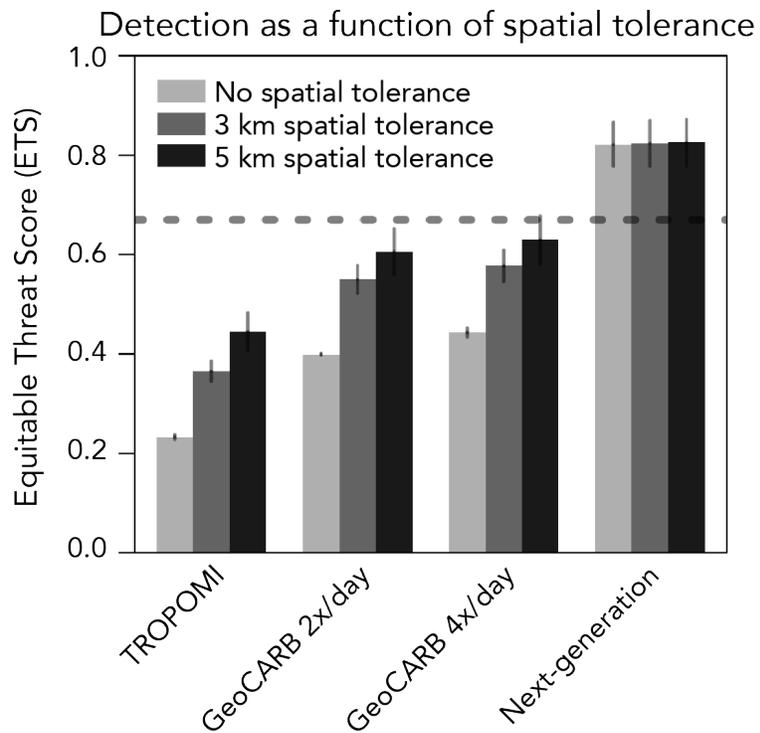


Figure 3.7: Effect of introducing spatial tolerance in the detection of high-mode emitters. Spatial tolerance is the radius within which a high-mode emitter must be located in order for a prediction to be called true positive (TP). The results are for an emission field with 100 production sites in the $50 \times 50 \text{ km}^2$ domain. Only results from the L-1 inversion method are shown. The dashed line represents the ETS success criterion.

tection, i.e., if TROPOMI and GeoCARB could benefit from an in situ supporting surface network and vice versa. This is addressed with a joint inversion of the satellite and surface observations, taking into account the error correlation between the two as described in Section .

Figure 3.8 shows the results for a field of 100 emitters. The already successful next-generation instrument shows no benefit from added surface sites. On the other hand, the surface sites provide greatly added value to TROPOMI and GeoCARB. Adding 10-20 surface monitors enables near-successful detection of the high-mode emitters. At the same time, TROPOMI and GeoCARB data add significantly to the performance of a surface observing system alone. We find that TROPOMI and GeoCARB perform similarly when added to surface sites, and that their main benefit is to decrease the FAR. Accounting for clouds would show more benefit for GeoCARB because the finer pixels allow for better cloud clearing (Sheng et al., 2018).

3.4 CONCLUSIONS

We performed observing system simulation experiments (OSSEs) to test the ability of near-future satellite instruments measuring atmospheric methane (TROPOMI, GeoCARB, next-generation geostationary) to detect anomalous point source emitters among a dense field of individual point sources, alone or supported by a surface monitoring network. We focused on the practical problem of detecting malfunctioning devices (anomalous high-mode emitters) in an oil/gas production field with a high

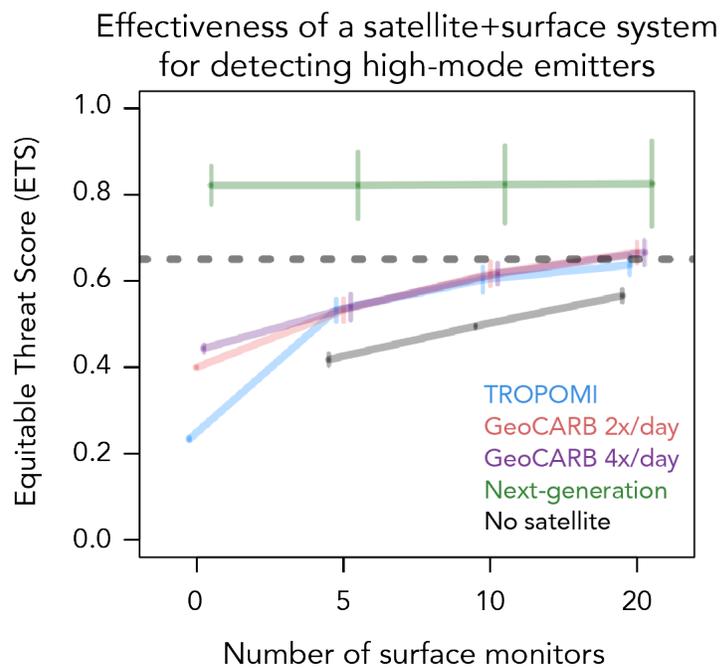


Figure 3.8: Effectiveness of a combined satellite and surface observing system for detecting high-mode emitters in an oil/gas field of 100 emitters over a 50×50 km² domain, as determined from joint inversion of the observations. The dashed line represents the ETS success criterion.

density of wells. Our results can be summarized usefully in terms of answers to questions that a field manager might have:

Can I rely on satellite data alone to detect anomalous emitters among the production sites in my oil/gas field? We find that TROPOMI and GeoCARB can detect anomalous emitters as long as the density of point sources is relatively small (20 sites within our $50 \times 50 \text{ km}^2$ domain, or a density of 0.008 km^{-2}). GeoCARB shows little difference in success rate for 2 or 4 overpasses per day. GeoCARB is close to successful for 50 sites (0.02 km^{-2}) but fails for 100 sites (0.04 km^{-2}). A next-generation geostationary satellite instrument with 1-km pixel resolution and hourly return time would deliver precise detection in dense fields up to 500 sites (0.2 km^{-2}). Allowing for a 5-km spatial tolerance, we find that GeoCARB comes close to successful detection in a field of 100 sites.

How should I analyze the satellite observations to detect anomalous emitters? Detection of anomalous emitters from the satellite observations is not a simple matter of flagging hot spots because the methane column enhancements are relatively small, even for high-mode emitters. Repeated observation is required including of the transported plume, and inverse analysis with an atmospheric transport model is needed. We find that an inversion with L-1 regularization produces much better results than L-2 regularization. This is expected since the L-1 regularization method is designed to recover sparse signals.

Can I usefully supplement satellite information with surface monitoring? Both

TROPOMI and GeoCARB add significantly to the information provided by a surface monitoring network of 10-20 sites within the $50 \times 50 \text{ km}^2$ domain, and both benefit from the added surface information. The combination of these satellite instruments with the surface monitors can deliver successful detection of high-mode emitters through a joint inversion. Adding surface sites provides no benefit to the next-generation geostationary instrument, which can successfully detect high-mode emitters on its own.

Bibliography

- Alexe, M., Bergamaschi, P., Segers, A., Detmers, R., Butz, A., Hasekamp, O., Guerlet, S., Parker, R., Boesch, H., Frankenberg, C., Scheepmaker, R. A., Dlugokencky, E., Sweeney, C., Wofsy, S. C., and Kort, E. A.: Inverse modelling of CH₄ emissions for 2010-2011 using different satellite retrieval products from GOSAT and SCIAMACHY, *Atmospheric Chemistry and Physics*, 15, 113–133, doi:10.5194/acp-15-113-2015, URL <https://www.atmos-chem-phys.net/15/113/2015/>, 2015.
- Arlot, S. and Celisse, A.: A survey of cross-validation procedures for model selection, *Statistics surveys*, 4, 40–79, 2010.
- Augustine, J. A. and Dutton, E. G.: Variability of the surface radiation budget over the United States from 1996 through 2011 from high-quality measurements, *Journal of Geophysical Research: Atmospheres*, 118, 43–53, 2013.
- Augustine, J. A., DeLuisi, J. J., and Long, C. N.: SURFRAD—A national surface radiation budget network for atmospheric research, *Bulletin of the American Meteorological Society*, 81, 2341, 2000.

- Bergamaschi, P., Houweling, S., Segers, A., Krol, M., Frankenberg, C., Scheepmaker, R. A., Dlugokencky, E., Wofsy, S. C., Kort, E. A., Sweeney, C., Schuck, T., Brenninkmeijer, C., Chen, H., Beck, V., and Gerbig, C.: Atmospheric CH₄ in the first decade of the 21st century: Inverse modeling analysis using SCIAMACHY satellite retrievals and NOAA surface measurements, *Journal of Geophysical Research: Atmospheres*, 118, 7350–7369, doi:10.1002/jgrd.50480, URL <http://dx.doi.org/10.1002/jgrd.50480>, 2013.
- Bond, T. C., Bhardwaj, E., Dong, R., Jogani, R., Jung, S., Roden, C., Streets, D. G., and Trautmann, N. M.: Historical emissions of black and organic carbon aerosol from energy-related combustion, 1850–2000, *Global Biogeochemical Cycles*, 21, 2007.
- Booth, B. B., Dunstone, N. J., Halloran, P. R., Andrews, T., and Bellouin, N.: Aerosols implicated as a prime driver of twentieth-century North Atlantic climate variability, *Nature*, 484, 228–232, 2012.
- Bousserez, N., Henze, D. K., Rooney, B., Perkins, A., Wecht, K. J., Turner, A. J., Natraj, V., and Worden, J. R.: Constraints on methane emissions in North America from future geostationary remote-sensing measurements, *Atmospheric Chemistry and Physics*, 16, 6175–6190, doi:10.5194/acp-16-6175-2016, URL <https://www.atmos-chem-phys.net/16/6175/2016/>, 2016.

- Brasseur, G. P. and Jacob, D. J.: Modeling of Atmospheric Chemistry, Cambridge University Press, 2017.
- Buchwitz, M., Reuter, M., Bovensmann, H., Pillai, D., Heymann, J., Schneising, O., Rozanov, V., Krings, T., Burrows, J. P., Boesch, H., Gerbig, C., Meijer, Y., and Löscher, A.: Carbon Monitoring Satellite (CarbonSat): assessment of atmospheric CO₂ and CH₄ retrieval errors by error parameterization, Atmospheric Measurement Techniques, 6, 3477–3500, doi:10.5194/amt-6-3477-2013, URL <https://www.atmos-meas-tech.net/6/3477/2013/>, 2013.
- Budyko, M. I.: The effect of solar radiation variations on the climate of the earth, Tellus, 21, 611–619, 1969.
- Butz, A., Orphal, J., Checa-Garcia, R., Friedl-Vallon, F., von Clarmann, T., Bovensmann, H., Hasekamp, O., Landgraf, J., Knigge, T., Weise, D., Sqalli-Houssini, O., and Kemper, D.: Geostationary Emission Explorer for Europe (G3E): mission concept and initial performance assessment, Atmospheric Measurement Techniques, 8, 4719–4734, doi:10.5194/amt-8-4719-2015, URL <https://www.atmos-meas-tech.net/8/4719/2015/>, 2015.
- Clough, S., Shephard, M., Mlawer, E., Delamere, J., Iacono, M., Cady-Pereira, K., Boukabara, S., and Brown, P.: Atmospheric radiative transfer modeling: a

summary of the AER codes, *Journal of Quantitative Spectroscopy and Radiative Transfer*, 91, 233–244, 2005.

- Darmenov, A. and da Silva, A.: The quick fire emissions dataset (QFED)–documentation of versions 2.1, 2.2 and 2.4, NASA Technical Report Series on Global Modeling and Data Assimilation, NASA TM-2013-104606, 32, 183, 2013.
- Dey, S., Di Girolamo, L., van Donkelaar, A., Tripathi, S., Gupta, T., and Mohan, M.: Variability of outdoor fine particulate (PM 2.5) concentration in the Indian subcontinent: a remote sensing approach, *Remote Sensing of Environment*, 127, 153–161, 2012.
- Diamond, H. J., Karl, T. R., Palecki, M. A., Baker, C. B., Bell, J. E., Leeper, R. D., Easterling, D. R., Lawrimore, J. H., Meyers, T. P., Helfert, M. R., et al.: US Climate Reference Network after one decade of operations: Status and assessment, *Bulletin of the American Meteorological Society*, 94, 485–498, 2013.
- Ding, A., Huang, X., Nie, W., Sun, J., Kerminen, V.-M., Petäjä, T., Su, H., Cheng, Y., Yang, X.-Q., Wang, M., et al.: Enhanced haze pollution by black carbon in megacities in China, *Geophysical Research Letters*, 43, 2873–2879, 2016.
- Dubovik, O., King, M. D., et al.: A flexible inversion algorithm for retrieval of aerosol optical properties from Sun and sky radiance measurements, *Journal of Geophysical Research*, 105, 20 673–20 696, 2000.

- Durre, I., Vose, R. S., and Wuertz, D. B.: Overview of the integrated global radiosonde archive, *Journal of Climate*, 19, 53–68, 2006.
- Eshel, G.: Recent Southern New York Climate Change: Observations, Mechanisms, and Spatial Context, *Journal of Climate*, 29, 209–226, 2016.
- Evgeniou, T., Pontil, M., and Poggio, T.: Regularization networks and support vector machines, *Advances in computational mathematics*, 13, 1, 2000.
- Fishman, J., Iraci, L. T., Al-Saadi, J., Chance, K., Chavez, F., Chin, M., Coble, P., Davis, C., DiGiacomo, P. M., Edwards, D., Eldering, A., Goes, J., Herman, J., Hu, C., Jacob, D. J., Jordan, C., Kawa, S. R., Key, R., Liu, X., Lohrenz, S., Manino, A., Natraj, V., Neil, D., Neu, J., Newchurch, M., Pickering, K., Salisbury, J., Sosik, H., Subramaniam, A., Tzortziou, M., Wang, J., and Wang, M.: The United States’ Next Generation of Atmospheric Composition and Coastal Ecosystem Measurements: NASA’s Geostationary Coastal and Air Pollution Events (GEO-CAPE) Mission, *Bulletin of the American Meteorological Society*, 93, 1547–1566, doi:10.1175/BAMS-D-11-00201.1, URL <https://doi.org/10.1175/BAMS-D-11-00201.1>, 2012.
- Frankenberg, C., Meirink, J. F., van Weele, M., Platt, U., and Wagner, T.: Assessing Methane Emissions from Global Space-Borne Observations, *Science*, 308,

1010–1014, doi:10.1126/science.1106644, URL <http://science.sciencemag.org/content/308/5724/1010>, 2005.

- Friedman, J., Hastie, T., and Tibshirani, R.: Regularization paths for generalized linear models via coordinate descent, *Journal of statistical software*, 33, 1, 2010.
- Gan, C.-M., Pleim, J., Mathur, R., Hogrefe, C., Long, C. N., Xing, J., Roselle, S., and Wei, C.: Assessment of the effect of air pollution controls on trends in short-wave radiation over the United States from 1995 through 2010 from multiple observation networks, *Atmospheric Chemistry and Physics*, 14, 1701–1715, 2014.
- Geng, G., Zhang, Q., Martin, R. V., van Donkelaar, A., Huo, H., Che, H., Lin, J., and He, K.: Estimating long-term PM 2.5 concentrations in China using satellite-based aerosol optical depth and a chemical transport model, *Remote Sensing of Environment*, 166, 262–270, 2015.
- Giglio, L., Csiszar, I., and Justice, C. O.: Global distribution and seasonality of active fires as observed with the Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) sensors, *Journal of geophysical research: Biogeosciences*, 111, 2006.
- Giglio, L., Randerson, J. T., and Werf, G. R.: Analysis of daily, monthly, and annual burned area using the fourth-generation global fire emissions database (GFED4), *Journal of Geophysical Research: Biogeosciences*, 118, 317–328, 2013.

- Giglio, L., Schroeder, W., and Justice, C. O.: The collection 6 MODIS active fire detection algorithm and fire products, *Remote Sensing of Environment*, 178, 31–41, 2016.
- Gupta, P. K., Sahai, S., Singh, N., Dixit, C., Singh, D., Sharma, C., Tiwari, M., Gupta, R. K., and Garg, S.: Residue burning in rice–wheat cropping system: Causes and implications, *Current science*, pp. 1713–1717, 2004.
- Gurjar, B., Ravindra, K., and Nagpure, A. S.: Air pollution trends over Indian megacities and their local-to-global implications, *Atmospheric Environment*, 142, 475–495, 2016.
- Guttikunda, S. K. and Jawahar, P.: Atmospheric emissions and pollution from the coal-fired thermal power plants in India, *Atmospheric Environment*, 92, 449–460, 2014.
- Hartigan, J. A. and Wong, M. A.: Algorithm AS 136: A k-means clustering algorithm, *Journal of the Royal Statistical Society. Series C (Applied Statistics)*, 28, 100–108, 1979.
- Hong, G., Yang, P., Minnis, P., Hu, Y. X., and North, G.: Do contrails significantly reduce daily temperature range?, *Geophysical Research Letters*, 35, 2008.

- Hoornweg, D. and Pope, K.: Socioeconomic pathways and regional distribution of the world's 101 largest cities, Tech. rep., Global Cities Institute Working Paper, 2013.
- Hsu, N., Jeong, M.-J., Bettenhausen, C., Sayer, A., Hansell, R., Seftor, C., Huang, J., and Tsay, S.-C.: Enhanced Deep Blue aerosol retrieval algorithm: The second generation, *Journal of Geophysical Research: Atmospheres*, 118, 9296–9315, 2013.
- Hu, H., Jochen, L., Rob, D., Tobias, B., de Brugh Joost, A., Ilse, A., André, B., and Otto, H.: Toward Global Mapping of Methane With TROPOMI: First Results and Intersatellite Comparison to GOSAT, *Geophysical Research Letters*, 0, doi: 10.1002/2018GL077259, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2018GL077259>, 2018.
- Jacob, D. J., Turner, A. J., Maasakkers, J. D., Sheng, J., Sun, K., Liu, X., Chance, K., Aben, I., McKeever, J., and Frankenberg, C.: Satellite observations of atmospheric methane and their value for quantifying methane emissions, *Atmospheric Chemistry and Physics*, 16, 14 371–14 396, doi:10.5194/acp-16-14371-2016, URL <https://www.atmos-chem-phys.net/16/14371/2016/>, 2016.
- Kaiser, J., Heil, A., Andreae, M., Benedetti, A., Chubarova, N., Jones, L., Morcrette, J.-J., Razinger, M., Schultz, M., Suttie, M., et al.: Biomass burning emis-

sions estimated with a global fire assimilation system based on observed fire radiative power, *Biogeosciences*, 9, 527, 2012.

- Kandlikar, M. and Ramachandran, G.: The causes and consequences of particulate air pollution in urban India: a synthesis of the science, *Annual review of energy and the environment*, 25, 629–684, 2000.
- Kaskaoutis, D., Kumar, S., Sharma, D., Singh, R. P., Kharol, S., Sharma, M., Singh, A., Singh, S., Singh, A., and Singh, D.: Effects of crop residue burning on aerosol properties, plume characteristics, and long-range transport over northern India, *Journal of Geophysical Research: Atmospheres*, 119, 5424–5444, 2014.
- Kim, P. S., Jacob, D. J., Fisher, J. A., Travis, K., Yu, K., Zhu, L., Yantosca, R. M., Sulprizio, M., Jimenez, J. L., Campuzano-Jost, P., et al.: Sources, seasonality, and trends of southeast US aerosol: an integrated analysis of surface, aircraft, and satellite observations with the GEOS-Chem chemical transport model, *Atmospheric Chemistry and Physics*, 15, 4111–4133, 2015.
- Koplitz, S. N., Mickley, L. J., Marlier, M. E., Buonocore, J. J., Kim, P. S., Liu, T., Sulprizio, M. P., DeFries, R. S., Jacob, D. J., Schwartz, J., et al.: Public health impacts of the severe haze in Equatorial Asia in September–October 2015: demonstration of a new framework for informing fire management strategies to reduce downwind smoke exposure, *Environmental Research Letters*, 11, 094023, 2016.

- Koster, R. D. and Suarez, M. J.: The components of a ?SVAT?scheme and their effects on a GCM's hydrological cycle, *Advances in Water Resources*, 17, 61–78, 1994.
- Kumar, P., Kumar, S., and Joshi, L.: Socioeconomic and environmental implications of agricultural residue burning: a case study of Punjab, India, Springer, 2015.
- Kumar, S., Kinter III, J., Dirmeyer, P. A., Pan, Z., and Adams, J.: Multidecadal Climate Variability and the ?Warming Hole? in North America: Results from CMIP5 Twentieth-and Twenty-First-Century Climate Simulations*, *Journal of Climate*, 26, 3511–3527, 2013.
- Kuze, A., Suto, H., Shiomi, K., Kawakami, S., Tanaka, M., Ueda, Y., Deguchi, A., Yoshida, J., Yamamoto, Y., Kataoka, F., Taylor, T. E., and Buijs, H. L.: Update on GOSAT TANSO-FTS performance, operations, and data products after more than 6 years in space, *Atmospheric Measurement Techniques*, 9, 2445–2461, doi:10.5194/amt-9-2445-2016, URL <https://www.atmos-meas-tech.net/9/2445/2016/>, 2016.
- Lan, X., Talbot, R., Laine, P., and Torres, A.: Characterizing Fugitive Methane Emissions in the Barnett Shale Area Using a Mobile Laboratory, *Environmental Science & Technology*, 49, 8139–8146, doi:10.1021/es5063055, URL <http://dx.doi.org/10.1021/es5063055>, PMID: 26148552, 2015.

- Leibensperger, E., Mickley, L. J., Jacob, D. J., Chen, W.-T., Seinfeld, J., Nenes, A., Adams, P., Streets, D., Kumar, N., and Rind, D.: Climatic effects of 1950–2050 changes in US anthropogenic aerosols–Part 1: Aerosol trends and radiative forcing, *Atmospheric Chemistry and Physics*, 12, 3333–3348, 2012.
- Leibensperger, E., Mickley, L. J., Jacob, D. J., Chen, W.-T., Seinfeld, J., Nenes, A., Adams, P., Streets, D., Kumar, N., and Rind, D.: Climatic effects of 1950–2050 changes in US anthropogenic aerosols–Part 2: Climate response, *Atmospheric Chemistry and Physics*, 12, 3349–3362, 2012.
- Levy, R., Mattoo, S., Munchak, L., Remer, L., Sayer, A., Patadia, F., and Hsu, N.: The Collection 6 MODIS aerosol products over land and ocean, *Atmospheric Measurement Techniques*, 6, 2989, 2013.
- Li, M., Zhang, Q., Kurokawa, J.-i., Woo, J.-H., He, K., Lu, Z., Ohara, T., Song, Y., Streets, D. G., Carmichael, G. R., et al.: MIX: a mosaic Asian anthropogenic emission inventory under the international collaboration framework of the MICS-Asia and HTAP, *Atmospheric Chemistry and Physics*, 17, 935, 2017.
- Lin, J., Gerbig, C., Wofsy, S., Andrews, A., Daube, B., Davis, K., and Grainger, C.: A near-field tool for simulating the upstream influence of atmospheric observations: The Stochastic Time-Inverted Lagrangian Transport (STILT) model, *Journal of Geophysical Research: Atmospheres*, 108, 2003.

- Liu, T., Marlier, M. E., DeFries, R. S., Westervelt, D. M., Xia, K. R., Fiore, A. M., Mickley, L. J., Cusworth, D. H., and Milly, G.: Seasonal impact of regional outdoor biomass burning on air pollution in three Indian cities: Delhi, Bengaluru, and Pune, *Atmospheric Environment*, 172, 83 – 92, 2018.
- Long, C. N. and Ackerman, T. P.: Identification of clear skies from broadband pyranometer measurements and calculation of downwelling shortwave cloud effects, *Journal of Geophysical Research: Atmospheres*, 105, 15 609–15 626, 2000.
- Long, C. N., Dutton, E. G., Augustine, J., Wiscombe, W., Wild, M., McFarlane, S. A., and Flynn, C. J.: Significant decadal brightening of downwelling shortwave in the continental United States, *Journal of Geophysical Research: Atmospheres*, 114, 2009.
- Lu, Z., Zhang, Q., and Streets, D. G.: Sulfur dioxide and primary carbonaceous aerosol emissions in China and India, 1996–2010, *Atmospheric Chemistry and Physics*, 11, 9839–9864, 2011.
- Lyon, D. R., Zavala-Araiza, D., Alvarez, R. A., Harriss, R., Palacios, V., Lan, X., Talbot, R., Lavoie, T., Shepson, P., Yacovitch, T. I., Herndon, S. C., Marchese, A. J., Zimmerle, D., Robinson, A. L., and Hamburg, S. P.: Constructing a Spatially Resolved Methane Emission Inventory for the Barnett Shale Region, En-

Environmental Science & Technology, 49, 8147–8157, doi:10.1021/es506359c, URL <http://dx.doi.org/10.1021/es506359c>, PMID: 26148553, 2015.

- Maasakkers, J. D., Jacob, D. J., Sulprizio, M. P., Turner, A. J., Weitz, M., Wirth, T., Hight, C., DeFigueiredo, M., Desai, M., Schmeltz, R., Hockstad, L., Bloom, A. A., Bowman, K. W., Jeong, S., and Fischer, M. L.: Gridded National Inventory of U.S. Methane Emissions, Environmental Science & Technology, 50, 13123–13133, doi:10.1021/acs.est.6b02878, URL <https://doi.org/10.1021/acs.est.6b02878>, PMID: 27934278, 2016.
- Mallia, D., Lin, J., Urbanski, S., Ehleringer, J., and Nehrkorn, T.: Impacts of upwind wildfire emissions on CO, CO₂, and PM_{2.5} concentrations in Salt Lake City, Utah, Journal of Geophysical Research: Atmospheres, 120, 147–166, 2015.
- Meehl, G. A., Arblaster, J. M., and Branstator, G.: Mechanisms contributing to the warming hole and the consequent US east–west differential of heat extremes, Journal of Climate, 25, 6394–6408, 2012.
- Melillo, J. M., Richmond, T. T., and Yohe, G.: Climate change impacts in the United States, Third National Climate Assessment, 2014.
- Menne, M., Durre, I., Korzeniewski, B., McNeal, S., Thomas, K., Yin, X., Anthony, S., Ray, R., Vose, R., Gleason, B., et al.: Global historical climatology network-daily (GHCN-Daily), Version 3, NOAA National Climatic Data Center, 2012.

- Mickley, L., Leibensperger, E., Jacob, D., and Rind, D.: Regional warming from aerosol removal over the United States: Results from a transient 2010–2050 climate simulation, *Atmospheric Environment*, 46, 545–553, 2012.
- Minnis, P., Nguyen, L., Duda, D. P., and Palikonda, R.: Spreading of isolated contrails during the 2001 air traffic shutdown, in: 10th Conference on Aviation, Range and Aerospace Meteorology, 2002.
- Minnis, P., Ayers, J. K., Palikonda, R., and Phan, D.: Contrails, cirrus trends, and climate, *Journal of Climate*, 17, 1671–1685, 2004.
- Minnis, P., Sun-Mack, S., Young, D. F., Heck, P. W., Garber, D. P., Chen, Y., Spangenberg, D. A., Arduini, R. F., Trepte, Q. Z., Smith, W. L., et al.: CERES edition-2 cloud property retrievals using TRMM VIRS and Terra and Aqua MODIS data?Part I: Algorithms, *IEEE Transactions on Geoscience and Remote Sensing*, 49, 4374–4400, 2011.
- Mishra, A. K. and Shibata, T.: Synergistic analyses of optical and microphysical properties of agricultural crop residue burning aerosols over the Indo-Gangetic Basin (IGB), *Atmospheric environment*, 57, 205–218, 2012.
- Mitchell, K. E., Lohmann, D., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., Cosgrove, B. A., Sheffield, J., Duan, Q., Luo, L., et al.: The multi-institution North American Land Data Assimilation System (NLDAS): Utilizing multiple

GCIP products and partners in a continental distributed hydrological modeling system, *Journal of Geophysical Research: Atmospheres*, 109, 2004.

- Mittal, S. K., Singh, N., Agarwal, R., Awasthi, A., and Gupta, P. K.: Ambient air quality during wheat and rice crop stubble burning episodes in Patiala, *Atmospheric Environment*, 43, 238–244, 2009.
- Mu, Q. and Liao, H.: Simulation of the interannual variations of aerosols in China: role of variations in meteorological parameters, *Atmospheric Chemistry and Physics*, 14, 9597–9612, 2014.
- Nagpure, A. S., Gurjar, B. R., and Martel, J.: Human health risks in national capital territory of Delhi due to air pollution, *Atmospheric Pollution Research*, 5, 371–380, 2014.
- Nain Gill, G.: A Green Tribunal for India, *Journal of Environmental Law*, 22, 461–474, 2010.
- Nehr Korn, T., Eluszkiewicz, J., Wofsy, S. C., Lin, J. C., Gerbig, C., Longo, M., and Freitas, S.: Coupled weather research and forecasting–stochastic time-inverted lagrangian transport (WRF–STILT) model, *Meteorology and Atmospheric Physics*, 107, 51–64, doi:10.1007/s00703-010-0068-x, URL <https://doi.org/10.1007/s00703-010-0068-x>, 2010.
- Network, I. C.: Water and atmospheric resources monitoring program, 2005.

- Oanh, N. T. K., Ly, B. T., Tipayarom, D., Manandhar, B. R., Prapat, P., Simpson, C. D., and Liu, L.-J. S.: Characterization of particulate matter emission from open burning of rice straw, *Atmospheric Environment*, 45, 493–502, 2011.
- O’Brien, D. M., Polonsky, I. N., Utembe, S. R., and Rayner, P. J.: Potential of a geostationary geoCARB mission to estimate surface emissions of CO₂, CH₄ and CO in a polluted urban environment: case study Shanghai, *Atmospheric Measurement Techniques*, 9, 4633, 2016.
- Overgaard, J. and Rosbjerg, D.: Energy-based land-surface modelling: new opportunities in integrated hydrological modelling, Technical University of Denmark-Danmarks Tekniske Universitet, Department of Hydrodynamics and Water Resources Strømningsmekanik og Vandressourcer, 2011.
- Pan, Z., Arritt, R. W., Takle, E. S., Gutowski, W. J., Anderson, C. J., and Segal, M.: Altered hydrologic feedback in a warming climate introduces a “warming hole?”, *Geophysical Research Letters*, 31, 2004.
- Pant, P., Shukla, A., Kohl, S. D., Chow, J. C., Watson, J. G., and Harrison, R. M.: Characterization of ambient PM_{2.5} at a pollution hotspot in New Delhi, India and inference of sources, *Atmos. Environ*, 109, 178–189, 2015.
- Park, R. J., Jacob, D. J., Chin, M., and Martin, R. V.: Sources of carbonaceous aerosols over the United States and implications for natural visibility, *Journal of*

Geophysical Research: Atmospheres, 108, 2003.

- Park, R. J., Jacob, D. J., Field, B. D., Yantosca, R. M., and Chin, M.: Natural and transboundary pollution influences on sulfate-nitrate-ammonium aerosols in the United States: Implications for policy, *Journal of Geophysical Research: Atmospheres*, 109, 2004.
- Petäjä, T., Järvi, L., Kerminen, V.-M., Ding, A., Sun, J., Nie, W., Kujansuu, J., Virkkula, A., Yang, X., Fu, C., et al.: Enhanced air pollution via aerosol-boundary layer feedback in China, *Scientific reports*, 6, 2016.
- Pinker, R. T., Tarpley, J. D., Laszlo, I., Mitchell, K. E., Houser, P. R., Wood, E. F., Schaake, J. C., Robock, A., Lohmann, D., Cosgrove, B. A., et al.: Surface radiation budgets in support of the GEWEX Continental-Scale International Project (GCIP) and the GEWEX Americas Prediction Project (GAPP), including the North American Land Data Assimilation System (NLDAS) project, *Journal of Geophysical Research: Atmospheres*, 108, 2003.
- Polonsky, I. N., O'Brien, D. M., Kumer, J. B., O'Dell, C. W., and the geoCARB Team: Performance of a geostationary mission, geoCARB, to measure CO₂, CH₄ and CO column-averaged concentrations, *Atmospheric Measurement Techniques*, 7, 959–981, doi:10.5194/amt-7-959-2014, URL <https://www.atmos-meas-tech.net/7/959/2014/>, 2014.

- Posada, D. and Buckley, T. R.: Model selection and model averaging in phylogenetics: advantages of Akaike information criterion and Bayesian approaches over likelihood ratio tests, *Systematic biology*, 53, 793–808, 2004.
- Randerson, J., Chen, Y., Werf, G., Rogers, B., and Morton, D.: Global burned area and biomass burning emissions from small fires, *Journal of Geophysical Research: Biogeosciences*, 117, 2012.
- Rayner, P. J., Utembe, S. R., and Crowell, S.: Constraining regional greenhouse gas emissions using geostationary concentration measurements: a theoretical study, *Atmospheric Measurement Techniques*, 7, 3285–3293, doi:10.5194/amt-7-3285-2014, URL <https://www.atmos-meas-tech.net/7/3285/2014/>, 2014.
- Registrar General, I.: Census of India 2011: provisional population totals-India data sheet, Office of the Registrar General Census Commissioner, India. Indian Census Bureau, 2011.
- Rella, C. W., Tsai, T. R., Botkin, C. G., Crosson, E. R., and Steele, D.: Measuring Emissions from Oil and Natural Gas Well Pads Using the Mobile Flux Plane Technique, *Environmental Science & Technology*, 49, 4742–4748, doi: 10.1021/acs.est.5b00099, PMID: 25806837, 2015.
- Rienecker, M. M., Suarez, M. J., Gelaro, R., Todling, R., Bacmeister, J., Liu, E., Bosilovich, M. G., Schubert, S. D., Takacs, L., Kim, G.-K., et al.: MERRA:

NASA's modern-era retrospective analysis for research and applications, *Journal of Climate*, 24, 3624–3648, 2011.

- Rodell, M., Houser, P., Jambor, U., Gottschalck, J., Mitchell, K., Meng, C., Arsenault, K., Cosgrove, B., Radakovich, J., Bosilovich, M., et al.: The global land data assimilation system, *Bulletin of the American Meteorological Society*, 85, 381–394, 2004.
- Rodgers, C. D.: *Inverse methods for atmospheric sounding: theory and practice*, vol. 2, World scientific, 2000.
- Sayer, A., Hsu, N., Bettenhausen, C., and Jeong, M.-J.: Validation and uncertainty estimates for MODIS Collection 6 "Deep Blue" aerosol data, *Journal of Geophysical Research: Atmospheres*, 118, 7864–7872, 2013.
- Schroeder, W., Oliva, P., Giglio, L., and Csiszar, I. A.: The New VIIRS 375m active fire detection data product: Algorithm description and initial assessment, *Remote Sensing of Environment*, 143, 85–96, 2014.
- Schubert, S. D., Suarez, M. J., Pegion, P. J., Koster, R. D., and Bacmeister, J. T.: On the cause of the 1930s Dust Bowl, *Science*, 303, 1855–1859, 2004.
- Seager, R. and Hoerling, M.: Atmosphere and Ocean Origins of North American Droughts*, *Journal of Climate*, 27, 4581–4606, 2014.

- Seibert, P., Beyrich, F., Gryning, S.-E., Joffre, S., Rasmussen, A., and Tercier, P.: Review and intercomparison of operational methods for the determination of the mixing height, *Atmospheric environment*, 34, 1001–1027, 2000.
- Sheng, J.-X., Jacob, D. J., Turner, A. J., Maasakkers, J. D., Sulprizio, M. P., Bloom, A. A., Andrews, A. E., and Wunch, D.: High-resolution inversion of methane emissions in the Southeast US using SEAC⁴RS aircraft observations of atmospheric methane: anthropogenic and wetland sources, *Atmospheric Chemistry and Physics*, 18, 6483–6491, doi:10.5194/acp-18-6483-2018, URL <https://www.atmos-chem-phys.net/18/6483/2018/>, 2018.
- Shindell, D. T., Schmidt, G. A., Miller, R. L., and Mann, M. E.: Volcanic and solar forcing of climate change during the preindustrial era, *Journal of Climate*, 16, 4094–4107, 2003.
- Singh, D., Gadi, R., Mandal, T., Dixit, C., Singh, K., Saud, T., Singh, N., and Gupta, P. K.: Study of temporal variation in ambient air quality during Diwali festival in India, *Environmental monitoring and assessment*, 169, 1–13, 2010.
- Singh, R. P. and Kaskaoutis, D. G.: Crop residue burning: a threat to South Asian air quality, *Eos, Transactions American Geophysical Union*, 95, 333–334, 2014.
- Skamarock, W., Klemp, J., Dudhia, J., Gill, D., Barker, D., Dudha, M., Huang, X., Wang, W., and Powers, Y.: A description of the advanced research WRF Ver. 30,

National Center for Atmospheric Research US, 2008.

- Stein, A., Draxler, R. R., Rolph, G. D., Stunder, B. J., Cohen, M., and Ngan, F.: NOAA's HYSPLIT atmospheric transport and dispersion modeling system, *Bulletin of the American Meteorological Society*, 96, 2059–2077, 2015.
- Stoffel, T.: Solar infrared radiation station (SIRS) handbook, Atmospheric Radiation Measurement Program Technical Report ARM TR-025, 2005.
- Suarez, M. J., Rienecker, M., Todling, R., Bacmeister, J., Takacs, L., Liu, H., Gu, W., Sienkiewicz, M., Koster, R., Gelaro, R., et al.: The GEOS-5 Data Assimilation System-Documentation of Versions 5.0. 1, 5.1. 0, and 5.2. 0, 2008.
- Tibshirani, R.: Regression shrinkage and selection via the lasso, *Journal of the Royal Statistical Society. Series B (Methodological)*, pp. 267–288, 1996.
- Tiwari, S., Srivastava, A., Bisht, D., Parmita, P., Srivastava, M. K., and Attri, S.: Diurnal and seasonal variations of black carbon and PM 2.5 over New Delhi, India: influence of meteorology, *Atmospheric Research*, 125, 50–62, 2013.
- Travis, D. J., Carleton, A. M., and Lauritsen, R. G.: Regional variations in US diurnal temperature range for the 11-14 September 2001 aircraft groundings: evidence of jet contrail influence on climate, *Journal of climate*, 17, 1123–1134, 2004.
- Turner, A. J., Jacob, D. J., Wecht, K. J., Maasakkers, J. D., Lundgren, E., Andrews, A. E., Biraud, S. C., Boesch, H., Bowman, K. W., Deutscher, N. M.,

Dubey, M. K., Griffith, D. W. T., Hase, F., Kuze, A., Notholt, J., Ohyama, H., Parker, R., Payne, V. H., Sussmann, R., Sweeney, C., Velazco, V. A., Warneke, T., Wennberg, P. O., and Wunch, D.: Estimating global and North American methane emissions with high spatial resolution using GOSAT satellite data, *Atmospheric Chemistry and Physics*, 15, 7049–7069, doi:10.5194/acp-15-7049-2015, URL <https://www.atmos-chem-phys.net/15/7049/2015/>, 2015.

- Turner, A. J., Jacob, D. J., Benmergui, J., Brandman, J., White, L., and Randles, C. A.: Assessing the capability of different satellite observing configurations to resolve the distribution of methane emissions at kilometer scales, *Atmospheric Chemistry and Physics Discussions*, 2018, 1–23, doi:10.5194/acp-2018-164, URL <https://www.atmos-chem-phys-discuss.net/acp-2018-164/>, 2018.
- United Nations: World population prospects: The 2015 revision, United Nations Publications, 2015.
- USEPA: Our Nation’s Air: Status and Trends Through 2010, 2012.
- Vadrevu, K. P., Badarinath, K., and Anuradha, E.: Spatial patterns in vegetation fires in the Indian region, *Environmental monitoring and assessment*, 147, 1–13, 2008.
- Vadrevu, K. P., Ellicott, E., Badarinath, K., and Vermote, E.: MODIS derived fire characteristics and aerosol optical depth variations during the agricultural residue

burning season, north India, *Environmental pollution*, 159, 1560–1569, 2011.

- Van Der Werf, G. R., Randerson, J. T., Giglio, L., Van Leeuwen, T. T., Chen, Y., Rogers, B. M., Mu, M., Van Marle, M. J., Morton, D. C., Collatz, G. J., et al.: Global fire emissions estimates during 1997–2016, *Earth System Science Data*, 9, 697, 2017.
- Varon, D., Jacob, D., McKeever, J., Xia, Y., and Huang, Y.: Quantifying Methane Point Sources from Fine-scale Satellite Observation of Individual Plumes: Application to GHGSat Observations, in: *AGU Fall Meeting Abstracts*, 2017.
- Wang, C.-C.: On the Calculation and Correction of Equitable Threat Score for Model Quantitative Precipitation Forecasts for Small Verification Areas: The Example of Taiwan, *Weather and Forecasting*, 29, 788–798, doi:10.1175/WAF-D-13-00087.1, URL <https://doi.org/10.1175/WAF-D-13-00087.1>, 2014.
- Wang, J., Wang, S., Jiang, J., Ding, A., Zheng, M., Zhao, B., Wong, D. C., Zhou, W., Zheng, G., Wang, L., et al.: Impact of aerosol–meteorology interactions on fine particle pollution during China’s severe haze episode in January 2013, *Environmental Research Letters*, 9, 094002, 2014.
- Wang, Y., Zhang, Q., He, K., Zhang, Q., and Chai, L.: Sulfate-nitrate-ammonium aerosols over China: response to 2000–2015 emission changes of sulfur dioxide, ni-

trogen oxides, and ammonia, *Atmospheric Chemistry and Physics*, 13, 2635–2652, 2013.

- Wecht, K. J., Jacob, D., Frankenberg, C., Jiang, Z., and Blake, D.: Mapping of North American methane emissions with high spatial resolution by inversion of SCIAMACHY satellite data, *Journal of Geophysical Research: Atmospheres*, 119, 7741–7756, doi:10.1002/2014JD021551, URL <https://agupubs.onlinelibrary.wiley.com/doi/abs/10.1002/2014JD021551>, 2014a.
- Wecht, K. J., Jacob, D. J., Sulprizio, M. P., Santoni, G. W., Wofsy, S. C., Parker, R., Bösch, H., and Worden, J.: Spatially resolving methane emissions in California: constraints from the CalNex aircraft campaign and from present (GOSAT, TES) and future (TROPOMI, geostationary) satellite observations, *Atmospheric Chemistry and Physics*, 14, 8173–8184, doi:10.5194/acp-14-8173-2014, URL <https://www.atmos-chem-phys.net/14/8173/2014/>, 2014b.
- Westervelt, D., Horowitz, L., Naik, V., Golaz, J.-C., and Mauzerall, D.: Radiative forcing and climate response to projected 21st century aerosol decreases, *Atmospheric Chemistry and Physics*, 15, 12 681–12 703, 2015.
- Wiedinmyer, C., Akagi, S., Yokelson, R. J., Emmons, L., Al-Saadi, J., Orlando, J., and Soja, A.: The Fire INventory from NCAR (FINN): a high resolution global

model to estimate the emissions from open burning, *Geoscientific Model Development*, 4, 625, 2011.

- Wood, E. F., Lettenmaier, D., Liang, X., Nijssen, B., and Wetzel, S. W.: Hydrological modeling of continental-scale basins, *Annual Review of Earth and Planetary Sciences*, 25, 279–300, 1997.
- Worden, J. R., Turner, A. J., Bloom, A., Kulawik, S. S., Liu, J., Lee, M., Weidner, R., Bowman, K., Frankenberg, C., Parker, R., and Payne, V. H.: Quantifying lower tropospheric methane concentrations using GOSAT near-IR and TES thermal IR measurements, *Atmospheric Measurement Techniques*, 8, 3433–3445, doi:10.5194/amt-8-3433-2015, URL <https://www.atmos-meas-tech.net/8/3433/2015/>, 2015.
- Xi, X., Natraj, V., Shia, R. L., Luo, M., Zhang, Q., Newman, S., Sander, S. P., and Yung, Y. L.: Simulated retrievals for the remote sensing of CO₂, CH₄, CO, and H₂O from geostationary orbit, *Atmospheric Measurement Techniques*, 8, 4817–4830, doi:10.5194/amt-8-4817-2015, URL <https://www.atmos-meas-tech.net/8/4817/2015/>, 2015.
- Xia, Y., Mitchell, K., Ek, M., Sheffield, J., Cosgrove, B., Wood, E., Luo, L., Alonge, C., Wei, H., Meng, J., et al.: Continental-scale water and energy flux analysis and validation for the North American Land Data Assimilation System project

phase 2 (NLDAS-2): 1. Intercomparison and application of model products, *Journal of Geophysical Research: Atmospheres*, 117, 2012.

- Yacovitch, T. I., Herndon, S. C., Pétron, G., Kofler, J., Lyon, D., Zahniser, M. S., and Kolb, C. E.: Mobile Laboratory Observations of Methane Emissions in the Barnett Shale Region, *Environmental Science & Technology*, 49, 7889–7895, doi:10.1021/es506352j, URL <http://dx.doi.org/10.1021/es506352j>, PMID: 25751617, 2015.
- Yu, S., Alapaty, K., Mathur, R., Pleim, J., Zhang, Y., Nolte, C., Eder, B., Foley, K., and Nagashima, T.: Attribution of the United States ?warming hole?: Aerosol indirect effect and precipitable water vapor, *Scientific reports*, 4, 6929, 2014.
- Zavala-Araiza, D., Lyon, D., Alvarez, R. A., Palacios, V., Harriss, R., Lan, X., Talbot, R., and Hamburg, S. P.: Toward a Functional Definition of Methane Super-Emitters: Application to Natural Gas Production Sites, *Environmental Science & Technology*, 49, 8167–8174, doi:10.1021/acs.est.5b00133, URL <http://dx.doi.org/10.1021/acs.est.5b00133>, PMID: 26148555, 2015.
- Zhang, R., Delworth, T. L., Sutton, R., Hodson, D. L., Dixon, K. W., Held, I. M., Kushnir, Y., Marshall, J., Ming, Y., Msadek, R., et al.: Have aerosols caused the observed Atlantic multidecadal variability?, *Journal of the Atmospheric Sciences*, 70, 1135–1144, 2013.