ESTIMATING PM2.5 CONCENTRATIONS USING MODIS AND METEOROLOGICAL MEASUREMENTS FOR THE SAN FRANCISCO BAY AREA

A thesis submitted to the faculty
of San Francisco State University
In partial fulfillment of
The Requirements for
The Degree

Master of Arts
In
Geography

by

Carlos A. Jennings

San Francisco, California

August 2013

CERTIFICATION OF APPROVAL

I certify that I have read *Estimating PM2.5 Concentrations Using MODIS and Meteorological Measurements for the San Francisco Bay Area* by Carlos A. Jennings, and that in my opinion this work meets the criteria for approving a thesis submitted in partial fulfillment of the requirements for the degree: Master of Arts in Geography.

Andrew Oliphant
Professor of Geography

Leonhard Blesius
Assistant Professor of Geography

ESTIMATING PM2.5 CONCENTRATIONS USING MODIS AND METEOROLOGICAL MEASUREMENTS FOR THE SAN FRANCISCO BAY AREA

Carlos A. Jennings
San Francisco, California
2013

Ground-level fine particulate matter (PM2.5) is a major component of urban air pollution with links to adverse health effects and is regulated by the EPA to meet federal standards. Aerosol Optical Depth (AOD) derived from Moderate Resolution Imaging Spectroradiometer (MODIS) satellite sensors can be used as an indirect measurement of PM2.5 due to the attenuation of atmospheric transmission of light by aerosols. This method has the potential to greatly increase spatial coverage of PM measurements and provide cost effective information for air quality decision-makers; however, regional differences alter the relationship between AOD and PM2.5. A multivariate analysis is presented in this study to predict PM2.5 concentrations in the San Francisco Bay Area from MODIS AOD and meteorological variables from ground-stations and Rapid Update Cycle (RUC) reanalysis products. Twelve PM2.5 monitoring stations are used for model training over a one year period. PM2.5 distributions were fairly uniform throughout the region with R^2 values between 0.06 and 0.28. PM2.5 was highest during the winter and at night when lower boundary layer heights and cold temperature inversions help to concentrate pollutants closer to the earth's surface. The results of this study confirm previous studies that found low correlations between AOD and PM2.5 in the western U.S.

I certify that the Abstract is a correct representation of the content of this thesis.

Chair, Thesis Committee Date

ACKNOWLEDGEMENTS

I would like to thank my thesis committee, Andrew Oliphant and Leo Blesius, for guidance and support. I would also like to thank Nancy Wilkinson for encouraging us to always see the big picture. Thanks to Seth Hiatt and Bill Goedecke for invaluable technical support, Kurt Malone from BAAQMD for providing data, and Bart Worley for spreadsheet and formatting help.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

1. Introduction

Particulate matter is one of the six criteria pollutants regulated by the Environmental Protection Agency (EPA 2013a). Fine particulate matter 2.5 micrometers or smaller (PM2.5) is the most harmful air pollutant to public health in the San Francisco Bay Area (BAAQMD 2013a). In order to meet the National Ambient Air Quality Standards (NAAQS), PM2.5 levels cannot exceed 35 $\mu g/m^3$ for the 24-hour mean and 12 $\mu g/m^3$ for the annual mean averaged over a three-year period. In 2009 California had the highest number of designated Nonattainment counties that did not meet annual and 24-hour PM2.5 National Ambient Air Quality Standards (EPA 2013b). Figure 1 displays a map of U.S. counties that have noncompliance status for exceeding annual PM2.5 standards as of 2012.

PM2.5 can be inhaled more deeply into the lungs due to its small size. The EPA designated fine particulate matter from coarse particulate matter in 1997 after reviewing health and epidemiological studies that linked fine particle exposure to cardiovascular symptoms such as cardiac arrhythmias and heart attacks as well as respiratory symptoms like bronchitis and asthma. Studies have also found a significant association between exposure to fine particulate matter and premature death from heart or lung disease. Fine particulate matter pollution results in a high cost to society in the form of increased emergency room visits and hospitalizations, work and school absences, and restricted outdoor activity. Older adults, children, and individuals with compromised immune systems and chemical sensitivities have a higher risk of negative health effects from exposure to PM2.5 (Dockery 2001; Kappos et al. 2004; Schwarze et al. 2006).

PM2.5 is currently monitored by ground stations which provide accurate point measurements for the immediate vicinity. Values are then extrapolated to give an assessment for the entire air basin. Satellite remote sensing for air quality assessment may provide more accurate estimations across a greater geographical area than can currently be sampled by ground stations alone (Martin 2008). The Moderate Resolution Imaging Spectroradiometer (MODIS) on board Terra, a polar-orbiting satellite operated by NASA, can measure characteristics of the atmosphere, including aerosol optical depth (AOD), a measurement of light extinction attributed to aerosols (NASA 2013a). MODIS AOD has been used in previous studies to estimate PM2.5 concentrations near the surface using regression analysis.

Multiple regression analysis will be used in this study to estimate ground-level PM2.5 concentrations using MODIS AOD and meteorological inputs. The goals of this study are to test if the low correlations from previous studies between AOD and PM2.5 in the western U.S. will also apply to the San Francisco Bay Area, and to determine if including meteorological parameters such as the planetary boundary layer (PBL), temperature, and relative humidity will improve the AOD-PM2.5 relationship in the study area.

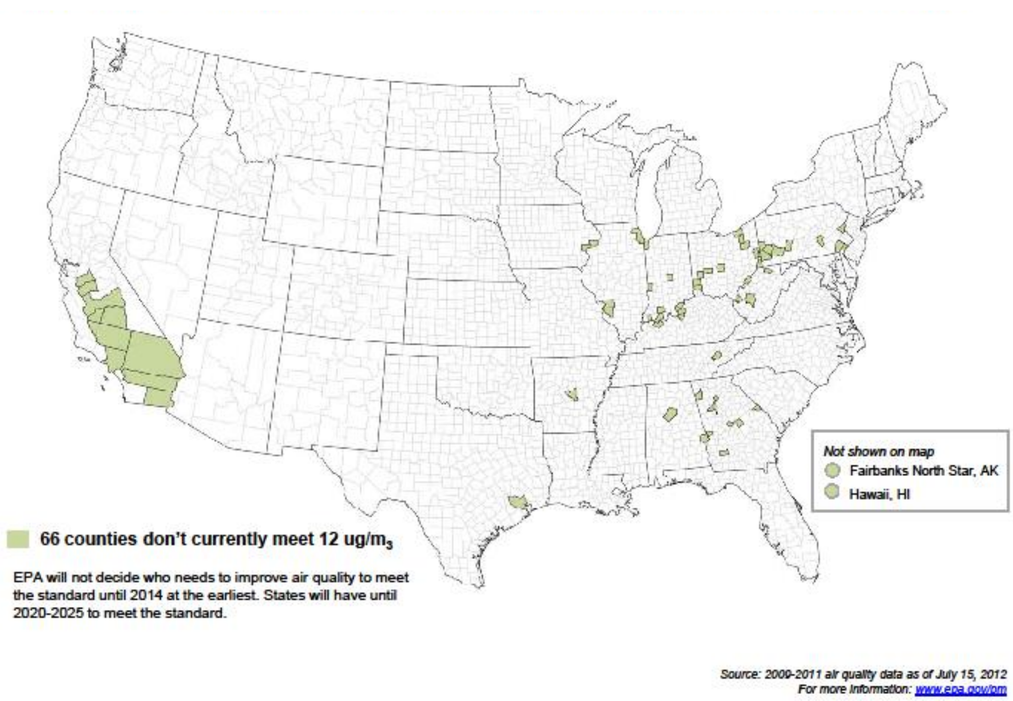


Figure 1. U.S. counties that are not in compliance with annual PM2.5 standard (EPA 2013b).

2. Literature Review

2.1 PM2.5 – Standards and Regulations

Particulate matter is a complex mixture of microscopic solids and liquid droplets suspended in the air that consists of a number of components including: nitrates and sulfates, metals, soil and dust particles, organic chemicals, fragments of pollen and mold spores, water, soot, smoke and tire rubber (EPA 2013c). Fine particulate matter (PM2.5) is defined as having a diameter of 2.5 micrometers and smaller. That is approximately 1/30th the diameter of a human hair (figure 2).

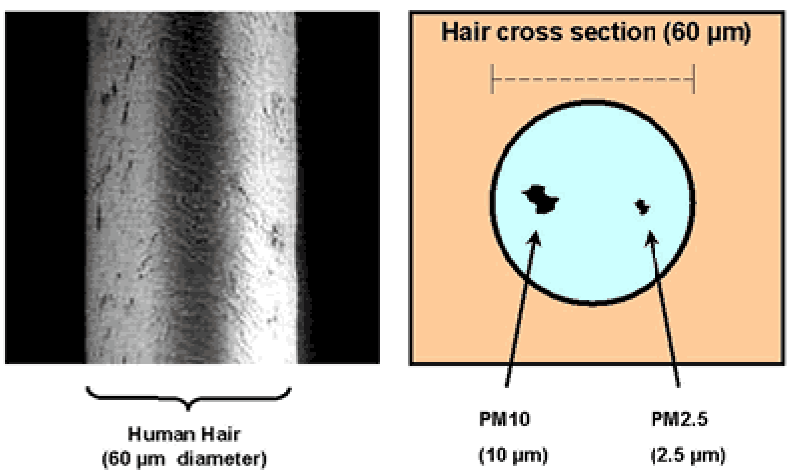


Figure 2. A size comparison of fine particulate matter (NYDEC 2013).

Fine particulate matter can be directly emitted or formed by secondary chemical reactions in the atmosphere. Sulfur dioxide, for example, is a type of secondary particle formed from industrial facilities and power plants. Nitrates are formed in a similar fashion from nitrogen oxides emitted from power plants, automobiles and other sources of combustion (EPA 2013c). When NAAQS were first implemented in 1971 particulate matter was measured by a high-volume sampler that measured particles between 25 and 45 micrometers. It was designated by the primary standard of total suspended particles (TSP). Annual regional standards were not to exceed 260 µg/m^3 within a 24-hour period more than once and with an annual geographic mean set at 75 µg/m^3.

The EPA subsequently reviewed the original standards in 1979 and implemented changes by 1987, when the primary indicator of particles was changed from TSP to PM10, categorically delineating particles small enough to penetrate deeper into the respiratory tract (thoracic particles) and therefore more prone to adversely affect public health. The new primary standard for PM10 was not to exceed 150 µg/m^3 more than once a year and with an annual mean of 50 µg/m^3 (EPA 2013d).

A second review began in 1994 that led to fine particulate matter 2.5 µg or smaller designated as a discreet category subject to regulation. The decision was based on studies that associated fine particles with serious health effects (Pope 1995; Laden et al. 2000; Harrison & Yin 2000; Schlesinger 2007; van Donkelaar et al. 2010). Standards for PM2.5 were set at 15 µg/m^3 (annual mean) and 65 µg/m^3 (24-hour average). The standards for PM_{10} were maintained at the previous levels set in 1987. The new regulations, which took effect in 1997, were challenged in court by several organizations, industries, and state governments. The U.S. Supreme Court ultimately upheld the new standards by a unanimous decision in 2001 affirming the EPA's authority to regulate air quality under the Clean Air Act and ruled that cost cannot be considered in setting standards. The EPA further tightened PM2.5 standards in 2006, reducing the 24-hour levels to 35 µg/m^3 from 65 µg/m^3 and again in 2012 when the annual mean standard was lowered to 12 µg/m^3 (EPA 2013d).

2.2 Remote Sensing of Particulate Matter

The use of satellite remote sensing to analyze atmospheric aerosols was first used in the 1970s. Satellite images are being attenuated due to the absorbing and scattering effect of particles in the atmosphere and algorithms were designed to correct for this distortion; however, it was realized that studying the backscatter radiation itself could reveal properties about aerosols in the atmosphere. (Veefkind et al. 1999).

The aerosol optical depth (AOD) parameter is a measurement of the degree of light extinction attributed to aerosols. AOD is a unitless measurement that describes the opacity of the atmosphere. For example, a very clear bright blue sky has an AOD of about 0.05. That is a very blue sky. As the sky becomes paler blue to murky white the AOD increases. AOD values approaching 1 are indicative of very hazy conditions. In extreme situations like dust storms, large fires or polluted urban areas, AOD can be 2.0 to 3.0. When AOD is 3.0 or above, one can look directly at the solar disk with the naked eye and the disk appears red. If AOD approaches 8.0, you cannot see the sun at all (NASA 2013a).

The presence of aerosols in an atmospheric column of air will prevent a certain amount of reflected light from being transmitted to the satellite sensor. It then follows that the optical depth (or thickness) of the atmosphere could be used as a proxy for estimating the amount of particulate matter near the surface. Table 1 lists previous research applying this method to obtain surface PM2.5 measurements from MODIS that show a strong correlation between satellite-measured AOD and ground-level PM2.5 measurements (Chu et al. 2003; Wang and Christopher 2003; Liu *et al*. 2004; Engel-Cox *et al*. 2004; Hutchison *et al*. 2005). Applied remote sensing for estimating and monitoring particulate matter and other types of air pollution is still in an early developmental phase. Many studies have used MODIS data to monitor pollution due to its daily global coverage and the 10 km^2 spatial resolution of its aerosol products.

Study	Date	Method	Instrument	Variables	Location	R^2
Liu et al.	2007	Multiple regression	MODIS/MISR	bl,temp,rh,ws	St. Louis, MO	0.71
Gupta & Christopher	2008	Linear regression	MODIS	AOD	Birmingham, AL	0.53
Li et al.	2009	Linear regression	MODIS	AOD	Southern U.S.	0.46
DiNicolantonio et al.	2009	Multiple regression	MODIS	bl, rh	Northern Italy	0.68
Tian & Chen	2009	Multiple regression	MODIS/GOES	bl,ws,temp,rh	So. Ontario	0.64
Wang et al.	2010	Multiple regression	MODIS	bl,temp,rh,ws	Beijing, China	0.66
Mansha & Ghauri	2011	Multiple regression	MODIS	temp, rh, ws	Karachi, Pakistan	0.65
Mao et al.	2011	Multiple regression	MODIS	AOD,land-use	Florida	0.58
Tsai et al.	2011	Multiple regression	MODIS/AERONET	bl, rh	Taiwan	0.67

Table 1. Recent studies correlating MODIS AOD and observed PM2.5.

MODIS derived AOD from both the Terra and Aqua satellites has been applied to improving estimates of the spatial distribution of PM2.5. on a regional and urban scale with varying degrees of success (Liu et al. 2007; Wallace & Kanaroglou 2007; Gupta & Christopher 2008; Kumar et al 2008; Christopher & Gupta 2009; Glantz et al. 2009, Li et al. 2009; 2010; Wang et al. 2010; Lee et al. 2011; Mansha & Ghauri 2011; Mao et al. 2011). AOD derived from MODIS cannot be used to replace ground-based measurement stations for year-round monitoring because of such limiting factors as cloud-cover contamination. However, it is effective for predicting ground-level PM2.5 for the hours of MODIS overpass from which PM2.5 at other hours can be interpolated (Tian & Chen 2009).

2.3 MODIS AOD and Particulate Matter

A number of studies since 2003 have applied MODIS AOD to estimate particulate matter concentrations near the surface (Gupta et al 2006; DiNIcolantonio et al. 2009; Schapp 2009; Tian & Chen 2009; Natunen et al. 2010). While other instruments, such as MISR, have a finer spatial resolution, they do not have a daily global revisit time. This makes MODIS more practical from a planning/management stance for estimating daily particulate matter. A quantitative relationship was first established using a simple linear regression to correlate AOD values with ground-based PM2.5 measurements and their variations over space and time (Chu et al. 1998, 2003; Engel-Cox 2004; Zhang et al. 2006; Gupta & Christopher 2008; Christopher & Gupta 2009; Li et al. 2009). Almost all of these studies were on a global or national scale and the strength of the correlations varied based on geographical location (see table 1). Nevertheless, the potential for applying MODIS AOD to monitor and predict pollution levels near the surface was promising enough to warrant further investigations.

2.3.1 Simple regression models

Chu et al. (2003) published one of the first papers to apply satellite derived AOD values to estimate particulate matter on a regional and urban scale. The study areas were eastern China and India, the eastern United State/Canada and Western Europe. These sites represent the most populous and industrialized regions on the planet. Three urban areas were also selected: Northern Italy, Greater Los Angeles, and Beijing, China. The study had two main objectives: to determine if AOD measurements from MODIS were as accurate as ground-based LIDAR measurements; and to investigate whether AOD data was robust enough to estimate particulate matter on multiple scales. Regression analysis was used to correlate LIDAR derived AOD measurements with ground-based measurements of PM10, and then AOD data from MODIS were compared with the LIDAR derived AOD for validation. Linear regression comparisons of MODIS and AOD resulted in an R correlation coefficient greater than 0.90, demonstrating that MODIS AOD were as robust as AOD measured by ground-based LIDAR measurements.

Chu et al. (2003) did not directly correlate MODIS AOD data with ground-level PM measurements, but regressed ground-based LIDAR AOD with PM2.5

ground measurements instead. In northern Italy, this yielded an R-value of 0.82, demonstrating the potential of using MODIS AOD for assessing and forecasting regional air quality. The time period of the study was from July 2000 thru May 2001. The study found a strong seasonal variation with AOD maximum occurring in the spring and summer and the minimum during the winter.

They applied different methods to each site and did not adequately account for every step of their research. For example, meteorological conditions, such as upper atmospheric winds and temperatures, were qualitatively compared with AOD levels but were not incorporated as parameters in their statistical analysis. The scope and scale of the paper is ambitious and perhaps would have been clearer if it focused either on testing MODIS AOD for accuracy, or for estimating regional aerosol levels. Nevertheless, Chu et al. (2003) laid the groundwork for subsequent research correlating MODIS AOD with surface PM levels.

Engel-Cox et al. (2004) compared qualitative true-color images and quantitative AOD data from MODIS with ground-based EPA monitoring networks to determine if satellite remote sensing is feasible for monitoring urban air quality. They expand on the concepts presented in Chu et al. (2003) by incorporating MODIS derived AOD data and ground-based PM2.5 and PM10 measurements in a regression analysis. The study was on a large scale and included the contiguous United States and had three main goals: to test how well MODIS can visualize distinct aerosol transport events; to correlate AOD from MODIS with EPA ground-based measurements; and to assess the capabilities and limitations of MODIS AOD data for EPA air quality monitoring. The time period was from April 1 thru September 30, 2002.

Regional differences in the effectiveness of AOD measurements from MODIS were detected. R-values ranged from near zero to as high as 0.9. Correlations were high (>.50) east of 100° W (Engel-Cox 2004). The relationship of ground measurements and satellite-based measurements differ regionally. Local terrain, weather and climate patterns are some of the factors that affect the AOD-PM2.5 relationship. Nevertheless, the authors conclude that MODIS can be used to assist agencies in monitoring air quality to meet EPA standards.

Zhang et al. (2006) further investigated the geographical and seasonal variations in the correlation between AOD and $PM_{2.5}$ over the contiguous United States. Two years of MODIS AOD data (2005-2006) from both the Terra and Aqua satellites were matched with ground-measurements from the ten EPA regions to determine the relationship between AOD and $PM_{2.5.}$ Their results showed a clear geographical and seasonal influence on the strength of the PM2.5 relationship. Good results were observed mostly over the eastern United

States in summer and fall. The southeastern U.S. had the highest correlation (r = 0.63, r^2 = 0.40) while the southwest region, Region 9, which includes California, had the lowest correlation (r = 0.26, r^2 = 0.07). No meteorological parameters were included in this study.

In contrast to these large scale studies, Gupta and Christopher (2008) used seven years of MODIS AOD data and $PM_{2.5}$ ground measurements from one site in Birmingham, Alabama to evaluate the effectiveness of MODIS in monitoring air quality and to better understand the monthly, seasonal, and inter-annual relationship between fine particulate matter and AOD. The correlation between AOD and PM2.5 resulted in an r-value of 0.52 using daily mean PM2.5 data. The r-value increased to 0.62 when hourly PM2.5 was used.

Natunen et al. (2010) also conducted a multiple-year study of the relationship between AOD and PM2.5 on a smaller regional scale in Finland. Data collected between 2000 and 2006 were obtained from ground-measurements of PM2.5 at four stations in Helsinki and from MODIS AOD on the Terra and Aqua satellites to investigate how temporal PM2.5 averaging and seasonality affect the correlation between AOD and PM2.5. They found that time averaging increased the correlation compared with using hourly PM2.5. Hourly PM2.5 was paired with the nearest satellite overpass time. An additional hour on either side of the hourly measurement up to 24 hours was then averaged and correlated with the AOD value. The time average with the highest correlation varied among the four sites with the best results found using 19, 15, 5 and 24-hour mean values. Monthly mean correlations ranged between 0.57 and 0.91.

In a similar study, Tian & Chen (2010) looked at the effects of time averaging on AOD-PM2.5 correlations from MODIS over southern Ontario, Canada. Hourly, 3-hour and daily mean ground-level PM2.5 was compared to find the best overall correlation. The 3-hour time window performed best with an r-value of .593. Tian & Chen also found that AOD values aggregated over 3 x 3 pixels from the Terra and Aqua satellites improved the correlation by .03 compared with using the single center pixel values.

2.3.2 Multivariate models

Subsequent research developed more complex empirical models that incorporate environmental factors such as meteorology and geographic data into multiple-regression analyses at regional and urban scales (Wallace & Kanaroglou 2007; Kumar et al. 2008; Li et al. 2009; Mao et al. 2011). One of the most important environmental parameters to consider when evaluating air quality from satellites is the height of the planetary boundary layer (PBL) because it affects the vertical distribution of aerosols. PBL is generally more developed during the day with a strong inversion at the top. Within this layer, anthropogenic aerosols are well mixed and confined. This is important because AOD measures the aerosol extinction for the entire atmospheric column of air from the ground up to the satellite sensor. AOD measurements will be the same whether the boundary layer is well developed or not. In the case of a high boundary layer with a weak inversion, ground-based measurements of PM2.5 may not correlate well with AOD values (Gupta & Christopher 2009). The AOD-PM2.5 relationship strongly depends on the height of the planetary boundary layer. This, along with meteorological parameters like relative humidity that affect particle formation and growth, is why simple two-variable regression analyses relating AOD with PM2.5 alone are not adequate for estimating ground-level PM2.5. A number of studies have incorporated PBL height in multiple-regression analyses (Liu et al. 2007; Green et al. 2009; Glantz et al. 2009; Di Nicolantonio et al. 2009; Schapp et al. 2009; Wang et al. 2010; Tsai et al. 2011).

One of the first to include meteorology in a multiple-regression analysis of particulate matter was conducted by Gupta et al. (2006). One year of AOD retrievals from MODIS aboard the Terra and Aqua satellites along with ground-measurements of PM2.5 were used to assess air quality over 26 locations in and around Sydney, Delhi, Hong Kong, New York City, and Switzerland. A correlation coefficient of 0.96 was obtained between bin-averaged daily mean AOD and ground-level PM2.5 when boundary layer height and wind speed were included in the model. The best results were obtained when boundary layer heights were less developed (100-200m) and with less than 25% cloud cover.

Tian and Chen (2009) used multiple-regression analysis to predict hourly concentrations of PM2.5 on a regional basis for southern Ontario, Canada. They sought to further enhance the already established relationship between AOD and PM2.5 concentrations on the ground by adding boundary layer and meteorological parameters. The goal of this study was to provide a cost-effective approach for supplementing ground-based monitoring stations. Data from 2004 was used to develop and validate the model and the model-predicted values were highly correlated with ground-based observations with an R^2 of 0.64.

DiNicolantonio et al. (2009) used MODIS AOD and simulated climate models to predict ground-level PM2.5 in the Po River Valley in northern Italy. The time span of the study was for the entire year of 2004, three summer months (May-July) in 2007, and three winter months (January-March) in 2008. Their predictive model showed good agreement with *in situ* PM2.5 measurements (R2 = 0.68 for MODIS/Terra and R=0.59 for MODIS/Aqua). PM2.5 levels were found to be higher for winter months and satellite-based concentrations of PM2.5

tended to underestimate values by ~20%. The authors did not discuss the possible causes and implications of the seasonal variability found in the study.

Wang et al. (2010) investigated whether including relative humidity from meteorological stations and LIDAR derived boundary-layer height would improve the AOD-PM2.5 correlation in Beijing, China over a 15-month time period (July, 2007 thru October, 2008). Adding the boundary layer height and relative humidity improved the model by 14% (R2 = 0.48 for 2-varialble study; R2 = 0.62 for multivariate study).

Lee et al. (2011) proposed a method to calibrate MODIS AOD data to accurately predict ground-level PM2.5 concentrations. They claim it is the first study to establish PM2.5 – AOD relationship on a daily basis. MODIS onboard the Terra and Aqua satellites was used to retrieve AOD data. The study area covered parts of Massachusetts and Rhode Island and was divided into 387 10 km^2 grids using ArcGIS version 9.3. Measured and predicted annual mean PM2.5 concentrations were statistically compared using multiple regression correlation coefficients. Lee et al. obtained a higher correlation ($R = 0.79$, $R^2 = 0.62$) using the multiple regression model when compared with a simple linear regression of the same data ($R^2 = 0.26$; $R = 0.5$) and conclude that the performance of the multivariate model to predict surface-level PM2.5 concentrations is superior to a simple two variable linear regression model and that this method would be suitable for both time-series and cross-sectional health effect studies.

Mao et al. (2011) developed and tested an enhanced land-use regression (LUR) model that incorporated monthly AOD data from MODIS Terra. Typically, LUR models use variables such as land use/land cover, point-source emission estimates, population, and traffic-information to predict the long-term intra-urban distribution of air pollutants. A problem with this method is that it lacks temporal variability and is often limited in representing the spatial variability of pollutants. No meteorological data was incorporated into the model. The predicted concentrations were higher during summer/fall and are consistent with previous studies from the eastern United States (Chu et al. 2003; Zhang et al. 2006). This is likely due to the high humidity, temperatures and strong insolation that cause atmospheric ions to react and form aerosol products.

Mansha & Ghauri (2011) investigated the seasonal and spatial variation of aerosol concentrations over Karachi, Pakistan using MODIS AOD from the Terra and Aqua satellites, sun photometers, and ground-based PM2.5 and meteorological measurements from 2008. Higher PM2.5 concentrations were recorded in the winter over summer, possibly due to lower boundary layer heights during winter and increased wind speed during the summer monsoon months. MODIS AOD, temperature, relative humidity, and wind speed variables predicted PM2.5 measurements with an R^2 between 0.47 and 0.67 when compared with the daily recorded mean, with the highest correlations during winter.

A similar study was conducted by Tsai et al. (2011) using three years (2006-2008) of surface particulate matter and MODIS AOD retrievals to assess the potential of satellite based air quality monitoring in Taiwan. Relative humidity and boundary layer heights were included as independent variables. Two sun photometers were used to validate the MODIS AOD retrievals (R^2 = 0.82 for Terra; R^2 = 0.69 for Aqua). The height of the boundary layer was found to be critical to the AOD-PM2.5 relationship as evident from the higher correlations (R^2 = 0.77 to 0.86) found in the fall and winter when a stable and well-mixed boundary layer predominates. Correlation did not exceed 0.67 during the spring and summer when monsoonal flows cause strong convective mixing and instability.

2.4 Summary

A review of the recent literature shows regional and seasonal variations in the relationship between AOD and PM2.5 with widely varying results. Better results were found in the eastern United States in the spring and summer, which is also when fine particulate concentrations are highest (Engle-Cox 2003; Zhang et al. 2006; Mao et al. 2011). Conversely, studies located in south Asia and north central Italy had the highest concentrations and correlations between AOD and PM2.5 during the fall and winter (DiNicolantonio et al. 2009; Mansha & Ghauri 2011; Tsai et al. 2011). The literature also shows that the inclusion of meteorological parameters such as PBL, temperature, and relative humidity, improved model results over those that only included AOD and PM2.5 in the regression models.

3. Study Area

This study builds on previous research to establish if MODIS AOD is a valid tool for estimating PM2.5 for an urban region in the western United States, and specifically for the San Francisco Bay Area (SFBA), a region with a heterogeneous topography, surface cover and climate. The SFBA in this case refers to the counties connected to San Francisco and San Pablo Bay in north central California. The SFBA is the fifth largest metropolitan area in the United States with a population of over 8 million (U.S. Census 2012). Few similar studies using MODIS AOD to estimate PM2.5 have been conducted in the west, and none for the SFBA. There is also a well-developed ground-based monitoring system in 8 of the 9 counties in the region maintained by the Bay Area Air Quality Management District (BAAQMD 2013b).

The San Francisco Bay area has a Mediterranean climate characterized by somewhat rainy winters and dry summers. During winter, cold fronts bring wind and rain, yet conditions can be quite stagnant with strong capping inversions between storm systems. In contrast, strong advection currents caused by the temperature gradient between warm inland locations and cool Pacific currents dominate the summer months and bring widespread fog to coastal and some bayside locations. Figure 3 shows a map of the study area and the locations of the 12 BAAQMD stations. The bounding coordinates of the study area are 38.5° N, 36.8° S, 121° W, and 122.9° W.

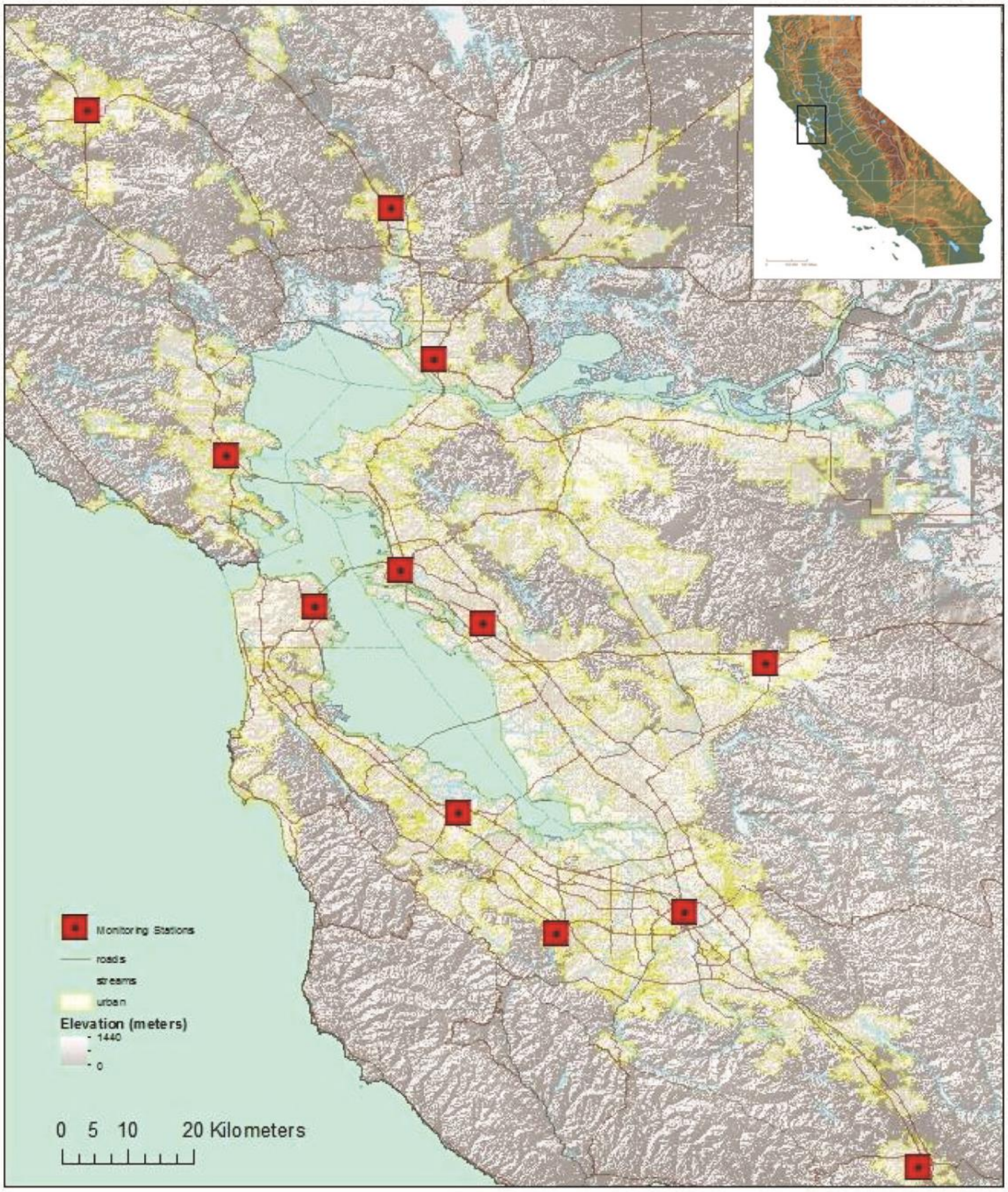


Figure 3. A map of the study region with locations for PM2.5 ground monitoring stations (BAAQMD 2013b).

4. Methods

The methods used to build a multiple regression model for estimating ground-level PM2.5 using MODIS will be discussed in this chapter. Data integration and multivariate statistical analysis will also be described. The main objective of this study is to test how well MODIS derived AOD can be used to estimate surface PM2.5 concentrations in the SFBA, as well as to evaluate if including meteorological parameters from NOAA's Rapid Update Cycle (RUC) improve the correlation between AOD and PM2.5. In North America, the bulk of research in estimating particulate matter using AOD on a regional scale has been for the central and eastern areas of the United States and Canada. These studies show that the correlation between AOD and PM2.5 varies according to geographic region and season (Gupta et al. 2006; Zhang et al. 2006). Furthermore, Engle-Cox et al. (2004) found that AOD- PM2.5 showed little to no correlation for Los Angeles, Portland, Oregon and Salt Lake City, Utah. They did not, however, incorporate meteorological parameters.

4.1 Data Acquisition & Integration

Information about the data, data sources, and software used will be discussed in the following sections. One year of data from 2011 was used for this study including: ground-level PM2.5 from the BAAQMD air quality dataset; AOD from MODIS aboard the Terra satellite; and archived PBL, temperature, and relative humidity from RUC. Ground point stations, satellite imagery, and meteorological measurements were co-located in space and time. PM2.5 from ground-based monitoring stations and PBL, relative humidity, and temperature from RUC for the nearest hour of MODIS/Terra flyover were matched with the coincident MODIS AOD for statistical analysis using ArcGIS 10.1. AOD files from MODIS were converted into GeoTIFF files and meteorological parameters from RUC were converted into NetCDF files in order to be compatible with ArcGIS. Ground stations were geo-located as points based on their geographic coordinates in ArcGIS. Coincident AOD values were then assigned to each ground station in a table format using the ‘extract multi-values to points’ tool. SPSS statistical software was used for the final regression analysis (see figure 4).

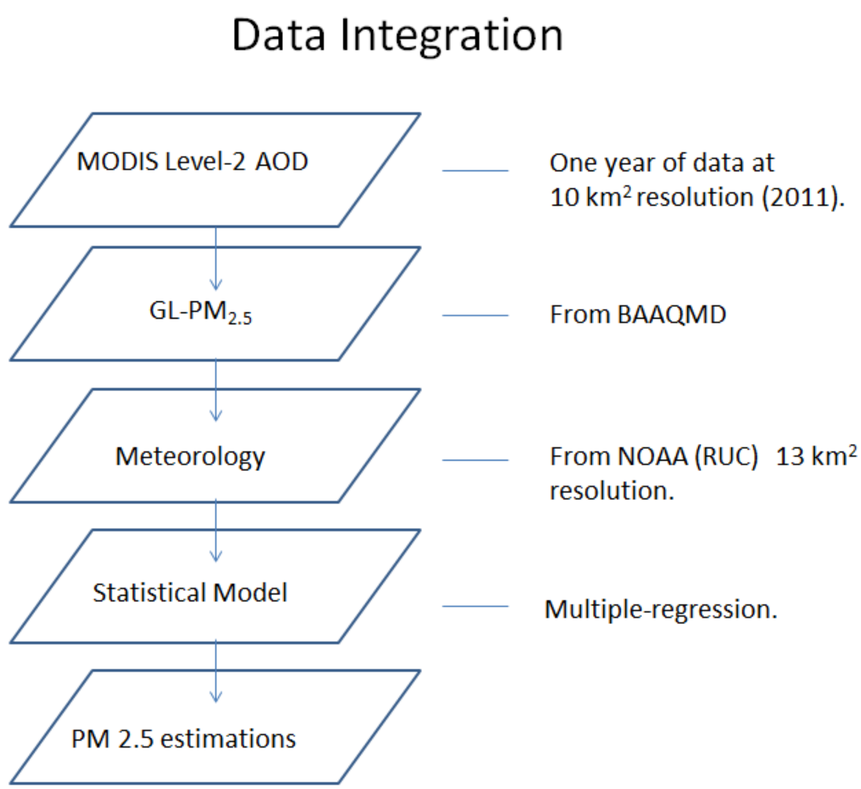


Figure 4. A procedural flowchart of methods and data integration.

The hourly PM2.5 measurements for 2011 that were obtained from BAAQMD were reformatted to only include the measurements for the hour nearest the satellite overpass times in order to build the statistical model. The overpass times ranged between 10 am and noon during standard time and between 11 am and 1 pm during daylight savings time. Days with missing data for the designated hour were first classified as 'No Data' in SPSS so that the results were not skewed. The sample size for the hours nearest the MODIS/Terra overpass was 3,475 with 184 missing values. The measured PM2.5 levels are the dependent variable in the regression analysis.

The AOD files from MODIS/Terra are the primary independent variable of interest for this study. There were a total of 1,673 valid inputs after all missing values were removed. The minimum recorded AOD value was 0 and the highest was 0.8. AOD is a unitless measurement and values approaching 1.0 or greater indicate hazy conditions. The remaining meteorological parameters: PBL height; temperature; and relative humidity from RUC are also included in the regression equation as independent variables.

The RUC data had 132 missing values; however, the number of days included in the statistical analysis was limited by the number of valid AOD retrieval days within the GIS grid cell containing a ground monitoring station.

These criteria limited the sample size of the RUC data to 1,541 out of the 3,660 valid samples.

4.1.1 PM2.5 ground measurements

Hourly PM2.5 data was obtained from BAAQMD, which operates a monitoring network of 12 air quality stations in eight SFBA counties. The stations measure a suite of air quality and meteorological parameters. The stations are located in Cupertino (37.32°N, 122.07°W); Gilroy (36.99°N, 121.57°W); Livermore (37.68°N, 121.78°W); Napa (38.31°N, 122.3°W); Oakland East (37.74°N, 122.17°W); Oakland West (37.81°N, 122.28°W); Redwood City (37.48°N, 122.2°W); San Francisco (37.77°N, 122.4°W); San Jose (37.35°N, 121.9°W); San Rafael (37.97°N, 122.52°W); Santa Rosa (38.44°W, 122.71°N); and Vallejo (38.1°N, 122.24°W).

The type of instruments used to measure PM2.5, as well as the frequency of sampling, is designated by the EPA and is referred to as the federal reference method (FRM). It specifies that small volume samplers with Teflon filters (46.2 mm diameter) be used to measure PM2.5 (Figure 5). BAAQMD uses the Rupprecht and Patashnick (R&P) Model 2025 Dichotomous Sequential Air Sampler and the Partisol Sequential Ambient Particulate Sampler in their ground monitoring stations (CARB 2013). Hourly PM2.5 data were collected from

BAAQMD in a spreadsheet for each of the 12 stations for the year 2011. Each station location was placed as a point feature on a map of the study site using ArcGIS 10.1 software and a table of the PM2.5 measurements for the nearest hour of MODIS/Terra flyover was linked to each station.

Figure 5. BAAQMD air monitoring station - Vallejo, CA (CARB 2013).

4.1.2 MODIS AOD

Aerosol Optical Depth (AOD) is an integral measurement of light extinction (absorption and scattering) caused by aerosols in a vertical column of air and is central to remote sensing of aerosols. AOD can be defined as:

$$AOD = \int_0^\infty \delta \exp(z) dz \tag{1}$$

where $\delta_{\exp}(z)$ is the aerosol extinction coefficient at elevation *z* above the ground and *dz* is the vertical optical path (Schäfer et al. 2008). AOD is computed from MODIS using the radiative transfer model to track the path of irradiance at the top of the atmosphere down to the surface and back (Tian & Chen 2009; Tsai et al. 2011). MODIS derived AOD products utilize algorithms that compare observed spectral reflectance with lookup tables in order to match conditions as closely as possible for the retrieved values from which associated aerosol properties are then derived (Tian & Chen 2009).

MODIS, onboard Terra, is in a sun-synchronous polar orbit at an altitude of 705 km and has a daily revisit time near 11:00 am (NASA 2013b). Raw MODIS data are processed into three levels with varying spatial and temporal resolutions. Level 1 data are comprised of individual daily MODIS scenes with 36 bands of image data and a spatial resolution of 0.25 to 1km. Level 2 products are derived from Level 1 images and have been atmospherically corrected for surface reflectance products with a resolution of 10 km^2. The parameter Corrected_Optical_Depth_Land was downloaded in hierarchal data format (HDF) from NASA's Level 1 and Atmospheric Archive and Distribution System (LAADS Web) for January 1, 2011 thru December 31, 2011 from MODIS collection 5.1 – Level 2 Aerosol data (NASA 2013c) using a file transfer protocol (FTP) server (NASA 2013c).

MODIS uses three channels, Band 3 at 470 nm (blue), Band 1 at 660 nm (red), and Band 7 at 2130 nm (Mid-Infrared) to interpolate AOD at 550 nm (green). 2130 nm is used because aerosols are nearly transparent at this wavelength which therefore can be used to estimate surface albedo. Then the scattering and absorption in the visible wavelengths of 470 nm (blue) and 660 nm (red) is used to calculate the amount of light extinction attributed to aerosols and total AOD value at 550 nm is interpolated (Remer et al. 2013).

Retrieved AOD values are reported at a 10 km^2 resolution, but MODIS initially uses a higher resolution of 0.5 km^2 to measure reflectance values. Then each of the 400 pixels (20 x 20) within the 10 km^2 area is examined for cloud contamination by a cloud-screening algorithm (Gupta & Christopher 2008). After the cloudy pixels are removed, if more than 12 cloud-free pixels remain, the mean reflectance of the remaining pixels is used to retrieve AOD. The measured mean reflectance is matched in a lookup table that contain the pre-calculated reflectance for various aerosol models to determine the conditions that most closely match the observed spectral reflectance to retrieve aerosol properties; including AOD (Schaap et al. 2009; Tian & Chen 2009). Because dust particles are significantly larger than urban/industrial pollution and bio-mass burning aerosols, MODIS can distinguish between coarse and fine aerosols. A priori assumptions are used based on geography varying with seasons to distinguish urban/industrial pollution from bio-mass burnings (Kumar et al. 2008).

4.1.3 Meteorological Parameters

Initial research tested the correlation between AOD and ground-level PM2.5 data to determine if AOD is a robust proxy for estimating particulate matter. Statistical results from Engle-Cox et al. (2004) varied widely, with R^2 values ranging from zero to 0.9. Subsequent research has attempted to improve the stability of the correlations by expanding the model to include meteorological parameters.

The majority of aerosols are located in the lower troposphere, especially within the PBL where they are more evenly mixed (Kaufman et al. 2003). At the top of the PBL is a capping inversion, a stable layer of air which effectively traps particulates. A well-developed PBL will contain a deep column of mixed particulates corresponding to a lower aerosol density for a given AOD value than a less well-developed (shallow) PBL (see figure 6.) Additionally, boundary layer humidity can affect AOD by adding to the optical extinction due to the fact that when air humidity is high, hygroscopic particles can grow exponentially in size (Liu et al. 2005; Tsai et al. 2011). This can result in an increase in the light extinction coefficients and an overestimation of aerosol particles.

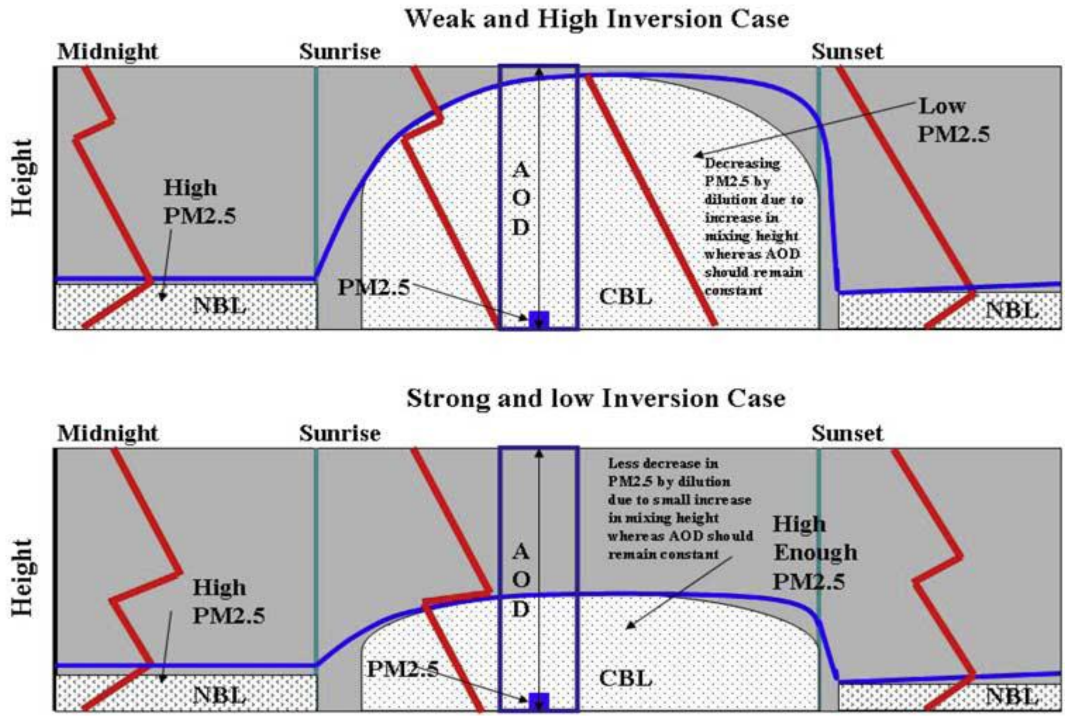


Figure 6. Development of planetary boundary layer height under weak and strong inversion conditions (Gupta & Christopher 2009)

PBL height, surface air temperature and relative humidity, from NOAA's Rapid Update Cycle (RUC) at a spatial resolution of 13 km was used in the regression analysis. RUC is a mesoscale meteorological model that produces short-range forecasts. It is operated by the National Centers for Envionmental Prediction (NCEP) and is initialized more frequently than any forecast model. Weather models are usually initialized at 0000 and 1200 UTC, however, RUC assimilates recent observations aloft and at the surface in order to provide accurate hourly forecasts (NOAA 2013). Hourly analysis can be particularly useful when coupled with recent satellite and radar images and with observations from surface stations and profilers.

Archived files in gridded binary format (GRIB2) for 2011 were downloaded from the FTP server at the Atmospheric Radiation Measurement (ARM) Climate Research Facility website. ARM is a U.S. Department of Energy scientific user facility that provides in situ and remote sensing observations from around the world to advance atmospheric research (USDE 2013). Geographic subsets of the study area were manually processed using Unidata's Integrated Data Viewer 3.1 (IDV) for the nearest Terra flyover times. PBL height, relative humidity at 2 meters above the ground, and temperature at 2 meters above the ground were the meteorological parameters selected as input variables for the regression analysis.

4.2 Statistical model

The regression equation for this study is expressed as follows:

$$PM_{2.5} = \sigma + \beta_1 * AOD + \beta_2 * PBL + \beta_3 * temp + \beta_4 * rh \qquad (2)$$

where σ represents the *y-axis* intercept and β_1 to β_4 represent the predictor coefficients for the independent variables of AOD, PBL, temp and rh. Multiple linear regression equations were formulated by applying the statistical methods outlined in previous studies that used AOD and selected meteorological parameters to estimate PM2.5 levels near the earth's surface. (Liu et al. 2007; Gupta & Christopher 2009; Tian & Chen 2009). PM2.5 is the dependent variable and represents the measured PM2.5 concentrations at the nearest hour of MODIS/Terra flyover. The strength of this model will be evaluated as follows: a T-test at the 95% confidence level to check for the statistical significance of each of the independent parameters; an R-squared coefficient of multiple determination (R^2), an F-test to indicate the significance of the entire regression; and a Durbin-Watson test to check for multiple autocorrelation.

Analysis was conducted for each individual ground station for the entire data set. Additionally, the complete dataset for all stations was divided into winter (December through February) and summer (June through August) in order to examine the PM2.5-AOD relationship as a function of season. Days with no data due to cloud contamination, mixed pixels, or systematic and random errors were not included in the regression analysis. The null and alternative hypotheses can be expressed as:

$$\begin{aligned} \mathrm{H}_0 &: \beta = 0 \\ \mathrm{H}_1 &: \beta > 0 \end{aligned} \tag{3}$$

where β is the coefficient for the primary predictor variable, AOD. If the AOD coefficient is equal to zero it means that an increase in AOD has no effect on PM2.5 levels and that AOD is not correlated with PM2.5 and has no predictive properties. AOD is a measure of light extinction attributed to aerosols, therefore, an increase in AOD should correspond to an increase in the amount of particulate matter. If so, it follows that AOD should have a positive coefficient. If β is greater than zero, the null hypothesis can be rejected in favor of the alternative hypothesis.

5. Results

Data from the PM2.5 ground-monitoring network operated by BAAQMD are used both to illustrate the spatiotemporal patterns of PM2.5 in the SFBA and to develop and validate the MODIS-AOD model. In this chapter, first the diurnal, seasonal and spatial distributions of PM2.5 are investigated using hourly data from all twelve stations for 2011. Second, multiple regression analyses using AOD, PBL, temperature, and relative humidity as predictors for PM2.5 are presented in order to determine if this method is a viable alternative for estimating PM2.5 concentrations on a regional scale.

5.1 PM2.5 Concentrations

In 2011 the mean annual PM2.5 for the complete SFBA dataset was 9 $\mu g/m^3$, 25% below EPA's National Ambient Air Quality Standard (NAAQS) of 12 $\mu g/m^3$. The annual mean was slightly higher for the hour closest to the MODIS/Terra flyover compared with all hours at 10.91 $\mu g/m^3$ and with a standard deviation of 7.64 $\mu g/m^3$. The EPA uses the highest annual and 24-hour mean measured at a single station as the designated value to determine compliance for the entire air basin. The highest annual mean of 10.1 $\mu g/m^3$ was located at the

Oakland East station and the highest 24-hour mean of 54.2 $\mu g/m^3$ was recorded at the Vallejo station: these were the designated values for 2011 in the San Francisco Bay Area (EPA 2013b). The annual mean was below the federal standard but the 24-hour mean was nearly 20 $\mu g/m^3$ higher than compliance levels. Eight of the twelve stations exceeded the 24-hour standard of 35 $\mu g/m^3$. Vallejo had 4 non-compliance days, followed by San Jose and Oakland East with 3 days, and San Francisco and Livermore had 2 days each with PM2.5 levels above the 24-hour standard. Figure 7 shows a map of PM2.5 frequency distribution with annual mean and 24-hour maximum mean for each station in 2011. The frequency distribution curve is similar for all stations and indicates that PM2.5 concentrations are fairly uniform throughout the SFBA. Approximately 40% of the observations were between 4 and 10 $\mu g/m^3$. All PM2.5 values above 30 $\mu g/m^3$ are represented by the last frequency bar and the apparent spike here does not show the actual tail of the distribution curve.

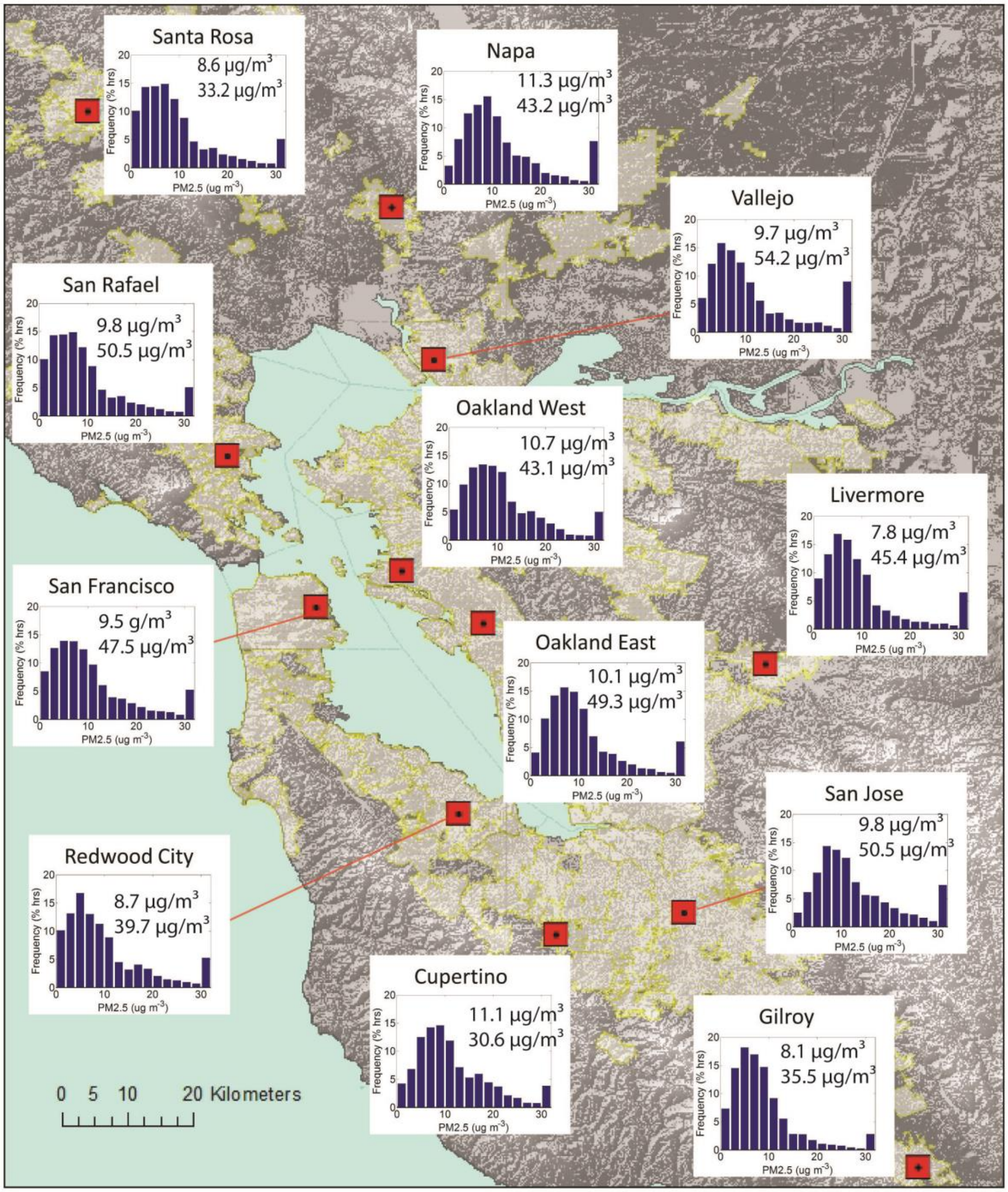


Figure 7. The annual PM2.5 frequency distribution in 2011 at the 12 stations operated by BAAQMD and with annual mean (upper) and maximum 24-hr mean (lower).

5.1.1 Diurnal and Seasonal PM2.5 Concentrations

Because PM2.5 concentrations were higher in winter, the dataset was divided into winter and summer to test how the model performs under different conditions. Figure 8 presents hourly ensemble averages for summer and winter at the 12 stations in 2011. PM2.5 levels were generally higher at night and all of the stations showed higher PM2.5 levels during the winter, except for Cupertino. Gilroy and Oakland East recorded higher PM2.5 at night with a secondary mid-morning peak around 10 am in both winter and summer. Oakland West, Redwood City, San Jose, and Santa Rosa had a similar pattern in winter, however during the summer; PM2.5 levels were highest in the morning. During the winter, Napa and Vallejo had the highest levels at night, but daytime levels were relatively flat and there was no mid-morning peak. However, during the summer at these two locations, PM2.5 was highest during the day around mid-morning. Livermore, San Francisco, and San Rafael showed a similar pattern of nighttime peaks and no secondary morning peak, with PM2.5 levels fairly stable throughout the day. Cupertino was the most anomalous with PM2.5 concentrations higher in the summer instead of winter and with peak levels occurring during the day between 10 am and 2 pm for both winter and summer.

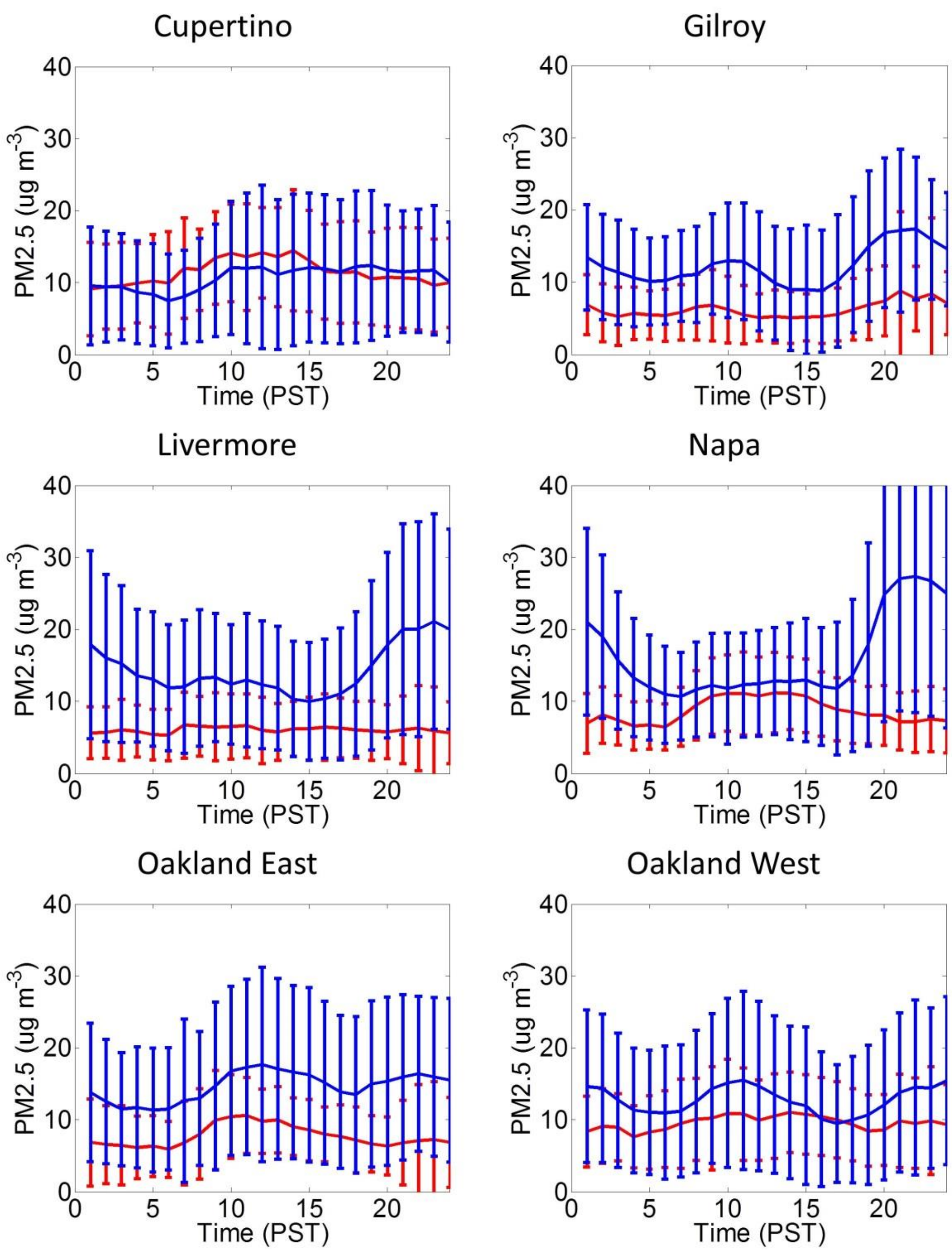


Figure 8 . Daily mean PM2.5 and standard deviation (bars) as a function of time (hour of day) and in winter (blue) and summer (red).

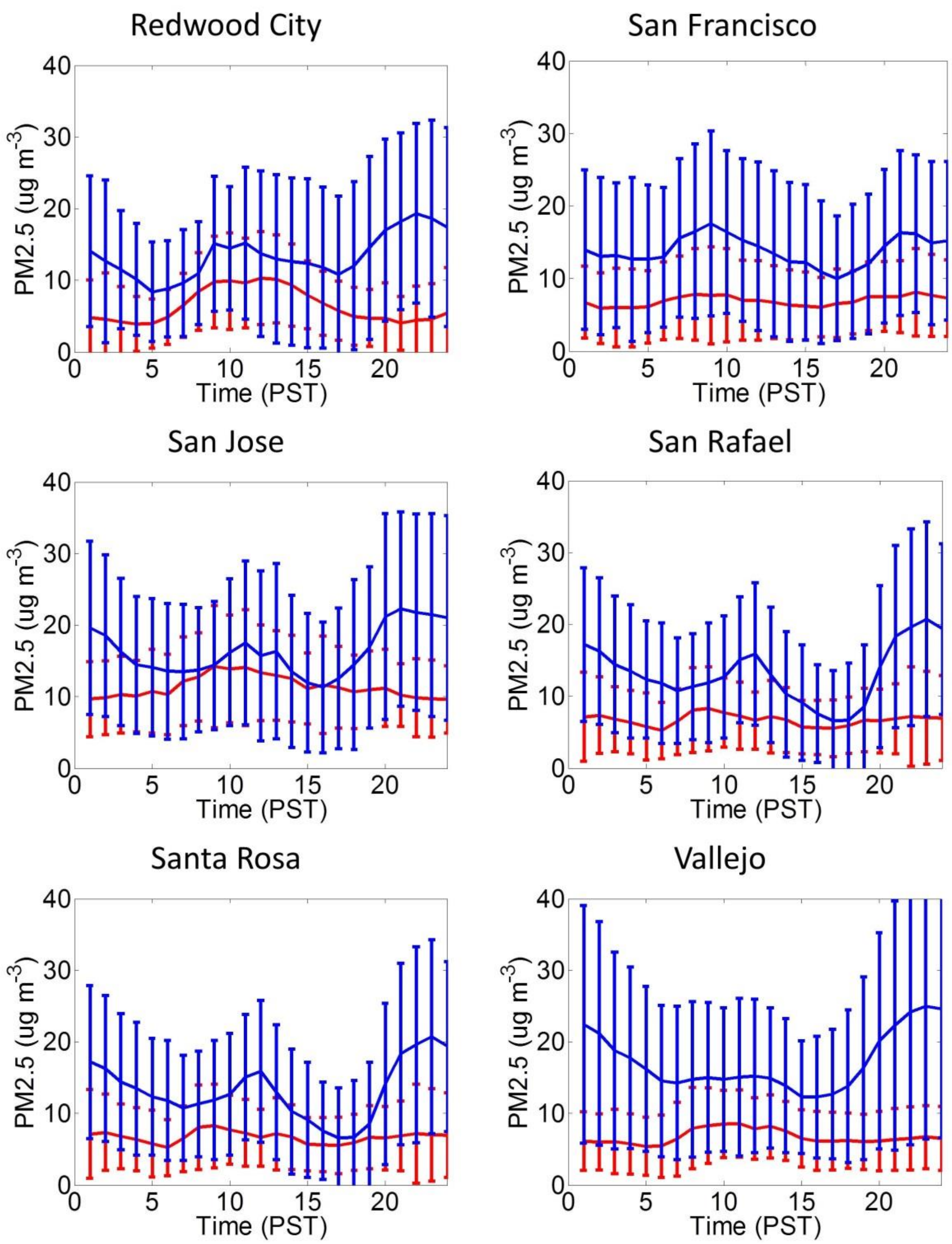


Figure 8 (Continued). Daily mean PM2.5 and standard deviation (bars) as a function of time (hour of day) and in winter (blue) and summer (red).

5.2 Multiple Regression Analysis

The multivariate model used to predict PM2.5 at the hour of MODIS/Terra overpass had an overall R^2 of 0.11 and a root mean squared error (RMSE) of 6.58 μg/m^3, which is 58.86% of the mean (11.18 μg/m^3). The performance of the independent variables in relation to PM2.5 will be discussed followed by an evaluation of the multiple regression equations and the spatial and seasonal variability of the model. Figure 9 displays a histogram of the standardized residuals, the difference between the predicted and observed values. They were approximately normally distributed; therefore, a multivariate analysis is valid. A Durbin-Watson test showed an independence of observations for the variables with a value of 1.5. A Durbin-Watson value close to 2 indicates that there is no correlation between the residuals. The F-statistic, the ratio of the mean sum of squares for the regression to the mean sum of squares for the residuals for the complete dataset was 32.4, which is greater than the critical value of 2.37 and shows that the regression model is a good fit for the data (Rogerson 2008). The slope coefficients for the independent variables for all of 2011 and for winter and summer are listed in table 2.

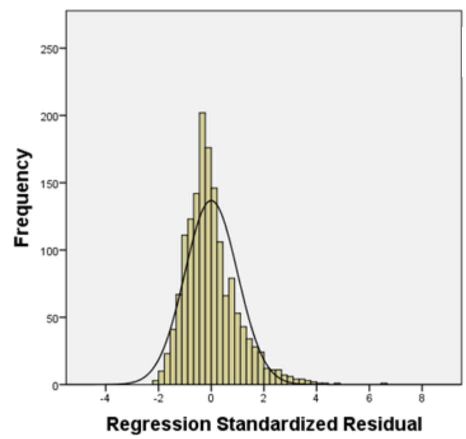


Figure 9. A histogram of multiple regression residiuals for the complete dataset.

5.2.1 MODIS AOD and PM2.5

The mean AOD value for the hour closest to the satellite flyover time was 0.11 with a standard deviation of 0.12. AOD contributed only 3% to the total model estimation of PM2.5 concentrations in the SFBA for 2011. However, it did qualify as a statistically significant independent variable at a 95% confidence level with a p-value less than 0.05. The unstandardized coefficient for AOD for the complete dataset was 6.286. Thus, the null hypothesis can be rejected in favor of the alternative hypothesis stating that the Beta coefficient for AOD is greater than zero and has a positive correlation with measured PM2.5. Although AOD was barely significant for the entire dataset, the positive sign of correlation is what was expected since an increase in AOD is equated with an increase in PM2.5 levels.

A preliminary simple regression using AOD as the independent variable found AOD to be a significant variable but the R^2 was very low at 0.01. This means that AOD predicted 1% the variability of PM2.5. The addition of PBL, temperature, and relative humidity increased the R^2 to 0.11. Mean values and model results for all datasets is presented in table 3.

Variable	2011			Winter			Summer		
	Unstandardized Coefficient	Standardized Coefficient	Sig.	Unstandardized Coefficient	Standardized Coefficient	Sig.	Unstandardized Coefficient	Standardized Coefficient	Sig.
AOD	6.185	0.108	0.001	26.217	0.204	0.001	8.768	0.196	0.001
PBL	-0.007	-0.289	0.001	-0.007	-0.227	0.001	-0.005	-0.2	0.001
Temp.	-0.136	-0.109	0.001	-0.343	-0.103	0.103	-0.181	-0.154	0.042
RH	-0.036	-0.094	0.001	0.117	0.224	0.001	-0.179	-0.505	0.001

Table 2. Unstandardized coefficients, standardized coefficients, and statistical significance at a 95% confidence interval for the independent predictor variables.

Dataset	PM2.5 µg/m³)	AOD (unitless)	PBL (meters)	Temp (°C)	RH (%)	R^2	RMSE (µg/m³)	N
All 2011data	11.18	0.1068	504.74	17.37	69	0.11	6.58	1509
Winter	15.01	0.0562	402.53	12.05	67	0.22	8.49	242
Summer	9.79	0.1208	437.92	20.98	72	0.08	5.31	551
PBL > 500 m	8.99	0.11	769.71	21.03	55	0.11	5.87	578
PBL < 500 m	11.86	0.1056	301.18	19.76	68	0.07	6.74	951

Table 3. Mean PM2.5, AOD, PBL, temperature, relative humidity and R-squared, RMSE, and number of samples for all data sets.

5.2.2 RUC parameters and PM2.5

PBL, temperature, and relative humidity were included in the multivariate statistical analysis to test if they would improve the model. The mean PBL height for the combined data set was 504.74 m with a standard deviation of 302.07 m. The minimum and maximum estimated PBL heights were 21 m on June 16 at Oakland West and 2,429 m on April 7 at San Jose. The mean temperature at 2 m height for 2011 was 17.37° C. The range of temperature was 6.4° C to 34.7° C and the mean relative humidity at 2 m height was 70.77% and ranged from between 14.83% and 100%. PBL was the most important predictor variable, accounting for 78% of the predictive power of the model (see figure 10). PBL was a statistically significant predictor in 8 of the 12 ground stations with coefficients between -.003 and -.015. The correlation with PM2.5 for the entire dataset was negative and had a coefficient of -0.008.

Temperature was the second most influential parameter affecting PM2.5 estimation, responsible for 13% of the model predictions. Temperature had a negative coefficient of -0.169, meaning that for every 1 degree increase in temperature the predicted PM2.5 levels decrease by 0.169 $\mu g/m^3$. Relative humidity and AOD were the least important predictor in this study, accounting for 6% and 3% of the model's predictions, respectively. RH had a regression coefficient of -.032, equating an increase in humidity with a small decrease in

PM2.5. High humidity can cause aerosols to grow exponentially in size, which increases the amount of light extinction that can then lead to an overestimation of PM2.5 based on AOD. However, relative humidity does not increase PM2.5.

Figure 11 shows the predicted vs. observed PM2.5 levels for the multivariate analysis and figure 12 shows the relationship between each independent variable and observed PM2.5 levels. The meteorological parameters of PBL height, temperature, and relative humidity were all more significant than AOD in the multivariate analysis to predict PM2.5 in this study. PBL contributed 78% of the predictive power of the model which shows the importance and necessity of including it as a variable in regression models to predict PM2.5 mass concentrations. A simple regression of PM2.5 and PBL gave an R^2 of 0.092, and the additional independent variables, including AOD, only increased the R^2 to 0.11. The graphs show that the high levels of PM2.5 are consistently underestimated.

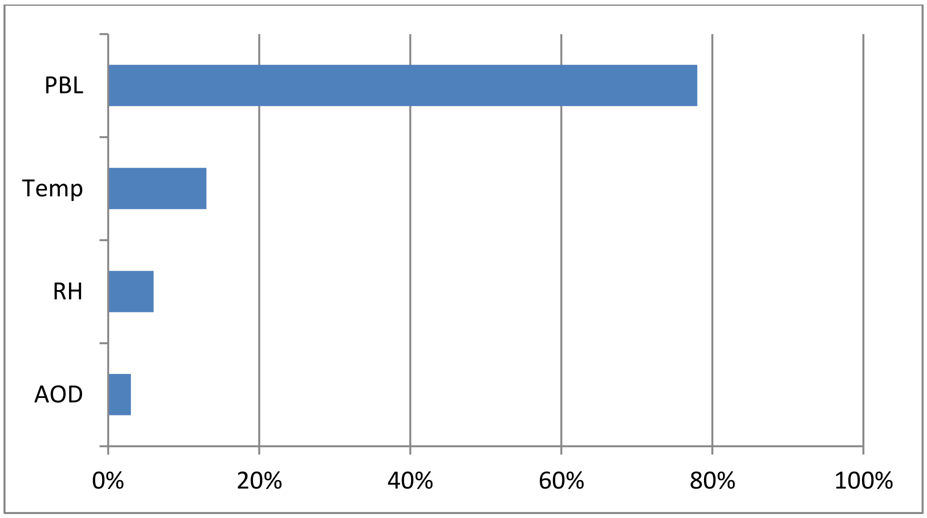


Figure 10. The contributions of the independent variables to the overall model results.

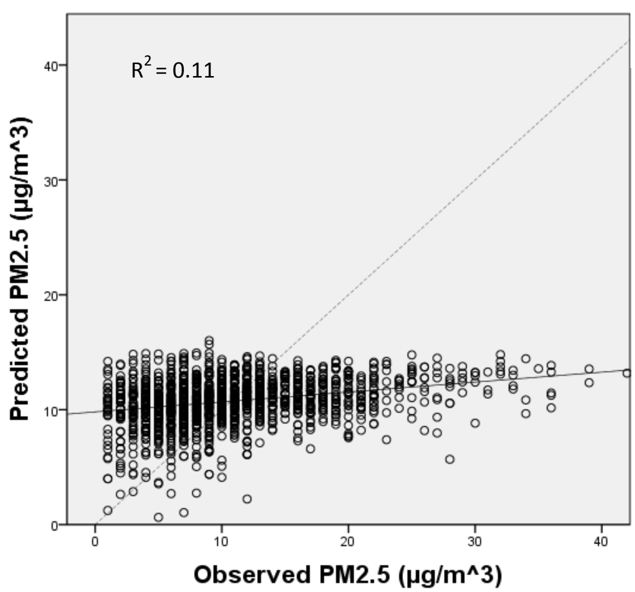


Figure 11. Predicted vs. Observed PM2.5 from the multivariate analysis.

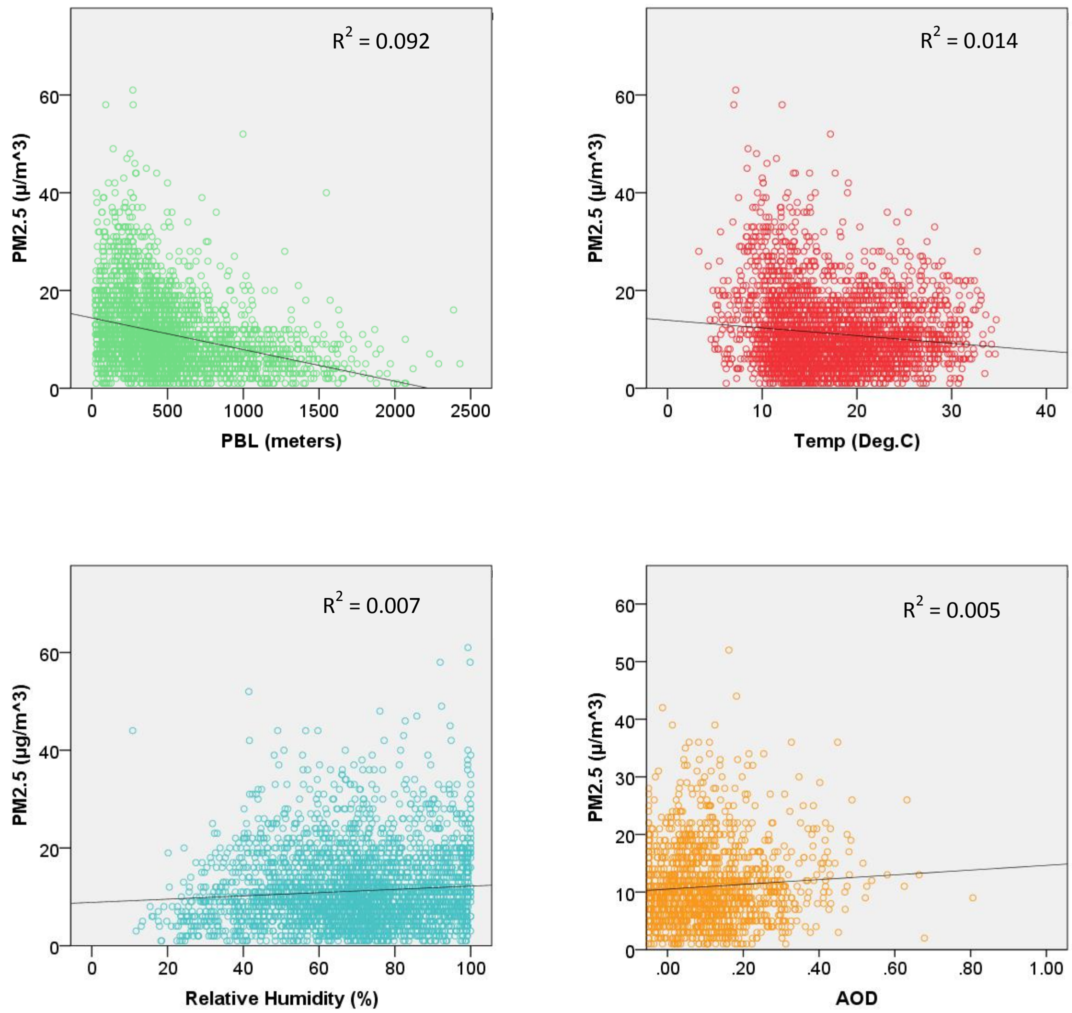
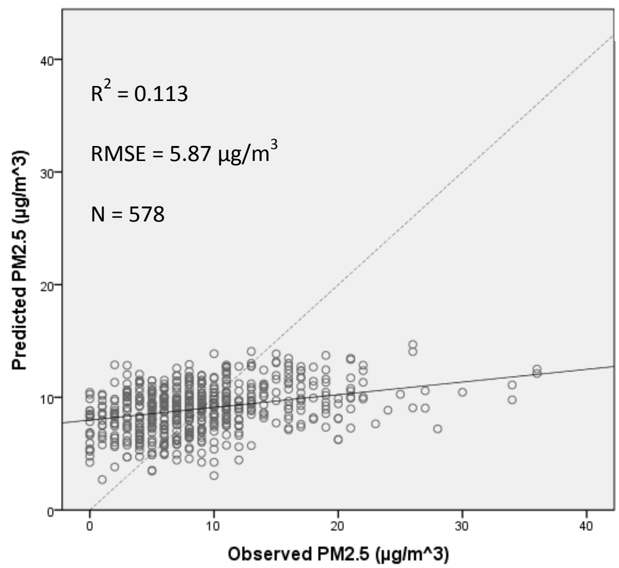


Figure 12. Scatterplots from simple regression of PM2.5 and independent variables

In order to test the influence of PBL height, the data was divided into two bins, one with PBL heights of 500 meters or less, and the other with PBL heights greater than 500 meters. The multivariate model was run for each PBL bin and there were 951 valid samples with PBL heights of 500 meters or less, and 578 samples with PBL heights greater than 500 meters. The model performed better under PBL conditions greater than 500 meters with an R^2 of 0.11, compared with 0.07 in summer. The higher correlation suggests that satellite based measurements of aerosols can estimate surface PM2.5 more accurately when PBL heights are more developed (>500 m). This is contrary to Gupta et al. (2006) whose model performed better when PBL heights were less than 500 m. Figure 13 depicts the observed and predicted PM2.5 concentrations for the two PBL bins. This figure also shows that under higher PBL, observed PM2.5 is lower, as shown by the negative relation in Figure 12a, and that the model also has generally lower values of PM2.5 modeled under higher PBL.

a.)



b.)

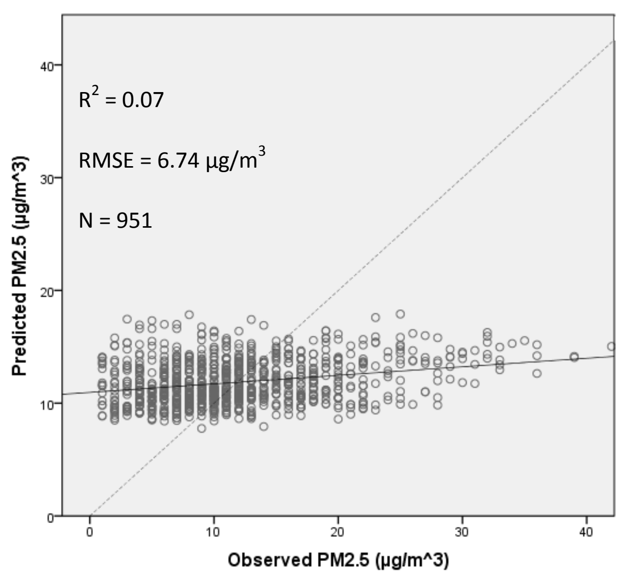


Figure 13. a.) Predicted v. Observed PM2.5 for PBL heights above 500 meters. **b.)** below 500 meters.

5.3 Spatial Variability

The model performed best in Oakland West and Oakland East with R^2 values of .281 and .247, respectively. However, they had two of the three smallest sample sizes. Oakland West had only 31 valid sample days for all of 2011 and Oakland East had a total of 90 sample days, or 25% of 2011. The remaining coastal locations: San Francisco, Redwood City and San Rafael, also had less than 100 sample days and R^2 values below 0.16. In contrast, stations with a more inland location had larger valid sample sizes and better overall model performance. Livermore, Santa Rosa, Vallejo and Gilroy had the greatest number of valid sample days.

Inland locations with larger temperature ranges had better model results on average than coastal and bayside locations. AOD measurements from MODIS are hampered by fog and cloud cover which are more prevalent in coastal locations of the SFBA due to a stronger maritime influence. The map in figure 14 shows the R^2 and root mean squared error (RMSE) for each station for the hour of MODIS/Terra flyover. Gilroy had the lowest RMSE of 4.6 $\mu g/m^3$ (57% of the mean) and San Jose had the highest RMSE at 8.1 $\mu g/m^3$ (83% of the mean).

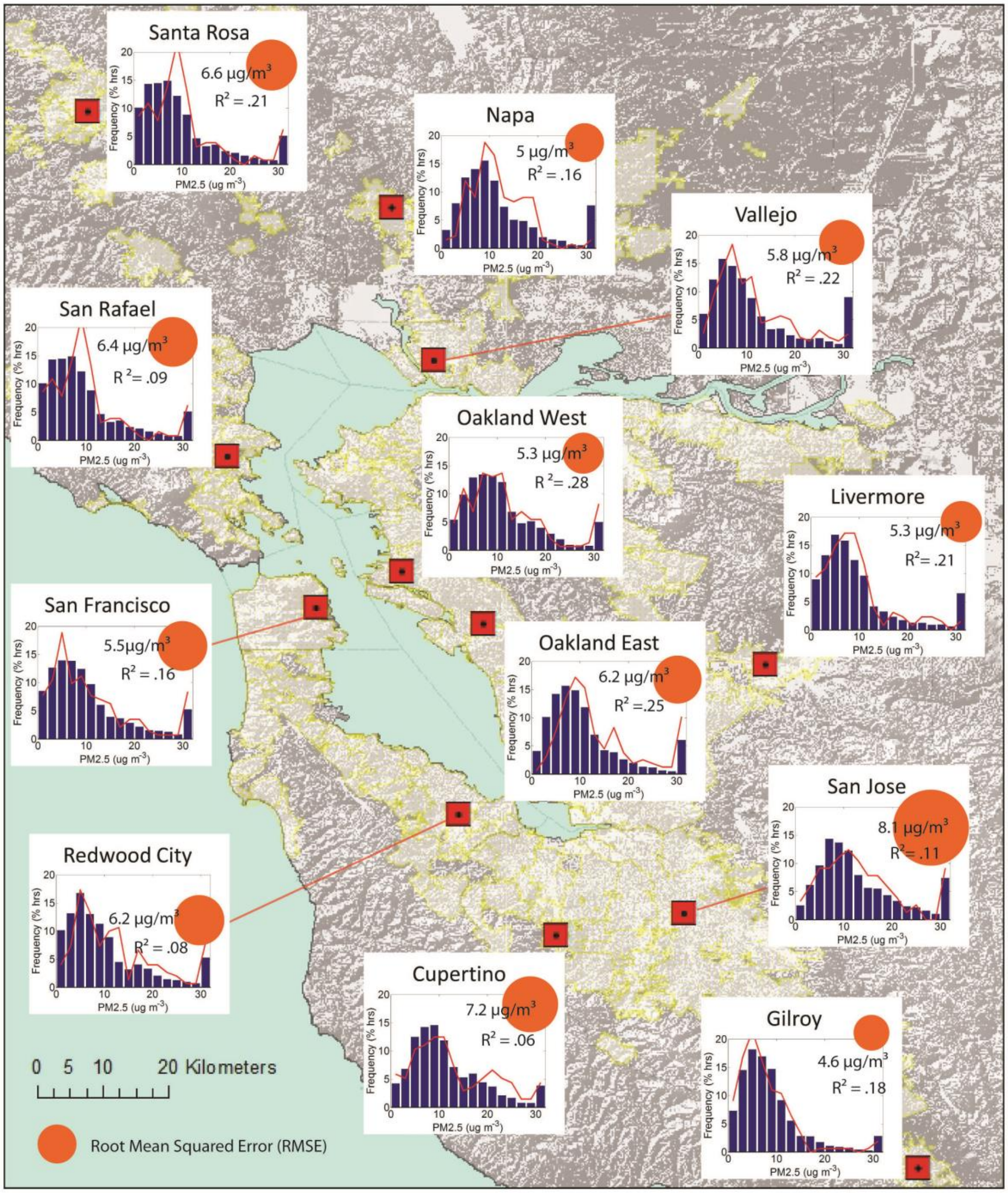


Figure 14. PM2.5 frequency distribution at BAAQMD stations for all hours and for the hour closest to MODIS/Terra flyover time (red line). Root mean squared error (RMSE) with proportional circles and R-squared values are included.

5.4 Seasonal Variability

PM2.5 concentrations were higher in winter than summer; therefore, analyses were conducted to see how the model performed under different conditions and PM2.5 concentrations. Winter months had an R^2 of 0.22 and 242 samples compared with an R^2 of 0.08 for summer with 551 samples. Mean PM2.5 levels were higher in winter, 15 $\mu g/m^3$ compared to 9.8 $\mu g/m^3$ in summer (see table 4). AOD was higher in the summer with a mean of .121 while the winter mean was .056. Higher AOD values in summer did not equate with elevated ground-level PM2.5 as concentrations were higher by 5 $\mu g/m^3$ in winter than in summer. Average PBL mixing heights were 402.53 m in the winter and 437.92 m in summer. The mean temperatures were 12.05° C in winter and 20.97 ° C in summer with a mean relative humidity of 68% in winter and 72% in summer. Table 5 shows observed mean PM2.5 levels for each station in winter and summer.

	Winter (D,J,F)	Summer (J,J,A)
PM2.5	15.01 $\mu g/m^3$	9.79 µg/m3
AOD (unitless)	0.056	0.121
PBL	402.53 m	437.92 m
Temp	12.05 ˚C	20.98 ˚C
RH	68%	72%

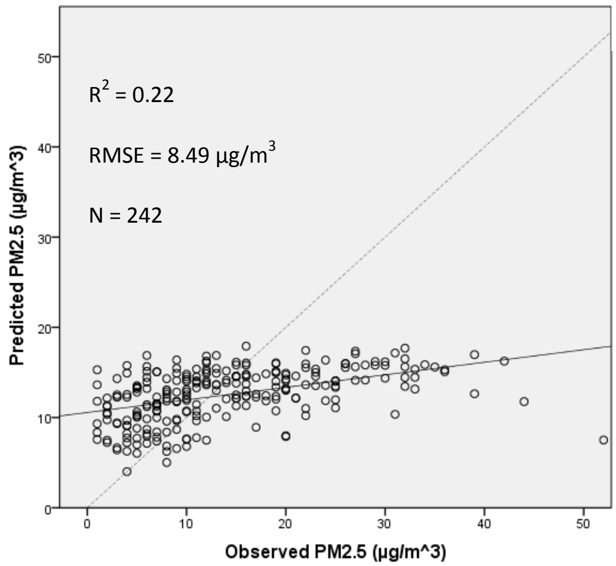
Table 4. Winter and summer mean values of PM2.5, AOD, PBL, temperature, and relative humidity for the hour of MODIS/Terra flyover.

Station	Winter (µg/m^3)	Summer (µg/m^3)
Cupertino	10.7	11.5
Gilroy	12.3	6.2
Livermore	16.2	8.7
Napa	14.4	6.1
Oakland E	14.7	7.7
Oakland W	12.8	9.6
Redwood City	13.5	6.4
San Francisco	13.9	7
San Jose	16.4	11.4
San Rafael	14.6	6.6
Santa Rosa	13.2	6.7
Vallejo	17.2	6.7

Table 5. Mean PM2.5 values for winter and summer at BAAQMD stations.

Figure 15 features scatterplots of observed and estimated PM2.5 mass concentrations for summer and winter. The sample size during summer was 551, almost twice as large as the number of valid sample days for winter at 242, most likely due to cloud cover. The root mean squared error (RMSE) was 8.49 µg/m^3 for winter (57% of the mean) and 5.31 µg/m^3 for summer (54% of the mean). The model performed better in the winter with an R^2 value of 0.22, an 14% improvement over the summer (R^2 = 0.08). The implications and sources of error for these results will be discussed in the following chapter

(a.)



(b.)

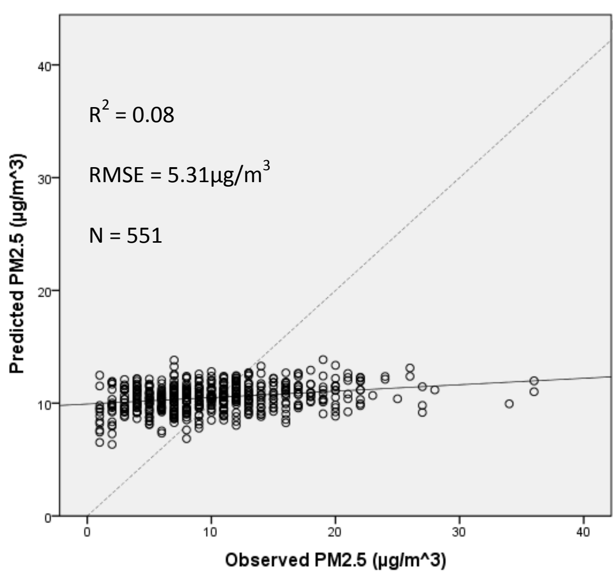


Figure15. Predicted vs. Observed PM2.5 for all stations, **(a.)** Winter (December through February), **(b.)** Summer (June through August), 2011.

6. Discussion and Conclusions

In the previous chapter, spatiotemporal patterns of PM2.5 concentrations and model performance were presented. In this chapter, the causes of these patterns are discussed and the performance of the statistical model is assessed. Section 6.1 analyzes the statistical performance of the model in a geographic and seasonal context. A summary and final conclusion, including areas of future research, is presented in section 6.2.

6.1 Model Evaluation

The statistical model for the entire dataset in this study yielded an R-squared coefficient of 0.11. This is comparable to a similar study of AOD and PM2.5 for the adjacent San Joaquin Valley that also had an R^2 of 0.11 (Justice et al. 2009). Individual stations had an R^2 between 0.06 in Cupertino and 0.28 at Oakland West (see figure 14). The model underestimates PM2.5 concentrations that are greater than 20 $\mu g/m^3$, which is a major weakness given that detecting unhealthy levels would be its main application. AOD, PBL, temperature and relative humidity were not robust predictors of PM2.5 in the study area, yet the parameters were statistically significant at a 95% confidence level. The sample size was reduced due to fog and cloud cover prevailing in coastal locations. San

Francisco, Oakland West and Oakland East had less than 90 sample days at the hour of the satellite overpass in 2011.

Christopher & Gupta (2009) compared monthly and annual PM2.5 measurements for the complete data set and for cloud-free conditions for the nearest hour of the satellite overpass for each EPA region and found the total mean differences to be less than 2.5 $\mu g/m^3$. They concluded that bias due to cloud cover is not a major problem when estimating PM2.5 from space-borne monitors on a monthly and annual basis. The results for this study were very similar with a mean difference of 1.6 $\mu g/m^3$ between the annual mean for all hours and for the satellite overpass time during cloud free conditions. While this may be useful, daily measurements from MODIS are reduced in some locations and are thus not reliable for monitoring PM2.5 on a daily basis.

6.1.1 Seasonal Analysis

Summer (Jun, Jul, Aug) and winter (Dec, Jan, Feb) months were extracted from the annual dataset for the purpose of seasonal analysis. The study area has a Mediterranean climate characterized by wet, cool conditions during winter and dry and warmer conditions during summer, although with cool coastal areas. The San Francisco Bay Area also exhibits more extreme temperatures with distance from the coast. Microclimates, formed by varied topography and the influence of the marine layer, add large climatic diversity to the area. A seasonal comparison can help determine if local climate and emissions affect PM2.5 concentrations and model performance, or if regional climate and emissions patterns are the dominant influence.

Mean PM2.5 was 6.35 μg/m^3 higher in winter (see table 2) and five of the stations exceeded the 24-hour standard of 35 μg/m^3 on December 25 and December 10, 2011. All but two of the remaining non-compliance days occurred within a day or two. This suggests that regional meteorological conditions contributed to the high levels of PM2.5 at this time. Conditions were clear with temperatures 3˚ to 7˚ C below normal. PBL heights were also low on these days, approximately 300 m. Figure 16 displays PBL heights at noon on December 25th. A positive correlation exists between temperature and PBL heights and capping inversions are more likely to form under cold, calm conditions with clear skies. There was probably an emissions component to the high concentrations from wood smoke as well. Residential wood burning is the single largest source of PM2.5 during winter when the SFBA experiences its highest PM2.5 levels (BAAQMD 2013a).

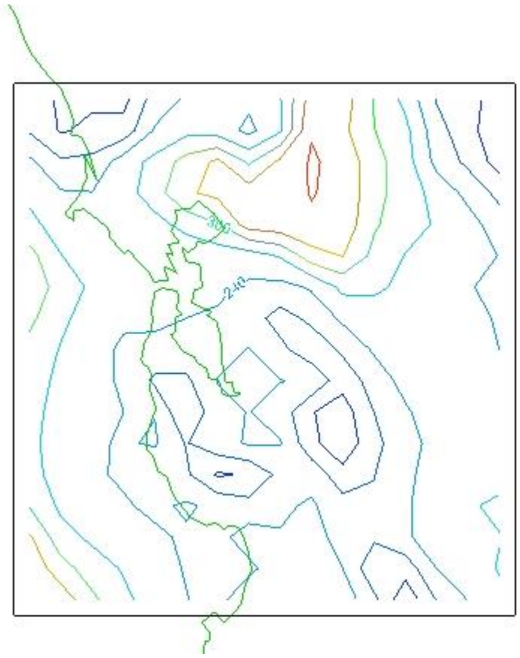


Figure 16. PBL heights at noon on December 25, 2011.
Contour interval: 30 meters (NOAA 2013).

Seasonal analysis resulted in an R^2 of 0.08 in summer with 551 samples and an R^2 of 0.22 in winter with 242 samples. A larger sample size can often improve the Pearson's correlation (R) and the coefficient of determination (R^2), yet the model performed 14% better in winter with fewer samples. In contrast, studies conducted in the eastern U.S. performed better during the summer than in winter (Zhang et al. 2006; Gupta & Christopher 2008; Green et al. 2009). Gupta and Christopher (2009), for example, attribute the poor performance of their model during the winter to low AOD values (<0.1) which are associated with high uncertainty (~65%).

However, the results for the SFBA were the opposite, with mean AOD values greater in the summer (0.1208) than in winter (0.0562), yet the model performed better in winter. A possible explanation for this is that the MODIS algorithm for AOD over land favors darker, vegetated surfaces. In the eastern U.S., surface reflectance is higher in the winter because forests have lost their leaves and because of snow and ice. The opposite; however, is true for the SFBA when vegetation, which was dry and brown during the summer, turns green from winter rains resulting in a darker, less reflective surface than in summer. This, along with higher PM2.5 concentrations, could explain why the model performed better in winter over summer.

Additionally, a more developed PBL during the summer did not improve the model. Gupta & Christopher (2009) state that lower PBL heights during winter in the southeastern U.S. do not allow as much vertical mixing, which contributes to the underestimation of fine particulate matter based on AOD. The findings for this study do not agree on this point since mean PBL heights were 35 m higher during summer but model results were better in winter when PM2.5 levels were greater.

In comparison, the southeast U.S., northern Italy, and the Netherlands had the best results during summer when there was a more developed PBL and greater PM2.5 levels (Gupta & Christopher 2009; DiNicolantonio et al. 2009; Schaap et al. 2009). Boundary layer heights were also 35 m higher on average during the summer for the SFBA, yet the models performance improved by 14% in the winter compared to summer. This is not in agreement with the previous studies where the AOD-PM2.5 relationship improved when boundary layer heights were more stable and well mixed (Gupta & Christopher 2009; DiNicolantonio et al. 2009; Schaap et al. 2009; Tsai et al. 2011).

6.1.2 Geographical Differences

There are geographic limitations linked to differences in surface reflectance, as well as meteorological and seasonal conditions that limit the ability of the MODIS algorithms to accurately measure AOD. The model results varied across the study area and generally performed better for inland locations. However, the overall regional differences within the Bay Area are small with annual mean PM2.5 concentrations between 7.8 and 10.1 $\mu g/m^3$ for all 12 stations. The low correlation of AOD and PM2.5 in the SFBA confirms the findings from previous studies for the western United States (Zhang et al. 2006; Donkelaar et al. 2006; Justice et al. 2009). This indicates that the conditions which hamper the model performance are not from local climate and emissions, but are caused by more regional conditions. PM2.5 was fairly uniform throughout the SFBA, with greater concentrations at night and during the day between 10 am and noon. The higher concentrations at night are driven by meteorological conditions but the mid-morning peak is primarily driven by emissions. Lower PBL at night traps and concentrates pollutants closer to the ground. The mid-morning peak is likely driven by emissions from vehicles, factories, ports, and manufacturing. The reason there is not an afternoon peak is because PBL heights are not yet fully developed in the morning.

Mean PM2.5 was significantly higher in winter (15.01 $\mu g/m^3$) over summer (9.79 $\mu g/m^3$). Colder, more stagnant conditions during the winter, coupled with an increase in residential wood combustion, lead to a buildup of PM2.5 during winter months (CARB 2013, BAAQMD 2013a). The majority of aerosols reside within the mixing layer so that when PBL heights are lower, aerosols are more concentrated near the surface. When the boundary layer height decreases, particulate matter levels near the surface increase. The observed negative correlation between PBL and PM2.5 is in agreement with previous studies (Gupta & Christopher 2009; Wang et al. 2010; Tsai et al. 2011).

The MODIS aerosol retrieval algorithm over land was designed for surfaces which are vegetated or have dark soils and with low surface reflectance values (Remer et al. 2005). The SFBA is situated in the western half of North

America, which is drier and with less vegetation than the eastern half of the continent. Brighter, arid regions will have more coarse-dominated dust aerosols which would also reduce the quality of the retrieved AOD (Levy et al. 2010). The SFBA also contains areas of dense urban development that have a greater surface reflectance than surrounding rural areas. These are some of the possible reasons the correlations between MODIS AOD and PM2.5 are weak in this region.

Additionally, 6 of the 12 PM2.5 monitoring stations are located in areas with mixed pixels (both land and water). Although the stations with mixed pixels had a higher R^2 (0.18), than those without mixed pixels (0.15), the mixed pixel locations had 53% less samples (see table 6). The MODIS land algorithm will still retrieve AOD for pixels identified as ocean, but that can decrease the quality of the retrieval (Remer et al. 2005). The ocean is darker than land and if there are a large number of ocean pixels there may not be enough valid “dark target” pixels remaining to calculate AOD for the reasons outlined below.

Each 10 km^2 box contains 400 0.5 km^2 pixels (20 x 20), and after the preliminary masking to remove pixels containing water, clouds, snow and ice, the surface reflectance of the remaining pixels are analyzed at 660 nm and the brightest 50% and the darkest 20% are discarded to eliminate contamination from cloud shadows on the dark end and residual cloud contamination on the bright end. There must be at least 12 “dark target” pixels remaining before AOD will be retrieved. Depending on the number of ocean pixels, there may not be enough valid pixels to calculate AOD, therefore mixed pixels can indirectly affect the number of valid samples for coastal locations (Remer et al. 2005; Levy et al. 2010).

Mixed Pixels			Non-mixed Pixels		
	R^2	N		R^2	N
Oakland East	0.247	90	**Cupertino**	0.059	143
Oakland West	0.281	31	**Gilroy**	0.179	185
Redwood City	0.082	97	**Livermore**	0.212	210
San Francisco	0.157	37	**Napa**	0.159	164
San Rafael	0.087	110	**San Jose**	0.106	131
Vallejo	0.223	168	**Santa Rosa**	0.211	175

Table 6. A comparison of PM2.5 stations with and without mixed pixels.

6.2 Summary and Conclusions

A multiple regression equation was developed to estimate PM2.5 from MODIS derived aerosol optical depth (AOD) and meteorological parameters (PBL, temperature, and relative humidity) for the hour closest to the satellite flyover time. A year of PM2.5 measurements from 2011 were obtained from the Bay Area Air Quality Management District (BAAQMD) as well as AOD from MODIS/Terra. BAAQMD operates 12 ground-stations throughout the Bay Area that continuously monitor PM2.5 on an hourly basis. Measured PM2.5 data are point measurements and the AOD values were extracted from grid cells that were coincident with each ground station and assigned to that station in ArcGIS. PBL, temperature and relative humidity values were assigned using the same method.

Based on the results of this study, the claim from previous studies (Zhang et al. 2006; Gupta & Christopher 2009) that multiple regression equations using satellite derived AOD products can assist air quality monitoring and decision-making are not true for the San Francisco Bay Area. The R-squared value for the complete data set was 0.11, meaning that 11% of the variance could be explained by the regression equation. These results are comparable to a study conducted in the San Joaquin Valley that also had an R^2 of 0.11 (Justice et al. 2009). The primary objective of this study was to test if the low correlation from previous studies in the western U.S. would also apply to the San Francisco Bay Area. The low R^2 (0.11) from this thesis confirms previous findings that show a poor correlation between AOD and PM2.5 in the western U.S. The R-squared value for individual stations ranged from 0.06 to 0.28, however, locations near the bay had valid AOD values for less than 30% of the year due to cloud cover and/or mixed pixels that contain both land and water.

The secondary objective of this study was to determine if including meteorological parameters, such as PBL height, would improve the AOD-PM2.5

relationship within the study area. Including these additional independent parameters improved the model R^2 to .11, a 10% increase over a simple two-variable approach using only PM2.5 and AOD (R^2 = .01). PBL height accounted for 78% of the models ability to estimate PM2.5 and therefore, PBL height should be included when utilizing AOD to estimate surface PM2.5 concentrations. Temperature was the next most important independent variable at 13%, followed by relative humidity at 6%. AOD was the least significant predictor at 3%. RUC meteorological data alone is more useful for estimating PM2.5 as AOD contributed virtually nothing to the statistical model.

The results of this research also verify previous studies showing that this method performs better over relatively dark, vegetated surfaces, like the eastern U.S. in summer, and performs poorly over regions with brighter, more reflective surfaces such as deserts, ice, and urban areas. AOD alone cannot be used as a proxy for surface PM2.5 levels. AOD and PM2.5 concentrations can vary within 10 km^2, therefore, the 10 km resolution of MODIS AOD may be too coarse to accurately measure urban scale aerosol features. To address this problem, NASA has recently introduced dark target algorithms for ocean and land that can estimate AOD at a 3 km resolution and may be better suited for studying aerosol distributions in urban regions (NASA 2013a). The finer resolution will allow for

AOD retrievals closer to coastlines and within patchy cloud fields (Munchak et al. 2013).

Utilizing MODIS AOD to estimate PM2.5 is not yet feasible in the San Francisco Bay Area; however, remote sensing holds the potential to help air quality decision makers monitor and regulate air pollution and to provide air quality data over a greater geographical area than could be covered by ground-stations alone, which are expensive to maintain and operate. It may also provide coverage for regions where ground monitoring stations are sparse or absent. A limitation of using MODIS to estimate air pollution is that it can only do so when there is no cloud cover, unlike ground monitors that can directly measure PM2.5 concentrations on an hourly basis regardless of cloud conditions (Christopher & Gupta 2009). Also, incorrect sensor calibration, false radiance estimates, and contamination by glint can cause errors (Christopher & Gupta 2009; Remer et al. 2005; NASA 2013b).

Further research is needed to develop a linear regression method incorporating AOD and meteorological parameters to provide air quality and health officials with a more accurate picture of surface PM2.5 concentrations. It would be simpler to operationalize compared to complex climatic and dispersion models that require many inputs that are difficult to quantify, such as point sources of pollution. Future studies using the MODIS Deep Blue algorithm, designed for AOD retrieval over brighter surfaces, and land use/land cover data may improve model performance in the western U.S. and in urban areas. Additionally, the development of algorithms specifically designed for air pollution studies and with a finer spatial resolution may assist air quality and health officials to better estimate PM2.5 over a larger area than the current use of ground-monitors alone can provide.

Literature Cited

BAAQMD - Bay Area Air Quality Management District. 2013a. *PM conditions in the Bay Area*. http://www.baaqmd.gov/Divisions/Planning-and-Research/Particulate-Matter.aspx#pmcba. (last accessed 8 August 2013).

______. 2013b. District Air Monitor Sites. http://hank.baaqmd.gov/tec/maps/dam_sites.htm (last accessed July 14, 2013).

CARB - California Air Resources Board. 2013. *PM2.5 Monitoring Network Design for California.* http://www.baaqmd.gov/Divisions/Planning-and-Research/Particulate-Matter.aspx#pmcba (last accessed 25 June 2013).

Christopher, S. A., and P. Gupta. 2009. Satellite Remote Sensing of Particulate Matter Air Quality: The Cloud Cover Problem. Submitted *Journal of Air and Waste Management.*

Chu, D. A., Y. J. Kaufman, G. Zibordi, J. D. Chern, J. Mao, C. Li, and B. N. Holben. 2003. Global monitoring of air pollution over land from the Earth Observing System-Terra Moderate Resolution Imaging Spectroradiometer (MODIS). *Journal of Geophysical Research* 108: D214661.

Chu, D.A., Y. J. Kaufman, L. A. Remer, and B . N. Holben. 1998. Remote sensing of smoke from MODIS airborne simulator during the SCAR-B experiment. *Journal of Geophysical Research* 103 (D24): 31 979-987.

Di Nicolantonio, W., A. Cacciari, and C. Tomasi. 2009. Particulate matter at surface: Northern Italy monitoring based on satellite remote sensing, meteorological fields, and *in-situ* samplings. *Journal of Selected Topics in Applied Earth Observations and Remote Sensing, IEEE* 2 (4): 284-292.

Dockery, D. W. 2001. Epidemiologic evidence of cardiovascular effects of particulate air pollution. *Environmental health perspectives* 109 (4): 483.

Engel-Cox, J. A., C. H. Holloman, B. W. Coutant, and R. M. Hoff. 2004. Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality. *Atmospheric Environment* 38: 2495-2509.

EPA - U.S. Environmental Protection Agency. 2013a. *What are the six common pollutants?* http://www.epa.gov/airquality/urbanair/ (last accessed 7 August 2013).

______. 2013b. *Area Designations for the 2012 Annual Fine Particle Standard.* http://www.epa.gov/airquality/particlepollution/designations/2012standards/index.htm last accessed 7 August 2013).

______. 2013c. *Particulate Matter.* http://www.epa.gov/airquality/particlepollution/ (last accessed 1 August 2013).

______. 2013d. *Historical PM NAAQS.* http://www.epa.gov/ttn/naaqs/standards/pm/s_pm_history.html (last accessed 7 August 2013).

Glantz, P., A. Kokhanovsky, W. von Hoyningen-Huene, and C. Johansson. 2009. Estimating PM2.5 over southern Sweden using space-borne optical measurements. *Atmospheric Environment* 43: 5838–5846.

Green, M., S. Kondragunta, P. Ciren, and C. Xu. 2009. Comparison of GOES and MODIS Aerosol Optical Depth (AOD) to Aerosol Robotic Network (AERONET) AOD and IMPROVE PM2.5 Mass at Bondville, Illinois. *Journal of Air & Waste Management Association* 59: 1082-1091.

Gupta, P., S. A. Christopher, J. Wang, R. Gehrig, Y. Lee, and N. Kumar. 2006. Satellite remote sensing of particulate matter and air quality assessment over global cities. *Atmospheric Environment* 40: 5880–5892.

Gupta, P., and S. A. Christopher. 2008. Seven year particulate matter air quality assessment from surface and satellite measurements. *Atmospheric Chemistry and Physics* 8: 3311–3324.

______. 2009. Particulate matter air quality assessment using integrated surface, satellite, and meteorological products: Multiple regression approach. *Journal of Geophysical Research* 114: D14205.

Harrison, R. M., and J. Yin. 2000. *Particulate matter in the atmosphere: which particle properties are important for its effects on health? Science of the Total Environment* 249 (1): 85-101.

Hutchison, K. D., S. Smith, and S. Faruqui. 2005. Correlating MODIS aerosol optical thickness data with ground-based PM2.5 observations across Texas for use in a real-time air quality prediction system. *Atmospheric Environment* 39: 7190–7203.

Justice, E., L. Huston, D. Krauth, J. Mack, S. Oza, A.W. Strawa, J.W. Skiles, M. Legg, and C. Schmidt. 2009. Investigating correlations between satellite-derived aerosol optical depth and ground pm2.5 measurements in California's San Joaquin Valley. *American Society of Photogrammetry and Remote Sensing Annual Conference*. Baltimore, MD March 9 – 13, 2009.

Kappos, A. D., P. Bruckmann, T. Eikmann, N. Englert, U. Heinrich, Höppe, and E. Koch. 2004. Health effects of particles in ambient air. *International Journal of Hygiene and Environmental Health* 207 (4): 399-407.

Kaufman, Y. J., D. Tanré, J-F. Léon, and J. Pelon. 2003. Retrievals of profiles of fine and coarse aerosols using lidar and radiometric space measurements." *Geoscience and Remote Sensing, IEEE Transactions on* 41 (8): 1743-1754

Kumar, N., A. Chu, and A. Foster. 2008. Remote sensing of ambient particles in Delhi and its environs: estimation and validation. *International Journal of Remote Sensing* 29 (12): 3383–3405.

Laden, F., L. M. Neas, D. W. Dockery, and J. Schwartz. 2000. Association of fine particulate matter from different sources with daily mortality in six US cities. *Environmental health perspectives* 108 (10): 941.

Lee, H.J., Y. Liu, B. A. Coull, J. Schwartz, and P. Koutrakis. 2011. A novel calibration approach of MODIS AOD data to predict PM2.5 concentrations. *Atmospheric Chemistry and Physics Discussions* 11: 9769–9795.

Levy, R. C., L. A. Remer, R. G. Kleidman, S. Mattoo, C. Ichoku, R. Kahn, and T. F. Eck. 2010. Global evaluation of the Collection 5 MODIS dark-target aerosol products over land. *Atmospheric Chemistry and Physics* 10 (21): 10399-10420.

Li, H., F. Faruque, W. Williams, M. Al-Hamdan, J. Luvall, W. Crosson, D. Rickman, and A. Limaye. 2009. Optimal temporal scale for the correlation of AOD and ground measurements of PM2.5 in a real-time air quality estimation system. *Atmospheric Environment* 43: 4303–4310.

Liu, Y., R. J. Park, D. J. Jacob, Q. Li, V. Kilaru, and J. A. Sarnat. 2004. Mapping annual mean ground-level PM2.5 concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States. *Journal of Geophysical Research* 109: D22206.

Liu, Y., J. A. Sarnat , V. Kilaru, D. J. Jacob, and P. Koutrakis. 2005. Estimating Ground-Level PM2.5 in the Eastern United States Using Satellite Remote Sensing. *Environmental Science & Technology* 39 (9): 3269-3278.

Liu, Y., M. Franklin, R. Kahn, and P. Koutrakis. 2007. Using aerosol optical thickness to predict ground-level concentrations in the St. Louis area: A comparison between MISR and MODIS. *Remote Sensing of the Environment* 107 (2007) :33-44.

Mansha, M., and B. Ghauri. 2011. Assessment of fine particulate matter (PM2.5) in metropolitan Karachi through satellite and ground-based measurements. *Journal of Applied Remote Sensing* 5 (1): 053546 1-12.

Martin, R. V. 2008. Satellite remote sensing of surface air quality. *Atmospheric Environment* 30: 1–21.

Mao, L., Y. Qiu, C. Kusano,and X. Xu. 2011. Predicting regional space–time variation of PM2.5 with land-use regression model and MODIS data. *Environmental Science Pollution Research* : DOI 10.1007.

Munchak, L. A., R. C. Levy, S. Mattoo, L. A. Remer, B. N. Holben, J. S. Schafer, C. A. Hostetler, and R. A. Ferrare. 2013. MODIS 3 km aerosol product: applications over land in an urban/suburban region. *Atmospheric Measurement Techniques Discussions* 6 (1): 1683-1716.

NASA – National Aeronautics and Space Administration. 2013a. *MODIS product description*. http://modis-atmos.gsfc.nasa.gov/MOD04_L2/ (last accessed 8 August 2013).

______. 2013b. *MODIS*. http://modis.gsfc.nasa.gov/about/ (last accessed 7 August 2013).

______. 2013c. *LAAD Archives*. http://ladsweb.nascom.nasa.gov/data/ (last accessed 23 July 2013).

Natunen, A., A. Arola, T Mielonen, J. Huttunen, M. Komppula, and K. E. J. Lehtinen. 2010. A multi-year comparison of PM2.5 and AOD for the Helsinki region. *Boreal Environment Research* 15: 544–552.

NOAA – National Oceanic and Atmospheric Administration. 2013. *The Rapid Update Cycle (RUC)* http://ruc.noaa.gov/ (last accessed 29 June 2013).

NYDEC – N.Y. Department of Environmental Conservation. *PM2.5 monitoring*. http://www.dec.ny.gov/chemical/8539.html (last accessed 2 August 2013).

Pope, C. A., D. W. Dockery, and J. Schwartz. 1995. Review of epidemiological evidence of health effects of particulate air pollution. *Inhalation Toxicology* 7 (1): 1-18.

Remer, L. A., D. Tanré, and Y. J. Kaufman. 2005. Algorithm for remote sensing of tropospheric aerosol from MODIS. *National Aeronautics and Space Admnistration.*

Remer, L. A., S. Mattoo, R. C. Levy, and L. Munchak. 2013. MODIS 3 km aerosol product: algorithm and global perspective. *Atmospheric Measurement Techniques Discussions* 6 (1): 69-112.

Rogerson, P. A. 2006. *Statistical Methods for Geography*. London, U.K. SAGE Publications Ltd.

Schaap, M., A. Apituley, R. M. A. Timmermans, R. B. A. Koelemeijer, and G. de Leeuw. 2009. Exploring the relation between aerosol optical depth and PM2.5 at Cabauw, the Netherlands. *Atmospheric Chemistry and Physics* 9: 909–925.

Schäfer, K., A. Harbuscha, S. Emeisa, P. Koepkeb, and M. Wiegner. 2008. Correlation of aerosol mass near the ground with aerosol optical depth during two seasons in Munich. *Atmospheric Environment* 42: 4036‑4046.

Schlesinger, R. B. 2007. The health impact of common inorganic components of fine particulate matter (PM2. 5) in ambient air: a critical review. *Inhalation Toxicology* 19 (10): 811-832.

Schwarze, P. E., J. Øvrevik, M. Låg, M. Refsnes, P. Nafstad, R. B. Hetland, and E. Dybing. 2006. Particulate matter properties and health effects: consistency of epidemiological and toxicological studies. *Human & experimental toxicology* 25 (10): 559-579.

Tian, J., and D. Chen. 2009. A semi-empirical model for predicting hourly ground-level fine particulate matter (PM2.5) concentration in southern Ontario from satellite remote sensing and ground-based meteorological measurements. *Remote Sensing of Environment* 10 (1016).

______. 2010. Spectral, spatial, and temporal sensitivity of correlating MODIS aerosol optical depth with ground-based fine particulate matter (PM2.5) across southern Ontario. *Canadian Journal of Remote Sensing* 36 (2): 119-128.

Tsai, T., Y. Jeng, D. A. Chu, J. Chen, S. Chang. 2011. Analysis of the relationship between MODIS aerosol optical depth and particulate matter from 2006 to 2008. *Atmospheric Environment* 45: 4777-4788.

U.S. Census Bureau. 2012. *Metropolitan and Micropolitan Areas.* http://www.census.gov/population/metro/ (last accessed 14 December 2012).

USDE - U.S. Department of Energy. 2013. *Atmospheric Radiation Measurement (ARM) Data Archive.* http://www.archive.arm.gov/armlogin/login.jsp (last accessed 7 August 2013).

van Donkelaar, A., R. V. Martin, M. Brauer, R. Kahn, R. Levy, C. Verduzco, and P. J. Villeneuve. Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: development and application. *Environmental health perspectives* 118 (6): 847.

Veefkind, J. P., G. D. Leeuw, P. A. Durkee, P. B. Russell, P. V. Hobbs, and J. M. Livingston. 1999. Aerosol optical depth retrieval using ATSR-2 and AVHRR data during TARFOX. *Journal of Geophysical Research* 104 (D2): 2253-2260.

Wallace, J., and P. Kanaroglou. 2007. An investigation of air pollution in southern Ontario, Canada, with MODIS and MISR aerosol data. *Geoscience and Remote Sensing Letters, IEEE* 1 4244-1212-9/07: 4311-4314.

Wang, J., and A. Christopher. (2003). Intercomparison between satellite-derived aerosol optical thickness and PM2.5 mass: Implications for air quality studies. *Geophysics Research Letters*, 30: 1–4.

Wang, Z., L. Chen, J. Tao, Y. Zhang, and L. Su. 2010. Satellite-based estimation of regional particulate matter (PM) in Beijing using vertical-and-RH correcting method. *Remote Sensing of Environment* 114: 50–63.

Zhang, H., R. M. Hoff, and J. A. Engel-Cox. 2006. The relation between Moderate Resolution Imaging Spectroradiometer (MODIS) aerosol optical depth and $PM_{2.5}$ over the United States: A geographical comparison by U.S.Environmental Protection Agency regions. *Journal of Air & Waste Management Association* 59: 1358–1369.