Community Air Monitoring of Particulate Matter in Imperial County, CA

Graeme Nowell Carvlin

A dissertation

submitted in partial fulfillment of the

requirements for the degree of

Doctor of Philosophy

University of Washington

2018

Reading Committee:

Edmund Seto, Chair

Michael Yost

Timothy Larson

Program Authorized to Offer Degree:

Environmental & Occupational Health Sciences

University of Washington

Abstract

Community Air Monitoring of Particulate Matter in Imperial County, CA

Graeme Nowell Carvlin

Chair of the Supervisory Committee:

Professor Edmund Seto

Environmental & Occupational Health Sciences

The Imperial Project was a collaboration between community organizations, academia, and state partners created to investigate levels of particulate matter, a criteria air pollutant, in the highly impacted community of Imperial County. Community-based participatory research (CBPR) techniques were used to ensure that community members were engaged in each step of the research process. To cost-effectively measure air quality levels at a high temporal and spatial resolution a new low-cost community air monitor was designed and deployed in a 40-monitor network. These monitors were calibrated to regulatory beta-attenuation monitors (BAMs) operated by the California Air Resources Board. The calibration equation was validated by siting a mobile version of the BAM monitor called an E-BAM at six sites throughout Imperial County. The relatively high spatial density of the community air monitoring network enabled the creation of a land use regression (LUR) model. The LUR model estimated $PM_{2.5}$, PM less than 2.5 μm in diameter, and PM_{coarse}, PM between 2.5 and 10 μm in diameter, using land use and meteorology. Back-trajectory analyses were performed to help explain which wind conditions lead to high PM in Imperial County. The combined effect of wind speed, wind direction, and seasonality were explored using polar plots. These analyses, along with the estimated emissions and other information detailed in the Imperial County State Implementation Plans, provide a comprehensive view of the sources of particulate matter in Imperial County. Real-time data from the monitoring network is displayed on a public website, IVAN Air, run by the community group Comite Civico del Valle (CCV). CCV and other project partners have taken the information learned from this project and engaged community residents in learning about their air quality and the actions they can take to reduce their exposure.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ACKNOWLEDGEMENTS

Funding for this project came from NIH grant R01ES022722.

Many thanks to my committee members: Edmund Seto, Mike Yost, Chris Simpson, Tim Larson, and Kelly Edwards.

Special thanks to my wife Katherine for her support and patience.

Chapter 1. Introduction

Particulate matter (PM), a mix of liquid and solid particles in the air, has been linked to negative health outcomes that vary based on particle size. $PM_{2.5}$, particulate matter less than 2.5 microns in diameter, has been linked to childhood asthma, cardiovascular disease, cancer, and an increase in overall mortality (Anderson et al., 2011; Brook et al., 2017; Khreis et al., 2017; Atkinson et al., 2016). PM_{coarse}, particulate matter between 2.5 and 10 microns, has been linked to chronic obstructive pulmonary disease (COPD), asthma and respiratory problems, and respiratory-based emergency room visits (Brunekreef & Forsberg, 2005). PM_{10}, PM less than 10 microns, and $PM_{2.5}$ are regulated in the United States under the Environmental Protection Agency's National Ambient Air Quality Standards (NAAQS). Imperial County, a primarily Hispanic agricultural region in southeastern California with a population of approximately 180,000, has exceeded the NAAQS for $PM_{2.5}$ since 2009 and PM_{10} since 1990 (CFR, 2018; U.S. Census Bureau, 2010). According to the California Department of Health Services (2014), Imperial County has the second highest rate of childhood asthma emergency room visits in the state. The Imperial Project (NIH R01ES022722) was a collaboration between academics at the University of Washington and UCLA; a local community group, Comite Civico del Valle (CCV); and community health researchers at the California Environmental Health Tracking Program (CEHTP; a program of the Public Health Institute in collaboration with the California Department of Public Health) to investigate levels of particulate matter in Imperial County (English et al., 2017).

Sources of PM in Imperial include unpaved road and windborne dust from the surrounding desert; agricultural activities such as tilling, burning, and cattle feedlots; and mobile sources such as cars, aircraft, and heavy-duty vehicles (Imperial County Air Pollution Control District, 2014). The 2013 Imperial County State Implementation Plan (SIP) for $PM_{2.5}$ lists unpaved road dust and fugitive windborne dust combined as contributing nearly an order of magnitude more PM per day than the next highest source. Anecdotal evidence from community residents suggests that dust storms, dust from vehicles on unpaved roads, heavy smoke from agricultural burning, emissions from agricultural vehicles, and emissions from industrial sources are visible sources of concern. External sources of PM include transport from Mexicali, a city with around 1 million residents that is adjacent to the US-Mexico border, and dust from the bed of the Salton Sea. The Salton Sea is a large inland lake that is drying up due to changes in water rights that lower the runoff from agricultural fields that was keeping the lake from evaporating. In particular, emissions from Mexicali have been cited by the Imperial County Air Pollution Control district as being the reason for Imperial County's $PM_{2.5}$ non-attainment status (Imperial County Air Pollution Control District, 2014)

The regulatory air monitoring network in Imperial County consists of five sites that measure PM_{10}, two of which also measure $PM_{2.5}$ (California Air Resources Board, 2016). These sites are located to support compliance with the NAAQS and to provide air quality information to the public and to air pollution studies. However, residents felt that the existing air monitoring network did not adequately measure their exposures and did not provide air quality information in a transparent and easy to understand manner. They wanted higher spatial and temporal

resolution data to help them choose when to spend time outdoors to help lower exposure. Community members were also interested in drawing attention to local sources that they felt were impacting the air and their health. It was important for them to have control over the data, to be involved in data processing, and to have data analysis products provided in a way that was easy for them to understand.

The Imperial Project was designed to engage the community to help address their concerns by using a technique called community-based participatory research (CBPR). CBPR includes the local community where the study takes place as an integral member in the research process (Minkler, 2010; Kondo et al., 2014). Community members are invited to help define the scientific questions for the research study, collect and interpret the data, and provide feedback for future collaboration. Through meetings with community members and a cooperative community mapping process, CCV and other project partners helped to create a list of locations that the community felt were important places to site air monitors (Wong et al., 2018). This process led to the selection of 20 sites, primarily schools, businesses, and resident houses. Later on, 20 more monitors were placed at locations suggested by researchers at UW based on spatial analysis and land use regression. These locations were also chosen to increase the geographic coverage of the network and monitor specific large sources, primarily the Salton Sea and the US-Mexico border.

In order to cost effectively monitor air quality at 40 sites a new low-cost community air monitor was designed by researchers at UW. The monitor consisted of a Dylos 1700 laser particle counter, an Arduino Yun microcontroller, an HIH 6130 temperature and humidity sensor, an a NEMA-6 rated enclosure. The Dylos reports particle counts per unit volume, specifically, # of particles per 0.01 ft^3, once every 10 seconds. The laser light is scattered by particles and detected by a photodiode, which creates an output voltage proportional to the amount of incident light. Larger particles reflect more light and lead to a higher voltage. Different voltage cutoffs provide cumulative particle size bins. The Dylos 1700 was customized by the manufacturer per our request to measure particles in four size bins: particles greater than (>) 0.5 µm in diameter, >1.0 µm, >2.5 µm, and >10 µm. Data from the Dylos were averaged to the 5-minute time scale then recorded onto an SD memory card on the microcontroller and sent over Wi-Fi, cellular, or Ethernet networking, to a database hosted at UW. The data was then downloaded for data analyses and, after applying quality control processes, sent to a public website for community members to be able to see their local air quality levels.

Chapter 2 describes the conversion of the Dylos particle count concentrations to mass concentrations and calibration of the community monitors by collocation. A community monitor was collocated with a federal equivalent method (FEM) $PM_{2.5}$ and FEM PM_{10} beta-attenuation monitor (BAM) at the Calexico-Ethel site operated by the California Air Resources Board (CARB). Data were collected for six months during 2015 and 2016. Data were examined during this period to develop a quality control process for the project. The sizing of particles from the Dylos were compared with reference $PM_{2.5}$ and PM_{10} data. A calibration model was developed from the colocation that relates particle count concentrations from the Dylos to the reference PM mass concentrations, adjusting for the effect of relative humidity. Finally, this model was

validated at six other sites with collocated environmental beta attenuation monitors (E-BAMs) serving as references.

Chapter 3 describes a land use regression analysis that explores the effect of land use on monthly $PM_{2.5}$ and PM_{coarse}. PM_{coarse}, the difference between PM_{10} and $PM_{2.5}$, was investigated instead of PM_{10} since it is a separate size fraction and may be linked with different sources, whereas PM_{10} overlaps with $PM_{2.5}$ (Eeftens et al., 2012). Creating a land use regression model would not have been possible with the sparse government network in Imperial County since coverage of a large amount of different land use combinations is necessary. In the land use regression model particulate matter was the response variable and type of land use, such as area of water or length of roads, was one of the explanatory variables. Other explanatory variables included latitude and longitude; meteorological variables including relative humidity, temperature, wind direction and speed, and planetary boundary layer height; traffic; distance to the US-Mexico border and the Salton Sea; and acres burned during permitted agricultural burning events. Varying buffer sizes for land use variables were explored. Leave-one-out cross-validation was used for model selection. For the best-performing model, variable importance was investigated and the most relevant variables for explaining $PM_{2.5}$ and PM_{coarse} were identified. Finally, the models were used to predict surface maps for $PM_{2.5}$ and PM_{coarse} for the valley.

Chapter 4 presents an analysis of the effect of wind direction and wind speed on $PM_{2.5}$ and PM_{coarse} using particle back trajectories and polar plots. The Hybrid Single Particle Lagrangian Trajectory (HYSPLIT) program was used through R using the R package splitR to create 24-hour back trajectories for all hours with high (top 5%), very high (top 0.1%), and median (middle 5% and 0.1%) $PM_{2.5}$ and PM_{coarse} concentrations. The Global Data Assimilation System (GDAS) 1° dataset from the Air Resources Laboratory (ARL) at the National Oceanic and Atmospheric Administration (NOAA) was used as the meteorological input to HYSPLIT. Median hours were matched to high and very high hours by site so that the median and high/very high categories each had the same number of trajectories from each site. Back trajectories were analyzed to determine the places where air parcels were more likely to pass through during periods of elevated PM, compared to periods of moderate PM. Polar plots were used to investigate the effect of wind speed and direction on PM in Brawley, a city in the center of the Imperial Valley, and Calexico, a city on the US-Mexico border.

While not described in individual chapters in the dissertation, there are many other aspects of the project that reflect the CBPR monitoring process. For instance, the calibrated data from the 40 community air monitors are displayed on a website run by CCV called IVAN air, https://ivan-imperial.org/air. The current air quality level can be seen along with summary statistics for the last 24 hours, 30 days, and 90 days. Community members can sign up for text alerts for individual monitors. CCV has done outreach campaigns with the air monitoring data from the network to help teach residents how to use IVAN air, how to understand what the different air quality levels mean, and what actions they can take to protect themselves and their families during high PM episodes. A majority of this outreach has been directed at schools and teaching children about air quality. The Imperial Project had a community steering committee

that was made up of various community stakeholders including students from local high schools. This steering committee was integral in the community-based participatory research process and in teaching others about the Imperial community air monitoring network.

As a final step in the CBPR process the air monitors themselves and the operation of the network were handed over to CCV. CCV was the primary group that sited the air monitors and performed maintenance on the monitors throughout the study. They are now actively working on setting up and maintaining the database that stores the data from the monitors and running the software that does QC validation of the monitoring data. They are also able to assemble and program the monitors themselves.

The Imperial Project members are now working on documenting and exporting the Imperial Project community air monitoring model to other highly impacted communities. This includes creating a manual that details how community residents can work with academic and government partners to monitor the air quality in their community. A community learning conference will be held to help interested communities learn more about community air monitoring.

Reference for Chapter 1

2017. Protection of Environment. Title 40. Part 50, National Primary and Secondary Ambient Air Quality Standards. *U.S. Code of Federal Regulations.,* CFR. https://www.ecfr.gov/cgi-bin/text-idx?tpl=/ecfrbrowse/Title40/40cfr50_main_02.tpl

2018. Protection of the Environment. Title 40. Subpart C to Part 81, 81.305 Attainment Status Designations - California. U.S. Code of Federal Regulations., CFR. https://www.ecfr.gov/cgi-bin/text-idx?SID=6ce471d8cc33d1695ff368b16ccc3bce&mc=true&node=se40.20.81_1305&rgn=div8

Anderson, J., Thundiyil, J., & Stolbach, A. 2011. Clearing the Air: A Review of the Effects of Particulate Matter Air Pollution on Human Health. *Journal of Medical Toxicology, 8*(2), 166-175.

Atkinson R.W., Butland B.K., Dimitroulopoulou C., Heal M.R., Stedman J.R., Carslaw N., Jarvis D., Heaviside C., Vardoulakis S., Walton H., Anderson H.R. 2016. Long-term exposure to ambient ozone and mortality: a quantitative systematic review and meta-analysis of evidence from cohort studies. *BMJ Open, 6*(2).

Brook R.D., Newby D.E., Rajagopalan S. 2017. The Global Threat of Outdoor Ambient Air Pollution to Cardiovascular Health: Time for Intervention. *JAMA Cardiol.*

Brunekreef B., Forsberg, B. 2005. Epidemiological evidence of effects of coarse airborne particles on health. *European Respiratory Journal, 26*, 309-318.

California Air Resources Board, 2016. Annual Network Plan. https://www.arb.ca.gov/aqd/amnr/amnr2016.pdf (accessed January 17, 2017).

California Department of Health Services, 2014. Asthma ED Visits, Children, 2014. http://www.californiabreathing.org/asthma-data/county-comparisons/edvisits-children (accessed February 12, 2017).

Eeftens, M., Tsai, M., Ampe, C., Anwander, B., Beelen, R., Bellander, T., . . . Hoek, G. (2012). Spatial variation of PM2.5, PM10, PM2.5 absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO2 – Results of the ESCAPE project. *Atmospheric Environment, 62*, 303-317. doi:10.1016/j.atmosenv.2012.08.038

English, P., Olmedo, L., Bejarano, E., Lugo, H., Murillo, E., Seto, E., Wong, M., King, G., Wilkie, A., Meltzer, D., Carvlin, G., Jerrett, M., & Northcross, A. 2017. The Imperial County Community Air Monitoring Network: A Model for Community-based Environmental Monitoring for Public Health Action. *Environmental Health Perspectives, 125*(7), 074501-1 to 074501-5.

Imperial County Air Pollution Control District, 2014. Imperial County 2013 State Implementation Plan for the 2006 24-hour PM2.5 Moderate Nonattainment Area. http://www.arb.ca.gov/planning/sip/planarea/imperial/Final_PM2.5_SIP_%28Dec_2,_2014%29_Approved.pdf (accessed January 24, 2017).

Khreis H., Kelly C., Tate J., Parslow R., Lucas K., Nieuwenhuijsen M. 2017. Exposure to traffic-related air pollution and risk of development of childhood asthma: A systematic review and meta-analysis. *Environ Int., 100*, 1-31.

Kondo, M., Mizes, C., Lee, J., Mcgady-Saier, J., O'Malley, L., Diliberto, A., Burstyn, I. (2014). Towards Participatory Air Pollution Exposure Assessment in a Goods Movement Community. *Progress in Community Health Partnerships: Research, Education, and Action, 8*(3), 291-304.

Minkler, M. (2010). Linking Science and Policy Through Community-Based Participatory Research to Study and Address Health Disparities. *American Journal of Public Health, 100*(S1).

Richards, L., Alcorn, S., McDade, C., Couture, T., Lowenthal, D., Chow, J., Watson, J. (1999). Optical properties of the San Joaquin Valley aerosol collected during the 1995 integrated monitoring study. *Atmospheric Environment, 33*, 4787-4795.

U.S. Census Bureau. State and County QuickFacts, Imperial County, CA. Washington, D.C.: Government Printing Office. U.S. Census Bureau. (2010). Retrieved from https://www.census.gov/quickfacts/fact/table/imperialcountycalifornia/PST045217#viewtop

Wong, M., Bejarano, E., Carvlin, G., Fellows, K., King, G., Lugo, H., Jerrett, M., Meltzer, D., Northcross, A., Olmedo, L., Seto, E., Wilkie, A., & English, P. (2018). Combining Community Engagement and Scientific Approaches in Next-Generation Monitor Siting: The Case of the Imperial County Community Air Network. *International Journal of Environmental Research and Public Health, 15*(523), 1-14.

Wu, C., Delfino, R., Floro, J., Samimi, B., Quintana, P., Kleinman, M., Liu, L. (2005). Evaluation and quality control of personal nephelometers in indoor, outdoor and personal environments. *Journal of Exposure Analysis and Environmental Epidemiology, 15*, 99-110.

Chapter 2. Development and Field Validation of a Community-Engaged Particulate Matter Air Quality Monitoring Network in Imperial, CA*

*Published as: Graeme N. Carvlin, Humberto Lugo, Luis Olmedo, Ester Bejarano, Alexa Wilkie, Dan Meltzer, Michelle Wong, Galatea King, Amanda Northcross, Michael Jerrett, Paul B. English, Donald Hammond & Edmund Seto (2017) Development and field validation of a community-engaged particulate matter air quality monitoring network in Imperial, California, USA, *Journal of the Air & Waste Management Association*, 67:12, 1342-1352, DOI: 10.1080/10962247.2017.1369471

2.1 Introduction

Particulate matter (PM), a collection of liquid and solid particles in the air, has been found to be associated with numerous adverse health outcomes across the life course, including adverse birth outcomes, incident childhood asthma, delayed lung function development, cardiovascular disease development, cancer incidence, and premature death (Brook et al., 2017; Anderson et al., 2011; Atkinson et al., 2016; Khreis et al., 2017). Asthmatic children are particularly at risk from high PM levels due to heightened airway responsiveness, such that air pollution can often cause exacerbation of symptoms (Schwartz, 2004). Imperial County, located in southeastern California, has had levels of $PM_{2.5}$ (PM under 2.5 μm in diameter) and PM_{10} (PM under 10 μm in diameter) that have repeatedly exceeded the U.S. Environmental Protection Agency's (EPA) National Ambient Air Quality Standards (Imperial County Air Pollution Control District, 2014; California Air Resources Board, 2015). The county also has the second highest rate of childhood asthma emergency department visits in the state (California Department of Health Services, 2014).

The regulatory air monitoring network in Imperial Valley serves to support compliance with ambient air quality standards, provide air pollution data to the public, and support air pollution research studies. The PM network consists of five sites that measure PM_{10}, of which two also measure $PM_{2.5}$ (California Air Resources Board, 2016). One site is designed to assess concentrations near the United States–Mexico border and the community of Calexico, whereas the other four measure air quality levels throughout the Imperial Valley, largely in a strip of rural farmland running down the middle of Imperial County where most of the population resides in small communities.

The Imperial County Community Air Monitoring Project (National Institutes of Health [NIH] R01ES022722) was designed as a community-engaged research study that partnered public health researchers from the California Environmental Health Tracking Program, environmental justice leaders from a local community-based organization, Comite Civico del Valle (CCV), and various academics with experience in air quality and health effects studies to assess environmental quality needs, conduct community-led air quality monitoring, and identify opportunities in which higher spatial resolution environmental data may affect policy and planning efforts in the valley.

Early in the study, community residents expressed concern that the existing regulatory network did not adequately measure their exposure to air pollution. They desired higher spatial and temporal resolution data to help make decisions on how to best protect themselves and their children during high-pollution events. Furthermore, there was interest in additional monitoring where sensitive subpopulations may be exposed to local pollution levels higher than those observed throughout the rest of the region. Eleven communities were chosen as priority areas for air monitoring by a community steering committee (CSC), and local community members volunteered to participate in a process to identify, map, and collect data on potential monitoring sites (English et al., 2017). In order to cost effectively monitor 40 sites, the study organizers developed a monitoring platform that consisted of low-cost technologies—a commercially available particle counter, additional sensors for temperature and humidity, and a wireless microcontroller. The networked PM monitors provided increased coverage, in terms of number of monitors compared with the regulatory network, by nearly an order of magnitude and allowed for real-time display of air quality levels.

In the past few years, there has been a dramatic increase in the number of low-cost next-generation air monitoring technologies (Jiao et al., 2016). However, the performance of these monitors in comparison with existing federal equivalent method (FEM) and federal reference method (FRM) air monitors is often poorly documented. The South Coast Air Quality Management District and the EPA have begun to perform laboratory and field experiments for some monitors and have found that performance varies widely (South Coast Air Quality Management District, 2015; Williams et al., 2014). Most low-cost PM monitors operate based on light scattering. So, depending on particle composition, uncorrected measurements can vary regionally and seasonally. Therefore, it is important to calibrate and validate monitors in the field under typical use conditions. A calibration equation for estimating particle mass concentrations using particle count concentrations produced by optical light scattering monitors can be developed by collocating monitors with FEM or FRM instruments and then validated at other sites.

For this study, we collocated a monitor at the California Air Resources Board (CARB) Calexico-Ethel site, located at the Calexico High School on East Belcher Street, which collects reference measurements of PM from both FEM beta-attenuation monitors (BAMs) and FRM filter–based gravimetric samplers. By comparing FEM and FRM data with data from the community air monitors, an equation for estimating mass concentrations from the Dylos particle count concentrations was developed. The calibration equation was validated by comparing calibrated monitor results with $PM_{2.5}$ measured by collocated reference instruments at six other sites. Both the Dylos and BAM mass measurements were also converted to an air quality index (AQI) and displayed on a Web site, allowing community partners, air quality stakeholders, and residents access to real-time calibrated data from the Imperial community air monitoring network (Appendix A, Figure 10).

2.2 Materials and Methods

The particle counter used in the community monitors was a modified Dylos 1700 (Dylos Corporation, Riverside, CA). The firmware was changed to increase the number of particle size bins from two to four (>0.5 μm, >1.0 μm, >2.5 μm, and >10 μm). A custom circuit board was designed to interface the Dylos with an Arduino Yun (Arduino LLC, Turin, Italy) to add networking capabilities. The circuit board also integrated a HIH 6130 temperature and relative humidity sensor (Honeywell International Inc., Morris Plains, NJ). Each Arduino Yun was then connected to either Wi-Fi, Ethernet, or a T-Mobile cellular modem. This enabled transmission of the Dylos, temperature, and relative humidity data to a database operated by the University of Washington information technology (IT) staff in collaboration with the Seto laboratory. The Dylos and networking hardware were installed into a National Electrical Manufacturers Association (NEMA) 6–rated enclosure (Figure 1). Not including labor to construct and maintain the monitor, the cost of the parts for the system was around $1,500.

The Dylos output readings in units of particle number per 0.01 ft^3 every 10 sec to the Arduino Yun, which collected these values and sent the 5-min average to the database. For this data analysis, the 5-min averages were downloaded from the database and averaged to 1-hr values. Any hour with less than 75% data completeness was flagged for quality control and removed. All hours with a particle count ≤ 30 were also removed. Values below this threshold indicated a problem with the Dylos monitor; most often that the photodiode was dirty and the monitor needed to be cleaned. All data analyses were performed using R statistical software (R Development Core Team, v. 3.4.0, April 2017).

The Calexico-Ethel site (EPA Air Quality Site [AQS] no. 060250005) was used for collocation and comparison of the Dylos with reference samplers because it had several PM monitors, including a $PM_{2.5}$ FRM sampler (R&P 2025; Thermo Fisher Scientific, Albany, NY), a PM_{10} FRM sampler (Sierra Anderson 1200; Thermo Fisher Scientific, Albany, NY), and two $PM_{2.5}$/PM_{10} FEM samplers (MetOne BAM-1020; Met One Instruments, Inc, Grants Pass, OR). The $PM_{2.5}$ FRM sampler used a 16.67 liters per minute (LPM) medium volume pump and a 46.2-mm Teflon filter, which was weighed before and after sampling to determine PM concentration. The PM_{10} FRM sampler used a 40 LPM high-volume pump and a 46.2-mm quartz filter, which was weighed before and after sampling to determine PM concentration. $PM_{2.5}$ and PM_{10} FEM samplers used a beta-particle source and detector to determine the amount of radiation absorbed by particles caught on a filter tape, which is proportional to particle mass. The Calexico-Ethel PM samplers are located on top of a trailer in a semiurban area of Calexico, California, away from local sources and nearby obstructions.

The primary $PM_{2.5}$ monitor is an FRM sampler operating on a daily schedule. During 2014 and 2015, in addition to the FRM, a pair of collocated $PM_{2.5}$ FEM monitors were operated at the site to collect hourly data. At the beginning of 2016, one of the FEM $PM_{2.5}$ samplers was converted to a FEM PM_{10} sampler, which replaced the existing FRM PM_{10} sampler. At the same time, the second FEM $PM_{2.5}$ sampler was converted to a non-FEM $PM_{2.5}$ sampler to continue hourly monitoring for local forecasting. $PM_{2.5}$ FRM collocation data were available for the entire study period, whereas FEM $PM_{2.5}$ data were available for 2015. The $PM_{2.5}$ comparison used data

collected from (month/day/year) 6/18/2015 to 12/21/2015. The PM_{10} comparison used data collected from 1/15/2016 to 7/12/2016.

A conversion equation was developed to transform the Dylos particle count concentrations into particle mass concentrations using data from the collocation with the FEM $PM_{2.5}$ and FEM PM_{10} BAMs at the Calexico-Ethel site. After the data quality control checks, the resulting data were averaged to 1-day values using two different techniques. First, the 5-min Dylos data were averaged directly to 1-day data for comparison with $PM_{2.5}$ and PM_{10} FRM gravimetric filters. Second, the 1-hr Dylos data from Calexico-Ethel were averaged to 1-day data for comparison with the $PM_{2.5}$ and PM_{10} FEM BAMs using only hours when both the Dylos and BAMs had data. This was done to ensure that missing data in either data set did not bias the comparison.

The primary and collocated measurements for both $PM_{2.5}$ FEM and FRM were compared to validate the precision and accuracy of CARB's samplers. Hourly data were averaged for the day, only including days that had at least 18 valid hours or 75% completeness. $PM_{2.5}$ FEM daily averages with matching primary and collocated days were compared from June 2015, the start of the Dylos collocation, until December 2015, the end of $PM_{2.5}$ FEM monitoring.

The first step toward creating a calibration equation was to investigate which Dylos particle size bin or combination of bins was most highly correlated with $PM_{2.5}$ and PM_{10} as measured by the BAMs and filters. A priori, it would be expected that the Dylos would best measure $PM_{2.5}$ as bin 1 − bin 3 (all particles >0.5 μm minus all particles >2.5 μm) and PM_{10} as bin 1 − bin 4 (all particles >0.5 μm minus all particles >10 μm). However, the correlations between reference PM data and all individual bins and bin combinations were examined using the coefficient of determination (R2), which measures the amount of variance in the dependent variable that is predicted by the independent variable. For $PM_{2.5}$, bin 1 and bin 1 − bin 3 were most highly correlated and there was only a slight difference between them. For PM_{10}, bin 3 and bin 3 − bin 4 were most highly correlated, suggesting that Dylos counts for particles larger than 2.5 μm and between 2.5 and 10 μm are most representative of PM_{10}. There was only a slight difference between the correlation for bin 3 and bin 3 − bin 4. The highest correlated bins (bin 1 for $PM_{2.5}$ and bin 3 for PM_{10}, respectively) were used in subsequent analyses.

Finally, a conversion equation was developed using the Dylos counts and FEM $PM_{2.5}$ and PM_{10} BAM data. Relative humidity and temperature were tested as possible covariates, and it was found that separately both improved the conversion equation. Relative humidity is known to change particle size due to the addition or subtraction of water from particles (Winkler, 1973). Past research has used a non-linear adjustment above a certain RH, anywhere from 60% to 85% (Richards et al., 1999; Wu et al., 2005). However, this would not have much applicability to the Imperial data since the max RH in the dataset used to develop the conversion equation was 66% with a mean of 19%. Since temperature and relative humidity were moderately correlated, only relative humidity was included in the conversion equation.

$PM_{2.5}$ conversion equation:

$$Dylos_{bin1} = \beta_0 + \beta_1 BAM_{PM2.5} + \beta_2 RH + e() \quad (1)$$

PM_{10} conversion equation:

$$Dylos_{bin3} = \beta_0 + \beta_1 BAM_{PM10} + \beta_2 RH + e() \quad (2)$$

where β_0 is the intercept, RH is the relative humidity as measured by the RH sensor on our custom circuit board, and e is the residual error. In these models, we were most interested in the error of the Dylos relative to the reference instrument. Thus, the BAM values appear on the right side of the equation, and the error term explains the residual deviation of the Dylos measurement from the BAM measurement after a constant offset (the intercept) and RH have been accounted for. We assumed that the Dylos will have greater error than the BAM because the BAM is a FEM instrument operated by CARB.

This equation was inverted to estimate BAM-measured mass concentrations in units of $\mu g/m^3$ from the Dylos measurements:

$$Dylos_{bin1} - \beta_2 RH_{Dylos} - \beta_0 = \beta_1 BAM_{PM2.5} \quad (3)$$

$$\left(\frac{1}{\beta_1}\right) Dylos_{bin1} - \left(\frac{\beta_2}{\beta_1}\right) RH_{Dylos} - \left(\frac{\beta_0}{\beta_1}\right) = BAM_{PM2.5} \quad (4)$$

To make this simpler to implement, we defined the following constants:

$$c_1 = \frac{-\beta_0}{\beta_1} \quad (5a)$$

$$c_2 = \frac{1}{\beta_1} \quad (5b)$$

$$c_3 = \frac{-\beta_2}{\beta_1} \quad (5c)$$

Furthermore, $BAM_{PM2.5}$ was our estimate of $Dylos_{PM2.5\_mass}$, therefore:

$$Dylos_{PM2.5\,mass} = c_1 + c_2 * Dylos_{bin1} + c_3 * RH_{Dylos} \quad (6)$$

And for PM_{10}:

$$Dylos_{PM10\ mass} = c_1 + c_2 * Dylos_{bin\,3} + c_3 * RH_{Dylos} \quad (7)$$

These equations were applied to data collected from all community monitors in our study.

To validate the community monitors' measurements and the conversion equation for $PM_{2.5}$, CARB deployed MetOne environmental BAMs (E-BAMs) with size inlets for PM2.5 at six of the community monitoring sites from 3/2/2016 to 7/19/2016 (Figure 2). All E-BAMs were operated according to CARB standard operating procedures, and only E-BAM hourly averaged mass concentration data that passed their quality assurance were used in our analyses. The monitors at the validation sites were located on the top of school buildings and private businesses. Monitors were sited away from local sources such as heating, ventilation, and air conditioning (HVAC) systems and nearby obstructions such as tall trees.

Dylos $PM_{2.5}$ mass concentrations were compared with the collocated E-BAM $PM_{2.5}$ mass measurements. One site, Kennedy, was only continuously online during March. As a result, data from the site after March were excluded from the analysis. Dylos counts were converted to $PM_{2.5}$ mass concentrations using eq 6. Dylos and E-BAM hourly data were averaged into daily measurements. Daily averages were paired when both values met at least 75% completeness, that is, 18 or more valid hours. The Dylos and E-BAM collocation sampling was nearly a 5-month period from March through July 2016. The R^2 value was calculated for the Dylos $PM_{2.5}$ mass measurements and the E-BAM $PM_{2.5}$ mass measurements. Then an assessment of precision and bias between the samplers was conducted using EPA air monitoring statistics for $PM_{2.5}$ daily values (U.S. Code of Federal Regulations, 2016). Paired daily averages with both values above 3 $\mu g/m^3$ were used for the assessment as required by the statistical procedure.

Additionally, separate site-specific conversion equations were created for each site, modeling the relationship between the collocated Dylos and E-BAM. These models used the same variables and inversion technique as the Calexico-Ethel conversion equation. The site-specific models were used to convert Dylos counts to $PM_{2.5}$ separately for each site. The R^2 for these Dylos $PM_{2.5}$ mass measurements and the E-BAMs was calculated and compared with the $PM_{2.5}$ Dylos and E-BAM R^2 values achieved using eq 6.

Finally, to provide the community with a high spatiotemporal resolution real-time map of air pollution in Imperial County, the CSC and project partners provided input and determined requirements for the display of the air monitoring data. With this guidance, an existing community environmental reporting Web site operated by CCV was enhanced to display data from the 40-monitor network (www.ivan-imperial.org/air; Appendix A, Figure 10).

2.3 Results and Discussion

2.3.1 Conversion

Daily averaged $PM_{2.5}$ FEM and FRM measurements from the Calexico-Ethel site were compared to assess the precision and accuracy of CARB's samplers and to ensure that data from these samplers could be used as a reference to compare with the Dylos. The average concentrations for the primary and collocated samplers were 12.1 and 11.6 $\mu g/m^3$, respectively, or 4.2% difference. The correlation for $PM_{2.5}$ FEM samplers was very good, with $R^2 = 0.899$, slope = 0.996, and intercept of 0.54 $\mu g/m^3$ (Appendix A, Figure 1). $PM_{2.5}$ FRM measurements were similarly compared for the same time period. The average concentrations for the primary and collocated samplers were 11.1 and 11.2 $\mu g/m^3$, respectively, or 0.90% difference. The correlation between replicate $PM_{2.5}$ FRM samplers was very good, with $R^2 = 0.953$, slope = 0.971, and intercept of 0.15 $\mu g/m^3$ (Appendix A, Figure 2). Matching days for the $PM_{2.5}$ FRM and FEM primary samplers were compared as a check on method accuracy. The average concentrations for the FRM and FEM were 10.7 and 11.9 $\mu g/m^3$, respectively, or 11% difference. The correlation for the methods was fair, with $R^2 = 0.727$, slope = 0.664, and intercept of 2.79 $\mu g/m^3$. The FEM continuous sampler had a slight positive bias over the FRM one during the study period at the Calexico-Ethel collocation site. The FEM and FRM results support the accuracy of both methods and therefore can be used to model the correction for the Dylos.

Correlation of the community air monitor's Dylos bins with BAM measurements was performed in order to determine which bin(s) most accurately measure $PM_{2.5}$ and PM_{10} (Appendix A, Table 1). It was found that bin 1 and the difference between the fourth and first bins (bin 1 − bin 4) were most correlated with $PM_{2.5}$, with Pearson correlations of 0.78 and 0.78, respectively, and bin 3 and the difference between the fourth and third bins (bin 3 − bin 4) were most correlated with PM_{10}, with Pearson correlations of 0.87 and 0.86, respectively. Bin 1 and bin 3 were chosen to represent Dylos $PM_{2.5}$ and PM_{10} counts, respectively.

The Dylos particle counts compared quite well with the FRM filter and FEM BAM for both $PM_{2.5}$ and PM_{10} (Table 1). The Dylos compared better with the BAM on the daily rather than hourly time scale. The addition of RH or temperature in the Dylos-BAM hourly models increased the adjusted R^2 by 0.03; however, including both only raised the adjusted R^2 by 0.01 over including either temperature or RH. The reason for this may be because observed temperature and RH values were moderately correlated ($r = -0.51$). Thus, only RH was included in the conversion equation.

The calculated constants from the regression models were as follows: $PM_{2.5}$: c1 = 4.790, $c2 = 7.879 \times 10^{-3}$, $c3 = -2.294 \times 10^{-1}$; PM10: c1 = 8.045; $c2 = 2.375 \times 10^{-1}$; $c3 = -9.661 \times 10^{-1}$. These were used to convert the Dylos count data to mass concentration. There were 3907 hourly data points for the $PM_{2.5}$ conversion and 4008 hourly data points for the PM_{10} conversion. The conversion produced some negative Dylos PM mass values, which were not omitted from the following analyses.

The resulting R^2 values between the $PM_{2.5}$ and PM_{10} BAM measurements and Dylos measurements at the CARB Calexico-Ethel site after conversion from particle number to particle

mass concentration were as follows: $PM_{2.5}$: 0.79 (hourly), 0.84 (daily); PM_{10}: 0.78 (hourly), 0.81 (daily). The data are plotted in Figure 3. The correlation for PM_{10} on the daily scale was lower than the pre-conversion correlation due to a few low-PM, high- RH days where the model did not perform well.

Previous comparisons of the Dylos with reference instruments have shown R^2 values ranging from 0.53 to 0.95. South Coast Air Quality Management District (2015) reported correlations of 0.63 (hourly) and 0.81 (daily) with a PM2.5 BAM; Northcross et al. (2013) reported correlations of 0.81–0.99 (hourly) with a TSI DustTrak (Model 8520, TSI Inc, Shoreview, MN); Williams et al. (2014) and Manikonda et al. (2016) found a correlation of 0.53 (hourly) with a Grimm model EDM180 (GRIMM Aerosol Technik, Ainring, Germany); and Steinle et al. (2015) found correlations of 0.70 and 0.90 (hourly) with a tapered element oscillating microbalance (TEOM). The R2 values for Dylos and BAM in this study, 0.79 for hourly $PM_{2.5}$ and 0.78 for hourly PM_{10}, are similar to the Dylos-BAM correlations reported in the literature.

A report by the South Coast Air Quality Management District showed good correlation between a Dylos 1700 and a Grimm, and the Dylos response seemed to be linear at concentrations relevant to field use (South Coast Air Quality Management District, 2017), i.e., less than ~70,000 particles per 0.01 ft^3, which is above the maximum value seen in this study. However, without calibration, the Dylos overestimated the Grimm. This highlights the importance of calibrating the Dylos before trying to interpret its measurements.

The Dylos $PM_{2.5}$ mass concentrations compared well with both the $PM_{2.5}$ FEM BAM and the $PM_{2.5}$ FRM filter. The correlation for Dylos $PM_{2.5}$ to the $PM_{2.5}$ FRM filter was $R^2 = 0.792$ (slope = 1.413 and intercept of −2.27 $\mu g/m^3$). The average concentrations for the Dylos $PM_{2.5}$ and $PM_{2.5}$ FRM filter, which were compared from 6/20/2015 to 7/12/2016, were 14.0 and 11.5 $\mu g/m^3$, respectively, or 20% difference. The correlation for 24-hr Dylos $PM_{2.5}$ to the $PM_{2.5}$ FEM BAM was $R^2 = 0.843$ (slope = 1.108 and intercept of −1.38 $\mu g/m^3$). The average concentrations for the Dylos $PM_{2.5}$ and $PM_{2.5}$ FEM BAM, which were compared from 6/20/2015 to 12/21/2015, were identical at 12.6 $\mu g/m^3$. The correlation for Dylos PM_{10} to the PM_{10} FRM filter was $R^2 = 0.808$ (slope = 1.572 and intercept of −27.50 $\mu g/m^3$). The average concentrations for the Dylos PM_{10} and PM_{10} FRM filter, which were compared from 6/23/2015 to 1/19/2016, were 52.2 and 50.7 $\mu g/m^3$, respectively, or 3% difference. The correlation for the 24-hr Dylos PM_{10} to PM_{10} FEM BAM was $R^2 = 0.808$ (slope = 1.094 and intercept of −5.21 $\mu g/m^3$). The average concentrations for the Dylos PM_{10} and PM_{10} FEM BAM, which were compared from 1/15/2016 to 7/12/2016, were 55.5 and 55.4 $\mu g/m^3$, respectively, or 0.18% difference.

2.3.2 Validation

Based on the conversion equation developed for the Calexico-Ethel data, the Dylos data from the six E-BAM collocation sites were converted from particle number to particle mass concentration. Some error may have been introduced in the following analyses by comparing the Dylos with an E-BAM instead of a BAM, which was the instrument used in the development of

the conversion equation. However, there were no extra BAMs available to help validate the conversion equation. In future studies, we would suggest either collocating an E-BAM with a BAM to understand how they compare or using more BAMs for validation.

Summary statistics calculated for each site for the Dylos and the E-BAMs revealed that in general, with one exception, Dylos averages tended to be lower than the E-BAM averages (Table 2). Also, nearly all Dylos had lower variability than the collocated E-BAMs. CARB's quality control protocol for this study invalidated hourly E-BAM data below −3 $\mu g/m^3$ to retain data reading near zero that may have variability while removing very low negative values that were suspect.

$PM_{2.5}$ precision of collocated samplers was measured with a coefficient of variation (CV) statistic and used a performance goal within ±10%. Although this EPA criterion is seldom met in national air monitoring networks and the Dylos-E-BAM monitoring was not using identical samplers, CV and bias values provide useful indicators of performance (Table 3). Bias was the average percent difference among daily pairs with one sampler being the audit or reference sampler. The E-BAM was used as the reference sampler; therefore, a positive bias value indicates that the Dylos was measuring higher than the E-BAM; conversely, a negative bias value indicates the Dylos was measuring lower than the E-BAM. The average bias of all six collocations was −4.7% with a range of 28.3% to −31.4%. The average precision of all six collocations was 25.2% with a range of 17.5% to −35.2%. As a comparison, the CARB Calexico-Ethel $PM_{2.5}$ FRM filter–based collocated precision levels were 22.8% and 8.2% in 2015 and 2016, respectively. Given that the FRM-FRM comparison was using identical samplers whereas the Dylos-E-BAM had different methodologies, and should be expected to be larger, the precision for the Dylos-E-BAM appears to be reasonable for good measurements.

Time-series plots, scatterplots, and residual plots were created for each site. In general, the Dylos measurements tracked similarly to the E-BAM measurements over time and the Dylos detected similar peaks as the E-BAM (Appendix A, Figures 3–8). As can be seen from the scatterplots, the relationship between the Dylos and E-BAMs was approximately linear; however, the Dylos tended to underestimate $PM_{2.5}$ mass at higher concentrations (Figure 4).

To evaluate the relationship between the Dylos conversion over time and the E-BAM measurements, the differences were plotted as a function of time and fit to a linear regression (Appendix A, Figure 9). Most sites had slightly negative slopes for the differences, which seems to point to a gradually increasing underestimation by the Dylos over the 5 months of the study. This may be due to dust accumulation on the photodiode, which causes the monitors to read lower. During this study, the Dylos were switched out at these sites when they began reading too low (consistently under 30 particles per 0.01 ft^3 in bin 1), but no preventative maintenance was performed. Based on our experience with the Dylos, a twice-yearly preventative maintenance would be beneficial.

The intercepts were not uniformly negative or positive, pointing to a site-specific variation in the Dylos- E-BAM relationship. In particular, the Calipatria site had a −4.1 $\mu g/m^3$ intercept and 0 slope, indicating that the Dylos consistently read 4.1 $\mu g/m^3$ lower than the E-

BAM, and the Seeley site had an intercept of 3.5 μg/m^3 and a slope of −0.001, indicating that the Dylos consistently read 3.5 μg/m^3 higher than the E-BAM. The nonzero intercepts seen at Calipatria and Seeley may be due to an instrument problem or the placement of the E-BAM and community monitor at those sites. At Calipatria, the E-BAM was located on a secondary roof approximately 10 ft higher than the roof on which the community monitor was placed. Whatever the reason, in future investigations we would like to have an extra Dylos and an extra E-BAM to move around to different sites during the study period in order to see if shifts like these are due to instrument specific variations. The use of a rover community monitor to collocate for a period of time at sites of interest could also be used to verify performance.

It was hypothesized that the relationship between the Dylos and E-BAM might change by site due to site-specific variations in particle composition. A particle's optical properties, and therefore the Dylos response, is dependent upon its composition. This was tested by comparing the correlation between E-BAM and Dylos mass concentration using the Calexico-Ethel conversion versus the site-specific conversion (Table 4). The first column lists the R^2 values for Dylos $PM_{2.5}$ mass (Calexico-Ethel model) and E-BAM $PM_{2.5}$ mass. There was heterogeneity between sites, with R^2 values ranging from 0.35 to 0.81. The average R^2 between the Dylos $PM_{2.5}$ mass (Calexico- Ethel model) and E-BAM $PM_{2.5}$ across all sites was 0.59. The lower R^2 values for Kennedy (0.37) and Calexico Alvarez (0.35) were due to the smaller range of mass concentrations seen at those sites (Figure 4 and Appendix A, Figures 7–8). The R^2 values for the four other sites when restricted to concentrations below 93 μg/m^3, the maximum concentration seen at Kennedy and Calexico Alvarez, were between 0.31 and 0.62. The second column contains the R^2 values for Dylos $PM_{2.5}$ mass (site-specific model) and E-BAM $PM_{2.5}$ mass. There was only a small increase in the correlation (0.01–0.03) for each site when using a site-specific conversion equation. This suggests that there was only a marginal improvement, if any, to be gained by employing a site-specific model. To our knowledge, the Dylos has never been compared with an E-BAM in the literature, but these correlations are within the range of correlations seen between the Dylos and other reference instruments mentioned previously.

2.3.3 Data Completeness

One of the limitations we encountered in this study was loss of data at some of our community monitoring sites. Approximately 40% of the Dylos data were lost across the collocation sites. Data loss occurred due to failure of the Dylos monitor and issues with instability of Wi-Fi connections, which often caused the microprocessor to freeze and fail to record data to the onboard SD card. This data loss ranged from <1% to 83% across the six sites (Appendix A, Table 2). The reason for differing data loss across sites was due to both environmental conditions and network type. Location was important to data loss because it determined whether or not the Dylos was exposed to the high-dust conditions that can occur in the north of the Imperial Valley (Figure 2). Both the Calexico-Ethel site used for the conversion model and the Calexico Alvarez site, which are located in southern Imperial County, did not have any problems with the Dylos needing to be cleaned. Three of the sites in the central or northern part of the valley needed to have the Dylos cleaned one or more times. As for network

connections, the Calexico-Ethel and Kennedy sites used a cellular modem to upload data and Seeley and Meadows used Ethernet; both were very reliable and, except for two short network outages at Calexico-Ethel, experienced basically no data loss. Although each monitor had an SD card for internal data storage, network outages would cause the microcontroller to crash and therefore not record any data. This is an issue with the specific type of microcontroller, which is no longer used in our newer monitors.

Generally, ongoing maintenance has been a recurring issue in the study and will remain an issue as we try to sustain community-led monitoring in the valley. Our field crew, a single part-time CCV staff member, has limited resources to dedicate to network maintenance. However, by implementing cellular connections and a more rigorous maintenance schedule, our total data loss dropped from 44% to 18% between the study period and 1/1/2017–3/1/2017. This exceeded our original data completeness goal of 75%.

2.4 Conclusion

Our study illustrates the importance of understanding the performance of low-cost next-generation air quality monitoring technologies under particular use scenarios. Calibration and validation of low-cost air monitors with regulatory instruments is crucial to ensure quality data. This is particularly important for community air monitoring, in which the data are displayed back to the community members. It should be considered best practice for other communities engaging in community air monitoring to do similar evaluations if possible.

Furthermore, augmenting basic air monitors with relative humidity data and network connectivity can allow for more robust conversion from particle count to particle mass and provide the possibility for real-time data quality control, analysis, and reporting. The results from this study suggest that measurements from the Dylos monitor, with added RH measurements, are correlated with multiple reference instruments and can be used for a networked community monitoring system to augment the existing regulatory network in Imperial County and provide higher spatial and temporal data, particularly for susceptible populations.

Although we are very clear that our community air monitoring data are nonregulatory, the interest that we received to collaborate to evaluate new instruments from regulatory agencies, like CARB, and other academic researchers suggests a desire to understand the performance limitations and possible use cases of new monitoring approaches. Ultimately, we hope that other communities and air quality stakeholders will increasingly embrace community-led efforts to augment regulatory monitoring.

References for Chapter 2

Anderson, J., J. Thundiyil, and A. Stolbach. 2011. Clearing the air: A review of the effects of particulate matter air pollution on human health. *J. Med. Toxicol.* 8:166–75. doi: 10.1007/s13181-011-0203-1.

Atkinson, R.W., B.K. Butland, C. Dimitroulopoulou, M.R. Heal, J.R. Stedman, N. Carslaw, D. Jarvis, C. Heaviside, S. Vardoulakis, H. Walton, and H.R. Anderson. 2016. Longterm exposure to ambient ozone and mortality: A quantitative systematic review and meta-analysis of evidence from cohort studies. *BMJ Open* 6:e009493. doi: 10.1136/ bmjopen-2015-009493.

Brook, R. D., D. E. Newby, and S. Rajagopalan. 2017. The global threat of outdoor ambient air pollution to cardiovascular health: Time for intervention. *JAMA Cardiol.* 2:353. doi: 10.1001/jamacardio.2017.0032.

California Air Resources Board. 2015. Chronology of state PM10 designations. http://www.arb.ca.gov/desig/changes/ pm10.pdf (accessed January 17, 2017).

California Air Resources Board. 2016. Annual Network Plan. https://www.arb.ca.gov/aqd/amnr/amnr2016.pdf (accessed January 17, 2017).

California Department of Health Services. 2014. Asthma ED Visits, Children, 2014. http://www.californiabreathing.org/ asthma-data/county-comparisons/edvisits-children (accessed February 12, 2017).

English, P., L. Olmedo, E. Bejarano, H. Lugo, E. Murillo, E. Seto, M. Wong, G. King, A. Wilkie, D. Meltzer, G. Carvlin, M. Jerrett, and A. Northcross. 2017. The Imperial County community air monitoring network: A model for community- based environmental monitoring for public health action. *Enviro. Health Perspect.* 125 (7): 074501-1–074501-5. doi: EHP1772/EHP1772s

Imperial County Air Pollution Control District. 2014. Imperial County 2013 State Implementation Plan for the 2006 24-hour PM2.5 Moderate Nonattainment Area. http:// www.arb.ca.gov/planning/sip/planarea/imperial/Final_ PM2.5_SIP_%28Dec_2,_2014%29_Approved.pdf (accessed January 24, 2017).

Jiao, W., G. Hagler, R. Williams, R. Sharpe, R. Brown, D. Garver, R. Judge, M. Caudill, J. Rickard, M. Davis, L. Weinstock, S. Zimmer-Dauphinee, and K. Buckley. 2016. Community Air Sensor Network (CAIRSENSE) project: Evaluation of low-cost sensor performance in a suburban environment in the southeastern United States. *Atmos. Meas. Tech.* 9:5281–92. doi: 10.5194/amt-9- 5281-2016.

Khreis, H., C. Kelly, J. Tate, R. Parslow, K. Lucas, and M. Nieuwenhuijsen. 2017. Exposure to traffic-related air pollution and risk of development of childhood asthma: A systematic review and meta-analysis. *Environ. Int.* 100:1– 31. doi: 10.1016/j.envint.2016.11.012.

Manikonda, A., N. Zikova, P. Hopke, and A. Ferro. 2016. Laboratory assessment of low-cost PM monitors. *J. Aerosol Sci.* 102:29–40. doi: 10.1016/j.jaerosci.2016.08.010.

Northcross, A., R. Edwards, M. Johnson, Z. Wang, K. Zhu, T. Allen, and K. Smith. 2013. A low-cost particle counter as a realtime fine-particle mass monitor. *Environ. Sci. Processes Impacts* 15:433–9. doi: 10.1039/C2EM30568B.

R Development Core Team. 2017 April. *R: A Language and Environment for Statistical Computing v. 3.4.0.* Vienna, Austria: R Foundation for Statistical Computing.

Richards, L., Alcorn, S., McDade, C., Couture, T., Lowenthal, D., Chow, J., Watson, J. (1999). Optical properties of the San Joaquin Valley aerosol collected during the 1995 integrated monitoring study. *Atmospheric Environment, 33*, 4787-4795.

Schwartz, J. 2004. Air pollution and children's health. *Pediatrics* 113:1037–43.

South Coast Air Quality Management District. 2015. Field Evaluation Dylos—DC1100 pro. http://www.aqmd.gov/docs/ default-source/aq-spec/field-evaluations/dylos-dc1100—fieldevaluation. pdf?sfvrsn=2 (accessed February 23, 2017).

South Coast Air Quality Management District. 2017. Laboratory Evaluation Dylos—DC1700 PM sensor. http://www.aqmd. gov/docs/default-source/aq-spec/laboratory-evaluations/dylos —lab-evaluation.pdf?sfvrsn=2 (accessed March 20, 2017).

Steinle, S., S. Reis, C. Sabel, S. Semple, M. Twigg, C. Braban, S. Leeson, M. Heal, D. Harrison, C. Lin, and H. Wu. 2015. Personal exposure monitoring of PM2.5 in indoor and outdoor microenvironments. *Sci. Total Environ.* 508:383– 94. doi: 10.1016/j.scitotenv.2014.12.003.

U.S. Code of Federal Regulations. 2016. Protection of the Environment. Title 40. Appendix A to part 58, 4.2.1: Collocated quality control sampler precision estimate for PM10, PM2.5 and Pb; 4.2.5: Performance evaluation programs bias estimate for PM2.5. https://www.ecfr.gov/ cgi-bin/text-idx?SID=cc05bf48bbd2cfd0a9fd4d b1710e92c6&mc=true&node=ap40.6.58_161.a&rgn=div9 (accessed March 18, 2017).

Williams, R., A. Kaufman, T. Hanley, J. Rice, and S. Garvey. 2014. *Evaluation of Field-Deployed Low Cost PM Sensors.* EPA/600/R-14/464 (NTIS PB 2015-102104). Washington, DC: U.S. Environmental Protection Agency.

Winkler, P. 1973. The growth of atmospheric aerosol particles as a function of the relative humidity—II. An improved concept of mixed nuclei. *J. Aerosol Sci.* 4:373–87. doi: 10.1016/0021-8502(73)90027-X.

Wu, C., Delfino, R., Floro, J., Samimi, B., Quintana, P., Kleinman, M., Liu, L. (2005). Evaluation and quality control of personal nephelometers in indoor, outdoor and personal environments. *Journal of Exposure Analysis and Environmental Epidemiology, 15*, 99-110.Chapter 2: Figures and tables

Table 1. Instrument comparison

Comparisons[a]	**$PM_{2.5}$ Adjusted R^2 (n)**	**PM_{10} Adjusted R^2 (n)**
Dylos ~ FRM Filter (day avg)	0.79 (300)	0.77 (32)
Dylos ~ FEM BAM (day avg)	0.84 (160)	0.87 (168)
Dylos ~ FEM BAM (hour avg)	0.78 (3907)	0.77 (4008)
Dylos ~ FEM BAM + RH (hour avg)	0.81 (3907)	0.80 (4008)
Dylos ~ FEM BAM + Temp (hour avg)	0.81 (3907)	0.81 (4008)
Dylos ~ FEM BAM + RH +Temp (hour avg)	0.82 (3907)	0.81 (4008)

[a] $y \sim x$ is used to signify a linear model with $y = \beta_0 + \beta_1 * x$.

Table 2. Summary statistics for hourly average paired data from E-BAM collocation sites

Site	**Dylos (**μg/m^3)[a]	**E-BAM (**μg/m^3)[a]	**n**
Seeley	10.4 (-3.83 to 191, 9.30)	9.01 (-3 to 261, 11.6)	3140
Kennedy	4.76 (-2.86 to 50.2, 4.54)	6.47 (-3 to 66, 6.99)	529
Westmorland	8.53 (-0.26 to 104, 7.17)	10.1 (-3 to 119, 10.4)	842
Meadows	7.31 (-3.10 to 301, 13.9)	9.55 (-3 to 514, 20.3)	1951
Calipatria	7.61 (-5.83 to 174, 12.2)	11.6 (-3 to 239, 15.7)	1688
Calexico Alvarez	13.6 (-1.67 to 92.8, 11.2)	14.1 (-3 to 82, 10.9)	1406

[a]mean (min to max, standard deviation)

Table 3. Precision and bias statistics for daily paired data from the collocation sites

Site	Sampler Pair	Number of Pairs	Collocated Precision CV(%)	Bias (%)
Seeley	E-BAM-Dylos	124	24.4	28.3
Kennedy	E-BAM-Dylos	13	35.2	-6.8
Westmorland	E-BAM-Dylos	35	23.5	-7.0
Meadows	E-BAM-Dylos	67	24.7	-14.2
Calipatria	E-BAM-Dylos	64	17.5	-31.4
Calexico-Alverez	E-BAM-Dylos	57	25.7	3.0
Average			25.2	-4.7
Calexico-Ethel	FRM-FRM 2015	26	22.8	-
Calexico-Ethel	FRM-FRM 2016	28	8.2	-

Table 4. Correlation between E-BAM and Dylos $PM_{2.5}$ mass concentration using the Calexico-Ethel conversion and a site-specific conversion

Site	**n**	**R^2 (Calexico-Ethel Model)**	**R^2 (Site-Specific Model)**
Seeley	3140	0.63	0.64
Kennedy	529	0.37	0.38
Westmorland	842	0.62	0.67
Meadows	1951	0.75	0.75
Calipatria	1688	0.81	0.83
Calexico Alvarez	1406	0.35	0.36

Figure 1. (a) Dylos particle counter inside NEMA enclosure. (b) Collocation with regulatory monitor.

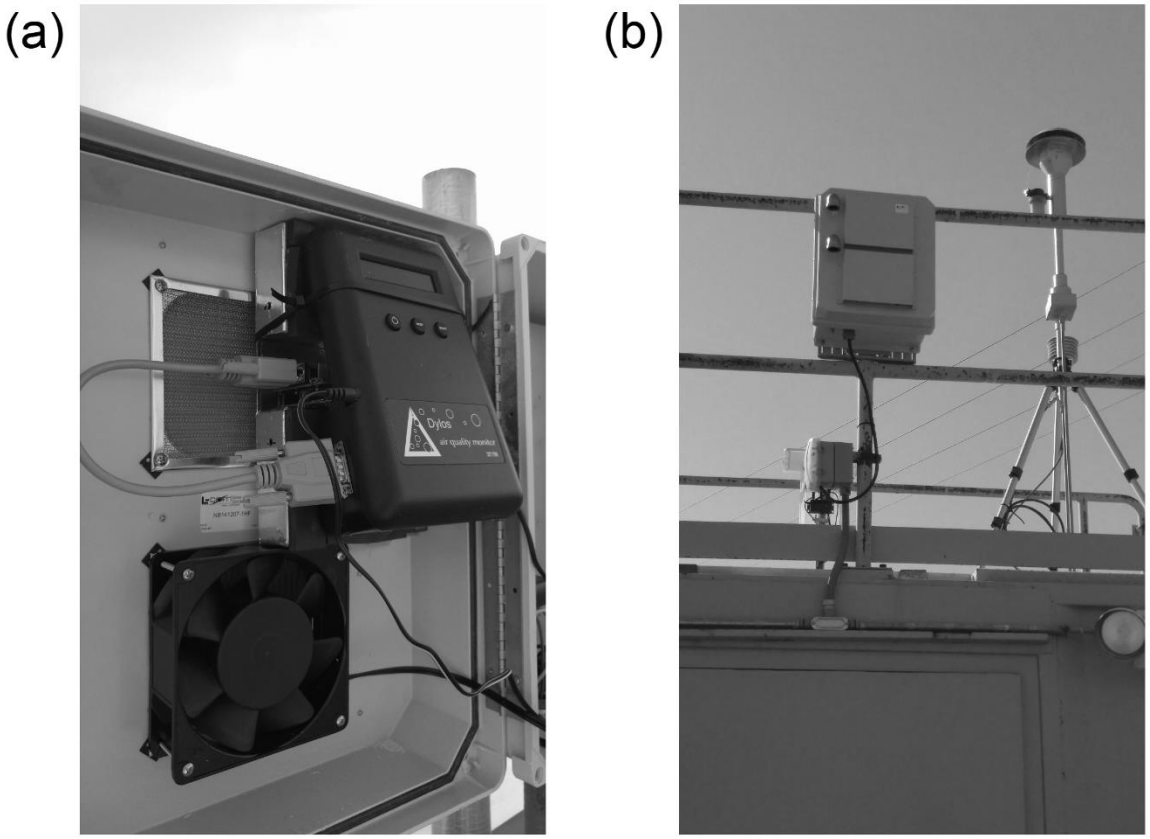


Figure 2. Collocations with E-BAMs were at: (1) Calipatria; (2) Westmorland; (3) Seeley; (4) El Centro Kennedy; (5) El Centro Meadows; (6) Calexico Alvarez. Collocation with FEM BAM and FRM gravimetric samples was at (7) Calexico-Ethel.

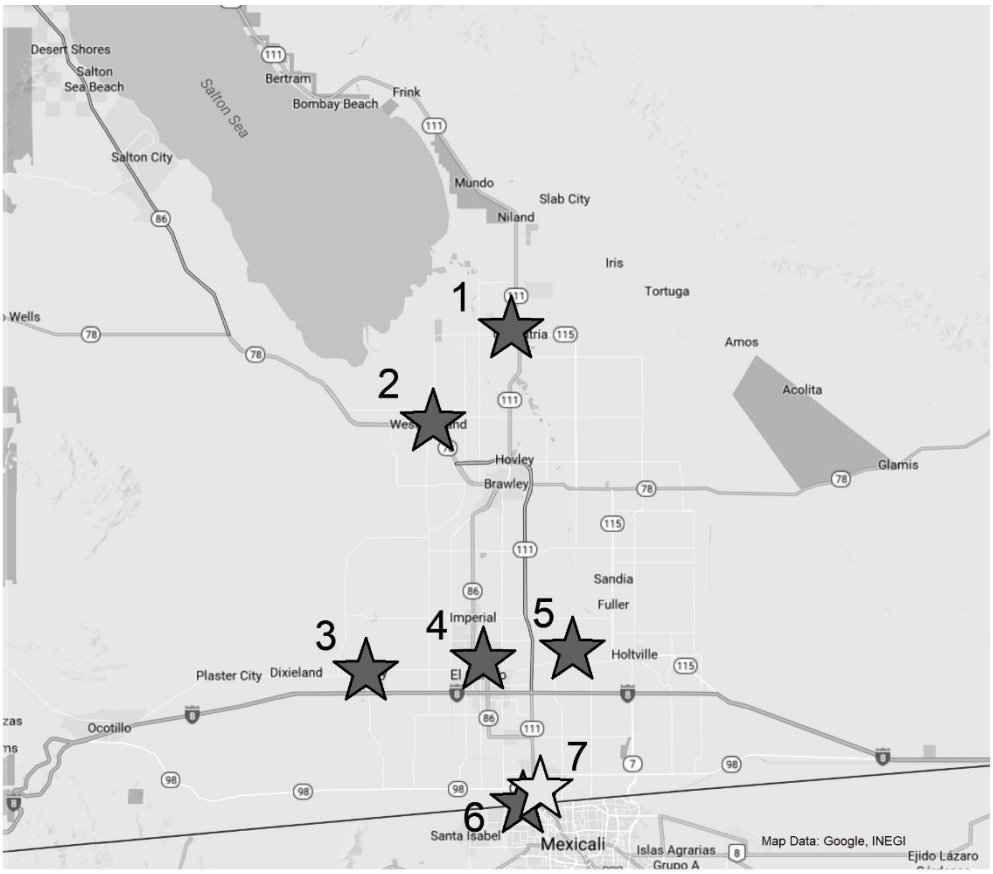


Figure 3. Comparison of Dylos-derived mass concentrations after count to mass conversion and BAM for (a) daily $PM_{2.5}$ and (b) daily PM_{10} with R^2. Dashed line indicates the 1:1 relationship.

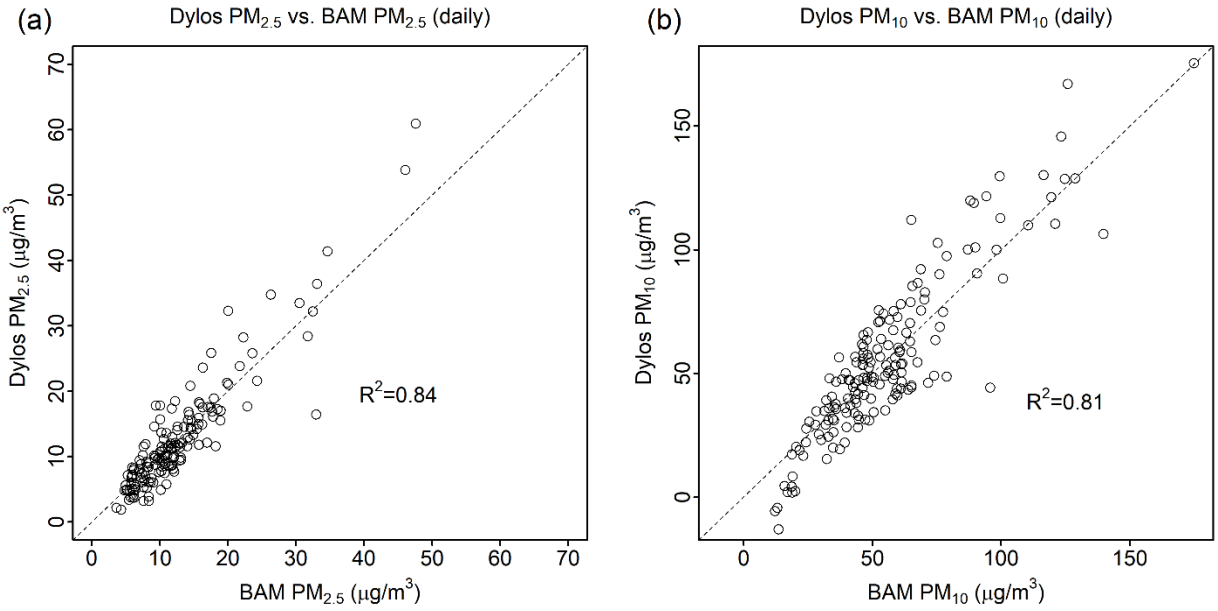
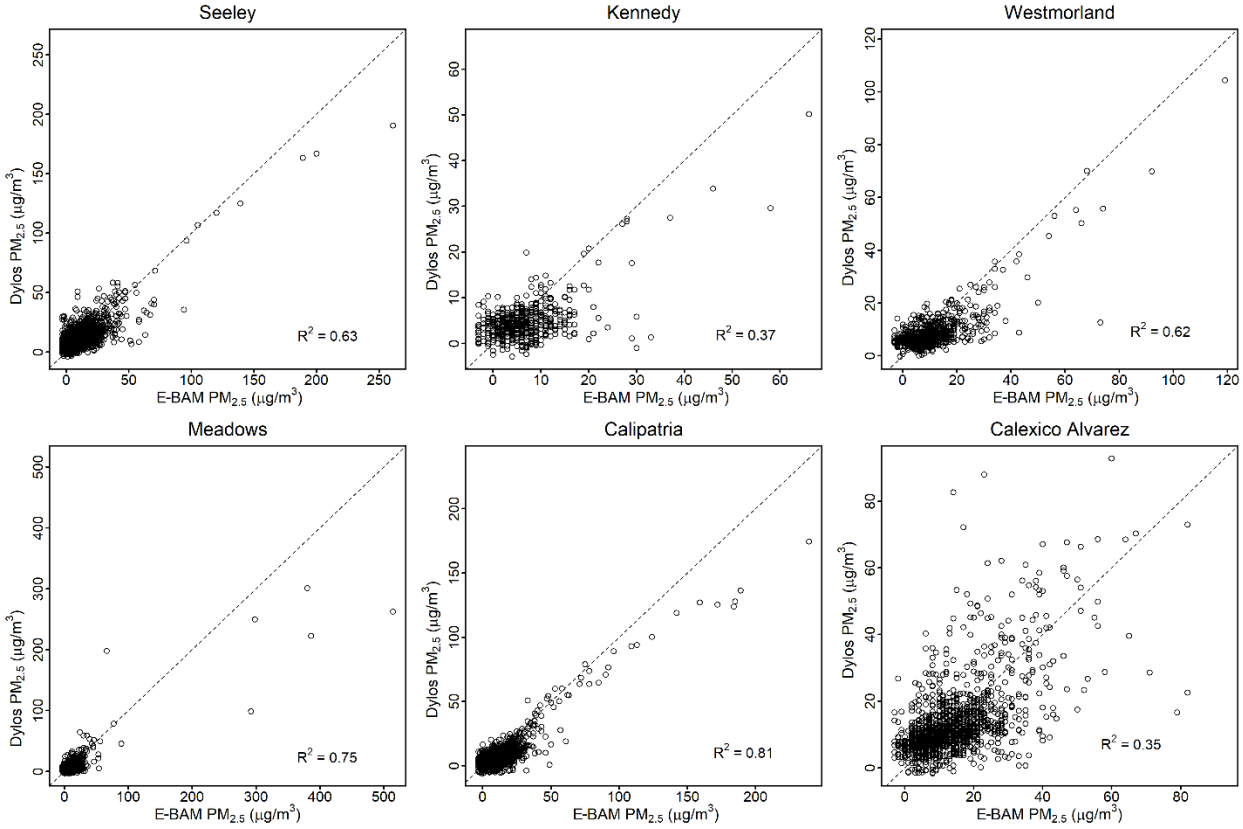


Figure 4. Scatterplots of E-BAM and Dylos conversion with R^2. Dashed line indicates the 1:1 relationship.



Chapter 3. Land Use Regression for $PM_{2.5}$ and PM_{coarse} in Imperial County, CA using Data from the Community Air Monitoring Network

Chapter 3.1 Introduction

Both $PM_{2.5}$ (particulate matter less than 2.5 μm in diameter) and PM_{10} (particulate matter less than 10 μm in diameter) have been linked to adverse health outcomes such as respiratory and cardiac disease, asthma, and increased mortality (Anderson et al., 2011; Brook et al., 2017; Khreis et al., 2017; Atkinson et al., 2016). PM_{coarse} is the difference between PM_{10} and $PM_{2.5}$ and as such represents a different size fraction than $PM_{2.5}$, whereas PM_{10} overlaps with $PM_{2.5}$. Studying PM_{coarse} rather than PM_{10} is therefore useful for understanding sources specifically associated with larger particle sizes.

PM differs spatially based on proximity to sources and transport. However, the ability to collect highly spatially resolved PM measurements is often limited by the number of monitors. This has led to the use of modeling to estimate PM concentrations at locations and, for temporally varying data, at times when measurements are not available. One technique used to estimate PM exposure for a population is called land use regression (LUR). LUR involves modeling PM using a combination of land use variables such as traffic, area of land types, and distance to known sources. LUR models produce a continuous surface across the study area and can provide accurate PM exposure estimates (Eftens et al., 2012; Hoek et al., 2008). Eeftens et al. (2012) used LUR modeling to predict $PM_{2.5}$ and PM_{coarse} across 20 European cities, using the resulting predicted pollution surfaces to estimate PM exposure for a cohort involved in the ESCAPE study. Hoek et al. (2008) reviewed 25 studies that used LUR models to predict air pollution levels including PM. Most studies employed a series of 2-week sampling periods and estimated annual averages. The authors suggest that continuous sampling could help inform shorter timescale models.

In this paper, a land use regression model is developed for Imperial County, CA, a rural agricultural community in Southeast California with a population of around 175,000 (U.S. Census Bureau, 2010). Imperial County has the second highest rate of childhood asthma emergency room visits in the state (California Department of Health Services, 2014). The primary sources of air pollution are wind-borne dust; agricultural burning; mobile emissions from cars, trucks, and heavy-duty vehicles; and transport from Mexicali, Mexico, a city adjacent to the US-Mexico border (Imperial County Air Pollution Control District, 2014; California Air Resources Board, 2015). The land use regression model uses data from the Imperial County Community Air Monitoring Network ("the network"). The network was created through a collaboration between a local community group, Comite Civico del Valle (CCV), the California Environmental Health Tracking Program (CEHTP) of the Public Health Institute, and academics at UCLA and the University of Washington (English et al., 2017). The network consists of 40 custom air quality monitors with Dylos particle counters, which measure both $PM_{2.5}$ and PM_{10}, and a relative humidity and temperature sensor. Carvlin et al. (2017) showed an R^2 between the

community monitor and regulatory monitors of 0.79 (hourly) and 0.84 (daily) for $PM_{2.5}$ and 0.78 (hourly) and 0.81 (daily) for PM_{10}.

The LUR model described herein used a combination of land use variables and meteorology. Leave-one-out cross-validated root mean square error (RMSE) was used to compare between four different model types. After the model was selected, variable importance statistics were calculated. The goal of this LUR model was to create a 2D air pollution surface at the monthly scale for both $PM_{2.5}$ and PM_{coarse}. The predictions were compared to PM measurements made by regulatory models as an independent test of the LUR model.

Chapter 3.2 Materials and Methods

Chapter 3.2.1 Model Building

Community monitoring $PM_{2.5}$, PM_{10}, relative humidity, and temperature data from 35 sites for a 12-month period, 10/1/2016 to 10/1/2017 (month/day/year), were used in the following analyses. The PM data were converted from particle counts to particle mass as detailed in Carvlin et al. (2017). However, the conversion equation was updated using data from the Calexico-Ethel site for the 12-month study period. The previous PM_{10} conversion equation used data from 1/15/2016 to 7/12/2016. The conversion constants used in this paper are PM_{10}: $c_1 = 9.35$, $c_2 = 0.216$, $c_3 = -0.344$ and $PM_{2.5}$: $c_1 = 5.41$, $c_2 = 0.00831$, $c_3 = -0.224$. PM_{coarse} was calculated as $PM_{10} - PM_{2.5}$.

As a part of the conversion process data were run through an automated quality control (QC) process; hours with less than 75% of data and data with particle counts less than 30 in Dylos bin 1 were discarded. After the automatic QC, a manual QC process was performed to identify time periods when the monitor response was slowly attenuated due to incremental dust build-up on the photodiode and when the monitor readings were oscillating rapidly between high and low, which resulted in a further 1.2% of data being dropped. The conversion equation can produce negative numbers, which were used as is in the following analyses unless otherwise noted. Hourly $PM_{2.5}$ and PM_{coarse} data were averaged to monthly data using a 50% data completeness cutoff. This left a total of 207 monthly data points across 33 monitors. Each monitor had 6 months of data on average, however some monitors had only a few months. The monitors that had the least amount of data were those near the Salton Sea. These monitors have poor cell reception and are subject to harsher environmental conditions, which leads to lower data completeness.

Regulatory data were downloaded from the California Air Resources Board (CARB) website. The regulatory network consists of five sites located near population centers in Imperial Valley that have Met One 1020 PM_{10} beta attenuation monitors (BAMs) (California Air Resources Board, 2017). Two of these sites also have Met One 1020 $PM_{2.5}$ BAMs. There are also five sites located around the Salton Sea that are operated by the Imperial Irrigation District that have $PM_{2.5}$ and PM_{10} Thermo Fischer Scientific Series 1405-D tapered element oscillating microbalances (TEOMs) (Imperial County Air Pollution Control District, 2010). These sites

were set up to monitor emissions from the Salton Sea as it recedes due to changes in water rights that have reduced the agricultural runoff that has kept the sea from evaporating. Only QC screened data were used. Values greater than 985 $\mu g/m^3$ for BAMs were excluded since this is above the range of the instrument (p. 7 and 41, Met One, 2008). No upper cutoff was used for TEOM measurements since all values were within the instrument range (p. 1-4, ThermoFisher Scientific, 2008). All negative values from regulatory instruments were kept as is, and were included in the analysis.

Land use variables and community monitor locations were loaded into ArcGIS (ESRI, v. 10.3.1). 250m, 500m, and 1000m buffers were created around each monitor. Land use parameters were sampled within each of these buffers. GIS and meteorological variables are listed in Table 5 along with their source, date, buffers, and averaging period. All data manipulation and analyses were performed using R statistical software (v. 3.3.3).

Agricultural burning records were received from the Imperial Air Pollution Control District. Acres burned was recorded on the daily level. This information was added to the model as acres burned within 5 km of a monitoring site within the last day. When multiple burns were recorded within 5 km of a site they were summed.

Other GIS variables were considered but rejected since all monitors had the same value or nearly the same value for that variable. Dropped variables included indicators of industrial PM emissions since none of our monitors were located near industrial sites that had permits to release PM. Satellite $PM_{2.5}$ was included, but was not predictive, perhaps because the data were 15 years old and satellite measurements are known to not perform well in desert areas.

Meteorlogical data completeness was less than ideal, especially for planetary boundary layer height. Because the models require a complete dataset hours that did not have meteorlogical data were dropped. This resulted in a limited amount of complete hours for October and November 2017 and therefore they are not included in the monthly and yearly PM maps.

Some monitors have buffers that cross the US-Mexico border. However, we had no land use data for Mexico. If only the US side of the land use was used, then the true value of that land use would be underestimated. To adjust for this, the percent area of the buffer within Mexico was recorded for each site. Then the variables were multiplied by 100/(100 - % area), which gives the land use for the whole buffer assuming the same distribution in Mexico as in the US.

Land use variables were converted from continuous to categorical or binary. This was done because the monitors do not cover the range of land use seen throughout the valley; in particular the range of land use sampled by the grid of points used for prediction, the fishnet. Therefore, if linear extrapolation were used then the fishnet predictions failed, becoming extremely small or large. Histograms of each variable were analyzed to determine whether the variable should be converted to binary or categorical. All categorical variables were given three categories. The cut point for the binary variables was the first quartile. Most of these variables had nearly all of the measurements around zero and just a few at much higher values. The cut points for the categorical variables were the first quartile, the median, and the third quartile.

After conversion from continuous to categorical and binary the fishnet predictions were much closer to the range of the monitoring measurements. However, the choice of cut points is dependent on the data and therefore limits the more general application of the models developed in this paper.

Three alternative models were considered. Categorical variables were converted to binary variables for use in models which cannot process categories. The models were a Bayesian additive regression trees (BART) model, a lasso model, and a partial least squares model (PLS). PLS is a modeling technique that reprojects the data in order to find the dimension in the input variable space that explains the most variance in the outcome. PLS has been used in PM modeling, in particular when there are a large number of variables (Sampson et al., 2013). Lasso is a penalized least squares method that reduces the number of variables in the model based on an alpha parameter. This parameter is chosen based on cross-validated testing. Mercer et al. (2011) and Knibbs et al. (2014) used lasso to help reduce the number of variables in PM LUR models. BART is a model sums individual regression trees using a Bayesian approach (Chipman, George, & McCulloch, 2010). It has been used to predict torrential rain and avalanches, and to relate vehicle trip duration to household characteristics (Wu, Huang, & Pan, 2010; Blattenberger & Fowles, p. 211-227, 2014; Chipman et al., 2010).

Models were compared by leaving out one site at a time and calculating the RMSE at that site using the rest of the sites. For each model, the RMSE was averaged across all sites. BART was found to be the best performing model for $PM_{2.5}$ and PM_{coarse}. A variable selection test for these models was performed to identify which variables had the most impact on the model. The variable selection for the BART models was done by dropping one variable at a time from the model and calculating the test R^2. The variables that led to the largest decrease in R^2 were those that had the most impact on the model. The 10 most important variables for the $PM_{2.5}$ and PM_{coarse} models were compared. In order to compare variable selection stability across models the top 10 BART and lasso variables were compared. Lasso variables were selected based on their standardized coefficient values, corrected by their standard deviations.

Chapter 3.2.2 Prediction

The same steps described in the Model Building section were followed for a grid of points, the fishnet, equally spaced 1 mile apart across all of Imperial County; i.e. sampling GIS parameters within buffers at each grid point, finding the nearest meteorological data, etc.

The BART models were used to predict monthly $PM_{2.5}$ and PM_{coarse} concentrations on the fishnet points. The residuals of the $PM_{2.5}$ and PM_{coarse} BART models were kriged and added to the model predictions before mapping. This helped the model account for purely spatial variability. These combined PM levels, at both the monthly and study-average timescale, were kriged and a 2D surface was created and compared to kriged PM measurements.

The fishnet predictions were then compared to regulatory PM concentrations. The fishnet point closest to the regulatory site was chosen for comparison. The longest distance

between a fishnet point and a regulatory site was <1km. Model predictions and regulatory measurements were compared using R^2 and RMSE on the monthly and yearly timescales.

Chapter 3.3 Results and Discussion

Table 6 presents summary statistics for the community monitors averaged over the 40 sites in the network compared to the regulatory monitors. Regulatory sites are located to provide coverage of populated areas in the Imperial Valley as well as to monitor the air quality around the Salton Sea and consist of a mix of BAM and TEOM instruments (California Air Resources Board, 2017). High hourly $PM_{2.5}$ and very high hourly PM_{10} and PM_{coarse} values were seen by both the community monitors and the regulatory monitors. The community monitors produced a similar county average to the regulatory monitors for $PM_{2.5}$. However, the community monitors measured slightly lower PM_{10} and PM_{coarse}.

Table 7 shows the result of the modeling process for $PM_{2.5}$. RMSE and R^2 are averaged across all of the leave-one-site-out variants. The lasso model chose 47 variables and the PLS model chose 24 variables. In terms of RMSE, Lasso and PLS performed similarly and BART performed the best. The average RMSE between model predictions and the test set ranged from 1.50 to 1.79 $\mu g/m^3$. Model R^2 values did not follow the same order as model RMSE. PLS had the highest R^2 (0.54) followed by PLS and BART (0.54 and 0.47 respectively). RMSE was used to select the top performing model since lower prediction error was deemed more important than linearity. Also, there were a variable number of months of data for each site and low sample size can dramatically change R^2, especially when the range of the data is small. BART was chosen as the top performing model based on RMSE.

The results of the modeling process for PM_{coarse} are presented in Table 8. Lasso chose 40 variables and PLS chose 7 variables. Lasso and PLS performed similarly and BART performed the best. The average RMSE for PM_{coarse} models ranged from 9.14 to 12.30 $\mu g/m^3$. Model R^2 values followed the order of model RMSE for PM_{coarse}. The R^2 value for the PM_{coarse} BART model was 0.65.

Looking to the literature, PM LUR models have shown variable performance depending on model location and variable selection. The $PM_{2.5}$ LUR models developed in Eeftens et al. (2012) had validation R^2 values, comparing the model to validation data, of 0.21 to 0.78 with a mean of 0.59. Hoek et al. (2008) compared $PM_{2.5}$ LUR models from eight different studies and found model R^2 values, comparison of model performance without external validation data, of 0.17 to 0.73 with a mean of 0.52 and RMSE values of 1.1 to 2.3 $\mu g/m^3$ with a mean of 1.5 $\mu g/m^3$. All of these models produced annual averages of $PM_{2.5}$ and therefore are not directly comparable to the monthly models described in this paper. That said, the R^2 and RMSE values for the $PM_{2.5}$ models presented in this paper are within the range of the annual models found in the literature.

Eeftens et al. (2012) created PM_{coarse} LUR models for 20 European cities as a part of the ESCAPE project. The range of validation R^2 seen across the 20 models was 0.03 to 0.73, with a median R^2 of 0.57. Wolf et al. (2017) used the ESCAPE data to create finer spatial resolution

PM_{coarse} LUR models for Germany, which had a validation R^2 of 0.49. Eeftens et al. (2016) modeled PM_{coarse} using a LUR model for eight areas in Switzerland. The average validation R^2 was 0.38. These studies were also annual averages. In general, PM_{coarse} LUR models seem to have similar or slightly lower performance compared to $PM_{2.5}$ LUR models. The PM_{coarse} models presented in this paper have R^2 values in the range of the annual models found in the literature.

Table 9 shows the results from the variable importance test for the $PM_{2.5}$ and PM_{coarse} models. The top variable for the $PM_{2.5}$ model is relative humidity and the top variable for the PM_{coarse} model is temperature. These variables vary monthly and are likely accounting for seasonal patterns. It should be noted the relative humidity may have a confounding effect since it was used to correct the Dylos measurements and was used in the model. However, the relative humidity data used in the model was a monthly average. The $PM_{2.5}$ model places more emphasis on meteorological variables, such as wind direction, wind speed, and planetary boundary layer height, and location while the PM_{coarse} model places more emphasis on land use variables for farmland, native vegetation, and roads. The California Department of Water Resources defined native vegetation as all desert land and non-riparian land near bodies of water.

For $PM_{2.5}$, the lasso model chose land use variables that represented nearly all categories of land use as well as a variable representing railroad length, whereas the BART model put more emphasis on meteorological variables. For PM_{coarse}, the lasso model and BART models chose similar variables with more emphasis on land use in the lasso model and more emphasis on transportation in the BART model. The PM_{coarse} lasso model also did not have any meteorological variables.

Figure 5 shows the kriged study-average residuals from the $PM_{2.5}$ and PM_{coarse} models. The $PM_{2.5}$ model underpredicted PM (-2.0 to -1.4 $\mu g/m^3$) in the area the west of the Salton Sea and overpredicted PM (0.3 to 1.3 $\mu g/m^3$) in the area east of the Salton Sea. The PM_{coarse} model underpredicted PM (-19.4 to -12.6 $\mu g/m^3$) in the area west of the Salton Sea. The PM_{coarse} model overpredicted PM (1.2 to 4.0 $\mu g/m^3$) in area east of the Salton Sea and in the south of Imperial Valley. Moran's I was used to see if there was spatial correlation in the residuals by month. $PM_{2.5}$ had two months where $p < 0.05$, May and August 2017, and PM_{coarse} had three months with $p < 0.05$; January, July, and August 2017.

The average measured $PM_{2.5}$ over the study period is shown in Figure 6a and the average predicted $PM_{2.5}$ in Figure 6b. The predicted $PM_{2.5}$ map is a combination of the $PM_{2.5}$ model and the kriged $PM_{2.5}$ residuals. The spatial structure of the predicted $PM_{2.5}$ map matches that of the measured $PM_{2.5}$ map. The highest average $PM_{2.5}$ levels were near the US-Mexico border and to the west of the Salton Sea near Salton City. This does not mean that the highest $PM_{2.5}$ peaks were at these locations, only that the model predicted higher average $PM_{2.5}$. Low $PM_{2.5}$ regions appeared primarily to the east of the Salton Sea. The predicted $PM_{2.5}$ map is biased low compared to the measured $PM_{2.5}$ map.

Figure 7a shows the average measured PM_{coarse} and Figure 7b shows the average predicted PM_{coarse} over the study period. The predicted PM_{coarse} map is a combination of the

PM_{coarse} model and the kriged PM_{coarse} residuals. The spatial structure of the PM_{coarse} map differs slightly from the measure PM_{coarse} map in that the predictions are lower to the southeast of the Salton Sea and higher in the southeast portion of the Imperial Valley. The highest PM_{coarse} to the west of the Salton Sea and the area east of the Salton Sea had low PM_{coarse} concentrations. The predicted PM_{coarse} map is biased low compared to the measure PM_{coarse} map.

Maps of measured and predicted monthly $PM_{2.5}$ are displayed in Figure 8. Only monitors that had data for that month were used in each monthly map. There was not enough data to create maps for October and November 2016. The predicted $PM_{2.5}$ map is a sum of the $PM_{2.5}$ model and the kriged $PM_{2.5}$ residuals. The monthly $PM_{2.5}$ predictions are spatially similar to the monthly $PM_{2.5}$ measurements, however the predictions are biased low. The highest $PM_{2.5}$ concentrations were seen in December 2016 and April through July 2017. $PM_{2.5}$ was high near the US-Mexico border throughout the year. $PM_{2.5}$ was high in the center of Imperial Valley near the city of Brawley in May 2017.

Figure 9 shows maps of monthly measured and predicted PM_{coarse}. The predicted PM_{coarse} map is a sum of the PM_{coarse} model and the kriged PM_{coarse} residuals. There was good spatial agreement between the PM_{coarse} measurements and predictions, however the predictions were biased low. The highest PM_{coarse} was seen around Brawley in May 2017 and in the southeastern part of Imperial Valley in September 2017.

Table 10 shows summary statistics comparing the fishnet predictions and data from the regulatory monitoring network for monthly and yearly averages. The yearly RMSE was 3.2 $\mu g/m^3$ for $PM_{2.5}$ and 17.7 $\mu g/m^3$ for PM_{coarse}.

Chapter 3.4 Conclusion

This paper helps to elucidate the effect of meteorological and land use parameters on particulate matter in Imperial County. The variable selection process pointed to seasonality, as measured by relative humidity; meteorology; and length of roads as prime drivers in monthly $PM_{2.5}$. Important variables for monthly PM_{coarse} were seasonality, based on temperature; land use variables for farmland; and length of roads. The predicted PM_{coarse} map showed areas of high PM_{coarse} around Brawley and to the west of the Salton Sea. The predicted $PM_{2.5}$ map showed high levels in a large area near the US-Mexico border and to the west of the Salton Sea. This may point to agriculture as a potential source of PM_{coarse}, cross-border transport as a potential source of $PM_{2.5}$, and windblown dust as a source of both sizes of PM.

References for Chapter 3

Anderson, J., Thundiyil, J., & Stolbach, A. 2011. Clearing the air: A Review of the Effects of Particulate Matter Air Pollution on Human Health. *Journal of Medical Toxicology, 8*(2), 166-175.

Atkinson R.W., Butland B.K., Dimitroulopoulou C., Heal M.R., Stedman J.R., Carslaw N., Jarvis D., Heaviside C., Vardoulakis S., Walton H., & Anderson H.R. 2016. Long-term Exposure to Ambient Ozone and Mortality: a Quantitative Systematic Review and Meta-analysis of Evidence from Cohort Studies. *BMJ Open, 6*(2).

Blattenberger G. & Fowles R. 2014. Avalanche Forecasting: Using Bayesian Additive Regression Trees (BART). In: Alleman J., Ní-Shúilleabháin Á., Rappoport P. (eds) Demand for Communications Services – Insights and Perspectives. The Economics of Information, Communication, and Entertainment (The Impacts of Digital Technology in the 21st Century). Springer, Boston, MA.

Brook R.D., Newby D.E., & Rajagopalan S. 2017. The Global Threat of Outdoor Ambient Air Pollution to Cardiovascular Health: Time for Intervention. *JAMA Cardiol.*

California Air Resources Board, 2015. Chronology of State PM_{10} Designations. http://www.arb.ca.gov/desig/changes/pm10.pdf (accessed October 20, 2017).

California Air Resources Board, 2017. Annual Network Plan. https://www.arb.ca.gov/aqd/amnr/amnr2017.pdf (accessed November 5, 2017).

California Department of Health Services, 2014. Asthma ED Visits, Children, 2014. http://www.californiabreathing.org/asthma-data/county-comparisons/edvisits-children (accessed October 20, 2017).

Chipman, H. A., George, E. I., Lemp, J., & McCulloch, R. E. 2010. Bayesian Flexible Modeling of Trip Durations. *Transportation Research Part B: Methodological, 44*(5), 686-698. Doi: 10.1016/j.trb.2010.01.007

Chipman, H. A., George, E. I., & McCulloch, R. E. 2010. BART: Bayesian Additive Regression Trees. *The Annals of Applied Statistics, 4*(1), 266-298. doi:10.1214/09-aoas285

Eeftens, M., Beelen, R., Hoogh, K. D., Bellander, T., Cesaroni, G., Cirach, M., . . . Hoek, G. 2012. Development of Land Use Regression Models for $PM_{2.5}$, $PM_{2.5}$ Absorbance, PM_{10} and

PM_{coarse} in 20 European Study Areas; Results of the ESCAPE Project. *Environmental Science & Technology*, *46*(20), 11195-11205. doi:10.1021/es301948k

Eeftens, M., Meier, R., Schindler, C., Aguilera, I., Phuleria, H., Ineichen, A., Davey, M., Ducret-Stich, R., Keidel, D., Probst-Hensch, N., Künzli, N., & Tsai, M. 2016. Development of Land Use Regression Models for Nitrogen Dioxide, Ultrafine Particles, Lung Deposited Surface Area, and Four Other Markers of Particulate Matter Pollution in the Swiss SAPALDIA Regions. *Environmental Health, 15*(53), 1-14. doi: 10.1186/s12940-016-0137-0

English, P., Olmedo, L., Bejarano, E., Lugo, H., Murillo, E., Seto, E., Wong, M., King, G., Wilkie, A., Meltzer, D., Carvlin, G., Jerrett, M., & Northcross, A. 2017. The Imperial County Community Air Monitoring Network: A Model for Community-based Environmental Monitoring for Public Health Action. *Environmental Health Perspectives, 125*(7), 074501-1 to 074501-5.

Hoek, G., Beelen, R., Hoogh, K. D., Vienneau, D., Gulliver, J., Fischer, P., & Briggs, D. 2008. A Review of Land-use Regression Models to Assess Spatial Variation of Outdoor Air Pollution. *Atmospheric Environment, 42*(33), 7561-7578. doi:10.1016/j.atmosenv.2008.05.057

Imperial County Air Pollution Control District, 2010. Quality Assurance Quality Control Plan for the Salton Sea Air Monitoring Network. http://www.water.ca.gov/saltonsea/docs/Appendix_10_FINAL_QA-QC_PLAN.pdf (accessed November 19, 2017).

Imperial County Air Pollution Control District, 2014. Imperial County 2013 State Implementation Plan for the 2006 24-hour $PM_{2.5}$ Moderate Nonattainment Area. http://www.arb.ca.gov/planning/sip/planarea/imperial/Final_PM2.5_SIP_%28Dec_2,_2014%29_Approved.pdf (accessed October 20, 2017).

Khreis H., Kelly C., Tate J., Parslow R., Lucas K., & Nieuwenhuijsen M. 2017. Exposure to Traffic-related Air Pollution and Risk of Development of Childhood Asthma: A Systematic Review and Meta-analysis. *Environ Int., 100*, 1-31.

Knibbs, L. D., Hewson, M. G., Bechle, M. J., Marshall, J. D., & Barnett, A. G. 2014. A National Satellite-based Land-use Regression Model for Air Pollution Exposure Assessment in Australia. *Environmental Research, 135*, 204-211. doi:10.1016/j.envres.2014.09.011

Mercer, L. D., Szpiro, A. A., Sheppard, L., Lindström, J., Adar, S. D., Allen, R. W., . . . Kaufman, J. D. 2011. Comparing Universal Kriging and Land-use Regression for Predicting Concentrations of Gaseous Oxides of Nitrogen (NO_x) for the Multi-Ethnic Study of Atherosclerosis and Air Pollution (MESA Air). *Atmospheric Environment, 45*(26), 4412-4420. doi:10.1016/j.atmosenv.2011.05.043

Met One, 2008. BAM 1020 Particulate Monitor Operation Manual Rev G. https://arb.ca.gov/airwebmanual/instrument_manuals/Documents/BAM-1020-9800%20Manual%20Rev%20G.pdf (accessed November 5, 2017).

Sampson, P. D., Richards, M., Szpiro, A. A., Bergen, S., Sheppard, L., Larson, T. V., & Kaufman, J. D. 2013. A Regionalized National Universal Kriging Model using Partial Least Squares Regression for Estimating Annual $PM_{2.5}$ Concentrations in Epidemiology. *Atmospheric Environment, 75*, 383-392. doi:10.1016/j.atmosenv.2013.04.015

ThermoFisher Scientific, 2008. TEOM 1405 Ambient Particulare Monitor Revision A. https://tools.thermofisher.com/content/sfs/manuals/EPM-TEOM1405-Manual.pdf (accessed November 5, 2017).

U.S. Census Bureau, 2010. State and County QuickFacts, Imperial County, CA. Washington, D.C.: Government Printing Office. U.S. Census Bureau. Retrieved from http://quickfacts.census.gov/qfd/states/06/06025.html (accessed October 20, 2017).

Wolf, K., Cyrys, J., Harciníková, T., Gu, J., Kusch, T., Hampel, R., Schneider, A., & Peters, A. 2017. Land Use Regression Modeling of Ultrafine Particles, Ozone, Nitrogen Oxides and Markers of Particulate Matter Pollution in Augsburg, Germany. Science of The Total Environment, 579, 1531-1540. doi: 10.1016/j.scitotenv.2016.11.160

Wu, J., Huang, L., & Pan, X. 2010. A Novel Bayesian Additive Regression Trees Ensemble Model Based on Linear Regression and Nonlinear Regression for Torrential Rain Forecasting. *2010 Third International Joint Conference on Computational Science and Optimization (CSO)*, Huangshan, Anhui, 2010, p.466-470. doi: 10.1109/CSO.2010.15

Chapter 3: Figures and tables

Table 5. Model variables

Data	**Source**	**Date**	**Buffers (m)**	**Averaging Period**
latitude, longitude	calculated			
distance to the border, Salton Sea	calculated			
urban, types of farmland, types of crops, native and riparian vegetation (area)	Agricultural Land Use Survey (California Department of Water Resources), Farmland Mapping and Monitoring Program (California Department of Conservation)	1997, 2012	250, 500, 1000	
small roads, large roads, railroads (length)	TIGER/Line (US Census Bureau)	2013	250, 500, 1000	
traffic	California Department of Public Health	2013	250, 500, 1000	annual average of traffic counts
agricultural burning (acres)	Imperial Air Pollution Control District	study period	sum within 5km	daily
RH, temperature	community monitors	study period		hourly
wind direction, wind speed	nearest government monitor (CARB)	study period		hourly
planetary boundary layer height	RAP 130 (NOAA)	study period		hourly
Sattelite $PM_{2.5}$	Randall Martin, Dalhousie University	2002-2004		3-year

Table 6. Descriptive statistics for $PM_{2.5}$, PM_{10}, and PM_{coarse}

Statistic[a]	**$PM_{2.5}$ (µg/m^3)**		**PM_{10} (µg/m^3)**		**PM_{coarse} (µg/m^3)**	
	Dylos	**Govt**	**Dylos**	**Govt**	**Dylos**	**Govt**
Min	-6.5	-3.0	-7.7	0.9	-22.2	-5
Median	6.4	6.8	23.9	29.4	17.1	21.0
Mean	7.9	7.8	32.4	37.5	24.3	28.2
Max	52.3	149.4	1523.1	739.0	1109.5	780.0

[a]Matched hourly values averaged across all sites during the study period

Table 7. Model selection $PM_{2.5}$

Model	**R^2**	**RMSE**
PLS	0.54[a]	1.79
Lasso	0.51	1.70
BART	0.47	1.50

[a]The PLS model was not able to run for every site. Sites that could not run were dropped.

Table 8. Model selection PM_{coarse}

Model	**R^2**	**RMSE**
Lasso	0.55	11.75
PLS	0.59[a]	11.48
BART	0.65	8.07

[a]The PLS model was not able to run for every site. Sites that could not run were dropped.

Table 9. Variable importance

	$PM_{2.5}$		PM_{coarse}	
Rank	**Name**	**Percent Difference in R^2 (%)[a]**	**Name**	**Percent Difference in R^2 (%)**
1	Relative humidity	16.1	Temperature	14.5
2	Wind speed	5.7	Area of native vegetation, bottom tertile, 250m	8.4
3	Wind direction, east	5.5	Area of fallow farmland, bottom 50%, 1000m	6.8
4	Wind direction, northeast	4.4	Length of large roads, top tertile, 250m	6.6
5	Length of large roads, middle tertile, 1000m	3.8	Area of good farmland, top tertile, 250m	6.3
6	Area of urban land, top tertile, 500m	3.7	Area of prime farmland, middle tertile, 1000m	6.2
7	Area of grain/hay crops, bottom tertile, 500m	3.6	Length of railroads, bottom 50%, 1000m	5.8
8	Planetary boundary layer height	3.4	Area of field crops, bottom tertile, 500m	5.6
9	Longitude	3.3	Length of small roads, top tertile, 250m	5.6
10	Length of small roads, top tertile, 1000m	3.3	Length of large roads, middle tertile, 1000m	5.5

[a]The difference in test R^2 for a model that drops the selected variable compared to the model with all variables.

Table 10. Comparison of predictions with regulatory monitors, by monitor

	$PM_{2.5}$		**PM_{coarse}**	
	R^2	**RMSE**	**R^2**	**RMSE**
Monthly	0.14	3.6	0.29	17.7
Yearly	0.51	3.2	0.36	16.3

Figure 5. (a) $PM_{2.5}$ residuals; (b) PM_{coarse} residuals

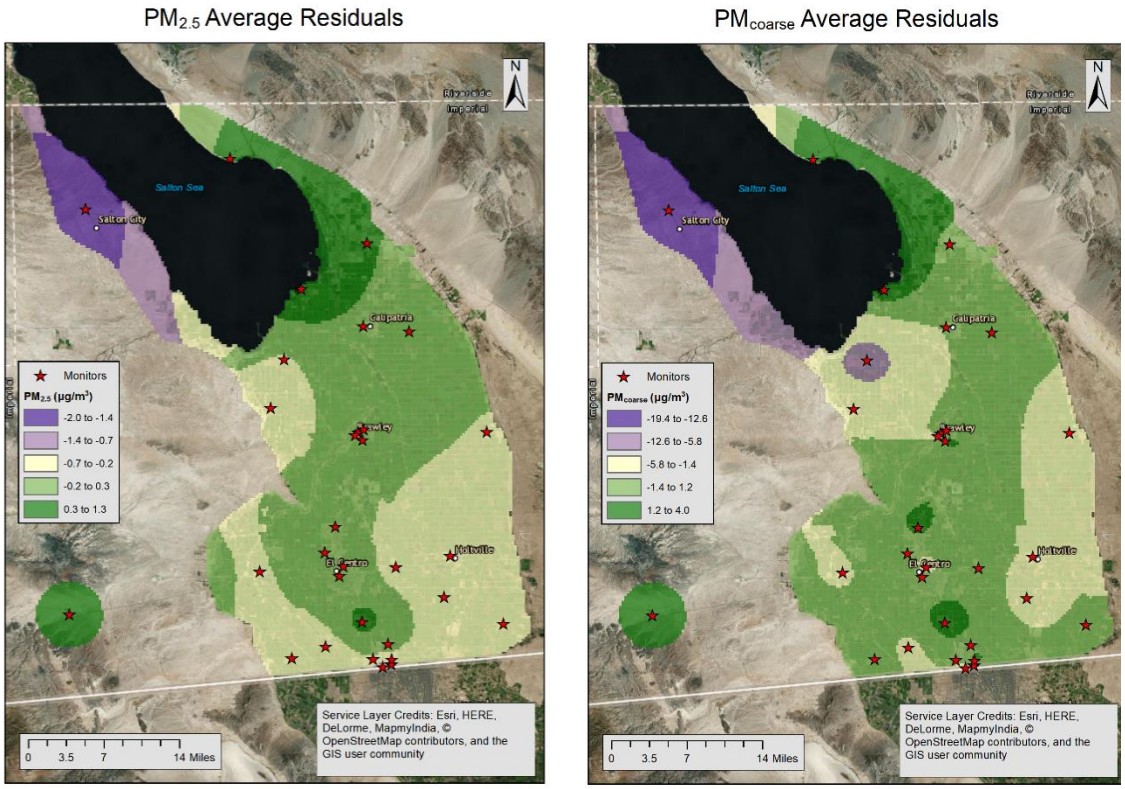


Figure 6. (a) Measured Average $PM_{2.5}$; (b) Predicted Average $PM_{2.5}$

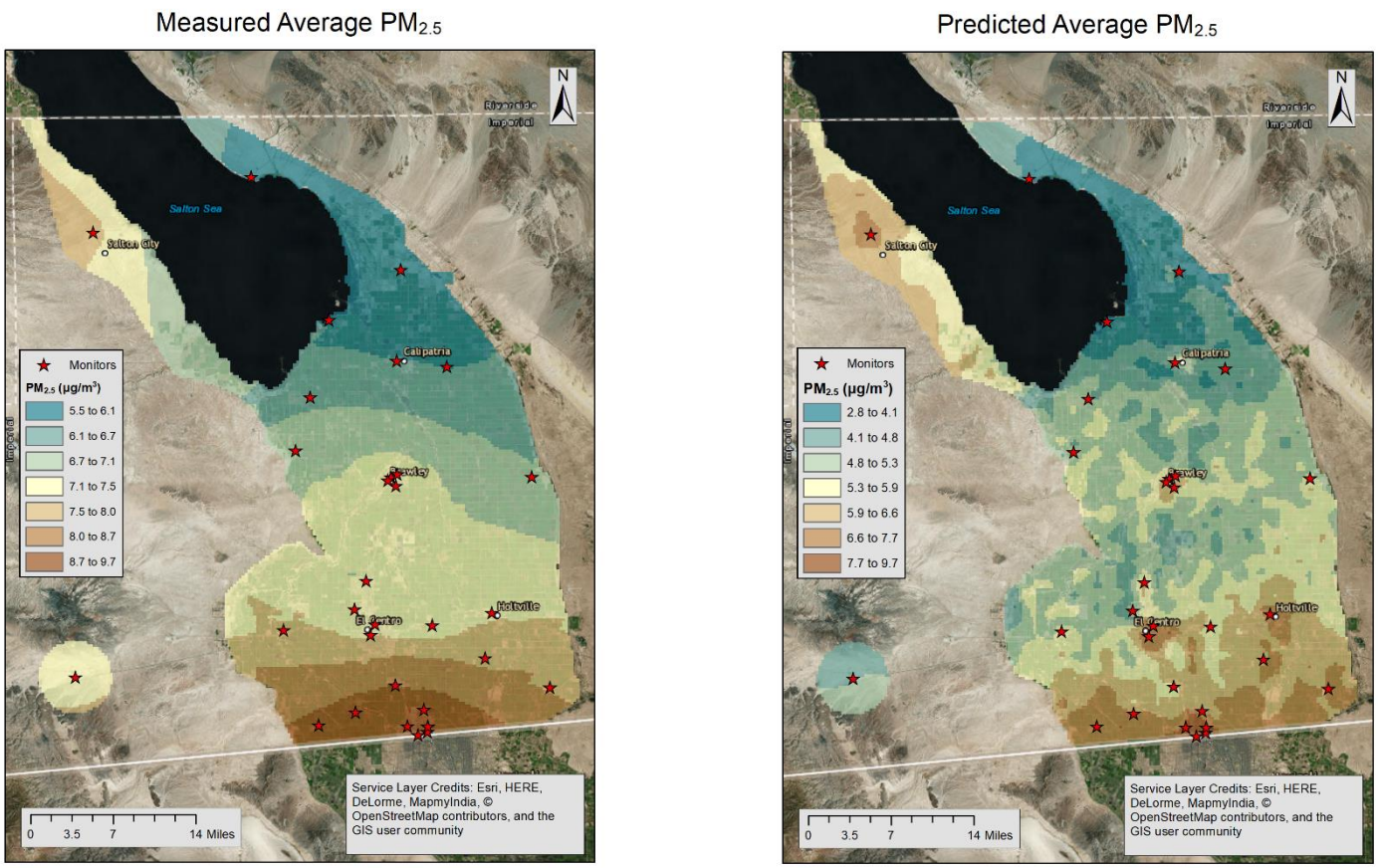


Figure 7. (a) Measured Average PM_{coarse}; (b) Predicted Average PM_{coarse}

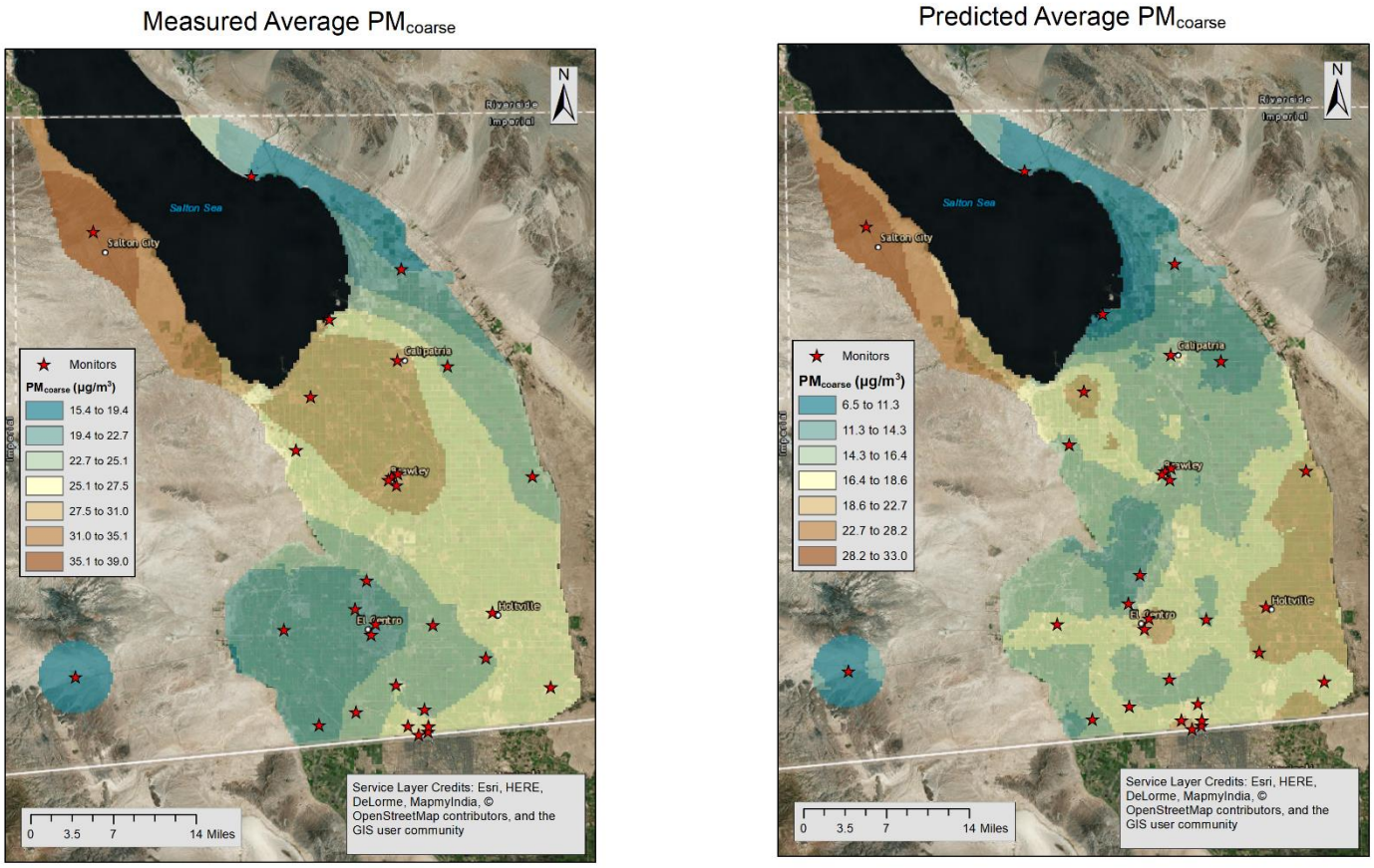


Figure 8. Monthly a) measured $PM_{2.5}$ and b) predicted $PM_{2.5}$

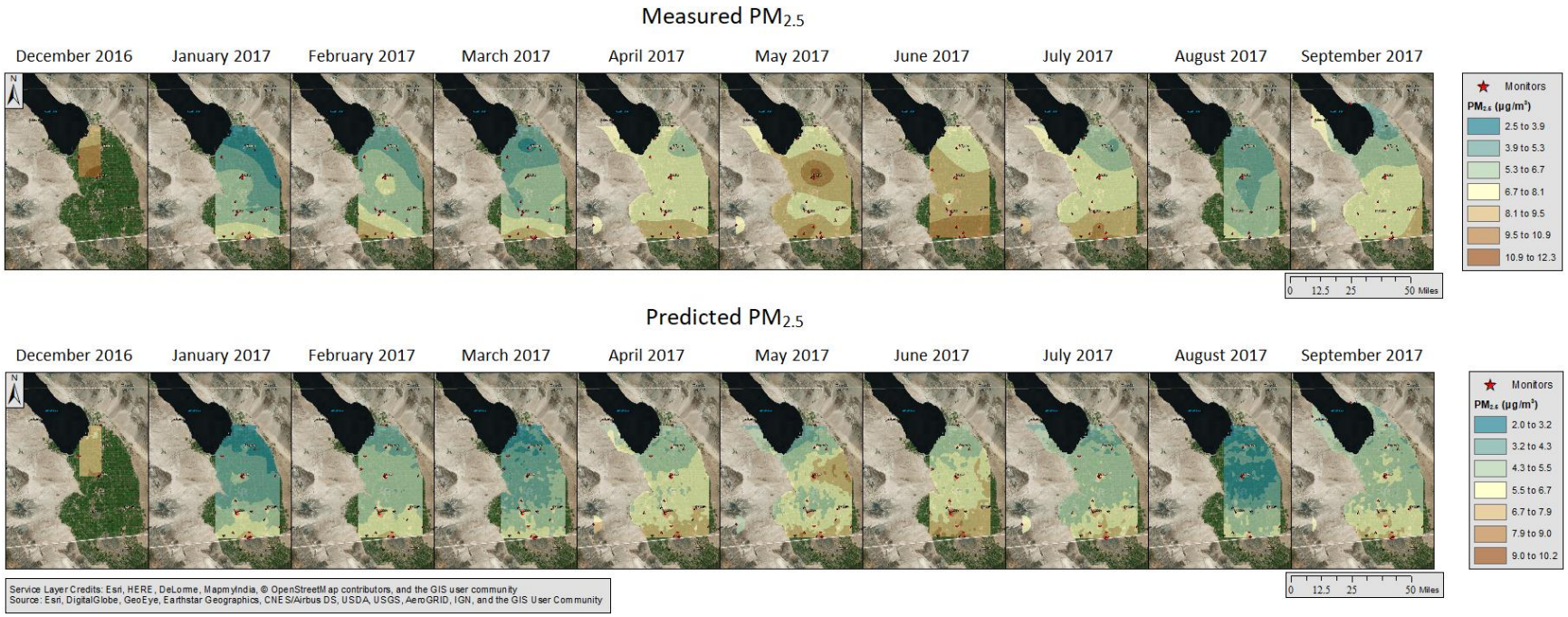
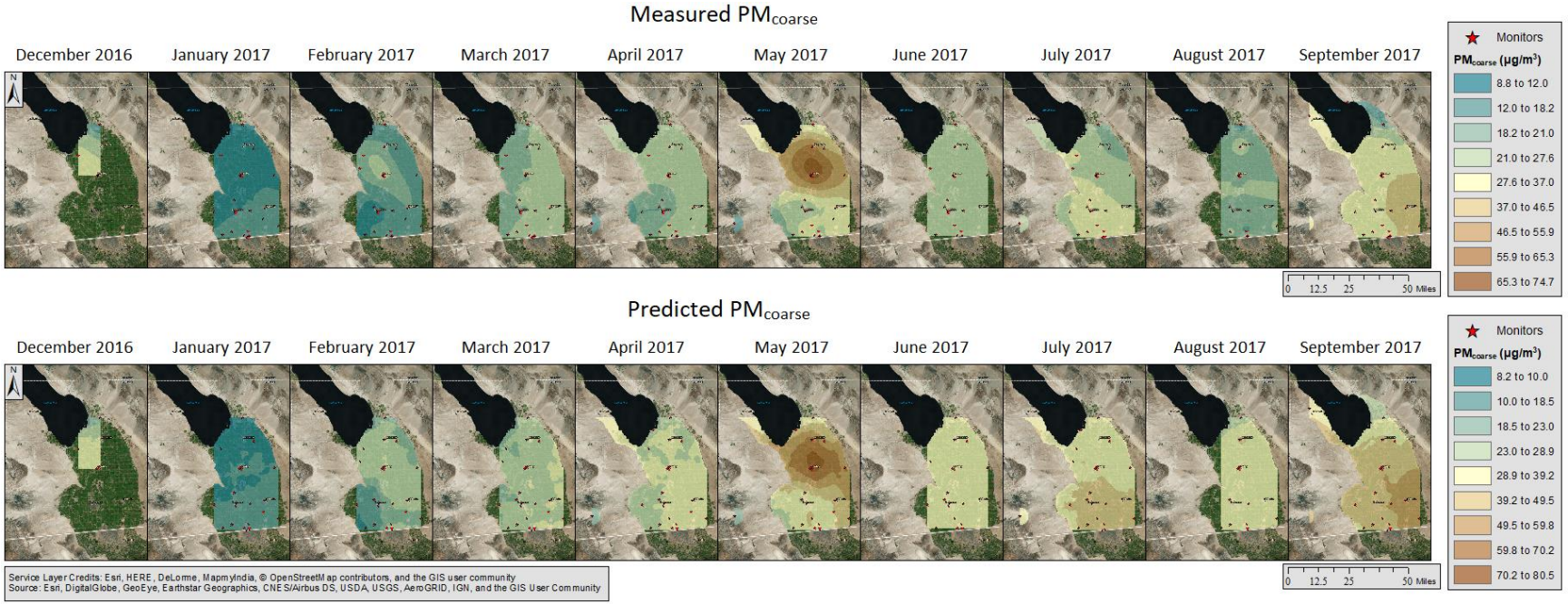


Figure 9. Monthly a) measured PM_{coarse} and b) predicted PM_{coarse}



Chapter 4. Back Trajectory Analysis of Particulate Matter Using the Imperial Community Air Monitoring Network

Chapter 4.1 Introduction

Imperial County is a primarily Hispanic community in the southeast corner of California with high rates of poverty and childhood asthma (U.S. Census Bureau, 2010; Imperial County Asthma Profile, 2015; California Department of Health Services, 2007). The Robert Wood Johnson Foundation (2015) ranked Imperial County as the second worst county in California for health, based on factors such as socioeconomic status, clinical care, and health behaviors. Concentrations of particulate matter (PM), a mix of liquid and solid particles in the air, in Imperial County have exceeded US EPA regulatory standards for PM_{10} (particulate matter with a diameter less than 10 μm) and $PM_{2.5}$ (particulate matter with a diameter less than 2.5 μm) for nearly a decade (California Air Resources Board, 2015; Imperial County Air Pollution Control District, 2014). Exposure to PM has been associated with many health outcomes including low birth weight, cardiovascular disease, childhood asthma, and premature death (Anderson et al., 2011; Brook et al., 2017; Khreis et al., 2017; Atkinson et al., 2016).

Significant sources of $PM_{2.5}$ and PM_{10} internal to Imperial County include unpaved roads and fugitive windborne dust (Imperial County Air Pollution Control District, 2009 & 2014). Agriculture, including tilling, cattle feedlots, and field burning, also contributes to PM in Imperial County. External sources include transport from Mexicali and San Diego and Tijuana (Mendoza, Pardo, & Gutierrez, 2010; Imperial County Air Pollution Control District, 2014).

PM_{coarse}, the difference between PM_{10} and $PM_{2.5}$ was used in this paper instead of PM_{10} since PM_{coarse} does not overlap with $PM_{2.5}$. Since PM_{coarse} is a separate size fraction of PM than $PM_{2.5}$, and oftentimes has a different source profile (Eeftens et al., 2012).

The Imperial Community Air Monitoring Network, referred to hereafter as "the Imperial network" or simply "the network", is comprised of 40 air pollution monitors that measure $PM_{2.5}$ and PM_{10}, relative humidity, and temperature in real time. The real-time data from the network is displayed on a community created and operated website, IVAN air: https://ivan-imperial.org/air. The network was created as a collaboration between a local community group, Comite Civico del Valle, the Public Health Institute, and academics at the University of Washington and the University of California at Los Angeles in response to community concerns over high PM levels in the valley and a desire for more spatially resolved data. A detailed description of the network and the calibration of the community monitors is presented in Carvlin et al. (2017).

This paper uses data from the Imperial network and back trajectory modeling to investigate the contribution of internal and external sources to air pollution in Imperial. Particle back trajectories, paths that particles may have taken to arrive at a receptor location, were calculated using the Hybrid Single Particle Lagrangian Trajectory (HYSPLIT) software and Global Data Assimilation System (GDAS) 1° wind field data from the Air Resources Laboratory

(ARL) at the National Oceanic and Atmospheric Administration (NOAA). HYSPLIT back trajectories have been used previously to determine the contribution of cross-county and cross-border transport to air pollution levels in Southern California and Mexico (Mendoza, Pardo, and Gutierrez, 2010; Shores et al., 2013). Mendoza, Pardo, and Gutierrez (2010) used HYSPLIT back trajectories to explore the contribution of long range transport to $PM_{2.5}$ concentrations in Mexicali during a winter and spring monitoring period in 2004/2005. They found that during the spring, trajectories often crossed through San Diego and Tijuana on their way to Mexicali. But during the high PM episode in the winter, the trajectories were coming from all directions. Shores et al. (2013) studied black carbon (BC) emissions and long-range transport from Tijuana using HYSPLIT. They concluded that BC from Tijuana could be a contributing factor to high PM in Calexico and Mexicali.

This paper presents analyses using gridded HYSPLIT trajectories. The difference in trajectory paths between high PM hours and median PM hours was investigated to determine what wind directions and specific locations might contribute to high PM. A grid was placed across Imperial County to identify points where an air parcel is more likely to pass on a high PM day compared to a median PM day. Polar plots were created to look at the combined effect of wind speed and wind direction on PM levels during summer and winter.

Chapter 4.2 Materials and Methods

$PM_{2.5}$ and PM_{coarse} converted hourly mass concentration data from 10/1/2016 to 10/1/2017 (month/day/year) were used in this analysis. PM_{coarse} was calculated as the difference between PM_{10} and $PM_{2.5}$. All analyses were done using the GMT time zone. See Carvlin et al. (2017) for an explanation of the conversion equation. 24-hour back trajectories were calculated using HYSPLIT for $PM_{2.5}$ and PM_{coarse} hourly measurements. Back trajectories were split up into very high, the top 0.1% of data across all sites, and high, the top 5% of data across all sites. High $PM_{2.5}$ hours were above 25 $\mu g/m^3$ and very high $PM_{2.5}$ hours were above 100 $\mu g/m^3$. High PM_{coarse} hours were above 78 $\mu g/m^3$ and very high PM_{coarse} hours were above 1.3 mg/m^3. For all back trajectories, the starting height was set at 5 m, the average height of the air pollution monitors. GDAS 1° wind field data from NOAA were used as meteorological inputs. Less than 1% of hours could not be run because there was no GDAS wind data for that location and time.

The number of high and very high PM hours, average hourly wind direction and average hourly wind speed were calculated by meteorological season. The wind data for these summary statistics was recorded by the California Air Resources Board at their El Centro monitoring site. The geographic distribution of high and very high hours is discussed.

Matched median 24-hour back trajectories were created for the grid trajectory analysis. Median PM hours were chosen so that there would be the same number of trajectories at each site for both the median and high runs. This was done to ensure that there could be fair comparison between median and high/very high trajectories. If one site records high PM levels more often than other sites and there is not a similar number of median trajectories, then the location of that site will show up with a high density on the grid analysis. All very high

trajectories were matched successfully. Two sites, Calexico-Alvarez and Mexicali, had a large number of high trajectories, possibly because the PM distribution at those sites was shifted higher than at other sites, and matches were not able to be found for approximately 3% of the high trajectories from those sites.

All trajectories were plotted and typed. Trajectory "typing" refers to assigning cardinal directions to the trajectories, which was done by finding the mean latitude and longitude for the trajectory, then finding which of the sixteen cardinal directions it lay between. Trajectories that lay within 10 km of Imperial County were given a separate type. The number of each trajectory type was compared between high/very high and median trajectories for $PM_{2.5}$ and PM_{coarse}.

A gridded trajectory comparison was performed. A 0.1° x 0.1° grid was placed across the southwestern United States and northeastern Mexico to cover the extent of most of the trajectories. The length of median trajectories and high trajectories within each grid cell was recorded. Median trajectories represent background wind patterns therefore, to get the length of high trajectories above background, the length of median trajectories was subtracted from the length of high trajectories. If this value was below 0 it was set to NA, indicating the absence of a value. The values were then normalized to a 0-1 scale. Values that were exactly zero were set to NA as well so as not to cover much of the map with white squares, thus increasing interpretability. The result is a map showing the locations, grid cells, where wind is more likely to pass through when that air parcel results in high PM conditions in Imperial County.

Polar plots, graphs that display wind speed, wind direction, and PM, were created to look at the effect of wind speed and wind direction on PM. Brawley and Calexico were compared during winter and summer months and the potential sources leading to high values were discussed.

In order to assess the uncertainty in the trajectories an ensemble method was applied to a randomly selected 10% subset of all trajectories. An ensemble of trajectories is a group of trajectories that is created by shifting the starting point of the original trajectory. For this analysis, the ensemble consisted of the original trajectory plus 9 other trajectories that had their starting point shifted by a certain amount. Four of the shifted trajectories were run at the height of the original trajectory, 5 m, and were shifted from the origin point by +/- 1 km latitude and longitude. The other five trajectories were run at 100 m. These consisted of the latitude and longitude of the original point with a height of 100 m and that point shifted by +/- 1 km latitude and longitude. This shifting of the starting point helps to assess the effect of uncertainties in the 1° wind field data. If there is a large difference in the path between the original trajectory and the trajectories with the shifted starting points, then we know that the original trajectory is sensitive its meteorological input.

In order to quantify the average spread of the trajectory ensemble, the distance between the furthest shifted trajectories and the original trajectory a trajectory envelope was calculated. The trajectory envelope can be visualized as a sort of cone extending from the origin point backwards in time towards the final trajectory point 24 hours prior to the start. The envelope is the average 3-dimensional distance between the original trajectory and the 9 shifted trajectories

and is calculated for each hour of the 24-hour back trajectory. The trajectory envelope was used to test the validity of the trajectory typing analysis by comparing the average width of the trajectory envelope to the width of a trajectory type. HYSPLIT ensembles have been used in the literature for both dispersion models and trajectory models to estimate uncertainty (Stein, 2007; Stein, 2015; Draxler, 2003).

Chapter 4.3 Results and Discussion

Chapter 4.3.1 Seasonality and Geographic Distribution

Hours with high $PM_{2.5}$ occurred more frequently during the winter (11% of all winter hours) and less frequently during the summer (2% of all summer hours) (Table 11). Hours with high PM_{coarse} occurred more often in the fall (7% of all fall hours) than the spring and winter (5% each) and the summer (3%). Very high $PM_{2.5}$ occurred more frequently in the spring and winter compared to the summer and fall. Very high PM_{coarse} had a similar pattern to very high $PM_{2.5}$. Sites near the US-Mexico border had more high $PM_{2.5}$ hours. High PM_{coarse} hours were concentrated near the US-Mexico border and around Brawley. Sites in Brawley, in the center of the Imperial Valley and sites on the west and north of the Salton Sea recorded more very high $PM_{2.5}$ and PM_{coarse} hours.

Average wind speed was highest in the spring (2.4 m/s) and lowest in the winter (1.7 m/s) (Table 12). The hourly wind direction was primarily from the west and northwest during the spring, fall and winter (55-62% of all hours). The wind was calm, less than 0.5 m/s, for 11% of hours during the winter. During the summer hourly winds came from the southeast (28% of all summer hours), west (20%), and north (18%).

Chapter 4.3.2 Trajectories

The trajectory images show a randomly selected subset of 100 trajectories for that class (Figure 10). Median $PM_{2.5}$ and PM_{coarse} trajectories seemed to have no obvious directionality. High $PM_{2.5}$ trajectories were denser in the west. High PM_{coarse} trajectories were clustered in the west, north, and south.

The ensemble testing resulted in an average uncertainty of 11°, which is only half the angle of a single directional type. High PM hours had similar trajectory patterns for both $PM_{2.5}$ and PM_{coarse} (Table 13). Median PM hours also had similar trajectory patterns for both $PM_{2.5}$ and PM_{coarse}. High and very high PM were associated with trajectories from the west while median PM levels were associated with trajectories from inside Imperial County and less so with trajectories from the north and west. PM_{coarse} was associated with trajectories internal to Imperial County more than $PM_{2.5}$.

For PM_{coarse}, wind came more frequently from WNW-NW (21% more) and W-WNW (12% more) on very high PM hours compared to median PM hours and less frequently from inside Imperial County (19% less) and NNW-N (6% less). High trajectories led to similar but

less pronounced results. For $PM_{2.5}$, wind came more frequently from WNW-NW (21% more) and W-WNW (12% more) on very high PM hours compared to median PM hours and less frequently from inside Imperial County (25% less) and SSW-SW (5% less). These differences were similar but less pronounced for high trajectories except there was a 6% increase when wind came from SSW-SW instead of a decrease. These results suggest that very high PM was driven primarily by westerly winds.

Chapter 4.3.3 Grid Analysis

The grid maps show the locations where wind is more likely to pass through when that air parcel results in high PM conditions in Imperial County. For $PM_{2.5}$, on the larger scale, grid cells that led to high PM were located near San Diego and the area between Mexicali and Tijuana (Figure 11a). On the smaller scale, areas that led to high PM included the US-Mexico border, in particular near the western edge of Imperial County, and the west side of the middle of Imperial County (Figure 11b). For PM_{coarse}, on the larger scale, areas that led to high PM included San Diego/Tijuana, the western part of Imperial County, the area from Mexicali to the Gulf of California, and north of Imperial County (Figure 12a). On the smaller scale, areas that led to high PM were spread out throughout Imperial County, particularly near Brawley and the US-Mexico border (Figure 12b).

The grid maps look different for $PM_{2.5}$ and PM_{coarse}. Areas that led to high $PM_{2.5}$ were focused mainly near the US-Mexico border and western Imperial County whereas areas that led to high PM_{coarse} were spread throughout Imperial County. Clusters of grid cells that led to high PM on the larger scale were more prominent in the south and north for PM_{coarse} and in the west for $PM_{2.5}$.

Chapter 4.3.4 Polar Plots

Figure 13 shows three polar plots, graphs which show the effect of wind speed and wind direction on PM. Wind direction is indicated by the X and Y location of the point cloud. Wind speed is indicated by the distance from the center point. Pollution level is indicated by the color scale. These maps show where pollution may have come from an hour before it arrived at the receptor. For example, a red point on the 10 m/s circle corresponds to an air parcel traveling at 10 m/s for an hour and having high PM concentration when it arrives at the receptor. Figures 13a and 13c are an average of all sites in Calexico and Figure 13b is an average of all sites in Brawley. There was not enough data from Brawley during the summer to create a polar plot. The PM scale is slightly different between the plots because of the range of PM concentration differs by location and season. The radius of the wind speed markers may be slightly different due to how the software interprets the different data sets. This was corrected for as much as possible.

High $PM_{2.5}$ levels in Brawley during the winter were associated with moderate wind speed, 5 to 10 m/s, and a S-SSE wind direction and a W-SW wind direction; and low wind

speeds, 2 to 3 m/s, and a certain small area NNE-NE of Brawley. The W-SW wind direction and moderate wind speed could be related to windborne dust from the nearby desert. The S-SSE wind direction and moderate wind speed suggests a source near the US-Mexico border. The low wind speed and NNE-NE direction may point to a nearby source of agricultural burning or other local emissions. High $PM_{2.5}$ levels in Calexico during the winter were associated with high wind speeds, 12 to 15 m/s, and a W-WSW wind direction. High speed W-WSW winds may be associated with windborne dust. Moderately high $PM_{2.5}$ concentrations from 40 $\mu g/m^3$ to 65 $\mu g/m^3$ were associated with low wind speed and no prominent wind direction. High $PM_{2.5}$ levels in Calexico during the summer were associated with low wind speeds, 3-5 m/s and a S wind direction. This suggests Mexicali as the main source.

Figure 14 shows four polar plots for PM_{coarse} at Brawley and Calexico. Note that the PM scale is different between the plots. There is higher maximum PM_{coarse} in the winter than the summer and at Calexico than Brawley. High PM_{coarse} levels in Brawley during both the summer and winter were associated with moderate wind speed, 5 to 10 m/s, and a W-SW wind direction; low wind speeds, 3 m/s, and a N-NE wind direction. The W-SW wind direction and moderate wind speed suggest windborne dust from the nearby desert as a possible source. The low wind speed and N-NE direction may point to a nearby source of agricultural burning. In the winter high PM_{coarse} levels in Brawley were also associated with moderate wind speed, 7 to 10 m/s, and a S-SSE wind direction. This may suggest transport from Mexicali as a possible source. High PM_{coarse} levels in Calexico were associated with moderate to high wind speeds, 13 to 15 m/s, and a W-WSW wind direction during the winter and low wind speed, 3 to 5 m/s, and a S wind direction during the summer. High speed W-WSW winds may be associated with windborne dust and low speed S winds may be associated with transport from Mexicali.

Chapter 4.4 Conclusion

Data from the Imperial Community Air Monitoring network were used to map the effects of wind direction and wind speed on PM in Imperial County. These results suggest potential sources and meteorological conditions that contribute to high and very high PM. Air parcels were more likely to pass through the western and southwestern part of Imperial County when leading to high compared to median levels of PM. High PM in Brawley, in the middle of Imperial County, was linked with high speed winds from the west and south and low speed winds from the northeast. High PM in Calexico, on the US-Mexico border, was associated with high speed westerly winds and low speed southerly winds. It is possible that the high PM from the high speed westerly winds was caused by windborne dust. The high PM associated with low wind speed conditions in Calexico and high speed southerly winds in Brawley may be linked to transport from Mexicali. The low wind speed high PM hours near Brawley were likely influenced by nearby sources such as agricultural burning and traffic.

References for Chapter 4

Anderson, J., Thundiyil, J., & Stolbach, A. 2011. Clearing the Air: A Review of the Effects of Particulate Matter Air Pollution on Human Health. *Journal of Medical Toxicology, 8*(2), 166-175.

Atkinson R.W., Butland B.K., Dimitroulopoulou C., Heal M.R., Stedman J.R., Carslaw N., Jarvis D., Heaviside C., Vardoulakis S., Walton H., Anderson H.R. 2016. Long-term exposure to ambient ozone and mortality: a quantitative systematic review and meta-analysis of evidence from cohort studies. *BMJ Open, 6*(2).

Brook R.D., Newby D.E., Rajagopalan S. 2017. The Global Threat of Outdoor Ambient Air Pollution to Cardiovascular Health: Time for Intervention. *JAMA Cardiol.*

California Air Resources Board. Chronology of State PM10 Designations. (2015). Retrieved from http://www.arb.ca.gov/desig/changes/pm10.pdf

California Department of Health Services. The Burden of Asthma in California: A Surveillance Report. (2007). Retrieved from http://www.californiabreathing.org/phocadownload/asthmaburdenreport.pdf

Draxler, R. R. (2003). Evaluation of an Ensemble Dispersion Calculation. *Journal of Applied Meteorology, 42*(2), 308-317. doi:10.1175/1520-0450(2003)042<0308:eoaedc>2.0.co;2

Eeftens, M., Tsai, M., Ampe, C., Anwander, B., Beelen, R., Bellander, T., . . . Hoek, G. (2012). Spatial variation of PM2.5, PM10, PM2.5 absorbance and PMcoarse concentrations between and within 20 European study areas and the relationship with NO2 – Results of the ESCAPE project. *Atmospheric Environment, 62*, 303-317. doi:10.1016/j.atmosenv.2012.08.038

Imperial County Air Pollution Control District. 2009 Imperial County State Implementation Plan for Particulate Matter less than 10 Microns in Aerodynamic Diameter. (2009). Retrieved from http://www.co.imperial.ca.us/AirPollution/Attainment%20Plans/Final%20IC%202009%20PM10%20SIP%20Document.pdf

Imperial County Air Pollution Control District. Imperial County 2013 State Implementation Plan for the 2006 24-hour PM2.5 Moderate Nonattainment Area. (2014). Retrieved from http://www.arb.ca.gov/planning/sip/planarea/imperial/Final_PM2.5_SIP_%28Dec_2,_2014%29_Approved.pdf

Imperial County Asthma Profile. (2015, February 2). Retrieved November 20, 2015, from http://californiabreathing.org/asthma-data/county-asthma-profiles/imperial-county-asthmaprofile

Khreis H., Kelly C., Tate J., Parslow R., Lucas K., Nieuwenhuijsen M. 2017. Exposure to traffic-related air pollution and risk of development of childhood asthma: A systematic review and meta-analysis. *Environ Int., 100*, 1-31.

Mendoza, A., Pardo, E. I., & Gutierrez, A. A. (2010). Chemical Characterization and Preliminary Source Contribution of Fine Particulate Matter in the Mexicali/Imperial Valley Border Area. *Journal of the Air & Waste Management Association, 60*(3), 258-270. doi:10.3155/1047-3289.60.3.258

Robert Wood Johnson Foundation. 2015 County Health Rankings, California. (2015). Retrieved from http://www.countyhealthrankings.org/sites/default/files/state/downloads/CHR2015_CA_0.pdf

Shores, C. A., Klapmeyer, M. E., Quadros, M. E., & Marr, L. C. (2013). Sources and transport of black carbon at the California–Mexico border. *Atmospheric Environment, 70*, 490-499. doi:10.1016/j.atmosenv.2012.04.031

Stein, A. F., Ngan, F., Draxler, R. R., & Chai, T. (2015). Potential Use of Transport and Dispersion Model Ensembles for Forecasting Applications. *Weather and Forecasting, 30*(3), 639-655. doi:10.1175/waf-d-14-00153.1

Stein, A. F., Isakov, V., Godowitch, J., & Draxler, R. R. (2007). A hybrid modeling approach to resolve pollutant concentrations in an urban area. *Atmospheric Environment, 41*(40), 9410-9426. doi:10.1016/j.atmosenv.2007.09.004

U.S. Census Bureau. State and County QuickFacts, Imperial County, CA. Washington, D.C.: Government Printing Office. U.S. Census Bureau. (2010). Retrieved from http://quickfacts.census.gov/qfd/states/06/06025.html

Chapter 4: Figures and tables

Table 11. High and very high hours by season[a]

	Spring	**Summer**	**Fall**	**Winter**	**Total**
$PM_{2.5}$ High	2255 (3%)	911 (2%)	2295 (3%)	6750 (11%)	12,211
PM_{coarse} High	3257 (5%)	1853 (3%)	4246 (7%)	2855 (5%)	12,211
$PM_{2.5}$ Very High	134 (0.2%)	2 (0.003%)	34 (0.06%)	75 (0.1%)	245
PM_{coarse} Very High	131 (0.2%)	8 (0.01%)	26 (0.05%)	80 (0.1%)	245

[a] Statistics are: number of hours (percent of hours in that season that are high/very high)

Table 12. Wind speed and direction by season

	Wind Speed (m/s)	**Wind Direction**
Spring	2.4	west (49%), northwest (13%), southwest (11%)
Summer	2.0	southeast (28%), west (20%), north (18%)
Fall	1.8	west (42%), northwest (13%), southwest (11%)
Winter	1.7	west (37%), northwest (12%), calm – no direction (11%)

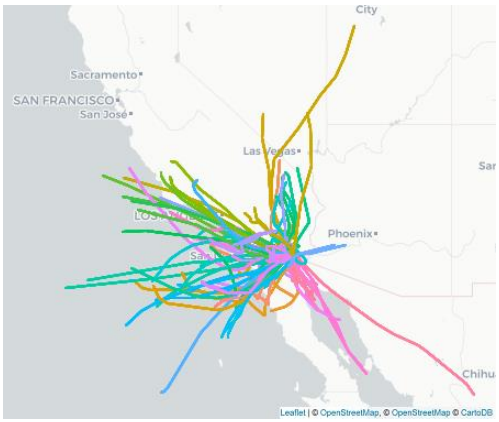
Table 13. Trajectory quadrants

	High Trajectories			**Very High Trajectories**		
	Top	**Second**	**Third**	**Top**	**Second**	**Third**
Elevated PM_{coarse}	WSW-W (21%)	W-WNW (15%)	IC[a] (14%)	W-WNW (29%)	WNW-NW (24%)	WSW-W (23%)
Median PM_{coarse}	IC (18%)	W-WNW (15%)	NNW-N (14%)	IC (27%)	WSW-W (21%)	W-WNW (17%)
Elevated $PM_{2.5}$	WSW-W (21%)	W-WNW (24%)	WNW-NW (22%)	WSW-W (31%)	W-WNW (24%)	WNW-NW (22%)
Median $PM_{2.5}$	IC (21%)	NNW-N (16%)	W-WNW (13%)	IC (29%)	WSW-W (20%)	W-WNW (18%)

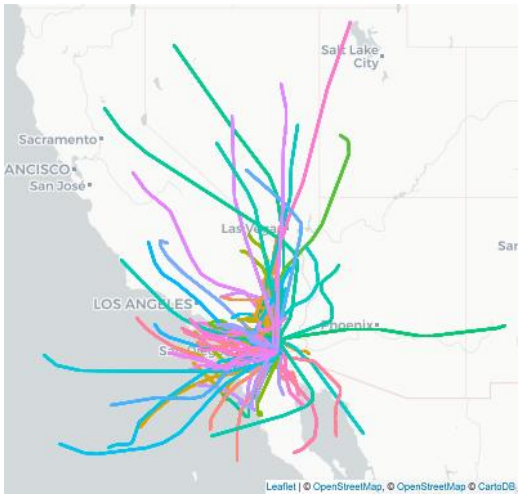
[a]IC = inside Imperial County

Figure 10. Trajectory images for a) the top 5% of $PM_{2.5}$ trajectories; b) the middle 5% of $PM_{2.5}$ trajectories; c) the top 5% of PM_{coarse} trajectories; d) the middle 5% of PM_{coarse} trajectories

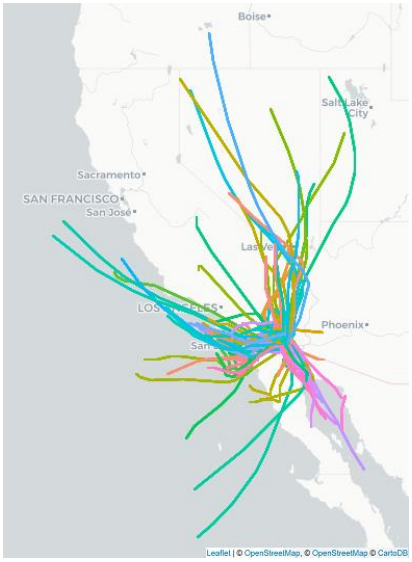
$PM_{2.5}$ Top 5% Trajectories



$PM_{2.5}$ Middle 5% (Median) Trajectories



PM_{coarse} Top 5% Trajectories



PM_{coarse} Middle 5% (Median) Trajectories

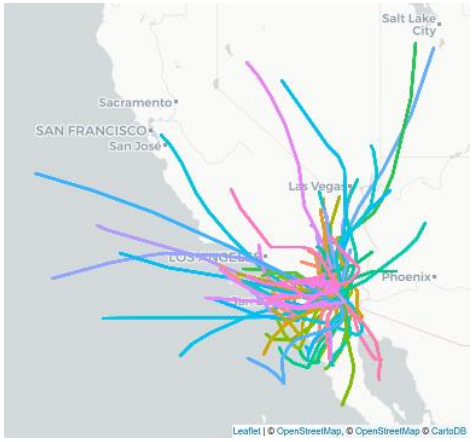


Figure 11. $PM_{2.5}$ trajectory difference grid map; a) large scale, b) small scale

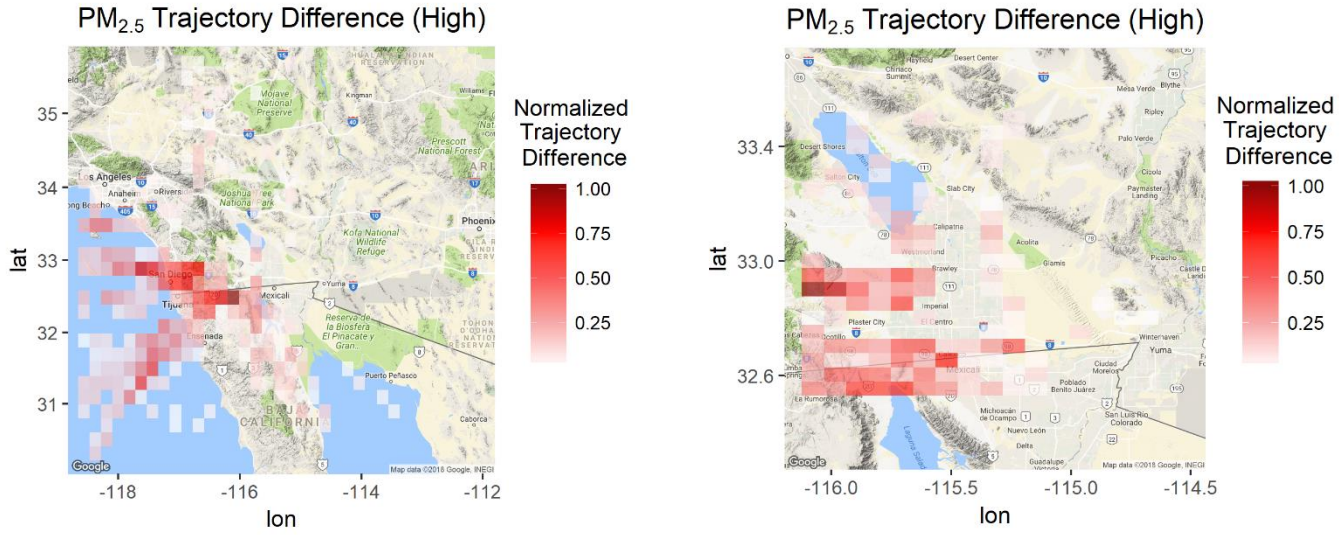


Figure 12. PM_{coarse} trajectory difference grid map; a) large scale, b) small scale

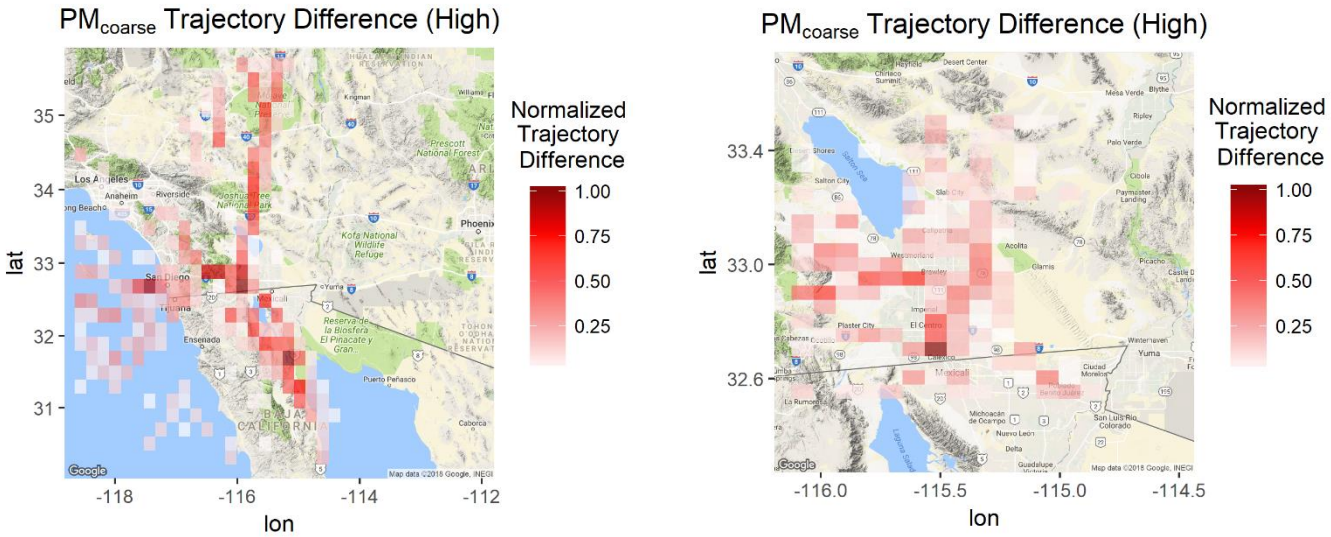
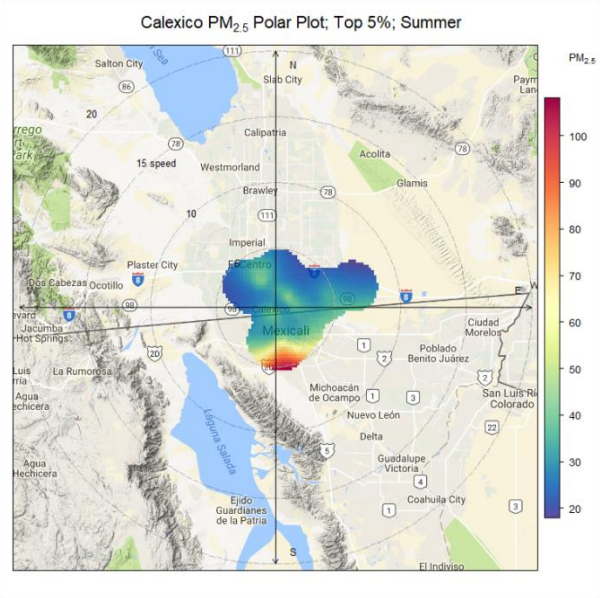
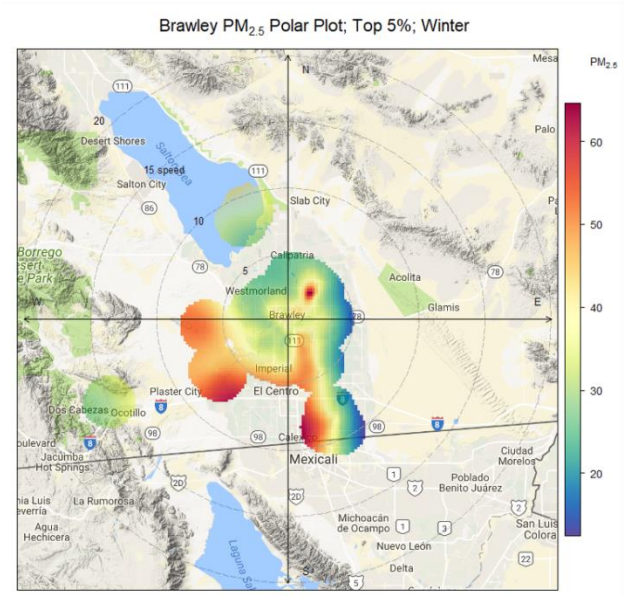


Figure 13. Polar plot of wind speed (m/s), wind direction, and high $PM_{2.5}$ at a) Calexico, summer; b) Brawley, winter; c) Calexico, winter





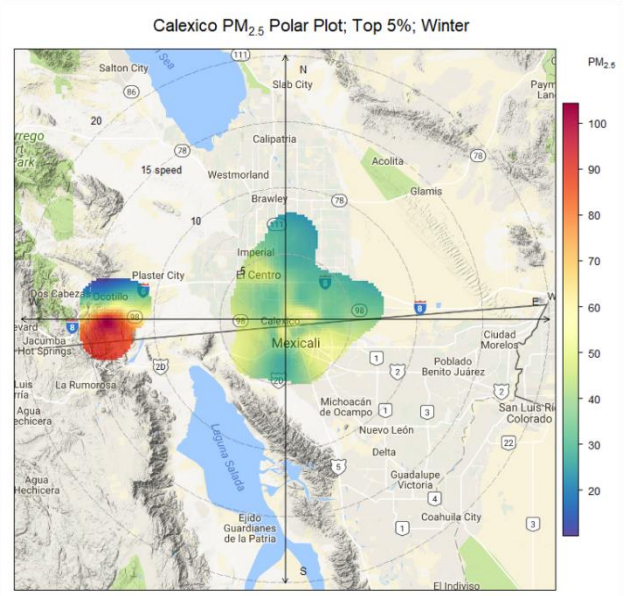
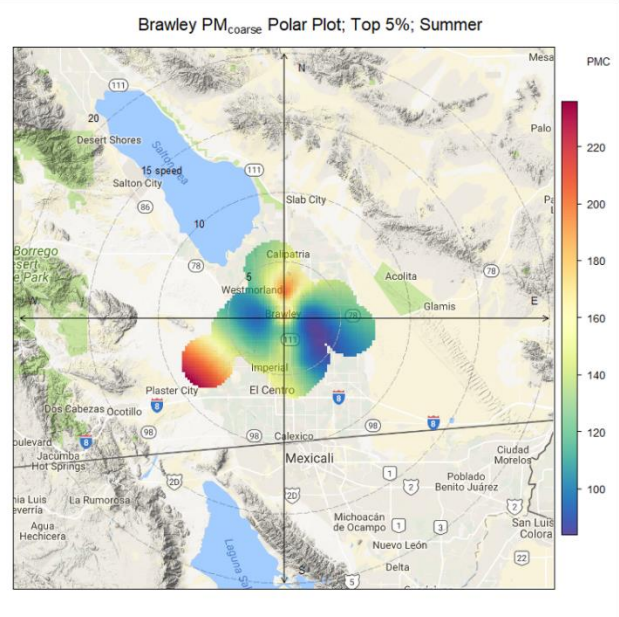
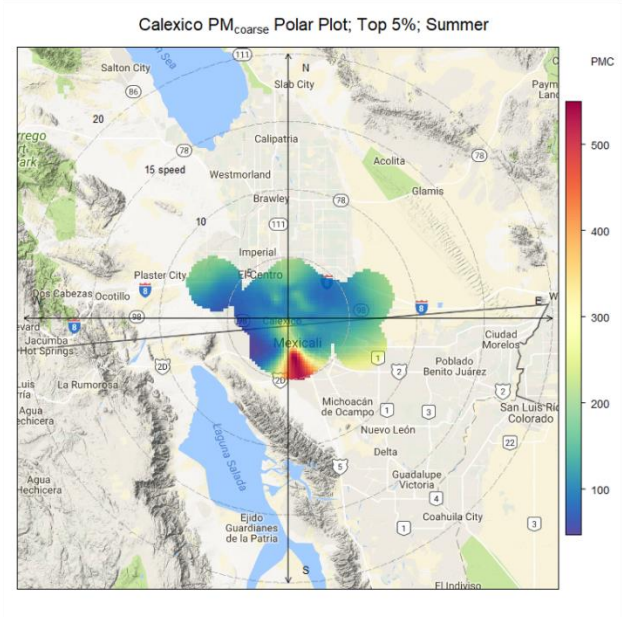
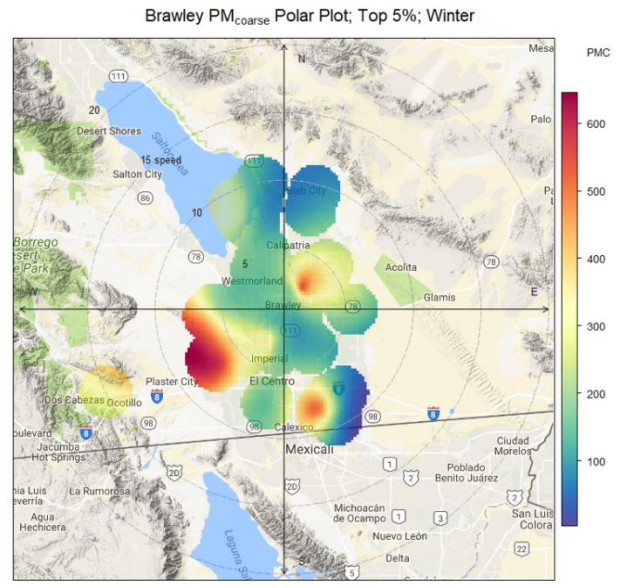
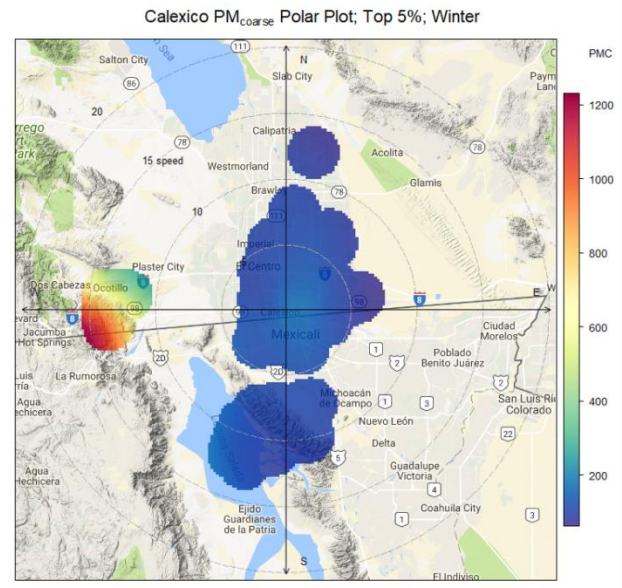


Figure 14. Polar plot of wind speed (m/s), wind direction, and PM_{coarse} at a) Brawley, summer; b) Calexico, summer; c) Brawley, winter; d) Calexico, winter









Chapter 5. Conclusions

This dissertation presents the calibration of novel community air monitors, their deployment in Imperial, CA as a 40-monitor network, and the use of the resultant data for land use regression and back trajectory modeling. Chapter 2 describes the development of the calibration model. A community monitor was collocated with regulatory instruments, $PM_{2.5}$ and PM_{10} beta-attenuation monitors (BAMs), at the Calexico-Ethel site operated by the California Air Resources Board (CARB). The calibration equation was a linear model that included relative humidity. The R^2 for the model was 0.79 for hourly $PM_{2.5}$ and 0.78 for hourly PM_{10}. The model was validated by comparing its performance at six sites that hosted environmental beta attenuation monitors (E-BAMs); average R^2 for hourly $PM_{2.5}$ was 0.59. Chapter 3 presents the development of a land use regression model for predicting monthly $PM_{2.5}$ and PM_{coarse}. Land use, meteorology, and temporal factors sampled within multiple buffer distances were the explanatory variables. Multiple model types were compared and the Bayesian Additive Regression Trees (BART) model performed the best for both $PM_{2.5}$ and PM_{coarse}; R^2 of 0.47 and RMSE of 1.50 $\mu g/m^3$ for monthly $PM_{2.5}$ and an R^2 of 0.65 and RMSE of 8.07 $\mu g/m^3$ for monthly PM_{coarse}. A variable selection procedure was performed that selected relative humidity, wind speed and direction, and road length as important factors influencing $PM_{2.5}$ and temperature, road length, and area of farmland as important factors for PM_{coarse}. Predicted $PM_{2.5}$ was higher near the US-Mexico border while predicted PM_{coarse} was elevated throughout the northern part of the agricultural region of the Imperial Valley. Both types of PM were elevated to the west of the Salton Sea. Chapter 4 discusses the use of HYSPLIT back trajectories and polar plots in determining potential sources contributing to high PM in Imperial County. Air parcels were more likely to pass through areas west and southwest of Imperial County when leading to high PM conditions inside the county. High PM in Brawley, a city in the middle of Imperial County, was associated with high speed wind from the west and south west, indicating wind borne dust as a potential source, and low speed winds from the northeast, indicating a nearby source. High PM in Calexico, a city on the US-Mexico border, was linked to high speed winds from the west, indicating wind borne dust as a potential source, and low speed winds from the south, suggesting a local source of transport from Mexicali, a city adjacent to Calexico.

In general, the results presented in this dissertation agree with the description of air pollution in Imperial County provided by the Imperial County Air Pollution Control District in their State Implementation Plans (SIPs). Both point to windblown dust and transport from Mexicali as significant sources of PM. However, two other large scale features of PM in Imperial County stand out in the analyses presented above. First, high PM to the west of the Salton Sea. While not explicitly mentioned in the SIPs, data from the California Air Resources Board's Air Quality and Meteorological Information System (AQMIS2) shows that Salton City has become the monitoring site with the highest number of PM exceedances in the last five years or so, taking over from sites in Calexico that have historically recorded the highest number of exceedances. Second, the high PM near Brawley, shown in Chapter 3 as high $PM_{2.5}$ and PM_{coarse} in the monthly measurements and predictions, reaching a high peak in May 2017, and in Chapter 4 as the local spike to the N/NE of Brawley.

Community air monitoring as a tool separate from regulatory air monitoring has provided a number of distinct benefits. First, the high spatial density allowed the creation of the land use regression models, a technique that can be used to estimate exposure for epidemiological studies. Second, as an independent check on the regulatory network and the air quality analyses provided by local and state agencies. While our results agreed with the regulatory network and Imperial County SIPs in general, there were notable exceptions. Local spikes near Brawley were seen in the land use regression and wind modeling analyses, but not mentioned in the SIPs. Also, concentrations of PM_{10} beyond the limit of the regulatory monitors were captured by the community monitors. This work was presented by Astrid Calderas, a post-doc and member of CCV, to CARB, which responded by changing the maximum limit on some of their BAMs from 1,000 $\mu g/m^3$ to 5,000 $\mu g/m^3$. They have already seen hourly values above 2,000 $\mu g/m^3$. Third, we, the members of the Imperial Project, have found community air monitoring to be an invaluable tool to help the community advocate for their air quality with local and state government. The effort that CCV has put into their environmental justice work has been immense and it has been awesome to watch them speak out for air quality.

During the five years of the Imperial grant, we learned many lessons, both technical and interpersonal. Technical lessons learned included networking choices, how to increase the resilience of our monitors to extreme environmental conditions, and individual component choices. Many places in Imperial County are rural and do not have any Wi-Fi/Ethernet connections and poor, if any, cellular connection. In these locations, primarily around the Salton Sea, we used cellular communications, attaining ~50% data completeness. If we had the time and funds it would have been good to use satellite or radio communications for these locations. During the initial deployment of the monitors we found cellular to be much easier to install and more reliable than Wi-Fi. This was due in part to poor Wi-Fi signal strength and the inability to use 5G and enterprise Wi-Fi networks. External Wi-Fi antennas were added at some sites, which helped somewhat. Imperial County has very hot summers, upwards of 120 °F, and occasional dust storms. A monitor that was located on top of a metal roof had the Dylos melt, something the manufacturer had never seen before. This unit was relocated and the fan in the enclosure was set to “always on” instead of using its internal thermostat. From this incident we found out that the internal thermostats did not work correctly in any of the enclosures and would only cool the enclosures occasionally. The dust storms, which happen more frequently around the Salton Sea and near the desert areas, would often either cause the Dylos to begin reading all 0s or read erratically. Affected units were sent to the manufacturer for cleaning and recalibration. We opened up one of the units that had been in a dust storm and could literally pour the sand out. We are working on adding sand traps to the monitors that are subject to frequent dust storms.

We have also learned many lessons regarding the practice of community-based participatory research (CBPR); especially since the goal from the beginning of the project was to turn over ownership of the network to the community after the grant ended. The biggest lesson was a lesson in communication and trust. Communities that are in locations with poor air quality often face other injustices, economic, social, linguistic, etc., and may have a history of negative interactions with government and academic institutions. This foundational mistrust of the government can make it challenging to foster productive interactions between regulatory air

quality stakeholders and communities in community monitoring studies. Furthermore, academic institutions can be seen as a sort of black box where the academics have total control over the results of the study and the process by which they arrive at those results is hidden behind a curtain of technical jargon. There has also been a history of "parachute science", where academics will perform a study in a community then leave without presenting the results to the community. From the government or academic perspective it can feel anxiety-provoking to hand over monitoring technology or study results to the community for fear that it will be misinterpreted and used to reinforce prior beliefs. The successful CBPR study must be able to address these concerns in a transparent and accountable way. As in many endeavors, open and honest communication is the foundation to a productive working relationship. Agreements for work to be done should be accompanied by a specific date and deliverable. In the Imperial study, check-in meetings were held every month over the phone with the whole project team. For pieces of work that didn't involve the whole team, separate calls were held. We also met in person in Imperial or Oakland, where the California Environmental Health Tracking Program is based, at least once a year. During these meetings we would discuss the status of the project and the ongoing operation and funding of the network. Meetings with community members and local politicians were held in Imperial to present our work to the public and advocate for funding after the end of the grant.

Transferring ownership of the monitoring network to CCV is an ongoing process that has, at times, been quite challenging. Part of that challenge is building technical capacity. CCV has hired a full-time programmer to work on their IVAN web platform for visualizing air quality and reporting environmental complaints. Their programmer has been instrumental in helping to enable CCV to take control of the network. Having a person with programming experience is necessary for installing and configuring all the code that runs the backend of the monitoring network. It is also good to know how to change the code on the microcontroller in case new features are desired or more bugs are found. Furthermore, it is necessary to know how to download data from the database and manipulate it into a flat file structure or use it for data analyses. All together this requires knowledge of databases (Kairos, SQL), server scripting (PHP), web scripting (HTML, JavaScript, CSS), microcontroller coding (C, C++, python), statistical software (R), and, if hardware changes are desired, circuit board creation, soldering, and mechanical design. These are the skills that I picked up during the project as I designed the community monitor. It is hard to find somebody with all of these skills at once since only a subset are needed for most IT/engineering jobs. That is why I have stayed on as a technical consultant. I am working with CCV to help them understand and improve on the current monitor and network design. It is recommended that communities that are interested in setting up a monitoring network with a custom air quality monitor find a technical consultant. However, most communities will not need this level of customization and there area handful of low-cost off-the-shelf monitoring solutions where monitors upload to a third-party database that includes a website for visualization and downloading of the data.

Comite Civico del Valle (CCV) has been doing environmental justice and advocacy work for over 25 years. Through the diligent pursuit of their advocacy work, CCV has been able carve out a seat for themselves in discussions with the local and state air quality agencies. Over the

duration of the project we have seen the approach of government air quality agencies towards interacting with community groups slowly change. In particular, the California Air Resources Board (CARB), with whom we collaborated on the Imperial project, and the South Coast Air Quality Management District (SCAQMD), with whom we did not work directly, have shown more interest in the knowledge and experience of the community and an eye towards investing in communities to help monitor their own air quality.

To this end, the California State Legislature passed Assembly Bill No. 617 in 2017. This bill has allocated funding for community air monitoring to be conducted statewide. CCV were involved with the creation of the bill and the Imperial network was used as a model of a successful community air monitoring project. The IVAN web platform developed by CCV and the structure of the community air monitoring project in Imperial are being adapted to use throughout California. Most of the initial partners on the Imperial project are involved with work stemming from AB617 and the effort to expand community monitoring across the state. The original Imperial project has applied for funding under AB617 to help pay for the costs of operating the network.

As more low-cost air sensors reach the market and are adopted by the public, there has been increasing effort from government and academic institutions to characterize their performance. SCAQMD's AQ-SPEC program has done lab and field testing for dozens of PM and gas sensors and provides a nice overview of the performance of currently available sensors. The US Environmental Protection Agency (EPA) has created a sensor scale, which provides guidelines for categorizing air pollution readings from sensors that oftentimes report data much more frequently than the regulatory monitors from which the EPA's Air Quality Index (AQI) was developed. The US EPA is also starting the discussion of performance standards for low-cost air sensors, formalizing the minimum requirements needed for a sensor to be considered acceptable to be used for a community monitoring project.

It is my hope that the lessons learned from the Imperial project can be used to foster increased collaboration between communities, government, and academia to create air monitoring networks that can help identify sources of poor air quality, educate citizens, and open the door to policy changes that will work to directly reduce air pollution emissions or exposure to poor air quality.

Appendix A. Supporting Information for Chapter 2

Table 1. R^2 for different size bins of the Dylos and BAM and filter data

	$PM_{2.5}$			**PM_{10}**		
	FEM BAM (hourly)	FEM BAM (daily)	FRM Filter	FEM BAM (hourly)	FEM BAM (daily)	FRM Filter
Bin 1 (>0.5 μm)	0.78	0.84	0.79	0.31	0.41	0.23
Bin 2 (>1.0 μm)	0.71	0.73	0.67	0.57	0.72	0.57
Bin 3 (>2.5 μm)	0.56	0.58	0.42	0.77	0.87	0.78
Bin 4 (>10 μm)	0.09	0.06	0.04	0.73	0.71	0.43
Bin 1 – Bin 2	0.68	0.81	0.74	0.20	0.30	0.13
Bin 1 – Bin 3	0.74	0.83	0.76	0.23	0.34	0.17
Bin 1 – Bin 4	0.78	0.84	0.79	0.31	0.41	0.23
Bin 2 – Bin 3	0.73	0.76	0.71	0.34	0.51	0.38
Bin 2 – Bin 4	0.71	0.73	0.68	0.57	0.72	0.57
Bin 3 – Bin 4	0.56	0.59	0.42	0.77	0.86	0.78
n[a]	3907	160	300	4008	168	32

[a]Total number of overlapping observations for the Dylos and the BAM/Filter

Table 2. Data Loss During the Study Period and After Implementing Lessons Learned

Site	**Location**	**Network Type[a]**	**Network Data Loss (%)**	**Dylos Data Loss (%)**	**Total Data Loss (%)**
Seeley[b]	Central	E, C	<1, 0	<1, <1	<1, <1
Kennedy[b]	Central	C, C	<1, 0	82, 0	83, 0
Westmorland[b]	North	W, C	49, 30	24, 19	73, 49
Meadows[b]	Central	E, E	<1, 0	37, <1	37, <1
Calipatria[b]	North	C, C	46, 22	<1, 0	46, 22
Calexico[b] Alvarez	South	W, W	24, 36	<1, 0	24, 36
Total (average)			19, 15	24, 3	44, 18
Calexico-Ethel[c]	South	C, C	9, <1	<1, 0	9, <1

[a]E = Ethernet; C = Cellular; W = Wi-Fi

[b]Study period (3/2/16 to 7/19/16), follow up period (1/1/17 to 3/1/17)

[c]From 6/18/15 to 7/12/16 (~13 months), follow up period (1/1/17 to 3/1/17).

Figure 1. Scatterplot of $PM_{2.5}$ BAM vs. $PM_{2.5}$ BAM replicate at Calexico-Ethel

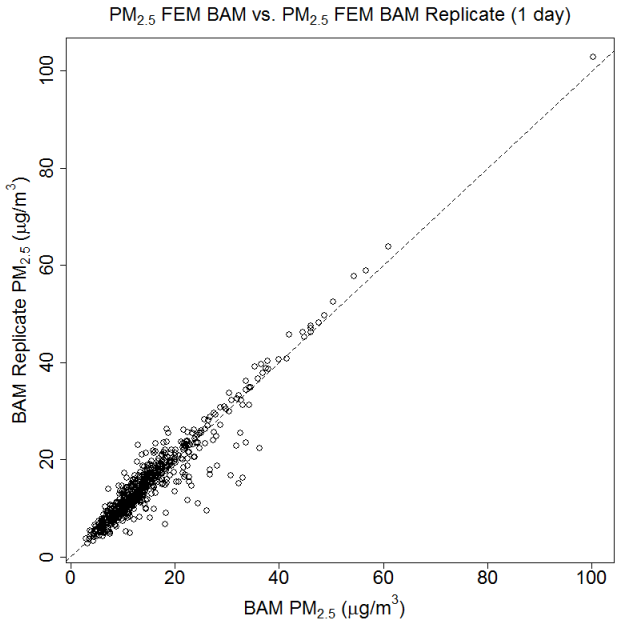


Figure 2. Scatterplot of $PM_{2.5}$ Filter vs. $PM_{2.5}$ Filter replicate at Calexico-Ethel

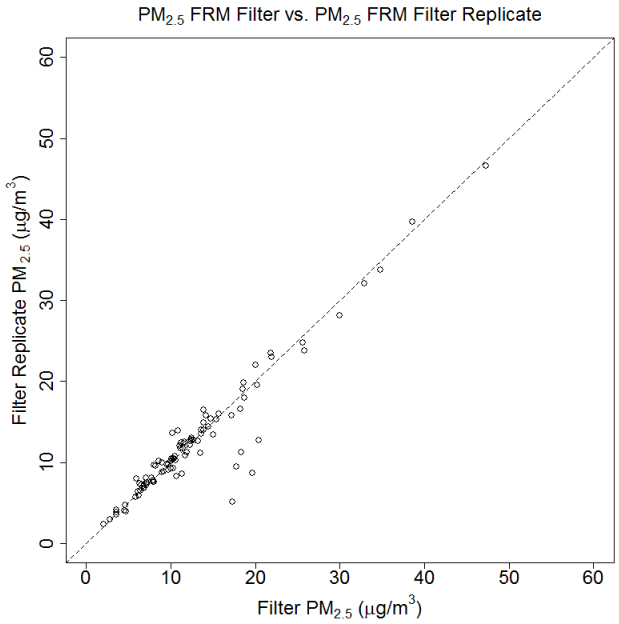


Figure 3. Seeley Time-series

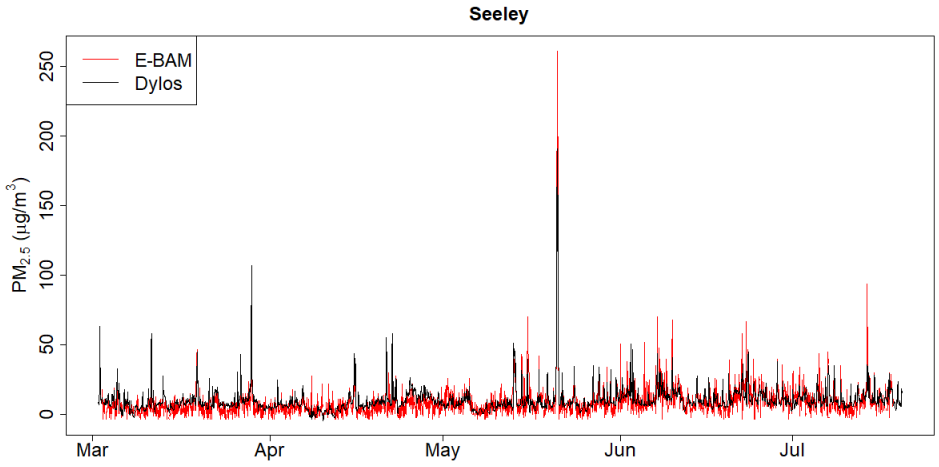


Figure 4. Calipatria Time-series

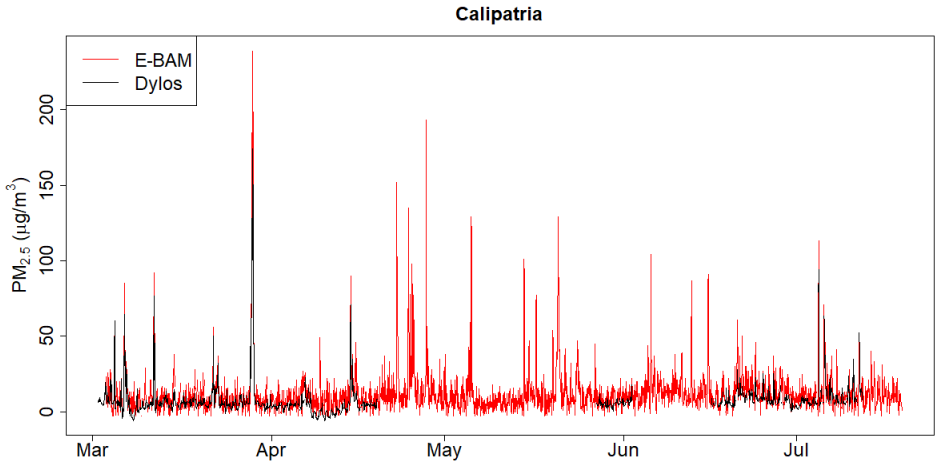


Figure 5. Westmorland Time-series

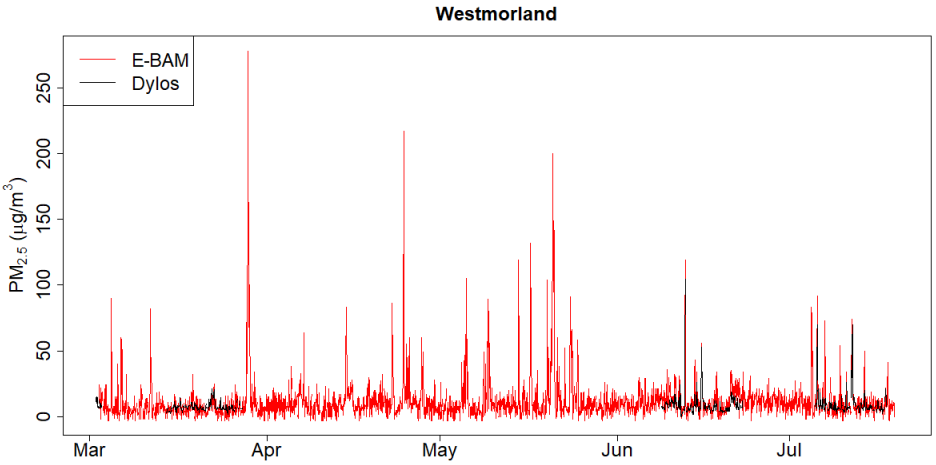


Figure 6. Meadows Time-series

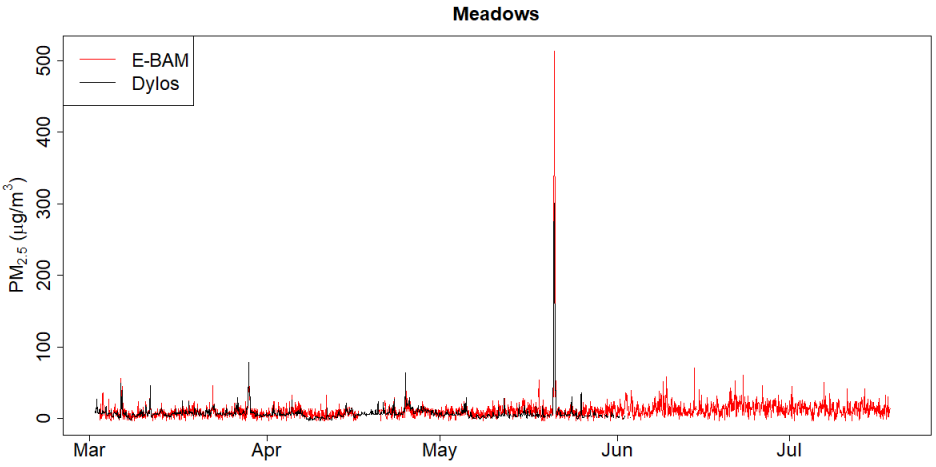


Figure 7. Kennedy Time-series

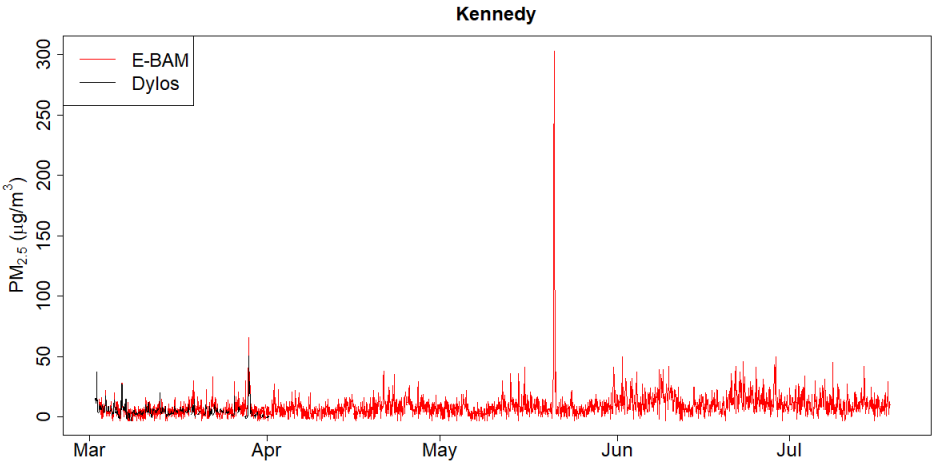


Figure 8. Calexico Alvarez Time-series

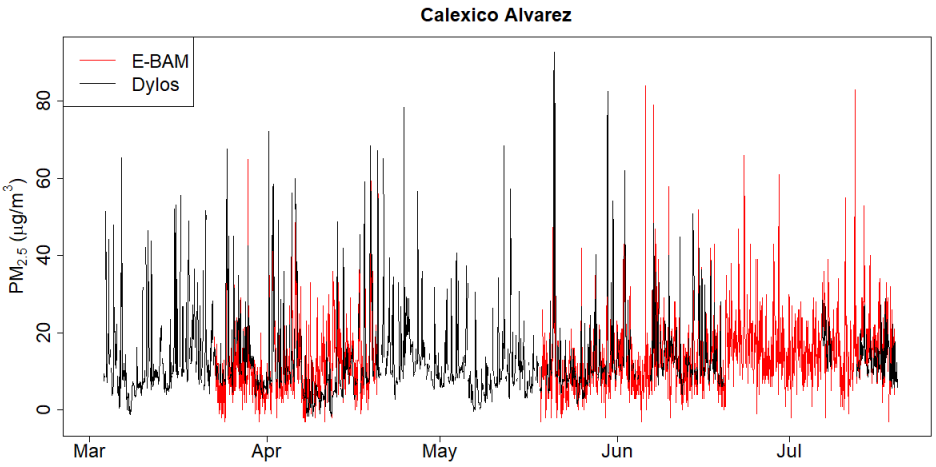


Figure 9. Difference between Dylos conversion and E-BAM

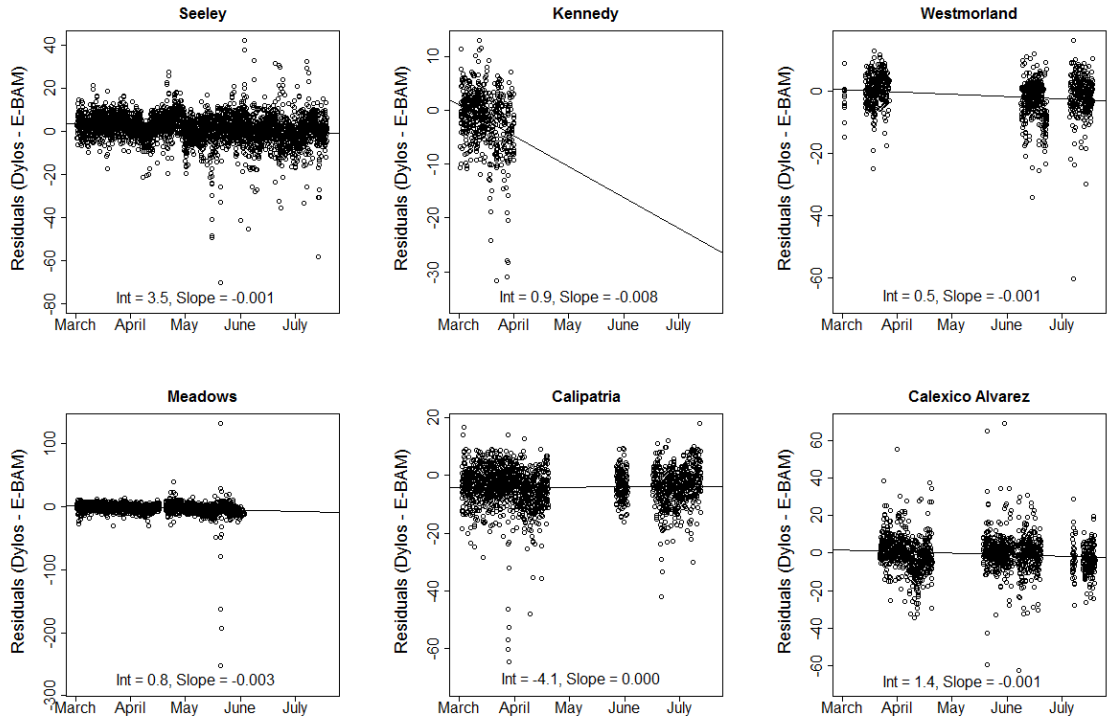
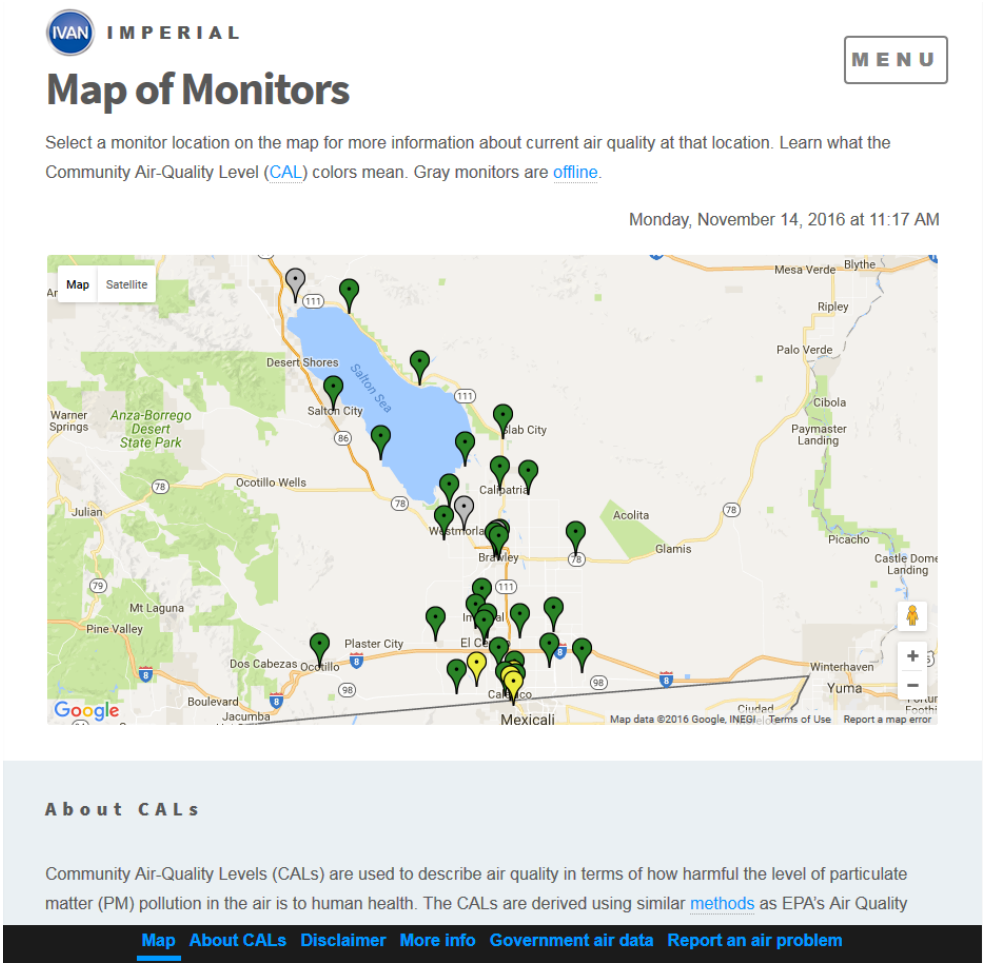


Figure A10. The Display of the Community Air Monitoring Network on the IVAN website



VITA

Graeme Carvlin got his BA in Environmental Studies at the University of Chicago and his MS in Chemistry at the University of California San Diego. During his PhD studies he has worked on numerous environmental monitor design projects, providing engineering, software design, and data analysis expertise. He now works at the Puget Sound Clean Air Agency as an Air Monitoring Specialist, helping to keep the regulatory air monitoring network running and developing new low-cost air monitors to use in community monitoring projects.