Thesis

USING OPERATIONAL HMS SMOKE OBSERVATIONS TO GAIN INSIGHTS ON NORTH AMERICAN SMOKE TRANSPORT AND IMPLICATIONS FOR AIR QUALITY

Submitted by

Steven J. Brey

Department of Atmospheric Science

In partial fulfillment of the requirements

For the Degree of Master of Science

Colorado State University

Fort Collins, Colorado

Fall 2016

Master's Committee:

Advisor: Emily V. Fischer

Elizabeth Barnes
Jeffery Pierce
Monique Rocca

ABSTRACT

USING OPERATIONAL HMS SMOKE OBSERVATIONS TO GAIN INSIGHTS
ON NORTH AMERICAN SMOKE TRANSPORT AND IMPLICATIONS FOR
AIR QUALITY

Wildfires represent a major challenge for air quality managers, as they are large sources of particulate matter (PM) and ozone (O_3) precursors, and they are highly dynamic and transient events. Smoke can be transported thousands of kilometers to deteriorate air quality over large regions. Under a warming climate, fire severity and frequency are likely to increase, exacerbating an existing problem. Using the National Environmental Satellite, Data and Information Service (NESDIS) Hazard Mapping System (HMS) smoke data for the U.S. and Canada for the period 2007 to 2014, I examine a subset of fires that are confirmed to have produced sufficient smoke to warrant the initiation of a National Weather Service smoke forecast. The locations of these fires combined with Hybrid Single Particle Lagragian Integrated Trajectory Model (HYSPLIT) forward trajectories, satellite detected smoke plume data, and detailed land-cover data are used to develop a climatology of the land-cover, location, and seasonality of the smoke that impacts the atmospheric column above 10 U.S. regions. I examine the relative contribution of local versus long-range transport to the presence of smoke in different regions as well as the prevalence of smoke generated by agricultural burning versus wildfires.

This work also investigates the influence of smoke on O_3 abundances over the contiguous U.S. Using co-located observations of particulate matter and the NESDIS HMS smoke data, I identify summertime days between 2005 and 2014 that Environmental Protection Agency

Air Quality System O_3 monitors are influenced by smoke. I compare O_3 mixing ratio distributions for smoke-free and smoke-impacted days for each monitor, while accounting for temperature. This analysis shows that (i) the mean O_3 abundance measured on smoke-impacted days is higher than on smoke-free days at 20% of monitoring locations, and (ii) the magnitude of the difference between smoke-impacted and smoke-free mixing ratios varies by location and is sensitive to the minimum temperature allowed for smoke-free days. For each site, I present the percentage of days when the 8-hr average O_3 mixing ratio (MDA8) exceeds 75 ppbv and smoke is present. When our most lenient temperature criteria are applied to smoke-free days, smoke-impacted O_3 mixing ratios are most elevated in locations with the highest emissions of nitrogen oxides. The Northeast corridor, Dallas, Houston, Atlanta, Birmingham, and Kansas City stand out as having smoke present 10-20% of the days when 8-hr MDA8 O_3 mixing ratios exceed 75 ppbv. Most U.S. cities maintain a similar proportion of smoke-impacted exceedance days when they are held against the new MDA8 limit of 70 ppbv.

LIST OF TABLES

Theres the smoke, wheres the fire? A regional analysis of smoke transport pathways based on 8 years of HMS smoke and fire location data

1.1. Introduction

North American fires represent a major source of atmospheric pollutants [1], and their contribution to elevated ground level ozone (O_3) and fine particulate matter ($PM_{2.5}$) in the U.S. has been documented repeatedly [2–6]. Exposure to wildfire smoke has been shown to have negative impacts on respiratory and cardiovascular health [7–9]. The relative importance of the contribution of smoke, particularly wildfire smoke, as both an O_3 precursor and as a direct source of $PM_{2.5}$ will grow as anthropogenic emissions decline [10, 11]. Smoke is also thought to make a major contribution to absorbing aerosols observed throughout the troposphere over North America [12–15]. Millions of hectares of forest burn in North America each year [16], the area burned by wildfires in the western U.S. has increased [17], and models predict this trend will continue in a warmer climate [18, 16, 19–21]. Agricultural fires are also a significant source of smoke in North America. Over 1.2 million hectares of cropland are burned in the contiguous U.S. each year [22]. Though often smaller in size and shorter in duration, these fires have been shown to increase the abundance of gas and aerosol species that can deteriorate air quality [23–25].

There are a growing number of case studies documenting instances of smoke traveling thousands of kilometers to affect atmospheric composition far downwind from the fire locations (*e.g.* [26–31]). Case studies demonstrate that regional and long-range transport

of smoke causes elevated column and surface concentrations of aerosols and trace gases over extensive (continental) regions of the U.S., and the smoke can have implications for both aerosol radiative forcing and air quality over relatively long temporal scales (weeks) [32, 2, 33, 29, 31]. Though case studies abound, it is challenging to assess the impact of fires on atmospheric composition and air quality in aggregate.

Past efforts to assess the impact of fires on atmospheric composition, or air quality specifically, have primarily relied on chemical transport models. A number of advances have recently been made in estimating the emissions inputs by improving burned area products (e.g [16]) and combining these with emission factors for a wider range of trace species [e.g. [25, 34]). However, incorporating the full suite of emitted species into models and simulating the rapid chemical evolution of smoke remains a challenge [35], as is proper treatment of injection height [36] and the timing of emissions [4]. Models are also subject to uncertainty associated with meteorological inputs [37]. Finally, running and analyzing a chemical transport model at the fine grid resolution appropriate to simulate all the individual smoke plumes of interest from North American fires over the scale of a decade is currently too computationally expensive to be practical. Thus other lenses are needed to examine how the smoke from North American fires is transported and dispersed in the atmosphere over seasonal and interannual timescales.

The two primary goals of this study are to 1) present the distribution and seasonality of fires that trigger National Weather Service (NWS) smoke forecasts, and 2) develop a regional climatology of smoke transport in the U.S. using operational data from the National Environmental Satellite, Data and Information Service (NESDIS) Hazard Mapping System (HMS) combined with forward trajectory calculations. Based on the subset of fires triggering smoke forecasts, HMS observations of smoke in the atmospheric column and trajectory
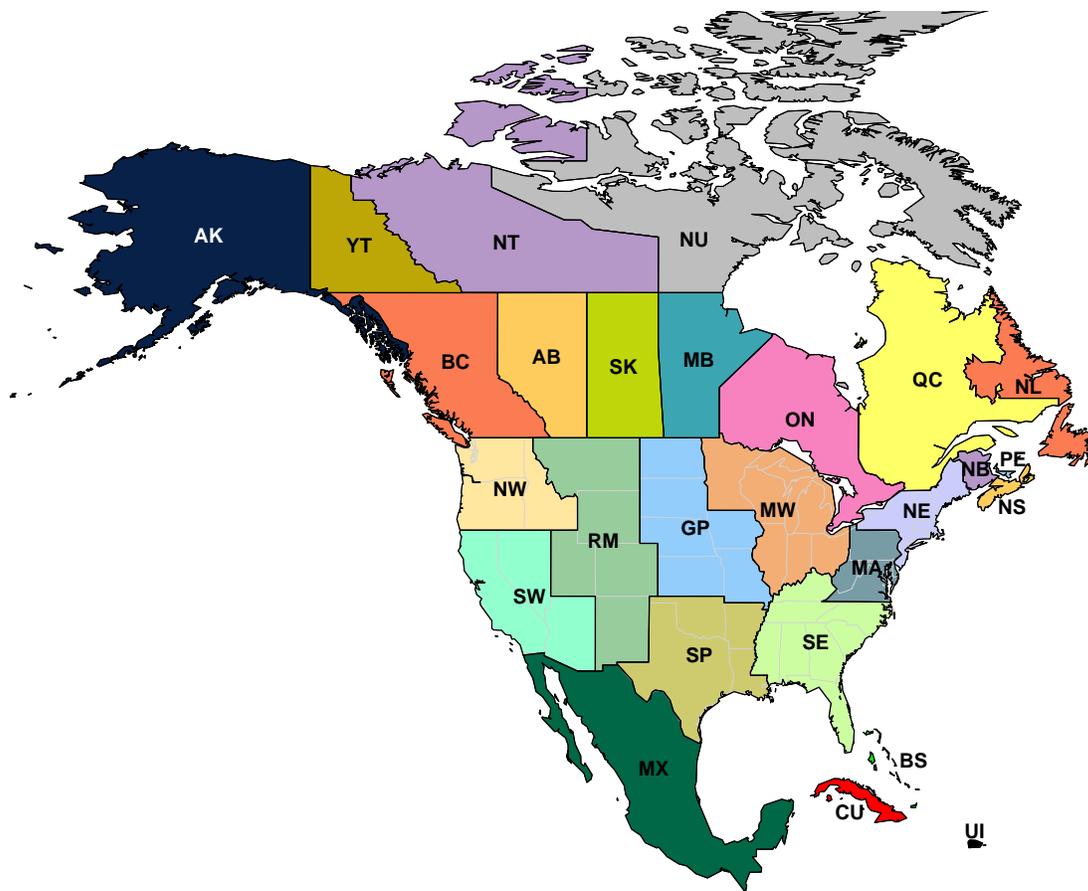
FIGURE 1.1. Smoke source and receptor regions used in this analysis. Northeast (NE), Mid Atlantic (MA), Southeast (SE), Midwest (MW), Southern Plains (SP), Great Plains (GP), Rocky Mountains (RM), Southwest (SW), Northwest (NW), Alaska (AK), US Islands (UI), Mexico (MX), Quebec (QC), Nova Scotia (NS), Saskatchewan (SK), Alberta (AB), Newfoundland and Labrador (NL), British Columbia (BC), New Brunswick (NB), Prince Edward Island (PE), Yukon Territory (YT), Manitoba (MB), Ontario (ON), Nunavut (NU), Northwest Territories (NT), Cuba (CU), and Bahamas (BS).

calculations, we present an estimate of the relative frequency that smoke observed over ten U.S. regions is associated with fires in each of the other U.S. regions or from North American regions (Canada, Mexico, Cuba, and the Bahamas).

1.2. METHODS

1.2.1. DESCRIPTION OF OPERATIONAL FIRE AND SMOKE PRODUCTS. The HMS (Hazard Mapping System) is an interactive environmental satellite image display and graphical

interface system that was developed by the National Environmental Satellite, Data, and Information Service (NESDIS). Trained satellite analysts use the HMS to generate a daily operational list of fire locations and outline areas of smoke. As a part of this process, analysts manually inspect automated fire detections against the mid-wave infrared (MWIR) images that they were produced from to ensure a fire exists [38]. Detections deemed to be false are removed and fires not detected are added manually. Visible satellite imagery is also used by analysts to identify fires that may be too small, not producing sufficient heat or obscured by a tree canopy to be automatically detected [39]. In these cases a smoke plume may be the only indication of a fire. The number of fire detections added manually can be significant. For example, over 50% of the total fire detections were added manually during a 12-month period in 2002-2003 examined by *Ruminski et al.* [38]. Land-cover data and power-plant locations compliment satellite imagery to help HMS analysts confirm whether automatic detections are fires [38]. The HMS office makes a distinction between all detected fires (hereafter HMS hotspots) and fires an HMS analyst has confirmed to produce a substantial amount of smoke (hereafter HYSPLIT points).

HYSPLIT points are a subset of the HMS hotspots; they are fire detections where an analyst also visually confirms the presence of smoke using visible satellite imagery. HYSPLIT points are human-vetted because they are used to initialize the NWS smoke forecasts [38, 39]. Each HYSPLIT point is assigned a latitude, longitude, date, time, and duration. The locations of HYSPLIT points are estimated to be accurate to within 2-3 km. The start time of smoke emissions is estimated to be accurate to within 1 hour. The accuracy of the duration is a bit more uncertain since many of the fires continue to generate smoke after sunset when visible imagery is no longer available. However, it is believed that the duration accuracy for most HYSPLIT points is within 2 hours. Single fires that produce notable amounts of

smoke are associated with a cluster of co-located, or nearly co-located, HYSPLIT points in proportion to the amount of smoke observed. The intended operational consequence of designating multiple HYSPLIT points for the largest smoke-producing fires is that it allows the NWS smoke forecast model to generate more smoke than for a single HYSPLIT point.

The HMS office does not make any distinctions between sources of smoke, and HYSPLIT points can be associated with agricultural burning, prescribed burning, or wildfires [38]. There is a relationship between the number of HYSPLIT points and the amount of smoke produced by fires, and the analyst determines this relationship. Large wildfires are represented by dozens of HYSPLIT points spread over many square kilometers. These are typically in the western U.S., Canada, and Alaska. The start times of HYSPLIT points can vary within the cluster of points. Operationally, this serves to represent the variability in the amount of smoke observed at different times of the day. For example, during a large wildfire event analysts may create several HYSPLIT points with a 24-hr duration starting at 08 UTC (middle of the night local time). They also create another set of HYSPLIT points in the vicinity that are assigned a shorter duration and a start time of 20 UTC (early afternoon local time). The operational intention of this strategy is to force the NWS smoke forecast model to produce more smoke in the afternoon and evening and less overnight to replicate observed diurnal trends. HYSPLIT points can also be proxies for unobservable smoke producing fires. When analysts see a large number of HMS hotspots but due to cloud cover, do not directly observe smoke, they create HYSPLIT points in order to initiate the NWS smoke forecast model. This occurs most frequently in Kansas, Oklahoma, the Northern Plains (Dakotas), and the lower Mississippi Valley (eastern Arkansas, eastern Louisiana, western Mississippi). The Servicio Meteorolgico Nacional (SMN, the Mexican National Weather Service) provides

most HYSPLIT points over Mexico. These locations are merged with the HMS product. Occasionally the HMS office will perform fire-detection analysis in parts of Northern Mexico. The archives of smoke and fire locations spans from August 2005 to present day; but the analysis presented here spans 2007–2014. These archives include text files for fire locations and GIS shapefiles for smoke.

This analysis uses data from 2007-2014 for two reasons. First, prior to April 2006 HYSPLIT points did not have duration or start-time estimates. The second reason is more complex. For 2015, we observed many more HYSPLIT points than each of the prior years, and there were many more instances of HYSPLIT points assigned durations of 24 hours. The reason for this is two-fold. I) HMS implemented a system that automatically generates HYSPLIT points in Northern Canada, Mexico, and Central America in the fall of 2014. The majority of fires in this area of Canada are wildfires in boreal forests, so each of these automated HYSPLIT points is assigned a duration of 24 hours whereas prior to this it would have received a mix of 24 hour and lesser durations. The intention of the automated system was to reduce the workload of analysts. A similar automated system was implemented for Alaska in 2009. However, this implementation did not lead to a significant change in the proportion of HYSPLIT points with durations of 24 hours analyzed in North America. II) Prior to these implementations each HYSPLIT point was entered manually, one at a time, with a mouse click. III) Northern Canada and Alaska had a very active wildfire season in 2015. Early in the season Alaska had its most active wildfire season on record; over the season, 2015 was the second most active in terms of acres burned. Prior to Fall 2014 the HYSPLIT points in Mexico and Central America were intended to be generated by SMN, but at some point in the years prior to 2014 SMN began performing HYSPLIT point analysis inconsistently. As a result the NESDIS HMS office developed an automated system based on

HMS hotspots. The durations for these HYSPLIT points are estimated using the difference between the latest and earliest time for hotspots aggregated into a 20km grid.

HYSPLIT points are sometimes analyzed at hours when no visible satellite imagery is available to confirm smoke production. This only occurs when the thermal signal in the MWIR imagery is significant in terms of intensity and duration (*e.g.* strong downslope winds at night can cause siginficant nighttime fire activity). Wildfires can burn for weeks or even months. For the long-lived wildfire scenario, HMS analysts will get two looks daily at a fire from a given polar satellite (NOAA-19 or MODIS-Aqua). One look overnight (in the Western U.S. between 08-10Z) and another in the afternoon (20-22Z). Analysts will add HYSPLIT points at both of those times. For the nighttime pass, fires are often not burning as actively (fewer hotspots). Analysts will add HYSPLIT points based on the fewer number of hotspots and typically assign them a 24-hr duration (*i.e* total hours smoke production observed), creating a baseline for emissions. In the afternoon, fires are more active and generate more smoke so analysts will add additional HYSPLIT points based on the afternoon satellite data. In this case, they only assign durations of 10 or 12 hours. The operational significance of these procedures is to attempt to account for the diurnal variations in smoke production in initializing smoke forecasts.

1.2.2. DESCRIPTION OF HMS SMOKE ANALYSIS. After identifying HYSPLIT points, HMS analysts use imagery from seven NOAA and NASA satellites to identify the geographic extent of smoke-plumes [38, 39]. Smoke detection is done with visible-band imagery occasionally assisted by infrared to distinguish between clouds and smoke when possible [38]. Geostationary GOES imagery, with its frequent refresh rate (typically every 15 minutes for each spacecraft), is used almost exclusively for smoke detection, although on rare occasions polar orbiting satellite imagery is used. Given the limitations of the satellite data (mostly

obscuration of smoke due to cloud cover), the number and extent of smoke plumes within this dataset represents a conservative estimate. Smoke is sometimes transported to areas with anthropogenic haze pollution. In some cases the smoke will mix with and become indistinguishable from the anthropogenic haze pollution. The greater the distance traveled by a smoke plume the more challenging it is to distinguish between smoke and anthropogenic haze. This challenge is particularly pronounced for aged smoke impacting the Southeast U.S. Due to the limitations of visible satellite imagery and time constraints, no information about the vertical location or extent of smoke plumes is provided. In 2006 HMS analysts began providing estimates of smoke-plume concentrations (5 μg m^{-3}, 16 μg m^{-3}, and 27 μg m^{-3}). These plumes of varying concentrations are often nested (*i.e.* 27 μg m^{-3}within 16 μg m^{-3}and 16 μg m^{-3}within 5 μg m^{-3}). Smoke plumes vary in size considerably; small plumes cover areas < 100 km^2 and others cover several Western States. Between 5 August 2005 and 21 December 2015 there are only 80 days ($\sim$2 %) where there are either no smoke plume GIS files available or no smoke-plumes analyzed. Most of these days occur during winter months. For this work, we use the archived GIS smoke-plume files available at the following URL `ftp://satepsanone.nesdis.noaa.gov/FIRE/HMS/GIS/`.

Many of the archived smoke polygons have straight-line edges particularly during the summer and over the ocean. The straight edges signify a boundary in which smoke-plume-detection analysis is performed. The smoke-plume-detection analysis is performed in five sectors. Each sector displays satellite imagery in a Lambert conic conformal projection. After analysis in all five regions has been performed, they are pieced together to form a single analysis. Strait edges of individual smoke GIS polygons occur when smoke plumes from different regions are pieced together. Not all sectors are analyzed year round. There is no analysis for Alaska or Northern Canada between November 1st and May 1st.

1.2.3. Correlation Between HYSPLIT Points and Smoke Plumes. There is not always a HYSPLIT point associated with every smoke plume and vice versa. Often (especially during the summer when there are many wildfires producing a large amount of smoke) analysts observe smoke that has drifted a long way and become detached from the fire that produced it. In this case the smoke plume is not associated with any HYSPLIT points on that day. For example, the wildfires in Alaska and northern Canada in 2015 produced smoke that drifted southeast into the Great Lakes and Mid Atlantic region and eventually reached Europe. This transport occurred over several weeks. When HMS analysts drew the smoke plumes as they traveled over the eastern part of the U.S., they did not associate the smoke with HYSPLIT points from the current day. There are also instances when HYSPLIT points are analyzed when no smoke plumes are analyzed. An example of how this can occur is when HYSPLIT points are analyzed where there are many small fires but no smoke plume analysis is done due to cloud cover.

1.2.4. Description of United States Geological Survey (USGS) Land Cover Characteristics Database. We assign each HYSPLIT point a land cover type using land-cover classifications from the 2002 North American Land-Cover Characteristics 1 km grid-spacing dataset, created by the National Center for Earth Resources Observation and Science (EROS) as part of the Global Land Cover Characterization Project [40]. Land-cover characteristics are assessed using 1 km Advanced Very High Resolution Radiometer (AVHRR) data between 1992 and 1993 using the methods described in [41]. There are three data types produced from the AVHRR [40]. For this work, we use the GeoTIFF file projected to latitude-longitude grid as a geospatial raster. The latitudinal extent is 18°N-72°N and 66°W-172°W. We assign land-cover classifications to HYSPLIT points based on the nearest grid-point center that is not classified as urban or water. When urban or water is the

closet grid-point, we substituted the assigned land-cover type with the most common other land-cover classification in the surrounding 0.06 degrees (∼5 km) land-cover grid-cells. If all surrounding grid cell land-cover classifications are urban or water, no land-cover assessment is made. No land-cover assignment is made for HYSPLIT points south of 18°N. These mainly occur for HYSPLIT points analyzed in Mexico in April and May. Data are available at: `https://www.sciencebase.gov/catalog/item/535fe572e4b078dca33ae61f`.

1.2.5. DESCRIPTION OF FORWARD TRAJECTORIES. The HYSPLIT (Hybrid Single Particle Lagrangian Integrated Trajectory) trajectory model simulates air parcel movement by wind advection using spatially and temporally gridded meteorology data, and it is used to establish source-receptor relationships [42, 43]. The model computational method is a hybrid between Lagrangian and Eulerian reference frames [43]. HYSPLIT has been used extensively to model the transport of smoke (*e.g.*, [43]). The model executable is available for download on the NOAA Air Resources Laboratory (ARL) webpage (`http://ready.arl.noaa.gov/hyreg/HYSPLIT_pchysplit.php`).

Based on numerical and physical limitations of the model, the error in the location of a trajectory is approximately 15-30% of the distance traveled by the air parcel [44]. The physical error is related to how well the numerical values (*e.g.* u and v winds) represent the true state of the atmosphere. The numerical error arises from integration error, truncation, and the fact that calculations of continuous variables are being done on a discrete grid [44]. We present results generated using the 32 bit Windows PC executable version of the trajectory model (hyts_std.exe) on a Debian Unix cluster at Colorado State University using the WINE (Wine Is Not an Emulator; `https://www.winehq.org`) compatibility layer. A Python based HYSPLIT manager system allows each available core to independently call the HYSPLIT trajectory executable. Six-day (144-hr) forward trajectories are calculated

for each HMS HYSPLIT point using meteorological data from the GDAS (Global Data Assimilation System) archive, which has a time step of 3-hours, horizontal grid spacing of 1° latitude by 1° longitude (∼120 km), and a vertical grid spacing of 23 pressure surfaces between 1000 and 20 hPa. Vertical layers 1-5 are separated by 25 hPa. All higher layers (with the exception of the top layer) are separated by 50 hPa [45]. The GDAS 1-degree archived data is available for download at: `ftp://arlftp.arlhq.noaa.gov/pub/archives/gdas1`. Six-day trajectories are also calculated for each HYSPLIT point in the data domain (excludes Alaska, parts of Canada and Mexico) using meteorological data from the EDAS (Eta Data Assimilation System) archive. This data has a 3-hour time step, horizontal grid spacing of 40 km x 40 km, and 26 pressure surfaces between 1000 and 50 hPa. EDAS40 archived data is available for download at: `ftp://arlftp.arlhq.noaa.gov/pub/archives/edas40`.

Each HYSPLIT point is assigned a duration by an analyst. Trajectories are initialized from each of the HYSPLIT points to represent possible smoke transport over the duration of time that the fire was observed to produce smoke. For HYSPLIT points assigned durations of 0-6 hours, one trajectory is initialized at the middle hour between the first and last hour. For HYSPLIT points with durations of 7-12 hours, two trajectories are initialized at the 25th and 75th percentile of the span of hours over the duration of the fire. For durations of 13-18 hours, three trajectories are initialized at the 20th, 50th, and 80th percentile of the span of hours between start time and end time. For durations of 19-24 hours, four trajectories are initialized at the 20th, 40th, 60th, and 80th percentile of the span of hours. Trajectories are not initialized at the start or finish hours because many of the HYSPLIT points have durations of 24 hours, last multiple days, and are redetected daily, so initializing trajectories at the time fires are detected would overweight early morning hours when HYSPLIT points

are first detected with visible satellite imagery. This is a concern because the most common duration for the HYSPLIT points is 24 hours (Figure 1.4).

Val Martin et al. [46] present a climatology of smoke-plume heights by land biome for North America for the fire seasons of 2002 and 2004—2007. The smoke-plume heights are estimated with the Multi-angle Imaging SpectroRadiameter (MISR) data. MISR has 9 cameras that provide multiple viewing angles, which enables the height of smoke plumes to be estimated. The mean height for all plumes observed is 922 m with a standard deviation of 586 m above mean ground level (amgl) [46]. A given fire can inject smoke into the atmosphere at many different altitudes. We initialize each HYSPLIT trajectory start hour at three different altitudes that span the range of injection heights presented in *Val Martin et al.* [46]: 500, 1500, and 2500 m amgl. For example, we run 12 trajectories for each HYSPLIT point with a duration of 24 hours, *i.e.* there are 4 different start hours with 3 different starting altitudes. In total, 3,925,932 trajectories are associated with the 517,214 HYSPLIT points analyzed between 2007 and 2014. On average there are 7.6 trajectories per HYSPLIT Point. For this analysis, trajectories are only considered for the hours that the calculated height above ground level is > 0, due to inability of trajectories to regain altitude after reaching 0 meters above ground [47]. Some HYSPLIT points have identical latitude, longitude, duration, and start-time information as a way to show that a particular location is producing more smoke. As a result slightly fewer unique trajectory calculations are required since repeated HYSPLIT points can be weighted by the number of times they are duplicated. A summary of the GDAS and EDAS trajectories used in this analysis can be seen in Tables 1.1 and 1.2.

1.2.6. COMBINED ANALYSIS OF HYSPLIT POINTS AND TRAJECTORIES. To build smoke source-receptor relationships for 10 U.S. regions, we define a smoke hour as an hourly

Table 1.1. Trajectories initialized for GDAS1 HYSPLIT runs following the strategy outlined in section 2.4. The third column is the percent of initiated trajectories that reach 0 m within the 6 day calculation. These trajectories are subset and only considered for the hours before they reach 0 m. The minimum number of trajectories that can be run for a HYSPLIT point is 3 (heights); the maximum is 12 (4 start times for a 24-hour duration at 3 initialization heights). The number of trajectories for each year includes all months.

Year	Unique trajectories (total with duplicates)	% reaching 0 m
2007	460209 (546377)	4.2
2008	433671 (482816)	4.9
2009	210089 (226872)	5.9
2010	379568 (417395)	8.2
2011	494375 (563040)	8.3
2012	481438 (563915)	9.7
2013	544463 (598368)	8.5
2014	494320 (527149)	8.0

Table 1.2. Same as Table 1.1 but for EDAS 40 km meteorology data. Fewer trajectories were run due to the limited domain of EDAS 40 km. Parts of Alaska, Northern Canada, and Southern Mexico are not included in this reanalysis dataset. The number of trajectories for each year includes all months.

Year	Unique trajectories (total with duplicates)	% reaching 0 m
2007	412930 (496592)	3.4
2008	356785 (397645)	3.1
2009	156219 (169415)	4.3
2010	255939 (287771)	7.6
2011	451102 (525644)	9.3
2012	427541 (506334)	19.7
2013	375993 (423159)	17.3
2014	215841 (231218)	15.6

latitude-longitude HYSPLIT trajectory location (hereafter trajectory point) that overlaps a HMS smoke plume. Our smoke transport climatology uses these smoke hours to represent the relative abundance of probable smoke transport pathways. Smoke plume overlap assessments are made using a two-step process. The two-step strategy reduces the number of trajectories included in the final analysis that are initialized at heights that do not represent the injection height of a given fire. Spatial overlap analysis is performed using the R Sp package over() function [48] (please see appendix for code and more information).

1) We confirm that at least one of the first 49 trajectory points (2 days + start hour) overlap a smoke plume analyzed for the same two dates. If none of the first 49 trajectory points overlap a smoke plume, the trajectory is immediately discarded. If any of the first 49 trajectory points overlap a smoke plume, plume-overlap analysis is performed for the entire trajectory (145 trajectory points). We validate a trajectory using two days of smoke plumes rather than one because smoke plumes only represent smoke perimeters during daylight hours. Validating the trajectories using only the first day of smoke would make it more likely that HYSPLIT points analyzed in the morning would overlap a plume than HYSPLIT points analyzed later in the day. By also including the next-dates smoke plumes in the test, fires that start in the afternoon are not evaluated more stringently than early morning HYSPLIT points because fires that start later in the day are less likely to overlap the matching dates smoke plumes then fires that start early in the day due to the visible daylight imagery limitations placed on analyzed smoke plumes (please see appendix for additional information). However, our results do not change significantly when validating trajectories with only the first-days smoke plumes.

2) When trajectories meet our first criteria, a point over polygon calculation is done for each matching date trajectory point and smoke plume. For example, if a trajectory starts on August 1, the trajectory points with that date will be tested for overlap with smoke plumes analyzed August 1, while the last trajectory points on August 6 are assessed using August 6 smoke plumes. One of the main weaknesses of this approach is that smoke-plume boundaries are only representative of smoke-plume perimeters during daylight hours, while trajectories exist at both day and night. Another weakness is that some analyzed smoke plumes on individual days are very large; during extreme-smoke events smoke plumes can cover most of the continental U.S., and thus this criteria is not always particularly stringent.

Our choice to only use a single-day overlap strategy is designed to reduce the number of smoke hours that result from a trajectory overlapping a smoke plume far downwind where the HYSPLIT point used to initialize the trajectory may not have contributed any smoke (discussed in greater detail in Section 1.3.5).

Each trajectory smoke hour is associated with the source region and land cover classification from its initialization point, allowing a source region and land classification analysis of the total number of smoke hours impacting or emanating from a region. Our methodology does not provide information about smoke concentration (μg m^{-3}), and we have not placed additional altitude constraints on the trajectories. We do not place a maximum altitude criterion on trajectories, so we can examine smoke transport pathways at all altitudes.

1.3. Results

1.3.1. HYSPLIT Point Analysis. Figure 1.2 presents the total number of HYSPLIT points by aggregated regions. The largest numbers of points are identified in the Southwest, Northwest, Northwest Territories and Southeast regions. Figure 1.3 presents the seasonality of North American HYSPLIT points (n = 517,214) between 2007-2014, and it shows that the majority of HYSPLIT points are identified between June and August. During these most active months, the dominant land cover classification for HYSPLIT points is evergreen needle leaf forest, followed by a nearly equal share of scrubland and mixed forest.
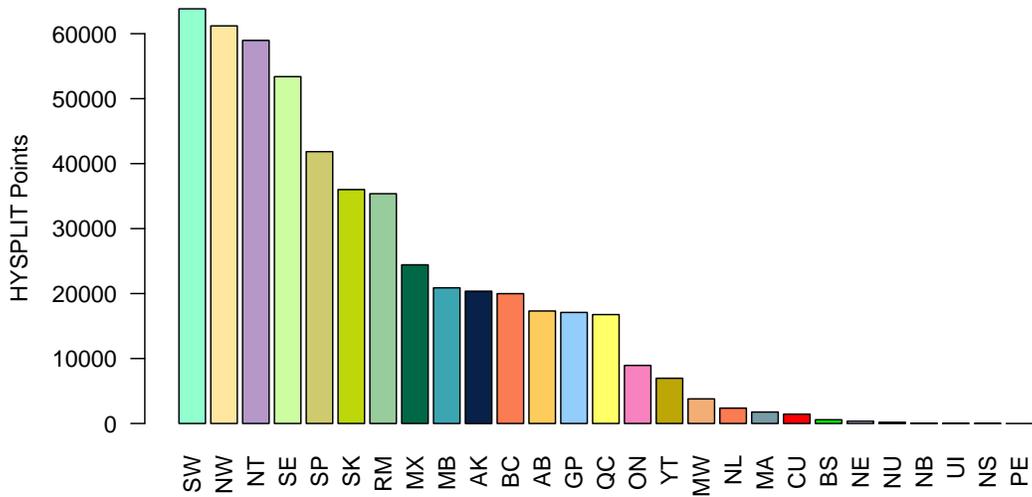
FIGURE 1.2. Total HMS HYSPLIT Point detections for all regions between 2007-2014. Region colors and abbreviations are defined in Figure 1.1.
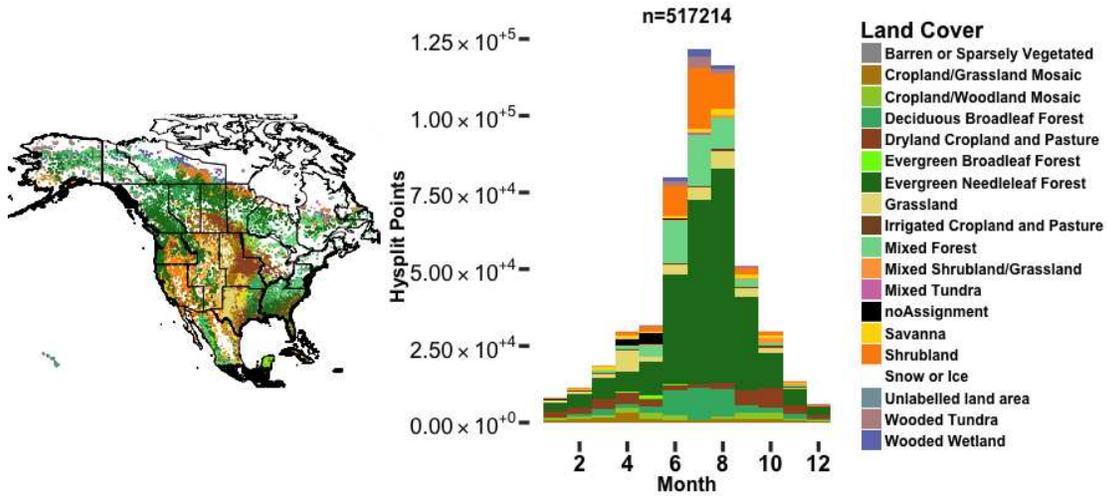
FIGURE 1.3. Location (left), and seasonality (right) and land cover classification assignment (color) for all HYSPLIT points analyzed over North America between 2007-2014. No land cover assignment is made for Southern Mexico (MX) and U.S. Islands (UI) due to the latitudinal range of land cover data.

The first panel of Figure 1.4 shows the locations, seasonality, and duration of HYSPLIT points in the ten U.S. regions in Figure 1.1 between 2007 and 2014. The same 3-panel Figure for all regions shown in Figure 1.1 are available in the appendix. Land cover classification is indicated by color using the same scale in Figure 1.3. Because of the way HYSPLIT points are

generated, the duration field is roughly proportional to the amount of smoke produced for a given HYSPLIT point. HYSPLIT points with durations of 24 hours represent large wildfires. HYSPLIT points with shorter durations represent smaller fires, and often occur on cropland. At the national scale shown in Figure 1.3, grassland and cropland land cover classifications make up a small proportion of the total HYSPLIT points throughout the year. However at the regional scale, grassland and cropland HYSPLIT points can represent a significant fraction or even dominate the total number of HYSPLIT points. This is the case for the Great Plains (Figure 1.4B), Midwest (Figure 1.4D) and Southern Plains (Figure 1.4H). These regions also have the fewest HYSPLIT points analyzed in the summer months, whereas regions dominated by evergreen forest have a minimum in the winter. The Southwest (Figure 1.4J) has more HYSPLIT points then any other U.S. region, followed by the Northwest (Figure 1.4F), Southeast (Figure 1.4H), and Southern Plains (Figure 1.4H). The Northeast (Figure 1.4E) has the fewest HYSPLIT points, which occur mostly on cropland in southern New Jersey. The Mississippi river valley has some of the most densely analyzed HYSPLIT points in the U.S, and these points are located most commonly on forest and cropland. These points are split between the Southeast (Figure 1.4H) and Southern Plains (Figure 1.4I) regions.

(A)

(B)

Land Cover
- Barren or Sparsely Vegetated
- Cropland/Grassland Mosaic
- Cropland/Woodland Mosaic
- Deciduous Broadleaf Forest
- Dryland Cropland and Pasture
- Evergreen Broadleaf Forest
- Evergreen Needleleaf Forest
- Grassland
- Irrigated Cropland and Pasture
- Mixed Forest
- Mixed Shrubland/Grassland
- Mixed Tundra
- noAssignment
- Savanna
- Shrubland
- Snow or Ice
- Unlabelled land area
- Wooded Tundra
- Wooded Wetland

18

Mid Atlantic
n=1765

(C)

Midwest
n=3787

(D)

Northeast
n=343

(E)

Land Cover

Barren or Sparsely Vegetated	Evergreen Needleleaf Forest	Savanna
Cropland/Grassland Mosaic	Grassland	Shrubland
Cropland/Woodland Mosaic	Irrigated Cropland and Pasture	Snow or Ice
Deciduous Broadleaf Forest	Mixed Forest	Unlabelled land area
Dryland Cropland and Pasture	Mixed Shrubland/Grassland	Wooded Tundra
Evergreen Broadleaf Forest	Mixed Tundra	Wooded Wetland
	noAssignment	

Northwest
n=61202

(F)

Rocky Mountains
n=35371

(G)

Southeast
n=53389

(H)

Land Cover

Barren or Sparsely Vegetated	Evergreen Needleleaf Forest	Savanna
Cropland/Grassland Mosaic	Grassland	Shrubland
Cropland/Woodland Mosaic	Irrigated Cropland and Pasture	Snow or Ice
Deciduous Broadleaf Forest	Mixed Forest	Unlabelled land area
Dryland Cropland and Pasture	Mixed Shrubland/Grassland	Wooded Tundra
Evergreen Broadleaf Forest	Mixed Tundra	Wooded Wetland
	noAssignment	

20

FIGURE 1.4. HYSPLIT points locations and land cover classification represented by color (left), total number of occurrence by month and land cover classification (middle), duration and land cover classification (right). Please refer to the legend in figure 1.3 for land cover classification. Non U.S. regions are available in the appendix.

There are more HYSPLIT points analyzed in the Southern Plains (Figure 1.4I) than there are in the Rocky Mountains region (Figure 1.4G). However, this does not indicate that the Southern Plains generate more smoke, because the number of points does not include information on fire duration (discussed in Section 1.2.1). The total smoke produced in a region is proportional to the number of HYSPLIT points multiplied by their respective

durations. Most of the fires in the Southern Plains have durations less then nine hours while the most-common duration of HYSPLIT points in the Rocky Mountains region is 24 hours. Alaska is an extreme example. Almost all of the HYSPLIT points in Alaska have 24-hour durations, there are very few HYSPLIT points on cropland and thus very few HYSPLIT points with durations less than 10 hours. The Hazard Mapping System uses visible satellite imagery that allows the detection of very small smoke-producing fires. The number of HYSPLIT points in the Southeast U.S. is qualitatively consistent with the view that small smoke-producing fires are common in the Southeast U.S. [22, 16].

1.3.2. SMOKE-TRANSPORT ANALYSIS. Figures 1.5 and 1.6 show the total number of smoke hours produced by and over each region. Years with elevated fire activity in each region can be easily identified using Figure 1.5. This figure provides context for isolated case studies of smoke transport associated with extreme periods including the summer 2013 Quebec wildfires [*e.g.* ([49]), the summer 2012 wildfires in the Rocky Mountain region (*e.g.* [29]), and the 2008 California wildfires (SW Region) (*e.g.* [50]). Figure 1.7 shows a summary of smoke transport between 2007 and 2014 for the months of June, July, August, and September. Each row of Figure 1.7 summarizes the smoke hours impacting a region as well as the geographic extent of the impact of the smoke hours produced by the region. Table 1.3 shows the total smoke hours produced by and over each U.S. region and the differences between using the GDAS and EDAS meteorology datasets for the trajectory calculations. The biggest difference can be seen in Alaska, since only the southeast portion of the state is within the EDAS domain. Regions that receive smoke from high latitudes generally have more smoke over them when the GDAS data is used (Mid Atlantic, Midwest, Northeast). Regardless of the meteorological dataset used for the trajectory calculations, the Northwest, Rocky Mountains, and Southwest are the largest smoke-source regions. They also each have

more smoke hours over their regions when the EDAS data is used. This indicates that the higher-resolution EDAS data does a better job overlapping smoke plumes within the first 49 hours, increasing the number of smoke hours that regions contributed to a region from fires within that region.

Total Smoke Hours Produced
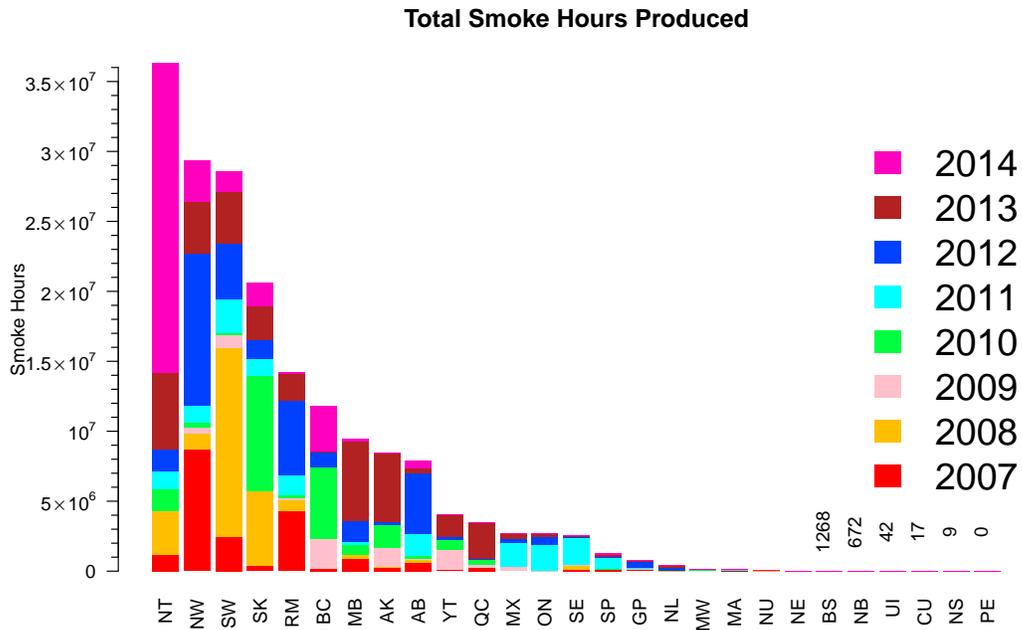
FIGURE 1.5. Total number of smoke hours produced (anywhere) by each region for months June-September between 2007 and 2014 using GDAS1 meteorology.

Total Smoke Hours Over Region
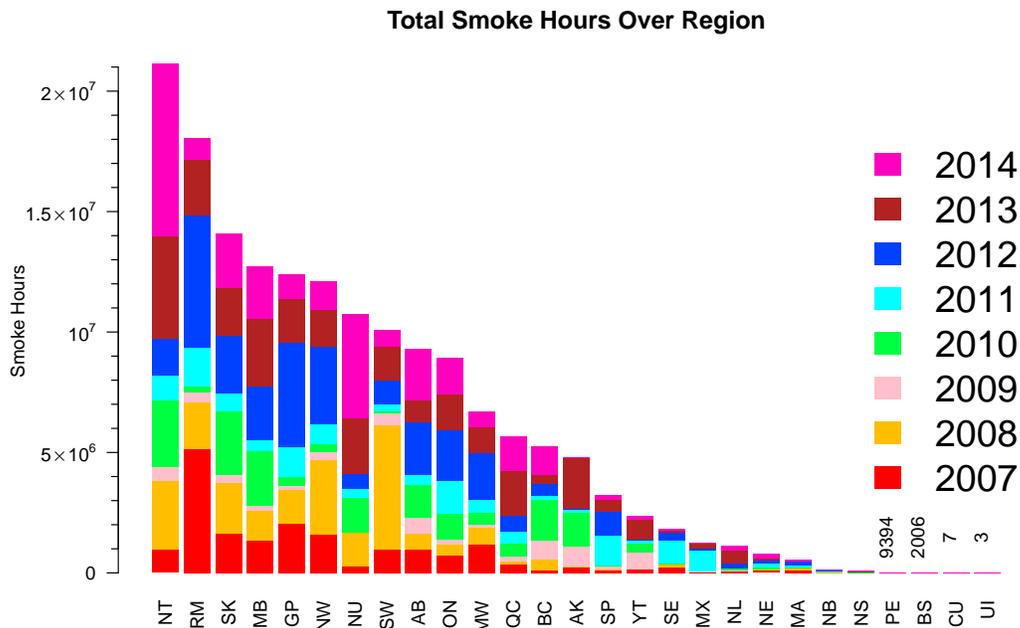
FIGURE 1.6. Total number of smoke hours over each region sorted from highest to lowest for months June-September between 2007 and 2014 using GDAS1 meteorology.

TABLE 1.3. Summary of smoke hours produced by and over each U.S. region for each meteorology dataset. Numbers show summer totals between 2007-2014.

Region	GDAS Produced	EDAS Produced	GDAS Over	EDAS Over
Alaska	8,412,383	2	4,785,421	10,685
Great Plains	781,704	757,094	12,382,290	10,539,018
Mid Atlantic	115,583	108,265	563,442	497,545
Midwest	143,343	145,691	6,684,312	5,282,670
Northeast	6,723	8,784	782,790	541,174
Northwest	29,360,026	29,120,925	12,083,168	12,890,475
Rocky Mt.	14,252,104	13,703,437	18,066,760	19,064,474
Southeast	2,553,898	2,378,645	1,799,372	1,745,907
Southern Plains	1,286,888	1,296,988	3,206,060	3,437,671
Southwest	28,552,926	28,255,271	10,069,858	10,084,839

1.3.3. RECEPTOR REGION SMOKE HOUR CLIMATOLOGY. The first column of Figure 1.7 shows the total smoke hours in a region broken down by the region of origin and land cover classification. The second column of Figure 1.7 shows the distribution of the age

(hour in HYSPLIT trajectory) of smoke hours in a region separated by the region of origin. Together these columns show what regions contribute smoke to each of the ten U.S. regions for summers 2007 to 2014. The total number of smoke hours (in millions) in each U.S. region in descending order is: the Rocky Mountains (18), Great Plains (12.4), Northwest (12.1), Southwest (10.1), Midwest (6.7), Alaska (4.8), Southern Plains (3.2), Southeast (3.5), Northeast (0.8), and Mid Atlantic (0.6). It is important to consider the age of smoke hours when assessing what potential they have for increasing the concentration of trace species in the atmospheric column. For example, older smoke hours are likely associated with lower $PM_{2.5}$ concentrations than smoke hours that are only a few hours old. Put into a regional context example; the Great Plains region (Figure 1.7B) has more smoke hours than the Northwest (Figure 1.7F) and Southwest (Figure 1.7J); however, the average age of smoke hours over the Great Plains is ∼one day older then in the Northwest and Southwest.

26

28
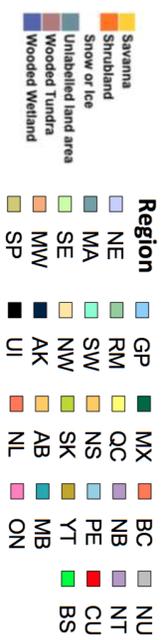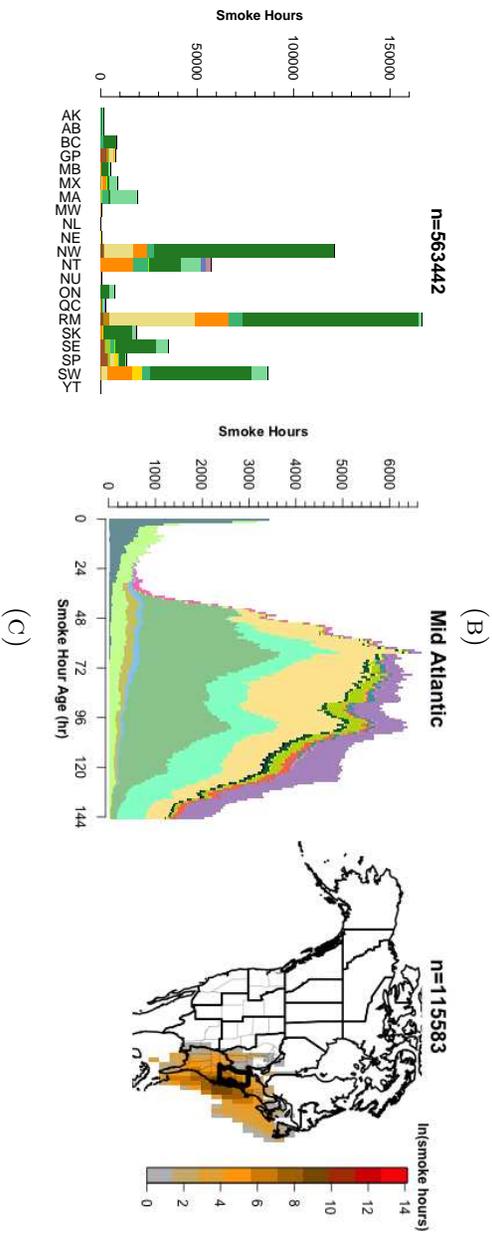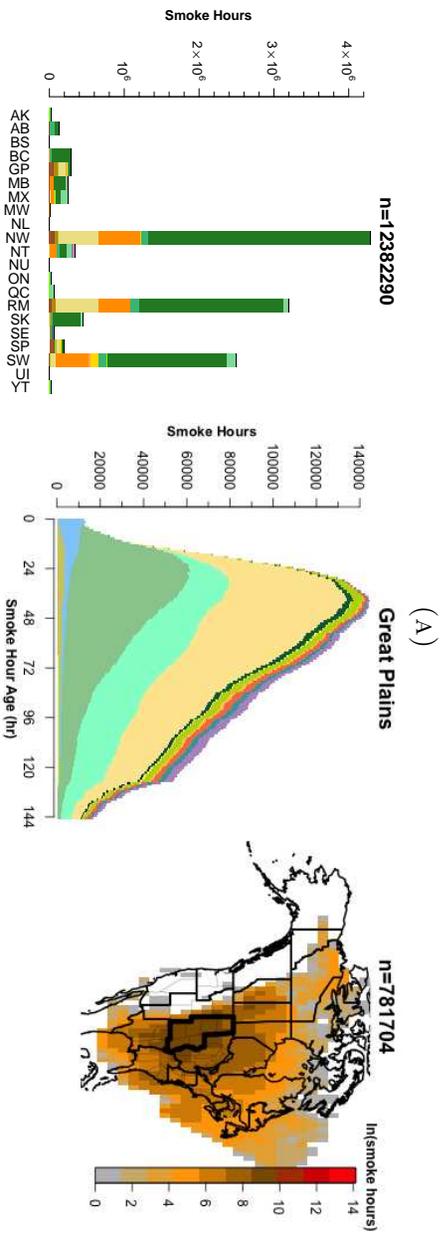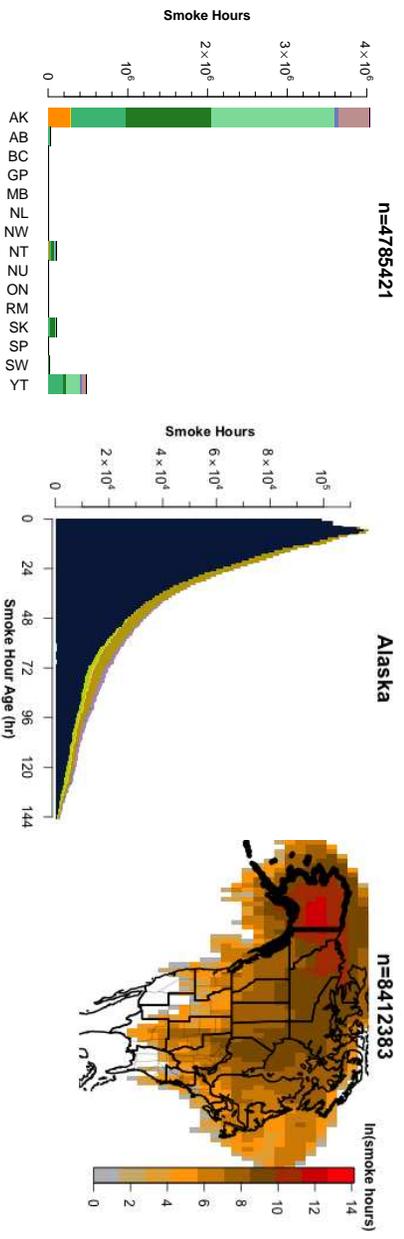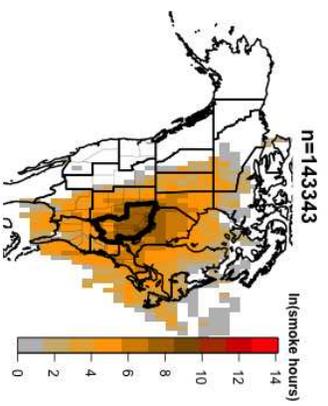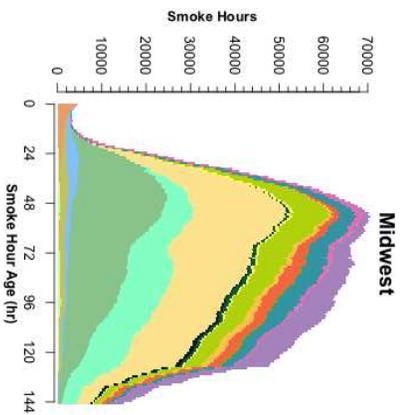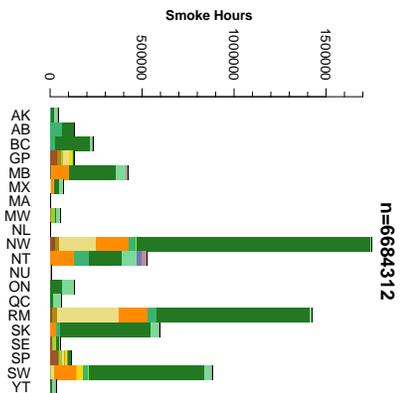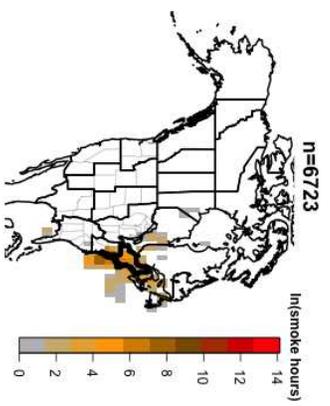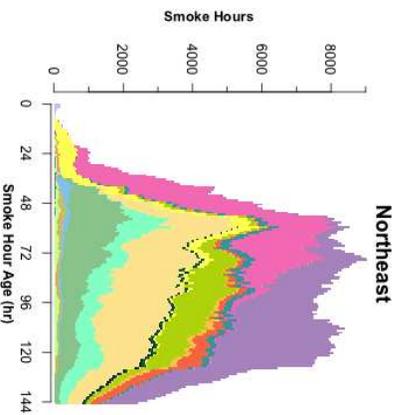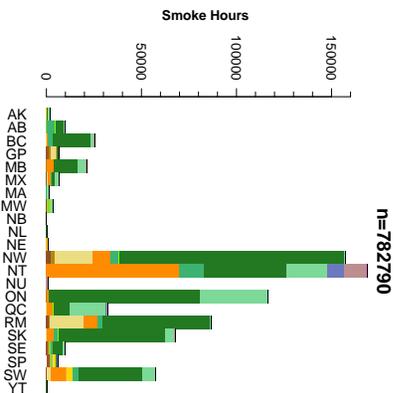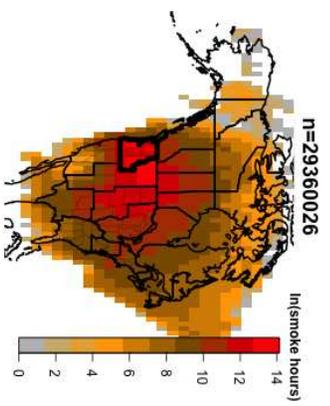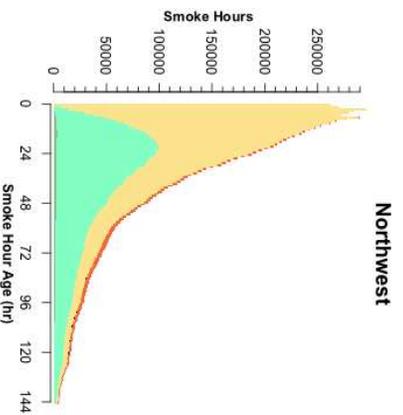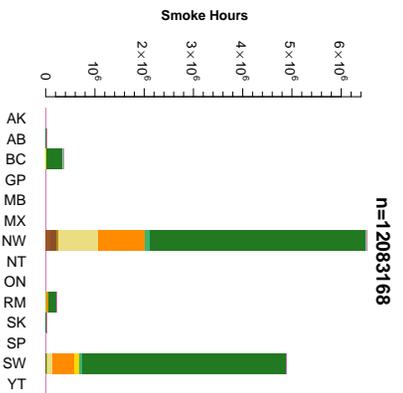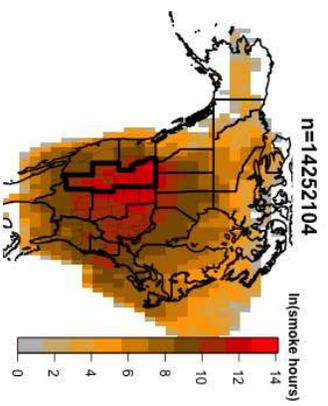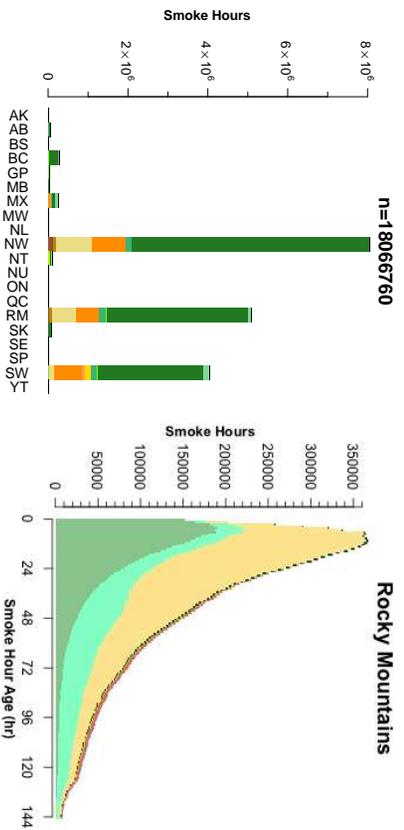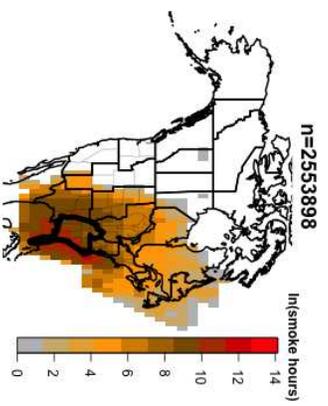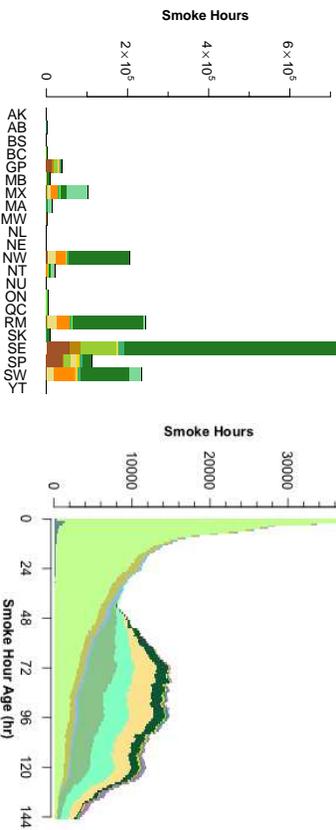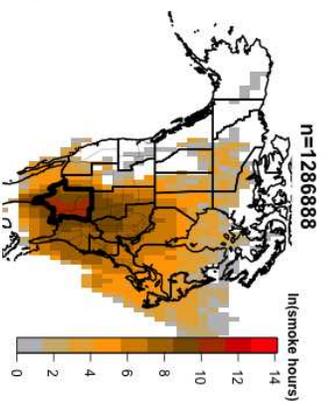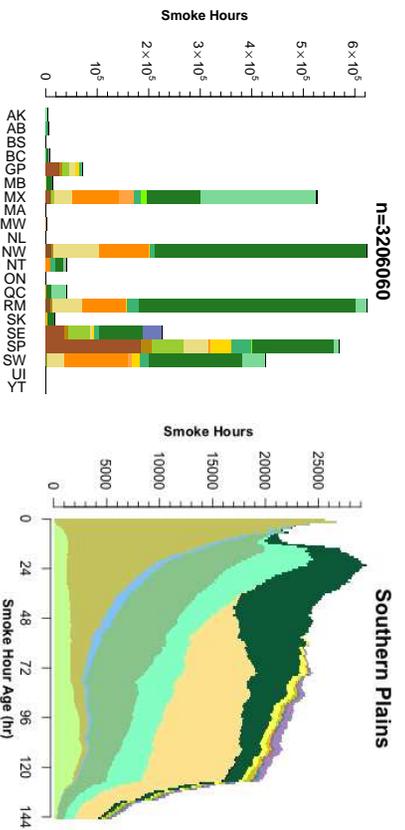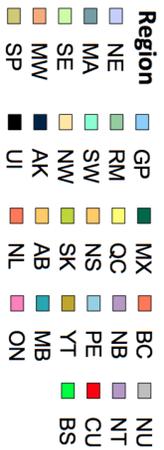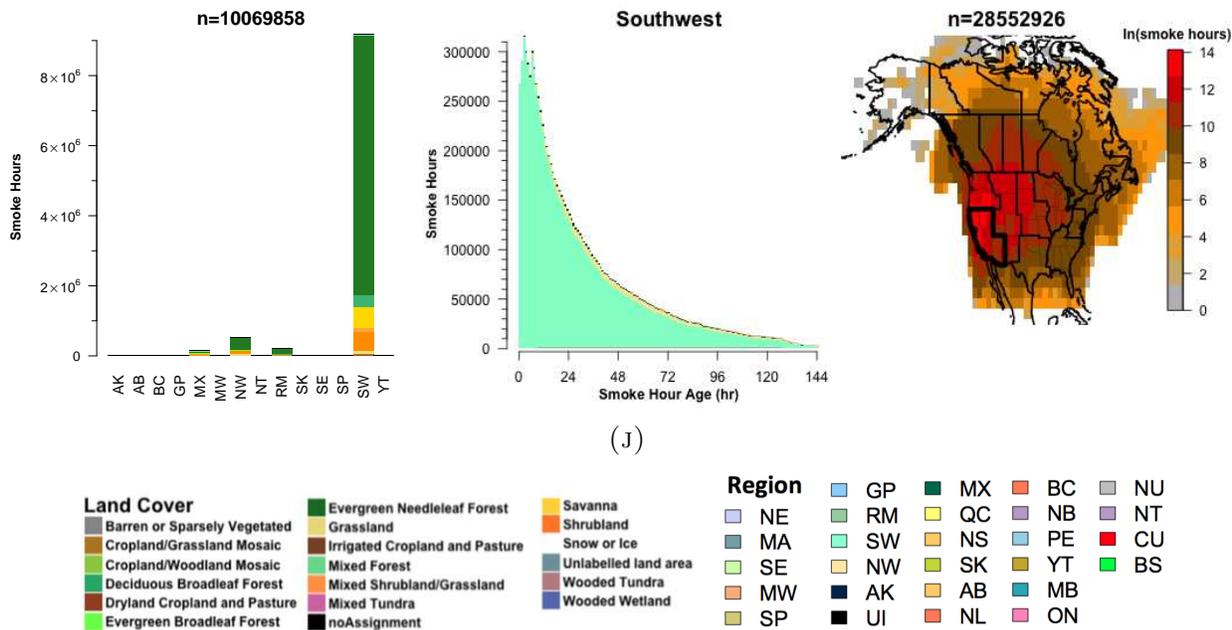
FIGURE 1.7. Summertime (June-September) regional smoke hour transport summaries. Column 1 (left) shows the number of smoke hours in the region by source region and land cover classification with the same color scheme as Figure 1.2. Only source regions with non-zero smoke hour contributions are shown. Column 2 (middle) shows the distribution of smoke age segregated by source region using the colors from Figure 1.1. Column 3 (right) shows the count of smoke hours produced by a region on a 2° by 2° degree grid with a consistent colorbar for all regions. The grid spans 18-180ºW and 18-90ºN, a domain covering all five sectors where HMS analyzes smoke plumes (only a subset plotted). Shaded values are the natural log of the number of smoke hours in each grid cell (min=1, max=1.2 million). All Figures generated using GDAS1 meteorology data for the months June-September 2007-2014.

Regions with the largest number of HYSPLIT points tend to contribute the largest proportion of the total smoke hours within their own borders. The smoke impacting the Southwest (Figure 1.7J) and Alaska (Figure 1.7A) originates almost exclusively within the regions themselves. Smoke hours over the Northwest (Figure 7F) are nearly equally likely to be associated with fires from the Northwest or Southwest. Regions with comparatively little local fire activity tend to have a diverse set of source regions contributing to their total smoke hour budgets; examples include the Southern Plains (Figure 1.7I), the Northeast (Figure 1.7E),

and the Midwest (Figure 1.7D). The Northwest, Southwest and Rocky Mountains dominate most other regions total smoke hours. The only regions where the three largest contributors of smoke hours are not the Southwest, Northwest, and Rocky Mountains, are Alaska (Figure 1.7A), the Southeast (Figure 1.7H), Northeast (Figure 1.7E), and the Southern Plains (Figure 1.7I). The only regions that contribute more smoke hours to their own total budget then any outside region are Alaska (Figure 1.7A), the Southwest (Figure 1.7J), the Northwest (Figure 1.7F) and the Southeast (Figure 1.7H). These regions may have more control over the total amount of smoke within their borders then other regions in this analysis.

We also produced a version of Figure 1.7 using trajectories run using 40 km grid EDAS reanalysis data, and the results are nearly identical to patterns presented in Figure 1.7 with the exception that smoke hours originating in the high latitude regions (Alaska, the Northwest Territories, and Nunavut) are eliminated due to the latitudinal range of EDAS. The version of this figure generated with EDAS meteorology is available in the appendix.

1.3.4. SMOKE-HOUR TRAJECTORY-PATHWAY CLIMATOLOGY MAPS. The third column of Figure 1.7 shows a map of the total smoke hours produced by each U.S. region for June to September between 2007 and 2014. The definition of the smoke hours presented in these maps is the same as in Section 1.2.6 except instead of counting the total smoke hours overlapping a region, individual 2° x 2° grid cells are used. The values plotted in Figure 1.7 are the natural logarithm of the total number of smoke hours within each grid box (hereafter smoke hour impact). These maps show the average smoke-transport pathways for fires in each source region (maps for all regions are available in appendix).

We focus our discussion of Figure 1.7 on the regions that have the largest numbers of smoke hours. Smoke produced by fires in California is transported over Northern California and Eastern Oregon (Southwest Region, Figure 1.7J, right column), based on our analysis of

the HYSPLIT points these fires occur primarily in evergreen needleleaf forests. Smoke originating from fires in the Northwest is transported most frequently over Eastern Washington, Eastern Oregon, Northern Idaho, and Montana (Figure 1.7F, right column). Smoke from fires in Alaska impact every U.S. and Canadian region. The dominant transport pathway for the smoke crosses Alaska, the Northwest Territories, and Yukon Territory (Figure 1.7A, right column). Smoke traveling from Alaska to Texas has been observed previously [51], but the right column of Figure 1.7A shows that this situation is relatively rare when viewed in an aggregate context. These maps look nearly identical whether they are produced using HYSPLIT trajectories driven by 40 km EDAS or lower resolution GDAS reanalysis data. The only significant difference between the maps produced using each dataset is the northward extent of the smoke-hour impact due to the 60°N cutoff of the EDAS domain. Similar maps for each region in Figure 1.1, individual years, and EDAS trajectories are available in the appendix.

1.3.5. SMOKE-HOUR FALSE POSITIVES. Our intent is for smoke hours to represent the advection of smoke; however, trajectory points can overlap smoke plumes that are not associated with the correct fires, and this creates smoke hours that are false positives. There are a number of possible scenarios that could lead to a smoke-hour false positive; to illustrate this we will highlight two possible scenarios here. 1) Trajectory points could be validated on the first two days when they overlap a smoke plume that has been advected from upwind. These trajectory points could then travel with that existing smoke plume and be incorrectly classified as smoke hours. This scenario does not necessarily mean that the fire associated with the trajectory does not produce smoke. It is possible that the trajectory misses the plume created by its associated HYSPLIT point due to being initialized at the wrong injection height. 2) Trajectory points could overlap smoke for the first day, then no smoke for a

31

couple of days, then overlap an unrelated smoke plume very far downwind. Both of these types of false positives have been observed in developing our definition for smoke hours.

Currently the best way we have of identifying false-positive smoke hours is sorting trajectories into smaller aggregates (individual years and seasons) and observing the heat maps (Figure 1.7, column 3) associated with these subsets of trajectories. The second type of false positive described above would be visible in these types of maps as disconnected smoke hours. The long-range transport of smoke hours originating in the Northwest during winter months provides ideal conditions to test the methods described in section 1.2.6. In winter months there are fewer smaller fires and fewer smoke plumes analyzed by HMS. Additionally, fires that occur in the winter have lower smoke-injection altitudes on average compared to their summer counterparts [46, 36]. All of these factors will tend to reduce the long-range transport of smoke during winter months. The Northwest has very-little fire activity for the first three months of the year (Figure 1.4F). In contrast, January, February, and March have a significant amount of local fire activity in the Southeast as indicated by the number of HYSPLIT points for these months (Figure 1.4H). These conditions create an ideal environment for trajectories that originate in the Northwest to travel to the Southeast without advecting any smoke and create false positive smoke hours. When we plot the heat map for the Northwest for these months we see two disconnected hot spots (Figure 1.8), which almost certainly represent false positives far downwind. Figure 1.8 shows how the Northwest contributes smoke hours to the Northwest and Southeast without impacting the Rocky Mountain region. This strongly suggests that the smoke hours over the Southeast are false positives. During the summer there is very little local fire activity in the Southeast (Figure 1.4H) so it is likely this problem does not dominate the summer smoke hour transport climatologies shown in Figure 1.7. This type of evidence for false positive smoke hours is

not apparent in summertime data; however, the very large smoke plumes analyzed during the summer may not allow for disconnection. Even for individual-year heat maps, we do not observe disconnected areas of smoke impact far downwind of source regions. Winter smoke-hour transport figures for all regions are available in the appendix.

HMS analysts observe that fires in the Great Plains, Southern Plains, and Midwest, generally produce short duration smoke plumes that quickly dissipate. It is unusual for smoke produced in these regions to persist long enough to reach areas of Northern Canada or the Canadian Maritimes. Thus the long range transport smoke hour impact from these regions shown in Figures 1.7B, 1.7D, 1.7H may be examples of trajectories overlapping HMS smoke plumes that originated in other source regions.
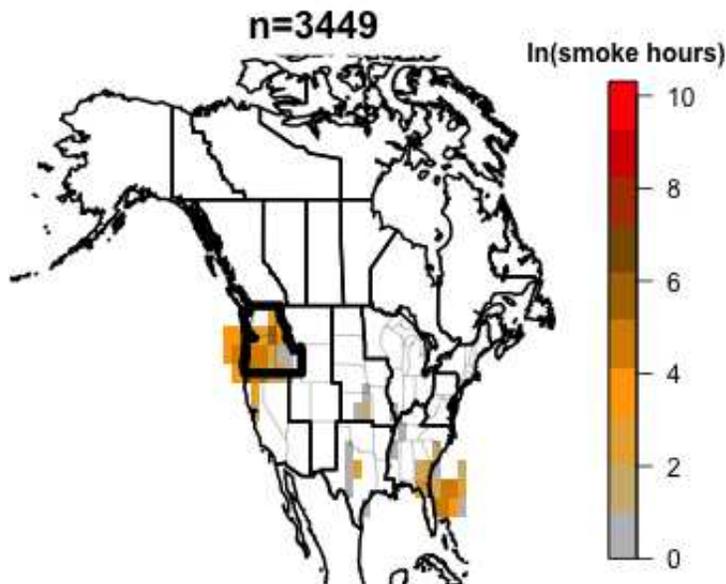
FIGURE 1.8. Total number of smoke hours produced by the Northwest region for months January-March between 2007 and 2014 using GDAS1 meteorology.

1.4. Conclusions

This work presents a new dataset, which previously only existed in a daily operational context. We define smoke hours, a quantity we believe to be proportional to total column smoke, by linking smoke plumes to fires using HYSPLIT trajectories. This work shows that the Southwest, Northwest, and Northwest Territories produce more smoke than any other North American region by measure of the number of HYSPIT points analyzed, the duration of those HYSPLIT points, and the total number of smoke hours produced (Figures 1.4, 1.5, and 1.7). This dataset shows that there is a significant amount of fire activity in the Southeast, particularly along the lower Mississippi valley. Unsurprisingly, the largest smoke source regions are located in the west, while receptor regions for smoke are primarily located in the east. A main distinction between source and receptor regions is the age of smoke present over a region. The majority of smoke located over source regions is less than 24 hours old. Receptor regions have very little smoke less than 24 hours old with most smoke in the column older than 48 hours. The Southeast is a unique exception. There is an abundance of fresh smoke (peak near zero hours in Figure 1.7H) due to the many small fires within the region, a minimum near 36 hours then a gradual increase in the frequency of aged smoke (>48 hours). We also show that the Northeast and Mid Atlantic receive more smoke from fires in Canada than regions in the U.S. Midwest.

We present a smoke transport climatology for the summer wildfire season. Based on our metric of smoke hours, the U.S. regions that produce the most smoke are the Northwest, Southeast, and Rocky Mountains. Heavily populated locations in the eastern U.S. (Northeast, Mid Atlantic) receive over 100,000 times more smoke hours than they produce. The upwind sources that contribute to the aged smoke (48 hours or older) present over the eastern U.S. are the Rocky Mountains, the Northwest, and the Northwest Territories.

The Southwest, Northwest, and Rocky Mountains are already the biggest contributors of smoke to the atmospheric column over the U.S. These are also the regions that are expected to experience the greatest increase in fire activity as the climate warms [17, 52, 20, 53–55]. Our analysis implies that predicted increases in fire activity in these regions have the potential to greatly impact air quality throughout the U.S. airshed.

CHAPTER 2

Smoke in the City: How Often and Where Does Smoke Impact Summertime Ozone in the United States?
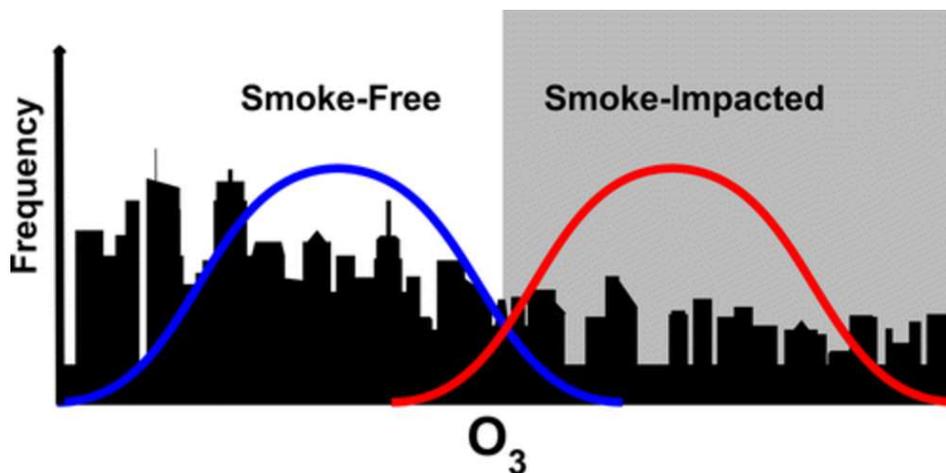
Figure 2.0. Cover Art for Environmental Science and Technology Publication DOI: 10.1021/acs.est.5b05218

2.1. Introduction

Ground-level O_3 is one of the six criteria pollutants the Environmental Protection Agency (EPA) regulates in order to comply with National Ambient Air Quality Standards (NAAQS) mandated under the Clean Air Act [56] [1]. Ground-level O_3 is associated with a suite of respiratory-related adverse health effects. Exposure to O_3 triggers asthma [57], decreases lung capacity, weakens the immune system and inflames lung tissue, and it can lead to increased mortality [58, 59]. Elevated O_3 also oxidizes plant tissue, reduces growth rates within plants, and causes visible damage to leaves and needles [60]. In 2009, 122 million

[1] This chapter contains published work. Citation: Smoke in the City: How Often and Where Does Smoke Impact Summertime Ozone in the United States? Steven J. Brey and Emily V. Fischer Environmental Science & Technology 2016 50 (3), 1288-1294 DOI: 10.1021/acs.est.5b05218

U.S. residents lived in regions that violated the O_3 standard, and it is clear that adverse affects from O_3 occur at exposures well below current air quality standards [61].

The contribution of wildfires to O_3 production in the U.S. is poorly understood. A handful of studies show that ground level O_3 abundance may be increased by fire activity in the U.S. Jaffe et al. [3] showed that western U.S. background O_3 mixing ratios are positively correlated with wildfire burn area. A number of case studies ([3, 62, 51, 63, 26, 64–66]) have observed enhanced O_3 production in specific wildfire smoke plumes.

There is evidence that the frequency, intensity, and total burn area of wildfires will continue to increase as the climate changes [67, 19, 68, 18, 17]. The expected change varies substantially by ecosystem type, contributing to regional differences and uncertainty regarding the impact on ground-level O_3 [20]. A century of fire exclusion policies and land use changes has also led to significant fire problems across the U.S. Suppression of natural fires has allowed excess fuel to accumulate in forests, and the fires that do occur are now larger and more intense [69–71, 17]. The impact of wildfires is not limited to western states since smoke plumes rich with O_3 precursors can travel thousands of kilometers [72, 73]. While they tend to be substantially smaller than their wildfire analogs, smoke plumes from agricultural burning also produce O_3 precursors [25].

The primary objectives of this chapter are to (i) identify the frequency that smoke impacts EPA Air Quality System (AQS) O_3 monitoring sites across the contiguous U.S. and (ii) estimate a location-specific conservative expected O_3 response to the presence of smoke. This work complements individual case studies because we show how EPA AQS O_3 monitoring sites are influenced by smoke nation-wide. Additionally, we show which U.S. cities have the largest number of MDA8 (defined as maximum daily 8-h average) O_3 standard exceedance days in the presence of smoke. We selected a subset of the data to examine issues related to

smoke concentration and temperature, and this level of analysis detail allows us to conclude that the O_3 response is related to the presence of smoke rather than these other factors.

2.2. DATA SOURCES AND METHODS

2.2.1. OVERVIEW OF EPA AIR QUALITY SYSTEM (AQS). We use O_3 mixing ratios and particulate matter ($PM_{2.5}$) concentrations archived by the EPA AQS. AQS is a central repository for ambient monitoring data collected by the EPA, state, and local environmental agencies. The goal of this repository is to log the concentrations of criteria pollutants designated by the National Ambient Air Quality Standards (NAAQS) [56].

2.2.2. EPA AIR QUALITY SYSTEM (AQS) OZONE DATA. Site-specific metadata for ultraviolet photometric O_3 analyzers is not stored in the AQS repository, but reported O_3 measurements must comply with calibration and quality assurance standards [56, 61, 60]. Current Federal Equivalent Methods (FEM) for ultraviolet photometric O_3 analyzers specify measurement noise be no greater than 5 ppbv and that 12 hr zero drifts be less than 20 ppbv [74]. *Ollison et al.* [75] showed that conventional ultraviolet photometric O_3 analyzers can overestimate O_3 mixing ratios when measurements are impacted by high concentrations of polycyclic aromatic hydrocarbons, which can be emitted by the incomplete combustion of wood [75, 76]. However, this overestimate was never larger than 6 ppbv for daily maximum 8-h average O_3 mixing ratios and was infrequently larger than 2 ppbv [75]. The 6 ppbv MDA8 difference was observed only once in the three month campaign. *Dunlea et al.* [77] investigated sources of interference in ultraviolet photometric O_3 analyzers in an urban environment. They concluded that particles are the most likely source of observed interferences. Using measurements of ambient hydrocarbons they showed that aromatic volatile organic compounds (VOCs) do not cause significant interference in ultraviolet photometric

O_3 analyzers, but they were not able to exclude the possibility that oxidized and/or nitrate aromatics cause interference. As discussed in Section 2.2.8, we tested the robustness of our results using different particle concentration thresholds and we find no evidence that smoke-driven interferences with the ultraviolet photometric analyzers are solely responsible for the patterns present in the O_3 data. Specifically, the differences we observe between smoke-impacted and smoke-free O_3 distributions (defined below) are larger than expected from interference driven by particles, and our results do not change when we set different particle concentration thresholds.

As of 2014, the EPA AQS contains daily MDA8 O_3 mixing ratios for over 1200 monitors. There is no standard start or end date for measurements made by the individual monitors in the data set. We include every monitor with at least two summers worth of MDA8 values in our analysis. There are locations where there are gaps in the daily data time series; however, our analysis does not require the variable time series be continuous (section 2.8). AQS only reports MDA8 values when at least 75% of the data are available to compute an average (6 of 8 h) [56]. We selected a subset of EPA AQS O_3 monitors to include only those that were co-located with $PM_{2.5}$ measurements, narrowing the number of unique monitors to 506.

2.2.3. EPA AIR QUALITY SYSTEM (AQS) $PM_{2.5}$ DATA. $PM_{2.5}$ measurements (`http://aqsdr1.epa.gov/aqsweb/aqstmp/airdata/download_files.html#Daily`) are made using 24-hr and 1-hr sample durations, which use gravimetric and beta-attenuation measurement techniques [56]. Both federal equivalent method and non-federal reference method sites are used in order to maximize $PM_{2.5}$ data availability. $PM_{2.5}$ concentrations are used to help determine the presence of smoke at the surface-monitoring site with the co-located O_3 monitors.

2.2.4. NESDIS Hazard Mapping System Fire and Smoke Product (HMS).
Please see section 1.2.2 for a detailed description of the HMS Fire and Smoke product.

2.2.5. In Situ Temperature Data: Temperature data was taken from the Automated
Surface Observing System (ASOS) stations located at over 900 locations across the U.S.
ASOS is jointly operated by the National Weather Service (NWS), the Federal Aviation
Administration (FAA), and the Department of Defense (DOD) [78]. Data are recorded at
1-hr time intervals. Daily maximum and average temperatures were used to estimate days
with conditions conducive to active photochemistry. Sky condition data (clear, partly cloudy,
cloudy, etc.) were used to estimate days with conditions conducive to active photochemistry.

2.2.6. ECMWF Reanalysis Data. The European Center for Medium-Range Weather
Forecasting (ECMWF) interim reanalysis 00Z 2-m temperature data are used as a second
way to assess the ambient temperature associated with the O_3 measurements. The temper-
ature data are available on a 0.75° by 0.75° (80 km) grid, and the data assimilation uses
4-dimensional variational analysis [79]. Gridded temperature data are used to test if the
results are sensitive to different temperature data sources. We find that the results are not
sensitive to whether in situ or reanalysis temperature data are used.

2.2.7. EPA National Emissions Inventory (NEI) 2008. The NEI contains emis-
sion estimates for NAAQS criteria gases on a 1° x 1° grid [56]. We use the 2008 nitro-
gen oxides (NO_x) emissions to distinguish O_3 monitors located in high-NO_x (>90 per-
centile nationwide) versus low-NO_x emission regions (everywhere else). There has been
a decrease in NO_x emissions in many regions of the U.S. over our study period [11]. We
do not use the 2008 NO_x emissions to subset the data beyond the 90th percentile cut-
off, as our goal is to simply provide an easily reproducible metric of urban regions versus

non-urban regions. We tested the results produced by the high-NO_x and low-NO_x designation by also separating O_3 monitors by urban areas GIS data from the U.S. Census Bureau (`https://www.census.gov/geo/maps-data/data/cbf/cbf_ua.html`), and we produced nearly identical results. Data availability precludes us from testing whether changes in NO_x emissions in certain regions has led to different O_3 responses to the presence of smoke over time.

2.2.8. SMOKE-IMPACTED VERSUS SMOKE-FREE O_3 DISTRIBUTIONS. The goal of this study is to quantify how O_3 abundance differs between days when surface O_3 chemistry is influenced by smoke (hereafter smoke-impacted day) versus days when surface O_3 chemistry is not likely influenced by smoke (hereafter smoke-free days). We include May through September in this analysis, as these are the most active wildfire months in North America, and this typically comprises the O_3 season from a monitoring perspective. In addition, we estimate how often smoke is present when O_3 exceeds the health standard at a given monitoring location. We only conducted each of these analyses at U.S. air quality monitoring sites where there is sufficient data to statistically test whether there is a real difference in O_3 mixing ratios between days that are smoke-free versus smoke-impacted.

We define a smoke-impacted day for a given monitor using the following criteria. (1) A smoke plume as outlined by the GIS polygons in the HMS data set overlaps a monitor. (2) The $PM_{2.5}$ concentration measured for that day is more than one standard deviation higher than the average summertime (May–August) $PM_{2.5}$ concentration measured at that location. When these two conditions are met, the day is flagged as smoke-impacted. Elevated carbon monoxide (CO) was also considered as a requirement for smoke-impacted days; however, using only locations with CO, $PM_{2.5}$, and O_3 measurements significantly limited the scope

of the analysis. The most likely reason for misclassifying a smoke-impacted day as a smoke-free day using this method is if HMS or $PM_{2.5}$ is not available on a smoke-impacted day. This could result in potentially mislabeling a smoke-impacted day as smoke-free.

There is a positive observed relationship between temperature and O_3 mixing ratios [3, 80]. Smoke-free days are defined using criteria designed to control for the influence of temperature. The smoke-free criteria are as follows: (i) the day was not flagged as smoke-impacted based on the description above, and (ii) the daily mean temperature was as warm or warmer than the mean plus one standard deviation (σ) of the smoke-impacted temperature distribution. This temperature requirement for smoke-free days is extremely stringent. Our discussion will focus on this criteria; however, we will also show results where smoke free days are warmer than the mean of smoke impacted days $- 0.5\,\sigma$, $+\,0\,\sigma$, and $+\,0.5\,\sigma$. By restricting the smoke-free days to dates where the temperature is warmer than the smoke-impacted days we are able to minimize the possibility that enhanced O_3 on smoke-impacted days is driven by higher temperature. After the smoke-free and smoke-impacted days are flagged for a given O_3 MDA8 time series, the mean MDA8 value for smoke-impacted and smoke-free days are computed. These mean values can be interpreted as the expected summertime MDA8 mixing ratio for smoke-free and smoke-impacted days at a given location. We find that the O_3 mixing ratios for smoke-free and smoke-impacted days are normally distributed, which allows us to test the significance of the difference of the means using a two-sided student t test (p-value <0.05) with a null hypothesis that there is no difference between the mean values of the samples.

To create the distributions and estimate the difference in MDA8 O_3 mixing ratios between the smoke-free and smoke-impacted data sets, we set a number of minimum data standards that narrow the number of eligible AQS O_3 monitors. For a monitor to be included in the

smoke-impacted data set, we require that $PM_{2.5}$ and O_3 instruments be co-located. This criterion reduces the number of eligible locations from over 1200 to 506. We further restrict monitors to those within 40 km of a temperature station. Gridded 0.75° by 0.75° ECMWF ERA-interim reanalysis temperature data is also used to control for the temperature of smoke-free days, yielding nearly identical results as the ground based temperature measurements. Finally, we restrict analysis to locations that meet our definition of smoke-impacted for at least 10 days in order to represent the variability in O_3 mixing ratios on smoke-impacted days. Our stringency limits the number of locations that have adequate data to create smoke-impacted and smoke-free O_3 distributions but grant confidence to the ability to estimate a difference between the two.

Issues with O_3 monitors under heavy smoke conditions are very relevant for this analysis. As noted in section 2.2.2, high concentrations of smoke can bias O_3 instrument MDA8 readings by up to 6 ppbv [75]. Some of the smoke-impacted days are days characterized by $PM_{2.5}$ concentrations over 100 μg m^{-3} while others are dilute with much smaller aerosol concentrations. To test whether the observed differences between smoke-impacted and smoke-free mixing ratios are likely a function of this type of instrumentation error, we set a threshold for the daily maximum hourly $PM_{2.5}$ value allowed in the smoke-impacted distributions. We tested three thresholds, hourly $PM_{2.5}$ < 100 μg m^{-3}, < 50 μg m^{-3}, and < 25 μg m^{-3}, and we observed no significant change in the difference between smoke-impacted and smoke-free mixing ratios. By controlling for extremely high $PM_{2.5}$ concentrations we are confident that the elevated smoke-impacted O_3 measurements are not driven solely by the bias documented in Ollison et al. [75].

2.3. Results and Discussion:

2.3.1. MDA8 Response to smoke: Figure 2.1 shows that sign of the difference in mean O_3 MDA8 mixing ratios between smoke-free and smoke-impacted days is sensitive to the temperature criteria applied to smoke-free days. Figure 2.1 panel A through D show the sign of the response in MDA8 mixing ratios as smoke-free day temperatures increase. Panel D shows the sign of the change when smoke-free days are warmer than the mean + 1 σ of smoke-impacted days temperature. With this highest stringency, 11% of O_3 monitors have lower MDA8 mixing ratios on smoke-impacted days, 20% have higher mixing ratios, and 68% cannot be distinguished at the 95% confidence level. Many U.S. cities show a decline in O_3 mixing ratios between 2005 and 2014 [11, 81]. A smaller number of locations show a small increase or no change. Linear trends in O_3 mixing ratios are subtracted from the data time series to ensure that trends in O_3 are not responsible for the observed difference between smoke-impacted and smoke-free mixing ratios. See the appendix for further details and an example. The average number of smoke-impacted days for locations where the distributions are significantly different is 34, compared to 29 smoke-impacted days where the distributions cannot be distinguished. This result suggests that locations where the means are statistically different are not likely due to the number of smoke-impacted days. Figure 2.1 also shows that the response of O_3 to the presence of smoke is positive most consistently in the Southeast U.S. and major metropolitan areas in Texas, despite very little local fire activity in these areas during summer months (Figure 1.4). Our smoke-free temperature criterion has the possibility of limiting our analysis since using only the warmest days does not represent the true variability in summertime MDA8 values on smoke-free days but rather the MDA8 values on days with ideal conditions for O_3 production.
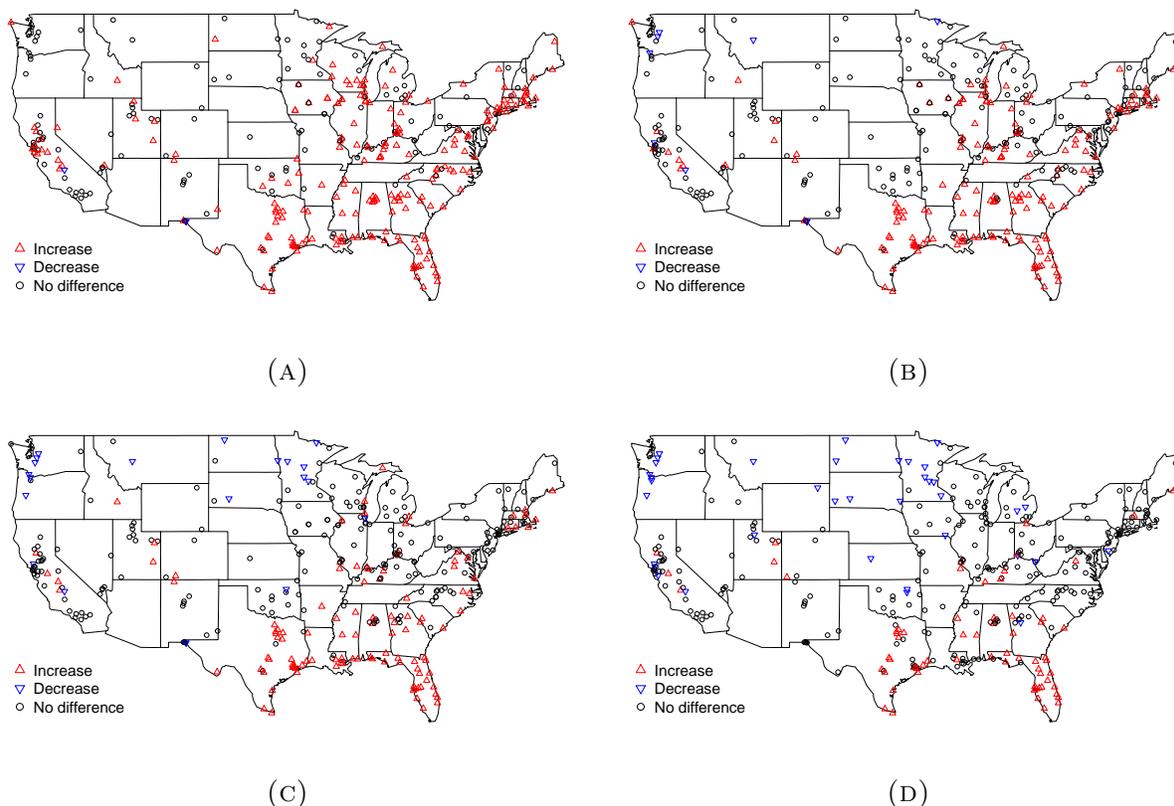
FIGURE 2.1. The sign of the change in MDA8 O$_3$ values on smoke-impacted vs. smoke-free days. ECMWF 0.75° x 0.75° temperature data was used to control for temperature. Monitors with < 10 smoke-impacted days are not shown. We show several options for temperature stringency: (A) Smoke-free days are warmer than the mean temperature of smoke-impacted days minus 0.5 standard deviation of the smoke-impacted days temperature values. This effectively removes the coldest days from the smoke-free subset of data. (B) Smoke-free days are warmer than the mean temperature of smoke-impacted days. (C) Smoke-free days are warmer than the mean temperature of smoke-impacted days plus 0.5 standard deviation of the smoke-impacted days temperature values. (D) Smoke-free days are warmer than the mean temperature of smoke-impacted days plus 1 standard deviation of the smoke-impacted days temperature values. Open black circles show monitors where the change is not statistically significant ($p < 0.05$).

Regional differences are observed when the magnitude of the differences between smoke-impacted day and smoke-free day O$_3$ MDA8 mixing ratios are shown. Figure 2.2 shows the variability in the differences in O$_3$ between the distributions across monitored locations for varying levels of smoke-free day temperature criteria. We do not observe an increase in the

difference between the means as the number of smoke-impacted days increases. In other words, the places with the most smoke-impacted days are not the same places where the difference between O_3 mixing ratios on smoke-impacted and smoke-free days are largest. The areas with the biggest positive differences are generally the areas with larger summertime MDA8 O_3 mixing ratios (East Coast urban corridor and cities along the Gulf of Mexico). Because the AQS O_3 monitors are used to make regulatory policy to meet the NAAQS, most measurements are made in or near urban areas. The smoke-impacted O_3 enhancements tend to be smaller or negative in the states with the most acres of wildfire burned area. On the basis of Federal Wildland Fire Occurrence data (`http://wildfire.cr.usgs.gov/firehistory/data.html`), between 2005 and 2013 Washington, Oregon, Idaho, Montana, Wyoming, and Colorado accounted for 40% of the lower 48 burned area.
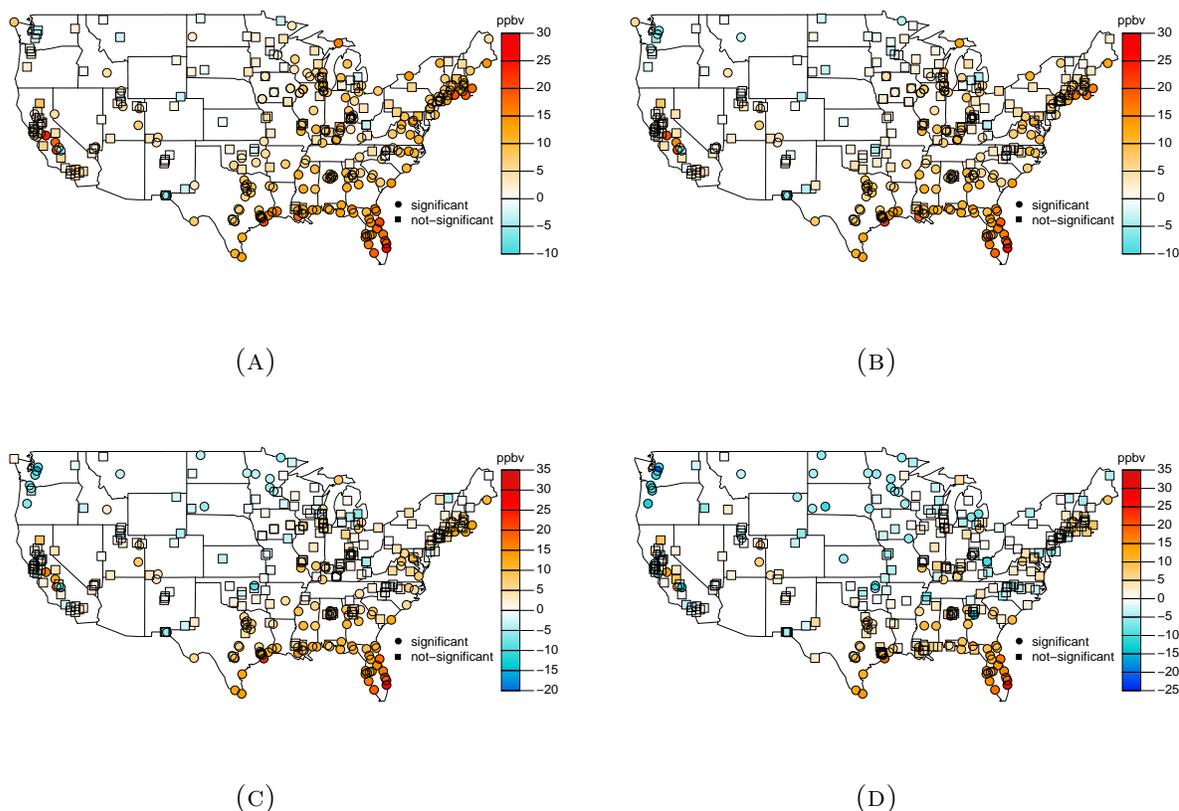
FIGURE 2.2. Difference between the mean smoke-impacted day MDA8 O_3 mixing ratio and the mean smoke-free day MDA8 O_3 mixing ratio. Temperature is controlled for using ECMWF reanalysis data. A minimum of 10 smoke-impacted days is required to estimate the difference between smoke-impacted and smoke-free days. At each station, the significance of the difference is indicated by the symbol shape. (A) Smoke-free days are warmer than the mean temperature of smoke-impacted days minus 0.5 standard deviation of the smoke-impacted days temperature values. (B) Smoke-free days are warmer than the mean temperature of smoke-impacted days. (C) Smoke-free days are warmer than the mean temperature of smoke-impacted days plus 0.5 standard deviation of the smoke-impacted days temperature values. (D) Smoke-free days are warmer than the mean temperature of smoke-impacted days plus 1 standard deviation of the smoke-impacted days temperature values.

The recent literature presents many patterns and case studies that are consistent with the patterns observed in Figure 2.2. *Jaffe and Wigder* [82] and references therein point out that the maximum O_3 production is often observed substantially downwind of the fire, after the smoke plumes have aged for several days. O_3 enhancement is possible in less photochemically

aged smoke plumes as well, and this appears to be important especially when smoke plumes interact with urban air masses [83, 84, 51, 85, 66]. As an extreme example, during the 2008 wildfire season in California *Singh et al.* [66] observed O_3 mixing ratios of 160 ppbv when wildfire smoke interacted with an urban air mass.

There is a strong relationship between temperature and observed O_3 mixing ratios, and there are several physical and photochemical reasons for this. On the basis of our temperature criteria alone, the average enhancement of O_3 on smoke-impacted days presented in Figure 2.2 is likely an underestimate since we have eliminated all but the warmest smoke-free days.

2.3.2. MDA8 O_3 SMOKE RESPONSE AND NO_x EMISSIONS. Case studies have shown that forest fire plumes can efficiently produce O_3 when mixed with urban pollution [66]. To investigate this on a national level, we separate O_3 monitors based on NO_x emissions using the 2008 EPA NEI summertime emission estimates. On average, O_3 monitors located in high NO_x emitting grid boxes, defined as grid boxes that exceed the 90th percentile of all grid box emission rates, show larger O_3 enhancements when smoke-impacted compared to regions below the 90th percentile. This result is only statistically significant at the 95% confidence level when our most relaxed temperature criteria for smoke-free days is used. Figure 2.3 shows the difference of the change in MDA8 mixing ratios between smoke-impacted and smoke-free days when smoke-free days are warmer than the mean minus one half σ of the smoke-impacted days temperature values.
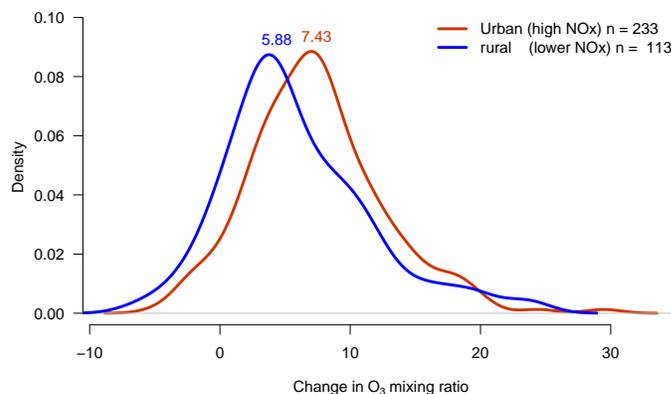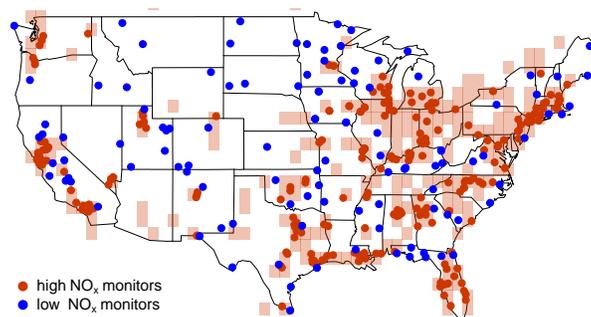
FIGURE 2.3. (Top) Monitors are distinguished based on high (orange boxes) and low NO_x emissions (blue boxes) using the 2008 NEI gridded emission inventories. High NO_x grid boxes include grid boxes above the 90th percentile for emissions and low emission grid boxes are all boxes below the 90th percentile. (Bottom) The kernel density estimates of the (smoke-impacted O_3 data sets) (smoke-free O_3 data sets) mean values for high and low NO_x emissions. Difference of mean smoke-impacted MDA8 O_3 mixing ratio - smoke-free MDA8 O_3 mixing ratios for urban and rural monitors as defined by NEI. Here, smoke-free days are required to have temperatures that exceed the mean - 0.5 σ of the smoke-impacted days temperature values.

2.3.3. SMOKE PRESENCE ON O_3 MDA8 EXCEEDANCE DAYS. The current MDA8 O_3 standard is 75 ppbv. Using AQS MDA8 O_3 time series, we calculate how many days O_3

exceeded the 2008 standard for each monitor between 2005 and 2014 during May–September. We also calculate the number of exceedance days that were smoke-influenced, and we present this as a percentage of days with smoke present during O_3 exceedance events (Figure 2.4).

Figure 2.4 shows that there is a large range in both the number of days exceeding the 75 ppbv standard, and the percentage of exceedance days that smoke is present at a given monitoring station. Cities where > 10 exceedance days occur per summer with > 10% of these days influenced by smoke (Figure 2.4 large green dots) include locations within the North East corridor, Dallas TX, Houston TX, Atlanta GA, Birmingham AL, and Kansas City MO.
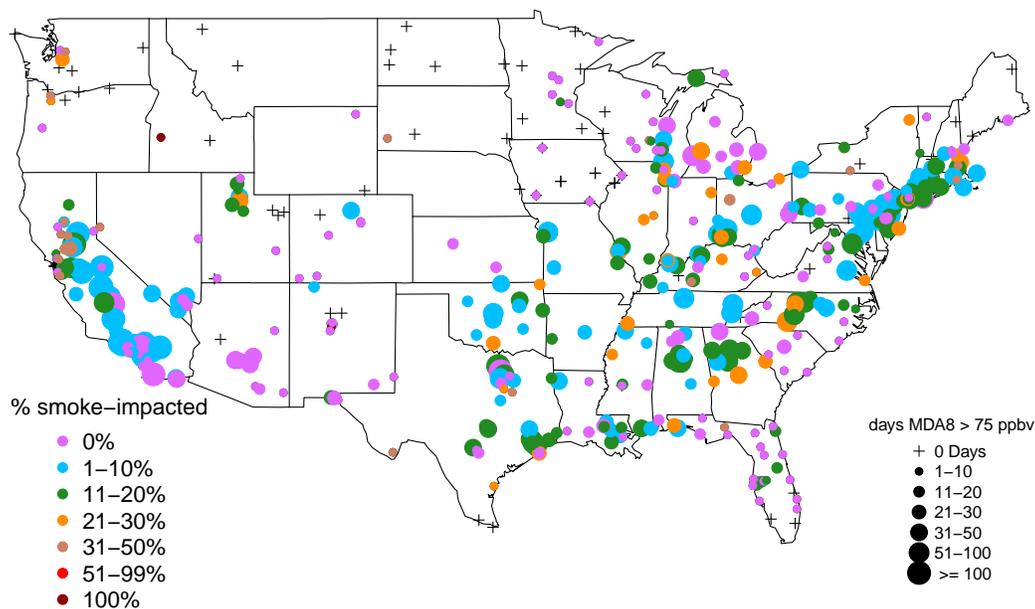
FIGURE 2.4. Percent (color) of days (dot size) the MDA8 O_3 exceeds 75 ppbv. 1263 summer days between 2005–2014 are used to calculate the number of MDA8 O_3 exceedance days. Cross symbol represents locations where the MDA8 O_3 standard is not violated.

There is evidence that lowering the MDA8 O_3 standard to 70 ppbv would result in significant gains for human and environmental health [59, 86, 61, 57, 85]. On 1 October 2015

the EPA announced a new MDA8 standard of 70 ppbv. If the 70 ppbv standard is applied to the 2005–2014 summer O_3 data, then the number of exceedance days increases at almost every location, but the percent of exceedance days influenced by smoke varies considerably by location (Figure 2.5). Most monitors experience small changes in the percent of smoke-influenced MDA8 exceedance days. A roughly equal number of monitors observe a –15% to 15% change in smoke-influenced MDA8 days (Figure 2.6 blue and green dots). La Crosse WI, Roosevelt UT, and Portland ME are some of the cities that experience a large increase in the percent of MDA8 exceedances that are smoke-influenced if held against a 70 ppbv standard. Locations where the increase in smoke-influenced exceedances is greatest are locations with few exceedance days at the 70 ppbv limit. This suggests that as the O_3 standard is lowered, the fraction of exceedance days impacted by smoke will remain nearly the same at most monitored locations (Figure 2.6).
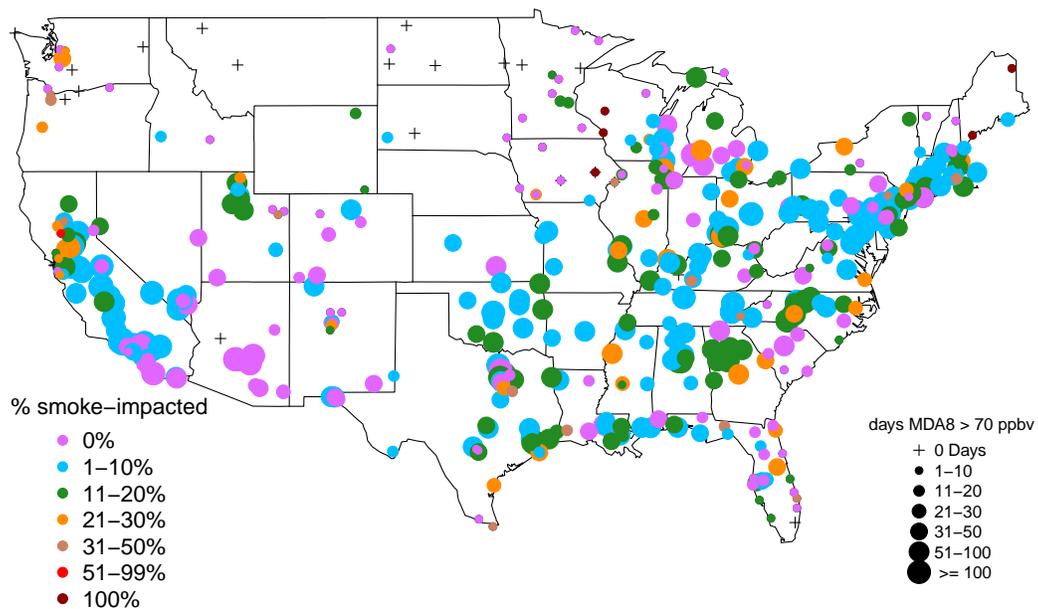
FIGURE 2.5. Percent (color) of days (dot size) the MDA8 O$_3$ exceeds 70 ppbv. 1263 summer days between 2005–2014 are used to calculate number of MDA8 O$_3$ exceedance days. Cross symbol represents locations where the MDA8 O$_3$ standard is not violated.
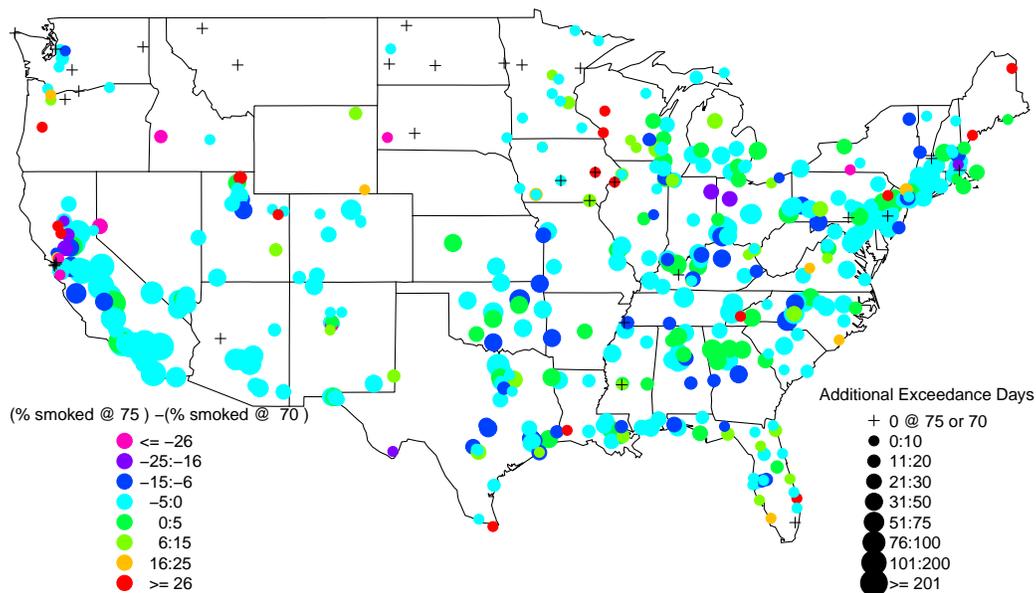
FIGURE 2.6. Change in the percent (color) of MDA8 violation days when the limit is changed from 75 to 70 ppbv and the additional exceedance days from lowering the standard (dot size). Cross symbol represents locations where the MDA8 O_3 standard is not violated.

In summary, we observe that the difference between O_3 MDA8 mixing ratios on smoke-free and smoke-impacted summer days varies considerably by location and that the difference is sensitive to the temperature criteria chosen for smoke-free days. We observe that MDA8 mixing ratios are higher on smoke-impacted days regardless of smoke-free day temperature for monitors in Florida and cities near the Gulf of Mexico. We find that monitors located in regions with high NO_x emissions are likely to see a larger O_3 enhancement on smoke-impacted days compared to monitors located in rural locations. When estimating the fraction of smoke-impacted exceedance days for O_3 monitors, we observed considerable spatial variability, with percentages ranging between 0 and 100. Generally locations with more exceedance days exhibit fewer than 20% impacted by smoke. The spatial variability and percent of

exceedance days that are impacted by smoke remains nearly constant when held against the 2008 75 ppbv MDA8 standard or the 2015 standard of 70 ppbv.

CHAPTER 3

FUTURE WORK

A practical next step for this work is to explore the ground level air quality implications of the smoke transport climatologies presented in Chapter 1. The smoke hour trajectory points could be subset by their height in the vertical based on the HYSPLIT trajectory height field. We could consider a height threshold (*e.g.* below the boundary layer) to identify smoke hours more likely to have air-quality health impacts. Ground level PM and CO measurements could also be used to understand the relative abundance of observed smoke plumes that impact the ground vs. well above the surface. Figures 1.2 through 1.7 could then be reproduced in a way that is more meaningful to ground-level air quality.

There are improvements to be made to the current definition of smoke hour. Another requirement we would like to implement is to require trajectory points that contribute to smoke hours be continuous. For example; the methods presented in Chapter 1 require that trajectories overlap a smoke plume in the first two days to be able to contribute smoke hours. Currently trajectories that overlap smoke plumes for the first two days, then don't overlap smoke plumes on days 3, 4, or 5, are able to produce smoke hours on day 6. It is unlikely that the overlap on day 6 represents the advection of smoke when overlap is absent for the three days before. Another possible improvement to make is to require more than one trajectory point to overlap a smoke plume within the first two days. If instead a certain percent of overlapping trajectory points are required in the first two days, the trajectories would be further subset to those with paths that best match smoke plumes.

Chapter 2 presents the observed differences in O_3 mixing ratios between smoke-free and smoke-impacted days without identifying the specific O_3 chemistry or smoke conditions that

lead to enhanced O_3 mixing ratios in the presence of smoke. Future work should combine the methods and data of Chapters 1 and 2 in a way that specifically informs the following research questions. For monitors and days where O_3 mixing ratios are enhanced by smoke, is O_3 production NO_x or VOC limited? How does this vary between monitored locations? These questions could be explored by looking at a subset of monitored locations that include PM, VOC, and NO_x measurements. The type of smoke that leads to increased O_3 mixing ratios could be informed by answering: What is the age of the smoke? How far has the smoke traveled? What type of fuel is responsible for the smoke? Another important unanswered question is whether O_3 production occurs in the smoke plume and is advected downwind or if ozone production occurs locally due to smoke plumes supplying the required O_3 precursors. The persistently elevated O_3 under smoke-impacted conditions in the Southeast U.S. could inform these questions. We know that there is very little local fire activity in the Southeast during summer months and that the smoke plumes that do reach the Southeast are on average two days old. O_3 production occurring in moving smoke plumes could be explored by making use of hourly O_3 measurements. For example; if a smoke plume is approaching the Southeast from the Northwest, we could track plume movement using HMS smoke plumes combined with geostationary satellite visible imagery. We could then use the expansive network of ground level O_3 monitors to see if O_3 mixing ratios increase as smoke passes. If the smoke is advecting O_3 all monitors would experience a similar increase in O_3 mixing ratios. If smoke is providing O_3 precursor species, we would expect to see different mixing ratio increases between monitors located in urban vs. rural locations. This concept could be explored for select case studies. If the case studies are promising the analysis could be done in aggregate. This would allow us to investigate the relative importance of O_3 production in plumes vs. local production when plumes arrive. Finally, if we find answers to these

questions it would be informative to understand how they change or remain the same in a changing climate. Fire is an important part of the earth system. Are there land or air quality management strategies that could be adopted that maximize both ground level air quality and forest health?

Knowing the altitude smoke is injected into the atmosphere is key to understanding both present and future smoke transport pathways. HYSPLIT points and smoke plumes could be used to develop a climatology of smoke-plume injection height by region and land biome. This could be done by running HYSPLIT trajectories initialized at HYSPLIT points at many more starting heights and keeping track of the number of trajectory points that overlap smoke plumes for each injection-height. We could assume that the initial heights associated with trajectories that more-frequently overlap observed smoke plumes are the most-probable injection height for a given region and biome. This climatology could directly inform the National Weather Service smoke forecast, which makes use of HYSPLIT trajectories and the HYSPLIT plume model.

In Chapter 1 we introduce a new dataset and explore what it can teach us about smoke transport pathways in North America. In order to maximize the utility of this dataset, we need to make it available to the smoke research community. To accomplish this, we need to continue coordinating with our co-author at the National Environmental Satellite, Data, and Information Service (NESDIS), Mark Ruminski, and make the quality vetted and documented HYSPLIT point and smoke plume data available on a public ftp site. This will enable other studies to use these data and provide the smoke-research community an opportunity to corroborate or refute the results discussed in this thesis.

References

[1] C. Wiedinmyer, B. Quayle, C. Geron, A. Belote, D. McKenzie, X. Zhang, S. ONeill, and K. K. Wynne, "Estimating emissions from fires in North America for air quality modeling," *Atmospheric Environment*, vol. 40, pp. 3419–3432, June 2006.

[2] R. J. Park, D. J. Jacob, and J. A. Logan, "Fire and biofuel contributions to annual mean aerosol mass concentrations in the United States," *Atmospheric Environment*, vol. 41, no. 35, pp. 7389–7400, 2007.

[3] D. Jaffe, D. Chand, W. Hafner, A. Westerling, and D. Spracklen, "Influence of fires on O-3 concentrations in the western US," *Environmental Science & Technology*, vol. 42, no. 16, pp. 5885–5891, 2008.

[4] P. E. Saide, D. A. Peterson, A. da Silva, B. Anderson, L. D. Ziemba, G. Diskin, G. Sachse, J. Hair, C. Butler, M. Fenn, J. L. Jimenez, P. Campuzano-Jost, A. E. Perring, J. P. Schwarz, M. Z. Markovic, P. Russell, J. Redemann, Y. Shinozuka, D. G. Streets, F. Yan, J. Dibb, R. Yokelson, O. B. Toon, E. Hyer, and G. R. Carmichael, "Revealing important nocturnal and day-to-day variations in fire smoke emissions through a multiplatform inversion," *Geophysical Research Letters*, vol. 42, p. 2015GL063737, May 2015.

[5] K. R. Baker, M. C. Woody, G. S. Tonnesen, W. Hutzell, H. O. T. Pye, M. R. Beaver, G. Pouliot, and T. Pierce, "Contribution of regional-scale fire events to ozone and PM2.5 air quality estimated by photochemical modeling approaches," *Atmospheric Environment*, vol. 140, pp. 539–554, Sept. 2016.

[6] S. J. Brey and E. V. Fischer, "Smoke in the City: How Often and Where Does Smoke Impact Summertime Ozone in the United States?," *Environmental Science & Technology*, vol. 50, pp. 1288–1294, Feb. 2016.

[7] M. Dennekamp and M. Carey, "Air quality and chronic disease: why action on climate change is also good for health," *New South Wales Public Health Bulletin*, vol. 21, pp. 115–121, June 2010.

[8] A. G. Rappold, S. L. Stone, W. E. Cascio, L. M. Neas, V. J. Kilaru, M. S. Carraway, J. J. Szykman, A. Ising, W. E. Cleve, J. T. Meredith, H. Vaughan-Batten, L. Deyneka, and R. B. Devlin, "Peat Bog Wildfire Smoke Exposure in Rural North Carolina Is Associated with Cardiopulmonary Emergency Department Visits Assessed through Syndromic Surveillance," *Environmental Health Perspectives*, vol. 119, pp. 1415–1420, Oct. 2011.

[9] A. Haikerwal, M. Akram, A. D. Monaco, K. Smith, M. R. Sim, M. Meyer, A. M. Tonkin, M. J. Abramson, and M. Dennekamp, "Impact of Fine Particulate Matter (PM2.5) Exposure During Wildfires on Cardiovascular Health Outcomes," *Journal of the American Heart Association*, vol. 4, p. e001653, July 2015.

[10] M. Val Martin, C. L. Heald, J.-F. Lamarque, S. Tilmes, L. K. Emmons, and B. A. Schichtel, "How emissions, climate, and land use change will impact mid-century air quality over the United States: a focus on effects at national parks," *Atmos. Chem. Phys.*, vol. 15, pp. 2805–2823, Mar. 2015.

[11] A. R. Russell, L. C. Valin, and R. C. Cohen, "Trends in OMI NO2 observations over the US: Effects of emission control technology and the economic recession," *Atmospheric Chemistry and Physics*, vol. 12, no. 6, pp. 15419–15452, 2012.

[12] Y. Liu, "A Regression Model for Smoke Plume Rise of Prescribed Fires Using Meteorological Conditions," *Journal of Applied Meteorology and Climatology*, vol. 53, pp. 1961–1975, Apr. 2014.

[13] H. Forrister, J. Liu, E. Scheuer, J. Dibb, L. Ziemba, K. L. Thornhill, B. Anderson, G. Diskin, A. E. Perring, J. P. Schwarz, P. Campuzano-Jost, D. A. Day, B. B. Palm,

J. L. Jimenez, A. Nenes, and R. J. Weber, "Evolution of brown carbon in wildfire plumes," *Geophysical Research Letters*, vol. 42, p. 2015GL063897, June 2015.

[14] J. Liu, E. Scheuer, J. Dibb, G. S. Diskin, L. D. Ziemba, K. L. Thornhill, B. E. Anderson, A. Wisthaler, T. Mikoviny, J. J. Devi, M. Bergin, A. E. Perring, M. Z. Markovic, J. P. Schwarz, P. Campuzano-Jost, D. A. Day, J. L. Jimenez, and R. J. Weber, "Brown carbon aerosol in the North American continental troposphere: sources, abundance, and radiative forcing," *Atmos. Chem. Phys.*, vol. 15, pp. 7841–7858, July 2015.

[15] R. A. Washenfelder, A. R. Attwood, C. A. Brock, H. Guo, L. Xu, R. J. Weber, N. L. Ng, H. M. Allen, B. R. Ayres, K. Baumann, R. C. Cohen, D. C. Draper, K. C. Duffey, E. Edgerton, J. L. Fry, W. W. Hu, J. L. Jimenez, B. B. Palm, P. Romer, E. A. Stone, P. J. Wooldridge, and S. S. Brown, "Biomass burning dominates brown carbon absorption in the rural southeastern United States," *Geophysical Research Letters*, vol. 42, p. 2014GL062444, Jan. 2015.

[16] J. T. Randerson, Y. Chen, G. R. van der Werf, B. M. Rogers, and D. C. Morton, "Global burned area and biomass burning emissions from small fires," *Journal of Geophysical Research: Biogeosciences*, vol. 117, p. G04012, Dec. 2012.

[17] A. L. Westerling, H. G. Hidalgo, D. R. Cayan, and T. W. Swetnam, "Warming and earlier spring increase western US forest wildfire activity," *Science*, vol. 313, no. 5789, pp. 940–943, 2006.

[18] M. Scholze, W. Knorr, N. W. Arnell, and I. C. Prentice, "A climate-change risk analysis for world ecosystems," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 103, no. 35, pp. 13116–13120, 2006.

[19] M. Keywood, M. Kanakidou, A. Stohl, F. Dentener, G. Grassi, C. P. Meyer, K. Torseth, D. Edwards, A. M. Thompson, U. Lohmann, and J. Burrows, "Fire in the Air: Biomass

Burning Impacts in a Changing Climate," *Critical Reviews in Environmental Science and Technology*, vol. 43, no. 1, pp. 40–83, 2013.

[20] X. Yue, L. J. Mickley, J. A. Logan, and J. O. Kaplan, "Ensemble projections of wildfire activity and carbonaceous aerosol concentrations over the western United States in the mid-21st century," *Atmospheric Environment (Oxford, England: 1994)*, vol. 77, pp. 767–780, Oct. 2013.

[21] M. D. Hurteau, A. L. Westerling, C. Wiedinmyer, and B. P. Bryant, "Projected Effects of Climate and Development on California Wildfire Emissions through 2100," *Environmental Science & Technology*, vol. 48, pp. 2298–2304, Feb. 2014.

[22] J. L. McCarty, S. Korontzi, C. O. Justice, and T. Loboda, "The spatial and temporal distribution of crop residue burning in the contiguous United States," *Science of The Total Environment*, vol. 407, pp. 5701–5712, Oct. 2009.

[23] A. Dennis, M. Fraser, S. Anderson, and D. Allen, "Air pollutant emissions associated with forest, grassland, and agricultural burning in Texas," *Atmospheric Environment*, vol. 36, pp. 3779–3792, Aug. 2002.

[24] J. Jimenez, C.-F. Wu, C. Claiborn, T. Gould, C. D. Simpson, T. Larson, and L. J. Sally Liu, "Agricultural burning smoke in eastern Washingtonpart I: Atmospheric characterization," *Atmospheric Environment*, vol. 40, pp. 639–650, Feb. 2006.

[25] S. K. Akagi, R. J. Yokelson, C. Wiedinmyer, M. J. Alvarado, J. S. Reid, T. Karl, J. D. Crounse, and P. O. Wennberg, "Emission factors for open and domestic biomass burning for use in atmospheric models," *Atmospheric Chemistry and Physics*, vol. 11, no. 9, pp. 4039–4072, 2011.

[26] G. G. Pfister, C. Wiedinmyer, and L. K. Emmons, "Impacts of the fall 2007 California wildfires on surface ozone: Integrating local observations with global model simulations," *Geophysical Research Letters*, vol. 35, no. 19, 2008.

[27] A. F. Stein, G. D. Rolph, R. R. Draxler, B. Stunder, and M. Ruminski, "Verification of the NOAA Smoke Forecasting System: Model Sensitivity to the Injection Height," *Weather and Forecasting*, vol. 24, pp. 379–394, Apr. 2009.

[28] D. J. Miller, K. Sun, M. A. Zondlo, D. Kanter, O. Dubovik, E. J. Welton, D. M. Winker, and P. Ginoux, "Assessing boreal forest fire smoke aerosol impacts on U.S. air quality: A case study using multiple data sets," *Journal of Geophysical Research: Atmospheres*, vol. 116, p. D22209, Nov. 2011.

[29] M. Val Martin, C. L. Heald, B. Ford, A. J. Prenni, and C. Wiedinmyer, "A decadal satellite analysis of the origins and impacts of smoke in Colorado," *Atmos. Chem. Phys.*, vol. 13, pp. 7429–7439, Aug. 2013.

[30] K. R. Baker, M. C. Woody, G. S. Tonnesen, W. Hutzell, H. O. T. Pye, M. R. Beaver, G. Pouliot, and T. Pierce, "Contribution of regional-scale fire events to ozone and PM2.5 air quality estimated by photochemical modeling approaches," *Atmospheric Environment*, vol. 140, pp. 539–554, Sept. 2016.

[31] J. M. Creamean, P. J. Neiman, T. Coleman, C. J. Senff, G. Kirgis, R. J. Alvarez, and A. Yamamoto, "Colorado air quality impacted by long range transport: A set of case studies during the 2015 Pacific Northwest fires," *Atmospheric Chemistry and Physics Discussions*, pp. 1–35, May 2016.

[32] R. J. Park, D. J. Jacob, M. Chin, and R. V. Martin, "Sources of carbonaceous aerosols over the United States and implications for natural visibility," *Journal of Geophysical Research: Atmospheres*, vol. 108, p. 4355, June 2003.

[33] S. Vedal and S. J. Dutton, "Wildfire air pollution and daily mortality in a large urban area," *Environmental Research*, vol. 102, no. 1, pp. 29–35, 2006.

[34] C. Wiedinmyer, S. K. Akagi, R. J. Yokelson, L. K. Emmons, J. A. Al-Saadi, J. J. Orlando, and A. J. Soja, "The Fire INventory from NCAR (FINN): a high resolution global model to estimate the emissions from open burning," *Geoscientific Model Development*, vol. 4, pp. 625–641, July 2011.

[35] M. J. Alvarado, C. R. Lonsdale, R. J. Yokelson, S. K. Akagi, H. Coe, J. S. Craven, E. V. Fischer, G. R. McMeeking, J. H. Seinfeld, T. Soni, J. W. Taylor, D. R. Weise, and C. E. Wold, "Investigating the links between ozone and organic aerosol chemistry in a biomass burning plume from a prescribed fire in California chaparral," *Atmos. Chem. Phys.*, vol. 15, pp. 6667–6688, June 2015.

[36] R. Paugam, M. Wooster, S. Freitas, and M. ValMartin, "A review of approaches to estimate wildfire plume injection height within large-scale atmospheric chemical transport models," *Atmos. Chem. Phys.*, vol. 16, pp. 907–925, Jan. 2016.

[37] F. Garcia-Menendez, Y. Hu, and M. T. Odman, "Simulating smoke transport from wildland fires with a regional-scale air quality model: Sensitivity to uncertain wind fields," *Journal of Geophysical Research: Atmospheres*, vol. 118, pp. 6493–6504, June 2013.

[38] M. Ruminski, S. Kondragunta, R. Draxler, and J. Zeng, "Recent changes to the Hazard Mapping System," *ResearchGate*, 2006.

[39] G. D. Rolph, R. R. Draxler, A. F. Stein, A. Taylor, M. G. Ruminski, S. Kondragunta, J. Zeng, H. C. Huang, G. Manikin, J. T. Mcqueen, and P. M. Davidson, "Description and Verification of the NOAA Smoke Forecasting System: The 2007 Fire Season," *Weather and Forecasting*, vol. 24, no. 2, pp. 361–378, 2009.

[40] J. Brown, "USGS Small-scale Dataset - North American Land Cover Characteristics - 1-Kilometer Resolution 200212 GeoTIFF - ScienceBase-Catalog," June 2016.

[41] J. R. Anderson, E. E. Hardy, J. T. Roach, and R. E. Witmer, "A land use and land cover classification system for use with remote sensor data," USGS Numbered Series 964, 1976.

[42] R. R. Draxler and G. D. Hess, "An overview of the HYSPLIT_4 modeling system of trajectories, dispersion, and deposition," *Australian meteorological magazine*, vol. 47, pp. 295–308, Dec. 1998.

[43] A. F. Stein, R. R. Draxler, G. D. Rolph, B. J. B. Stunder, M. D. Cohen, and F. Ngan, "NOAAs HYSPLIT Atmospheric Transport and Dispersion Modeling System," *Bulletin of the American Meteorological Society*, vol. 96, pp. 2059–2077, May 2015.

[44] R. Draxler, "NOAA - Air Resources Laboratory - FAQ - How do I estimate the absolute (in km) and relative (%) errors when using the HYSPLIT trajectory model?," Sept. 2008.

[45] M. Kanamitsu, "Description of the NMC Global Data Assimilation and Forecast System," *Weather and Forecasting*, vol. 4, pp. 335–342, Sept. 1989.

[46] M. Val Martin, J. A. Logan, R. A. Kahn, F.-Y. Leung, D. L. Nelson, and D. J. Diner, "Smoke injection heights from fires in North America: analysis of 5 years of satellite observations," *Atmos. Chem. Phys.*, vol. 10, pp. 1491–1510, Feb. 2010.

[47] A. Stein, "Trajectory "Hits" ground - HYSPLIT Forum: hysplitbbs.arl.noaa.gov," Apr. 2014.

[48] E. Pebesma, R. Bivand, B. Rowlingson, V. Gomez-Rubio, R. Hijmans, M. Sumner, D. MacQueen, J. Lemon, and J. O'Brien, "sp: Classes and Methods for Spatial Data," Apr. 2016.

[49] Q. Laffineur, A. Delcloo, H. De Backer, M. Adam, and D. Klugmann, "Observation of an intercontinental smoke plume over Europe on June 2013: some ambiguity in the determination of the source," vol. 16, p. 2173, May 2014.

[50] M. Gyawali, W. P. Arnott, K. Lewis, and H. Moosmller, "In situ aerosol optics in Reno, NV, USA during and after the summer 2008 California wildfires and the influence of absorbing and non-absorbing organic coatings on spectral light absorption," *Atmos. Chem. Phys.*, vol. 9, pp. 8007–8015, Oct. 2009.

[51] G. A. Morris, S. Hersey, A. M. Thompson, S. Pawson, J. E. Nielsen, P. R. Colarco, W. W. McMillan, A. Stohl, S. Turquety, J. Warner, B. J. Johnson, T. L. Kucsera, D. E. Larko, S. J. Oltmans, and J. C. Witte, "Alaskan and Canadian forest fires exacerbate ozone pollution over Houston, Texas, on 19 and 20 July 2004," *Journal of Geophysical Research-Atmospheres*, vol. 111, no. D24, 2006.

[52] A. Westerling, T. Brown, T. Schoennagel, T. Swetnam, M. Turner, and T. Veblen, "Briefing: Climate and wildfire in western U.S. forests," pp. 81–102, 2014.

[53] M. Val Martin, C. L. Heald, J.-F. Lamarque, S. Tilmes, L. K. Emmons, and B. A. Schichtel, "How emissions, climate, and land use change will impact mid-century air quality over the United States: a focus on effects at national parks," *Atmos. Chem. Phys.*, vol. 15, pp. 2805–2823, Mar. 2015.

[54] J. C. Liu, L. J. Mickley, M. P. Sulprizio, F. Dominici, X. Yue, K. Ebisu, G. B. Anderson, R. F. A. Khan, M. A. Bravo, and M. L. Bell, "Particulate air pollution from wildfires in the Western US under climate change," *Climatic Change*, pp. 1–12, July 2016.

[55] A. L. Westerling, "Increasing western US forest wildfire activity: sensitivity to changes in the timing of spring," *Phil. Trans. R. Soc. B*, vol. 371, p. 20150178, June 2016.

[56] US EPA, "AirData website Basic Information page,"

[57] K. Moore, R. Neugebauer, F. Lurmann, J. Hall, V. Brajer, S. Alcorn, and I. Tager, "Ambient ozone concentrations cause increased hospitalizations for asthma in children: An 18-year study in Southern California," *Environmental Health Perspectives*, vol. 116, no. 8, pp. 1063–1070, 2008.

[58] M. L. Bell, A. McDermott, S. L. Zeger, J. M. Samet, and F. Dominici, "Ozone and short-term mortality in 95 US urban communities, 1987-2000," *Jama-Journal of the American Medical Association*, vol. 292, no. 19, pp. 2372–2378, 2004.

[59] M. Jerrett, R. T. Burnett, C. A. Pope, K. Ito, G. Thurston, D. Krewski, Y. L. Shi, E. Calle, and M. Thun, "Long-Term Ozone Exposure and Mortality.," *New England Journal of Medicine*, vol. 360, no. 11, pp. 1085–1095, 2009.

[60] "Ozone monitoring protocol for the national park service, https://www.nature.nps.gov/air/monitoring/docs/finalozoneprotocol.pdf," 2004.

[61] "Ozone air quality standards: Epas proposed january 2010 revisions in service," *Congresional Research Service: Washington, DC*, 2010.

[62] H. B. Singh, W. Viezee, Y. Chen, J. Bradshaw, S. Sandholm, D. Blake, N. Blake, B. Heikes, J. Snow, R. Talbot, E. Browell, G. Gregory, G. Sachse, and S. Vay, "Biomass burning influences on the composition of the remote South Pacific troposphere: analysis based on observations from PEM-Tropics-A," *Atmospheric Environment*, vol. 34, no. 4, pp. 635–644, 2000.

[63] E. Real, K. S. Law, B. Weinzierl, M. Fiebig, A. Petzold, O. Wild, J. Methven, S. Arnold, A. Stohl, H. Huntrieser, A. Roiger, H. Schlager, D. Stewart, M. Avery, G. Sachse, E. Browell, R. Ferrare, and D. Blake, "Processes influencing ozone levels in Alaskan forest fire plumes during long-range transport over the North Atlantic," *Journal of Geophysical Research-Atmospheres*, vol. 112, no. D10, 2007.

[64] K. J. Bein, Y. J. Zhao, M. V. Johnston, and A. S. Wexler, "Interactions between boreal wildfire and urban emissions," *Journal of Geophysical Research-Atmospheres*, vol. 113, no. D7, 2008.

[65] S. J. Oltmans, A. S. Lefohn, J. M. Harris, D. W. Tarasick, A. M. Thompson, H. Wernli, B. J. Johnson, P. C. Novelli, S. A. Montzka, J. D. Ray, L. C. Patrick, C. Sweeney, A. Jefferson, T. Dann, J. Davies, M. Shapiro, and B. N. Holben, "Enhanced ozone over western North America from biomass burning in Eurasia during April 2008 as seen in surface and profile observations," *Atmospheric Environment*, vol. 44, no. 35, pp. 4497–4509, 2010.

[66] H. B. Singh, C. Cai, A. Kaduwela, A. Weinheimer, and A. Wisthaler, "Interactions of fire emissions and urban pollution over California: Ozone formation and air quality simulations," *Atmospheric Environment*, vol. 56, pp. 45–51, 2012.

[67] M. D. Hurteau, A. L. Westerling, C. Wiedinmyer, and B. P. Bryant, "Projected Effects of Climate and Development on California Wildfire Emissions through 2100," *Environmental Science & Technology*, vol. 48, no. 4, pp. 2298–2304, 2014.

[68] M. A. Moritz, M. A. Parisien, E. Batllori, M. A. Krawchuk, J. Van Dorn, D. J. Ganz, and K. Hayhoe, "Climate change and disruptions to global fire activity," *Ecosphere*, vol. 3, no. 6, 2012.

[69] J. K. Agee and C. N. Skinner, "Basic principles of forest fuel reduction treatments," *Forest Ecology and Management*, vol. 211, no. 1-2, pp. 83–96, 2005.

[70] M. A. Cochrane, C. J. Moran, M. C. Wimberly, A. D. Baer, M. A. Finney, K. L. Beckendorf, J. Eidenshink, and Z. Zhu, "Estimation of wildfire size and risk changes due to fuels treatments," *International Journal of Wildland Fire*, vol. 21, no. 4, pp. 357–367, 2012.

[71] M. P. Thompson, N. M. Vaillant, J. R. Haas, K. M. Gebert, and K. D. Stockmann, "Quantifying the Potential Impacts of Fuel Treatments on Wildfire Suppression Costs," *Journal of Forestry*, vol. 111, no. 1, pp. 49–58, 2013.

[72] P. R. Colarco, M. R. Schoeberl, B. G. Doddridge, L. T. Marufu, O. Torres, and E. J. Welton, "Transport of smoke from Canadian forest fires to the surface near Washington, D. C.: Injection height, entrainment, and optical properties," *Journal of Geophysical Research-Atmospheres*, vol. 109, no. D6, 2004.

[73] L. J. DeBell, R. W. Talbot, J. E. Dibb, J. W. Munger, E. V. Fischer, and S. E. Frolking, "A major regional air pollution event in the northeastern United States caused by extensive forest fires in Quebec, Canada," *Journal of Geophysical Research-Atmospheres*, vol. 109, no. D19, 2004.

[74] "Ambient air monitoring reference and equivalent methods, title 40: Protection of environment," 2015.

[75] W. M. Ollison, W. Crow, and C. W. Spicer, "Field testing of new-technology ambient air ozone monitors," *Journal of the Air & Waste Management Association*, vol. 63, no. 7, pp. 855–863, 2013.

[76] M. Y. Arshinov, B. D. Belan, O. A. Krasnov, V. K. Kovalevskii, V. A. Pirogov, A. P. Plotnikov, G. N. Tolmachev, and A. V. Fofonov, "Comparison of ultraviolet and chemiluminescent ozonometers," *Atmos. Oceanic Opt*, vol. 15, pp. 656–658, 2002.

[77] E. J. Dunlea, S. C. Herndon, D. D. Nelson, R. M. Volkamer, B. K. Lamb, E. J. Allwine, M. Grutter, C. R. R. Villegas, C. Marquez, S. Blanco, B. Cardenas, C. E. Kolb, L. T. Molina, and M. J. Molina, "Technical note: Evaluation of standard ultraviolet absorption ozone monitors in a polluted urban environment," *Atmospheric Chemistry and Physics*, vol. 6, pp. 3163–3180, 2006.

[78] "Automated surface observing system (asos), http://www.nws.noaa.gov/ost/asostech.html."

[79] "Ecmwf era-interim dataset (january 1979 to present), http://www.ecmwf.int/en/forecasts/datasets/era-interim-dataset-january-1979-present."

[80] D. J. Jacob, J. A. Logan, G. M. Gardner, R. M. Yevich, C. M. Spivakovsky, S. C. Wofsy, S. Sillman, and M. J. Prather, "Factors Regulating Ozone over the United-States and Its Export to the Global Atmosphere," *Journal of Geophysical Research-Atmospheres*, vol. 98, no. D8, pp. 14817–14826, 1993.

[81] O. Cooper, D. Parrish, J. Ziemke, N. Balashov, M. Cupeiro, I. Galbally, S. Gilge, L. Horowitz, N. Jensen, J. Lamarque, V. Naik, S. Oltmans, J. Schwab, D. Shindell, A. Thompson, V. Thouret, Y. Wang, and R. Zbinden, "Global distribution and trends of tropospheric ozone: An observation-based review," *Elementa Science of the Anthropocene*, vol. 2, no. 29, 2014.

[82] D. Jaffe and N. Wigder, "Ozone production from wildfires: A critical review," *Atmospheric Environment*, vol. 51, pp. 1–10, 2012.

[83] V. Junquera, M. M. Russell, W. Vizuete, Y. Kimura, and D. Allen, "Wildfires in eastern Texas in August and September 2000: Emissions, aircraft measurements, and impact on photochemistry," *Atmospheric Environment*, vol. 39, no. 27, pp. 4983–4996, 2005.

[84] S. A. McKeen, G. Wotawa, D. D. Parrish, J. S. Holloway, M. P. Buhr, G. Hubler, F. C. Fehsenfeld, and J. F. Meagher, "Ozone production from Canadian wildfires during June and July of 1995," *Journal of Geophysical Research-Atmospheres*, vol. 107, no. D14, pp. 7–1:7–25, 2002.

[85] H. B. Singh, B. E. Anderson, W. H. Brune, C. Cai, R. C. Cohen, J. H. Crawford, M. J. Cubison, E. P. Czech, L. Emmons, H. E. Fuelberg, G. Huey, D. J. Jacob, J. L. Jimenez,

A. Kaduwela, Y. Kondo, J. Mao, J. R. Olson, G. W. Sachse, S. A. Vay, A. Weinheimer, P. O. Wennberg, A. Wisthaler, and A. S. Team, "Pollution influences on atmospheric composition and chemistry at high northern latitudes: Boreal and California forest fire emissions," *Atmospheric Environment*, vol. 44, no. 36, pp. 4553–4564, 2010.

[86] J. I. Levy, T. J. Carrothers, J. T. Tuomisto, J. K. Hammitt, and J. S. Evans, "Assessing the public health benefits of reduced ozone concentrations," *Environmental Health Perspectives*, vol. 109, no. 12, pp. 1215–1226, 2001.

Chapter 1 Supplemental Information

A.1. HYSPLIT point seasonality and location figures for non-US regions:

In chapter one figure 1.4 shows only U.S. regions. Here we provide versions of these figures for all regions shown in figure 1.1. All figures were generated using the same methodology described in chapter 1.

(A)

(B)

(C)

(D)

(E)

Land Cover

Barren or Sparsely Vegetated	Evergreen Needleleaf Forest	Savanna
Cropland/Grassland Mosaic	Grassland	Shrubland
Cropland/Woodland Mosaic	Irrigated Cropland and Pasture	Snow or Ice
Deciduous Broadleaf Forest	Mixed Forest	Unlabelled land area
Dryland Cropland and Pasture	Mixed Shrubland/Grassland	Wooded Tundra
Evergreen Broadleaf Forest	Mixed Tundra	Wooded Wetland
	noAssignment	

72

(F)

(G)

(H)

New Brunswick
n=49

(I)

Prince Edward Island
n=0

Month

Duration (hr)

(J)

Yukon
n=6962

(K)

Land Cover
- Barren or Sparsely Vegetated
- Cropland/Grassland Mosaic
- Cropland/Woodland Mosaic
- Deciduous Broadleaf Forest
- Dryland Cropland and Pasture
- Evergreen Broadleaf Forest
- Evergreen Needleleaf Forest
- Grassland
- Irrigated Cropland and Pasture
- Mixed Forest
- Mixed Shrubland/Grassland
- Mixed Tundra
- noAssignment
- Savanna
- Shrubland
- Snow or Ice
- Unlabelled land area
- Wooded Tundra
- Wooded Wetland

74

(L)

(M)

(N)

Land Cover

- Barren or Sparsely Vegetated
- Cropland/Grassland Mosaic
- Cropland/Woodland Mosaic
- Deciduous Broadleaf Forest
- Dryland Cropland and Pasture
- Evergreen Broadleaf Forest
- Evergreen Needleleaf Forest
- Grassland
- Irrigated Cropland and Pasture
- Mixed Forest
- Mixed Shrubland/Grassland
- Mixed Tundra
- noAssignment
- Savanna
- Shrubland
- Snow or Ice
- Unlabelled land area
- Wooded Tundra
- Wooded Wetland

FIGURE A.1. HYSPLIT Points locations and land cover classification represented by color (left), total number of occurrence by month and land cover classification (middle), duration and land cover classification (right). Please refer to the legend in figure 1.3 for land cover classification for non U.S. regions.

A.2. Overlapping trajectories with HMS smoke plumes:

HMS analyzes smoke plumes using visible satellite imagery. During the summer months most latitudes of North America experience daylight hours that occur in two different UTC dates. For example, on the summer solstice in Seattle (47 degrees North) the sun sets at approximately 9:11 PM local time (PDT) which is 04:11 UTC the next date. However, operationally HMS stores daily smoke plume data for a single date relative to North America, though technically these files contain smoke plumes from two different UTC dates. In order to make the daily overlap analysis more relevant to the daylight hours over North America (when the smoke is observed), our overlap analysis shifted HYSPLIT trajectory times backwards by 6 hours. Had we not done this fires that start in the late afternoon could never overlap that dates smoke plumes since those trajectories occur after 24 UTC (next date). The adjustment allows more trajectories UTC dates to better overlap the daylight analysis date of the smoke plumes (including early morning). The results presented in Chapter 1 are highly insensitive to whether or not the 6 hour adjustment was applied. Previous versions of the figures and analysis prior to the 6 hour adjustment are nearly identical.

A.3. Smoke hour transport figures for non-US regions:

Here we present the GDAS smoke hour transport Figures (same as Figure 1.7) for regions outside of U.S. States. The U.S Islands and Cuba are not included due to limited smoke-hours over these regions.

(A)

(B)

81

Smoke Hours

n=8917716

Ontario

(L)

Nunavut

(M)

Smoke Hours

n=10709674
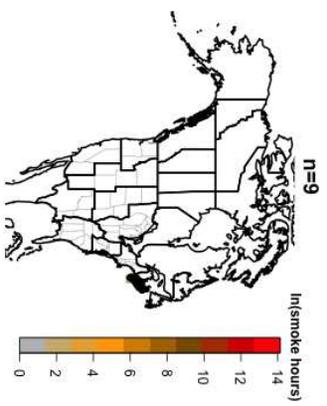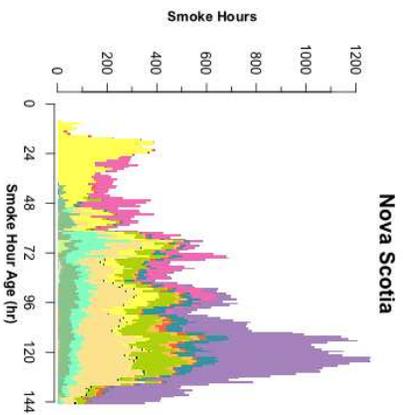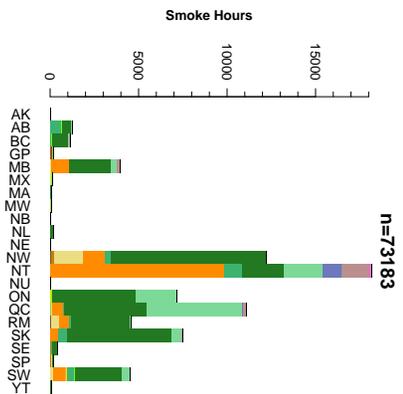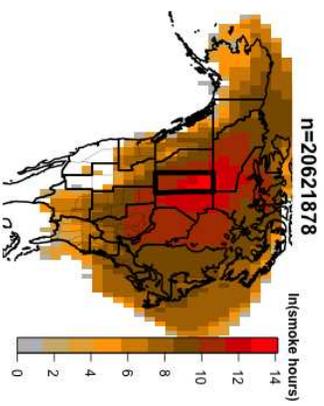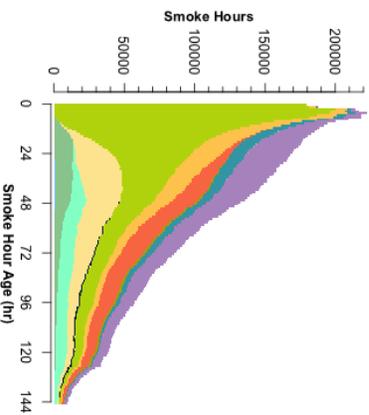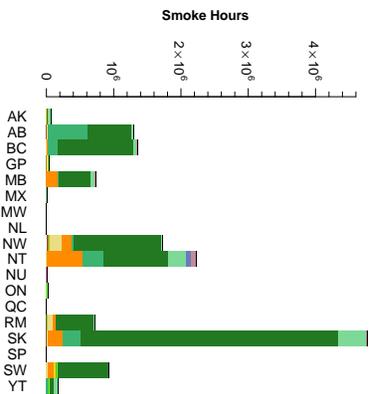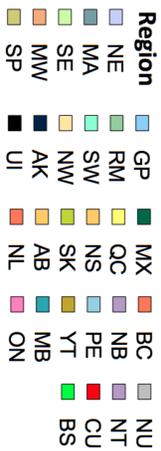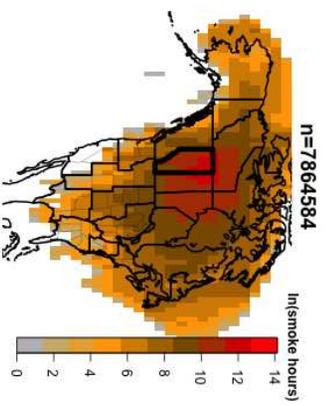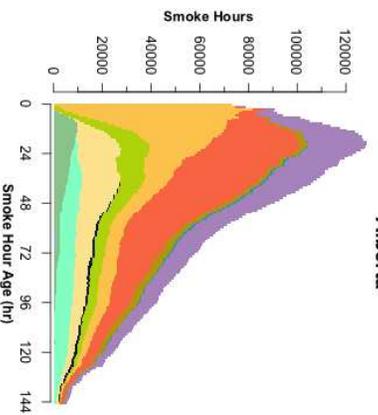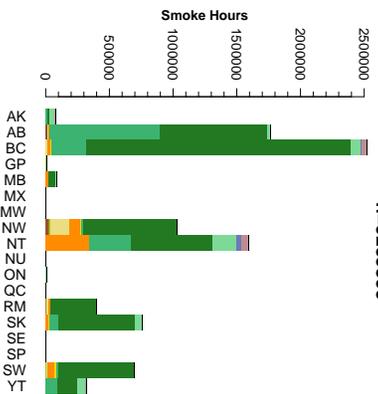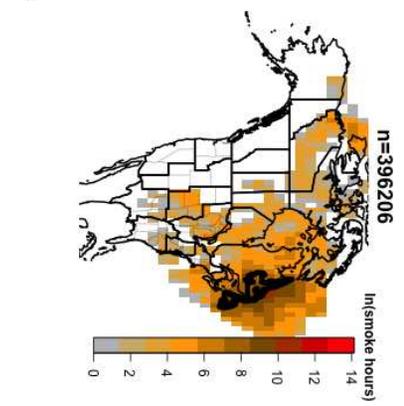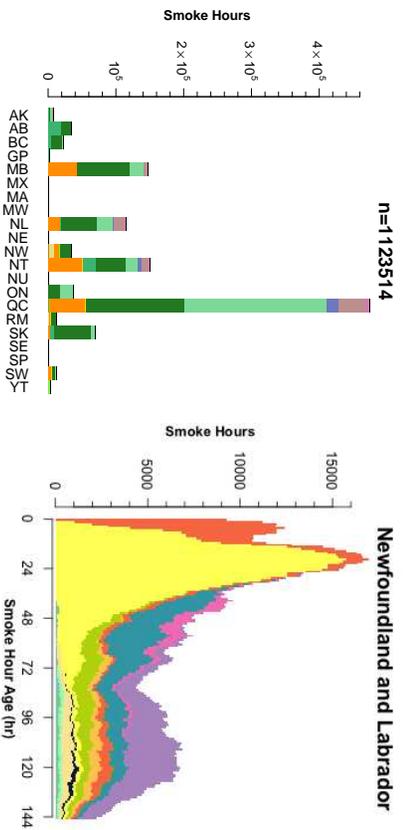
Land Cover

Region

FIGURE A.2. Summertime (June-September) regional smoke hour transport summaries for non-U.S regions. Column 1 (left) shows the number of smoke hours in the region by source region and land cover classification with the same color scheme as Figure 1.2. Only source regions with non-zero smoke hour contributions are shown. Column 2 (middle) shows the distribution of smoke age segregated by source region using the colors from Figure 1.1. Column 3 (right) shows the count of smoke hours produced by a region on a $2°$ x $2°$ degree grid with consistent colorbar for all regions. The grid spans 18 - $180°$W and 18 - $90°$N, a domain covering all five sectors where HMS analyzes smoke plumes (only a subset plotted). Shaded values are the natural log of the number of smoke hours in each grid cell. All figures generated using GDAS1 meteorology data for the months June-September 2007-2014.

Figures 1.5, 1.6, and 1.7 are all driven by GDAS1 meteorology data. This section provides these figures when EDAS40 meteorology data is used.

Total Smoke Hours Produced
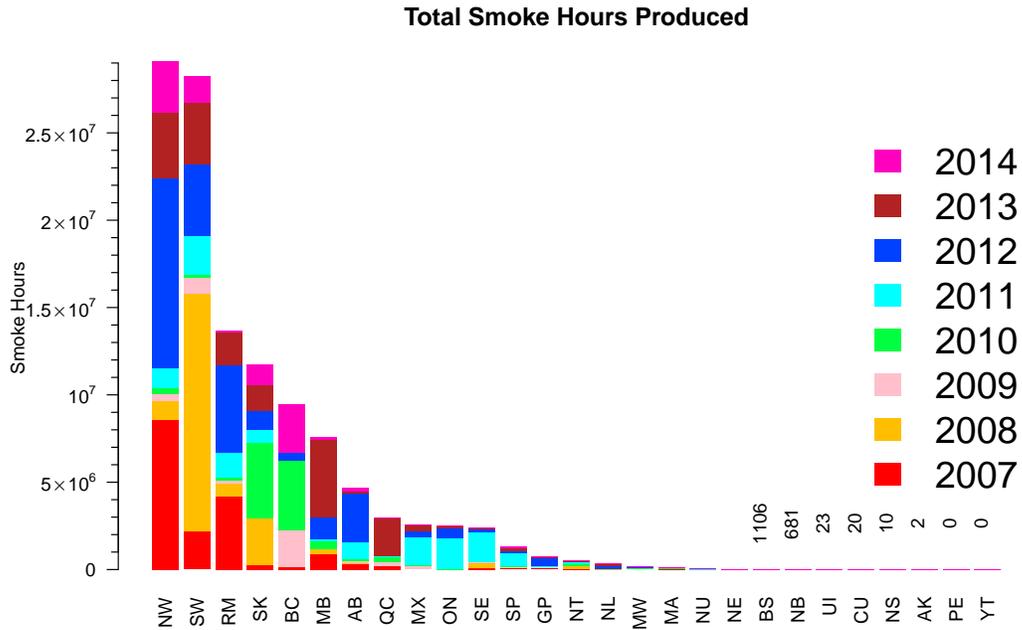
FIGURE A.3. Total number of smoke hours produced (anywhere) by each region for months June-September between 2007 and 2014 using EDAS40 meteorology.
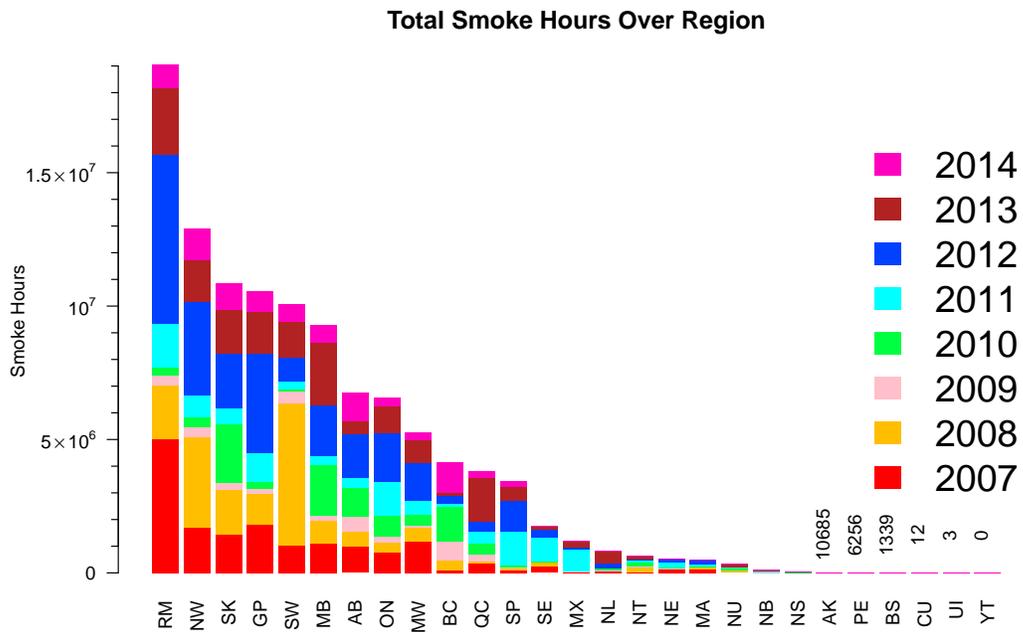
FIGURE A.4. Total number of smoke hours over each region sorted from highest to lowest for months June-September between 2007 and 2014 using EDAS40 meteorology.

(A)

(B)

(C)
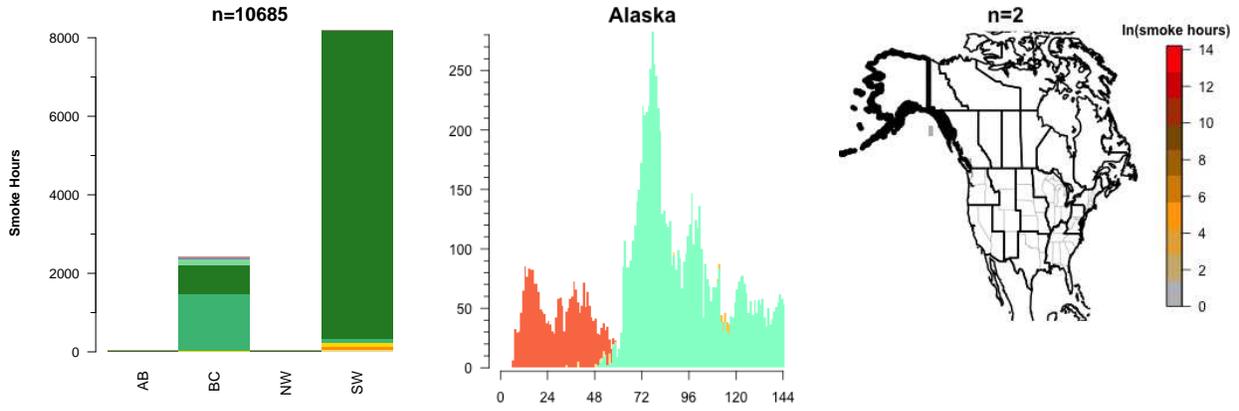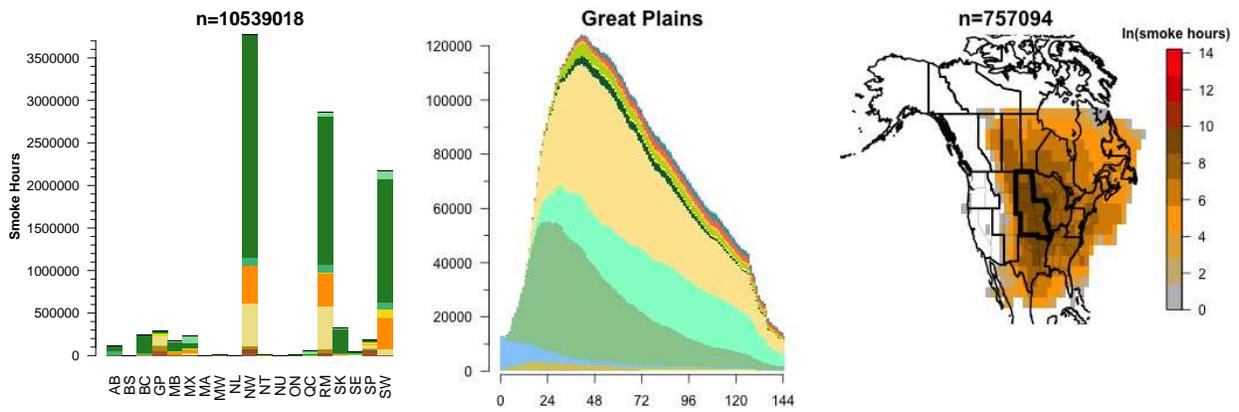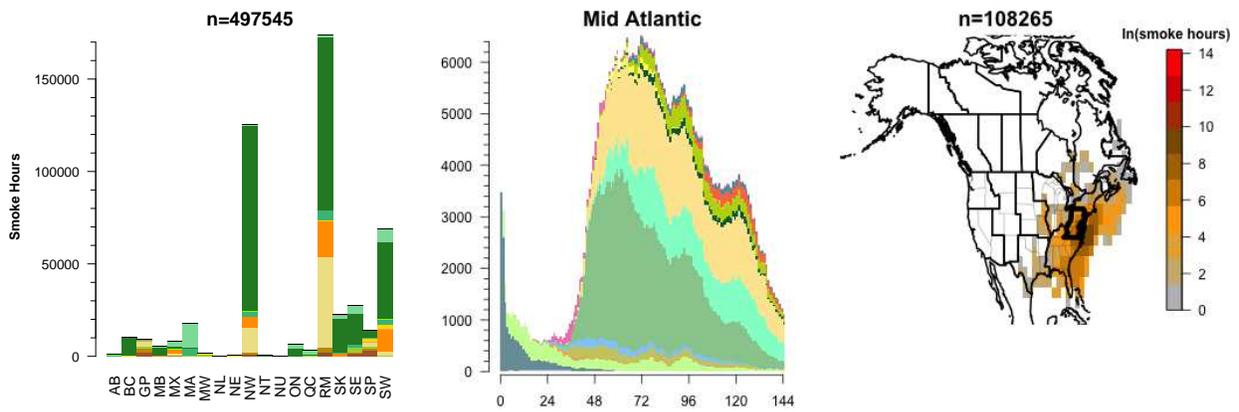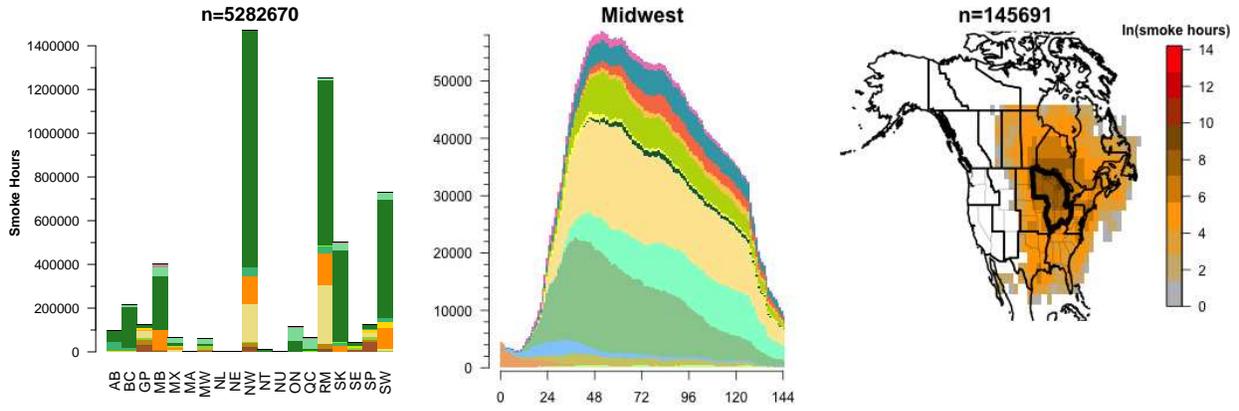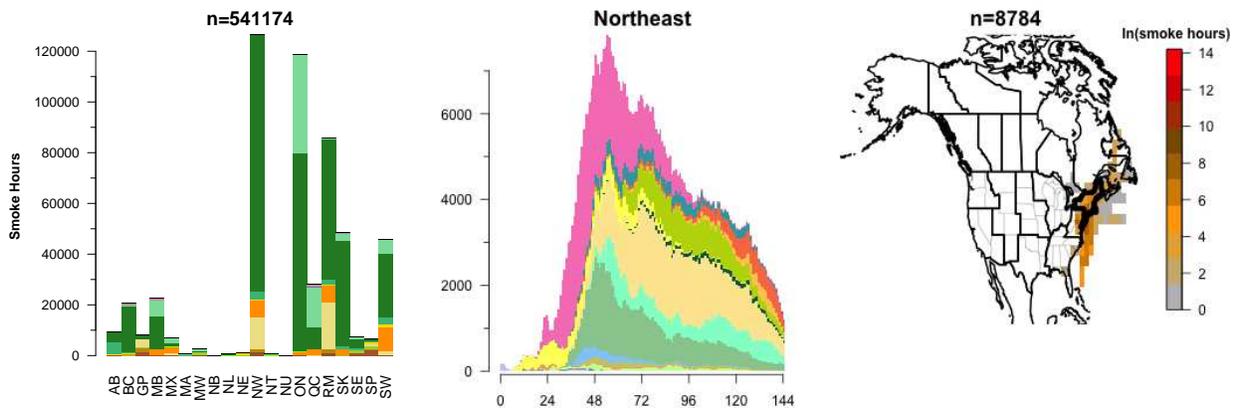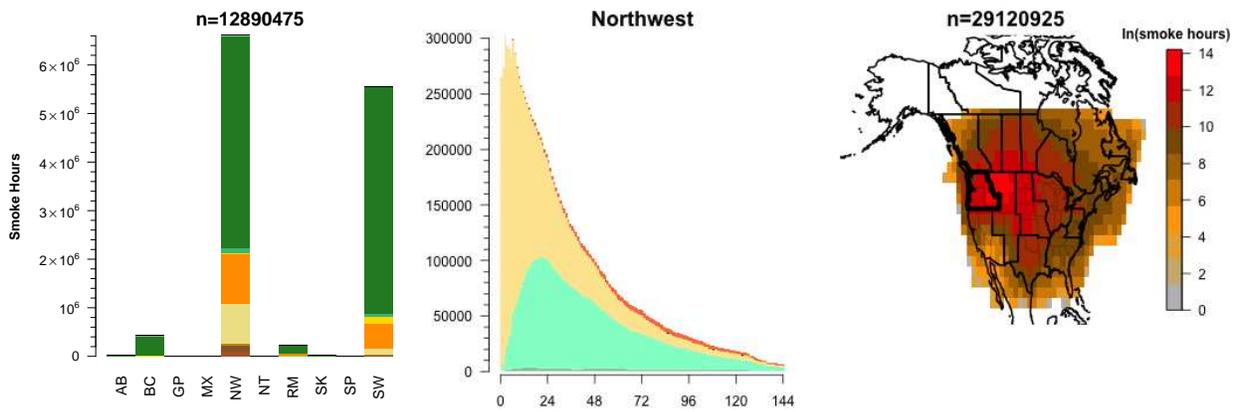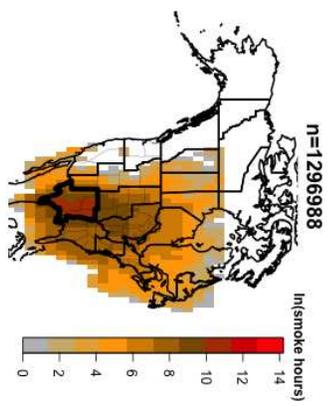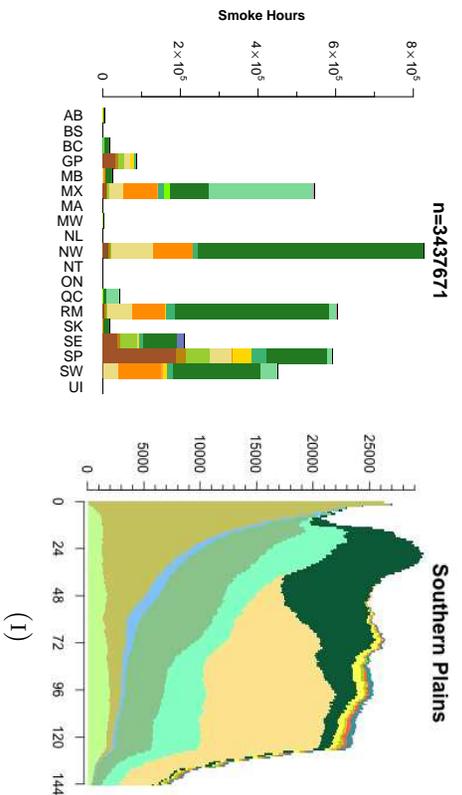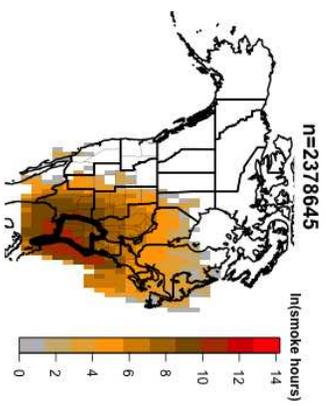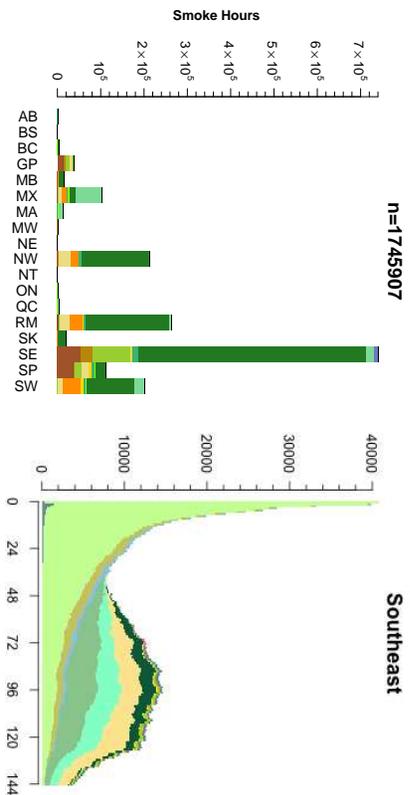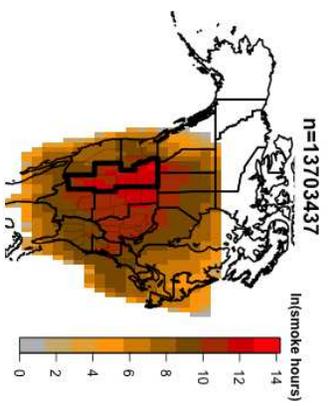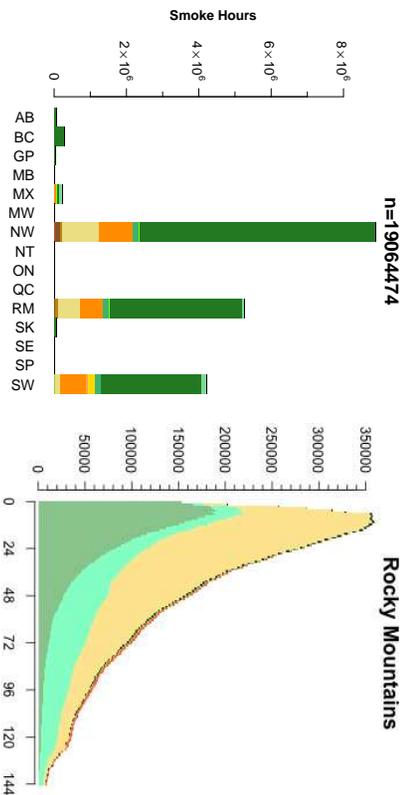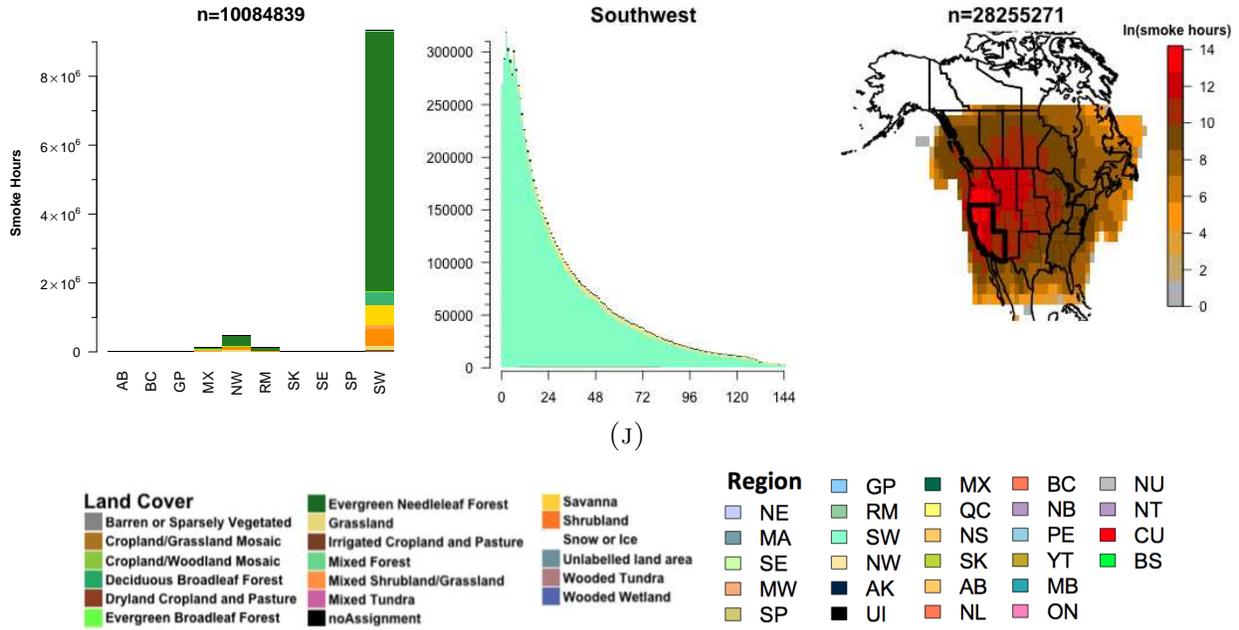
Figure A.5. Summertime (June-September) regional smoke hour transport summaries. Column 1 (left) shows the number of smoke hours in the region by source region and land cover classification with the same color scheme as Figure 2.1. Only source regions with non-zero smoke hour contributions are shown. Column 2 (middle) shows the distribution of smoke age segregated by source region using the colors from figure 1.1. Column 3 (right) shows the count of smoke hours produced by a region on a 2° x 2° degree grid with consistent colorbar for all regions. The grid spans 18 - 180ºW and 18 - 90ºN, a domain covering all five sectors where HMS analyzes smoke plumes (only a subset plotted). Shaded values are the natural log of the number of smoke hours in each grid cell. All figures generated using EDAS meteorology data for the months June-September 2007-2014.

A.5. Winter heat maps for all regions:

This section contains smoke hour heat maps for all regions with non-zero smoke hours produced in the months January through March years 2007 through 2014. These figures are identical to Figure 1.8 described in Section 1.3. The intention of these figures is to highlight possible smoke-hour false alarms, which are indicated by disconnected regions of smoke hours.

(A)

(B)

(C)

90
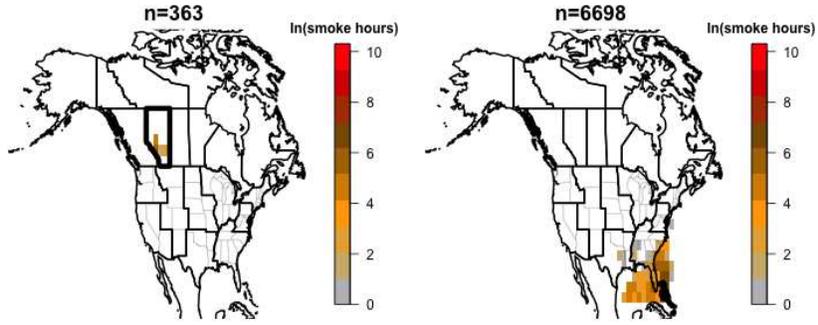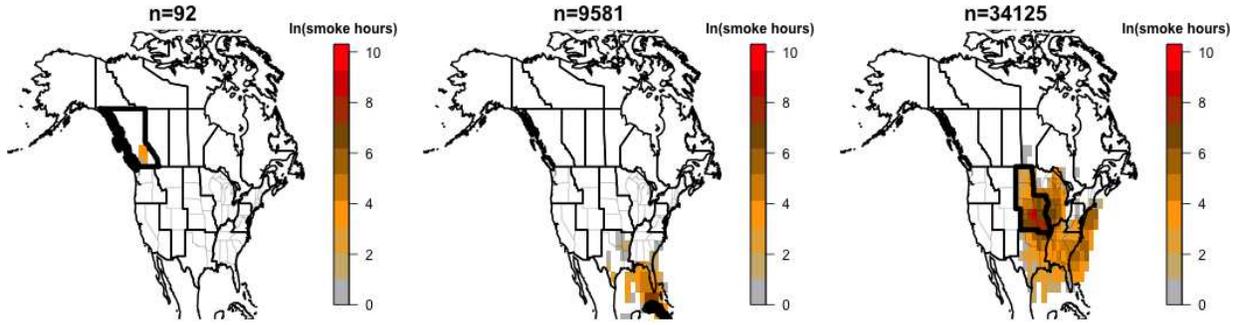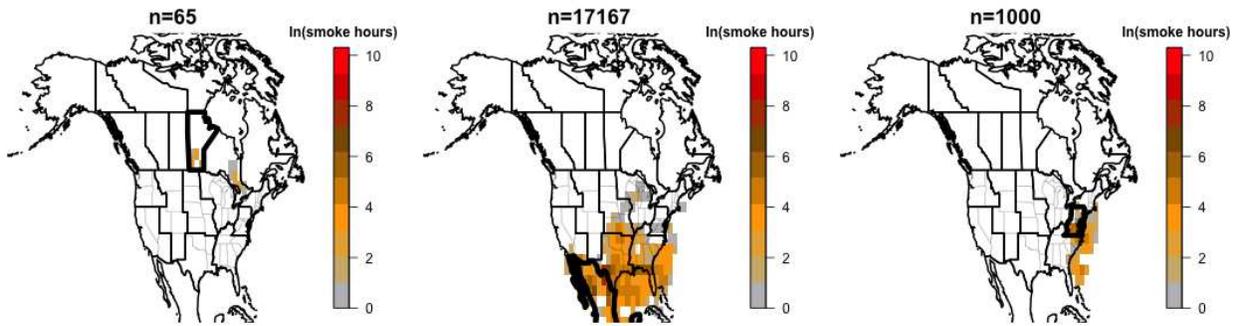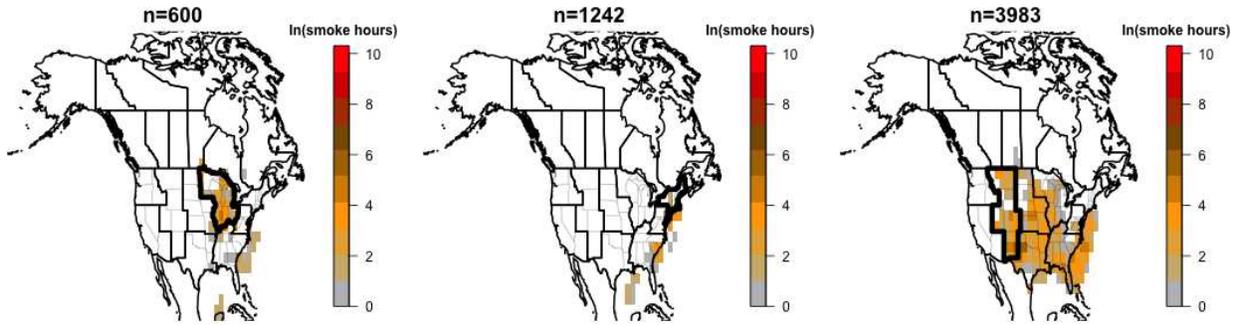
(D)
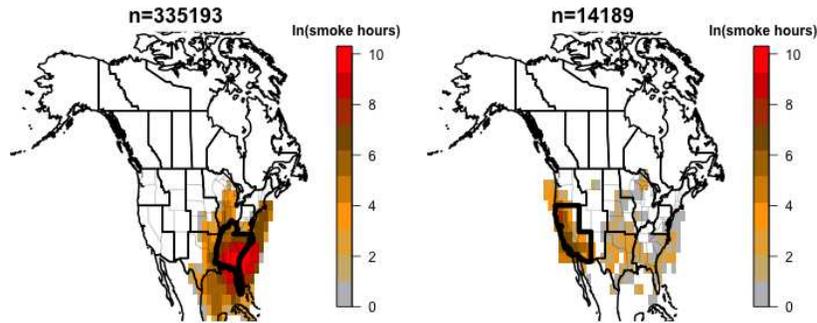
(E)
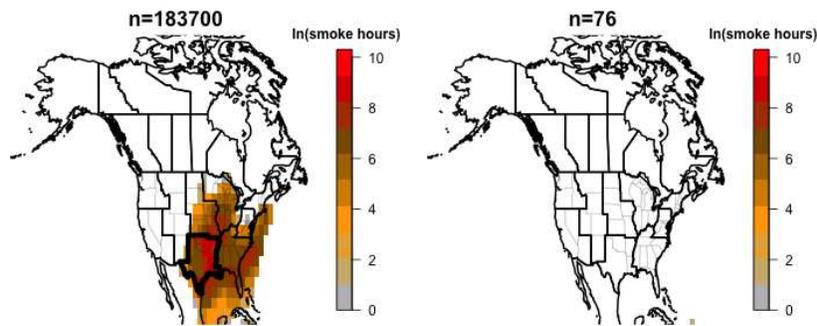
(F)

FIGURE A.6. Total number of smoke hours produced by regions for months January-March between 2007 and 2014 using GDAS1 meteorology. Only regions with more than zero smoke hours produced are shown. Region labels are as follows; row A: (left to right) Alberta, Bahamas. Row B: British Columbia, Cuba, Great Plains. Row C: Manitoba, Mexico, Mid Atlantic. Row D: Midwest, Northeast, Rocky Mountains. Row E: Southeast, Southwest. Row F: Southern Plains, U.S. Islands.

A.6. DETAILED DESCRIPTION OF OBSERVED WEAKNESSES IN LAND COVER

CLASSIFICATION ASSIGNMENTS:

I observed the following weaknesses when comparing the land cover assignment made using the methods described in section 1.2.4 to visible imagery provided by Google Earth. I extensively audited the quality of the land cover assignments made by plotting the locations in Google Earth and using the visual imagery to make my own human assessment of the land cover. I cannot determine species using this method but, I can tell things like crops, grass, forest, water, urban and crops apart.

- Cropland and grassland seem to get mixed up in dry places. Heavily irrigated (green in visible Google Earth imagery) farmland appears to more regularly be classified as cropland. I observed that cropland in Eastern Colorado are often classified as grassland.

- The summit of Mount Rainier and Mount Baker are mixed forest. That is nonsense as these are heavily glaciated peaks and this dataset does have glaciers and snow cover. It correctly assigned snow and ice to the summit of a mountain in the Alaskan Range.

- There is also a consistent issue with assigning forestland cover classifications to agriculture in Western Washington. Because of the age of this dataset it is possible that this was indeed forest in the early 90s and has since been converted to cropland.

- At the interface between grass, shrubs and forests it takes a considerable distance to transition assignments to forest. For example, the data assign Arthurs Rock in Lory State Park as grassland.

- There are around 10,000 HYSPLIT points ($\sim 1.3\%$) that are not given a land type assignment because the land cover data is convinced they are in urban areas or

water, which my current methods do not allow. One example of when my methods fail is the HYSPLIT point that occurred on 2005-09-29 at (34.195, -118.259), near the middle of the Verdugo Mountains in Southern California. These mountains are less then 5 km across and are surrounded by expansive heavily developed cities that include Pasadena, Glendale, and Burbank California.

Overall the land cover assignments seem to make sensible assignments and match the MODIS land cover classification used in WRF-Chem. For the purpose of distinguishing crops vs. forest vs. shrubs this dataset and methods of 1.2.4 seems to deliver.

A.7. Code repository:

In an effort to be as open, transparent, and reproducible as possible all work associated with this project is stored in the following subversion repository: `http://salix.atmos.colostate.edu/svn/smokeSource/`. This repository includes every version of all code, figures, and writing associated with this project. Please direct questions about this repository to sjbrey@rams.colostate.edu.

APPENDIX B

Chapter 2 Supplemental Information:

B.1. Removing Trends From O_3 MDA8 Time Series:

In this analysis we account for upward and downward trends in O_3 mixing ratios by subtracting the linear trends from the data such that there is no change in the average summertime O_3 mixing ratio between 2005 and 2014. For example, if the average summertime MDA8 O_3 mixing ratio was 4 ppbv higher in 2005 than 2014, the result of the de-trending is to make this difference 0. This strategy makes it appropriate to compare smoke-impacted MDA8 values between different summers, even at locations where there is a positive or negative trend in O_3. In this analysis we required 10 smoke-impacted MDA8 days at a location to compare smoke-impacted vs. smoke-free days. Infrequent high O_3 events were not excluded from the O_3 data. This was done for two reasons. 1) If a high O_3 event was on a smoke-impacted day the elevated value may be because of enhanced O_3 production due to the O_3 precursors in the smoke plume. 2) High O_3 events on smoke-free days would tend to make the smoke-free O_3 distributions look like the smoke-impacted distributions. The result of this strategy is to make a conservative estimate of the magnitude of the difference between the mean of the smoke-impacted and smoke-free distributions.

Though we believe removing linear trends in O_3 data is sufficient to account for the irregular timing of smoke-impacted days, we did manually inspect the occurrences of smoke-impacted days in the O_3 time series. $\sim$98% of the monitors we investigated that were subject to the minimum data requirements to compare smoke-impacted and smoke-free MDA8 values (500 of 506), have smoke-impacted days that occur in multiple summers between 2005 and 2014.
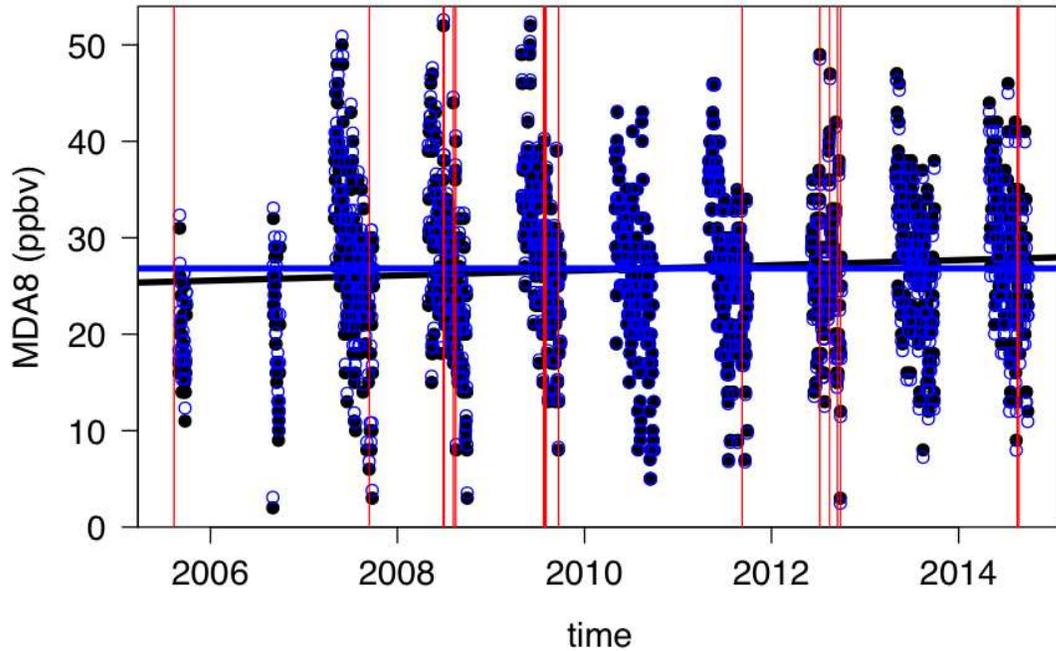
FIGURE B.1. Example MDA8 O_3 time series from a monitor located in Seattle, WA (ID 530330080). The solid black dots show the original O_3 MDA8 values. The open blue dots show the data after the ordinary least square fit regression (black line) are subtracted from the data. In order to plot side-by-side, the mean value of the time series is added to de-trended values (open blue dots). The blue line is the slope of the de-trended data (0 ppbv/year). The vertical red lines show the dates when smoke is present. At this location there is a slight upward trend in O_3. However, removing the linear trend allows us to compare smoke-impacted MDA8 values from 2005 and 2014 without the overall trend biasing the 2014 smoke-impacted MDA8 value upwards.

B.2. CODE REPOSITORY:

In an effort to be as open, transparent, and reproducible as possible all work associated with this project is stored in the following git repository: `https://github.com/stevenjoelbrey/SmokeInTheCity`. This repository includes every version of all code written after the correction was written for the originally published version of the manuscript. Please direct questions about this repository to sjbrey@rams.colostate.edu.