Machine learning-derived daily wildfire and non-wildfire $PM_{2.5}$ concentration estimates over the western US, 2008-2018

Colleen E. Reid[1*], Ellen M. Considine[2], Melissa M Maestas[2], Gina Li[1,2]

November 11, 2019

1. Department of Geography, University of Colorado Boulder, Boulder, Colorado, USA and 2. Cooperative Institute for Research in Environmental Sciences, Earth Lab *corresponding author: Colleen Reid (Colleen.Reid@Colorado.edu)

Abstract

Fine particulate matter ($PM_{2.5}$) levels are declining in many areas of the US due to policies and enforcement of the Clean Air Act. However, in much of the western US, $PM_{2.5}$ concentrations have been increasing, likely due to the increased presence of wildfires in this region. There is growing evidence of various health impacts of $PM_{2.5}$ exposures, even at levels below the federal standard. Health studies of $PM_{2.5}$ in the western US are limited by spatial sparseness of monitoring data. To improve population exposure assessment of $PM_{2.5}$, researchers are increasingly using statistical methods to "blend" information from multiple data sources to better estimate $PM_{2.5}$ in space and time. Some studies have created daily fine-resolution estimates of $PM_{2.5}$ for the whole US, but they perform poorly in the western US. We have tailored a machine learning model to the western US, combining satellite, meteorological, monitoring, land use and other spatiotemporal data to estimate daily $PM_{2.5}$ estimates at the census tract, ZIP code, and county levels during 2008-2018. Our methods improve upon previous models by: use of a more extensive monitoring station network, which captures more spatial locations and proximity to wildfires; use of ensembles of machine learning algorithms, which have been shown to improve model performance; and coverage of a longer period of time. We are making our data publicly available for use in future studies of the health impacts of fine particulate air pollution in the western US.

Background & Summary

Fine particulate matter ($PM_{2.5}$) air pollution is increasingly associated with numerous adverse health outcomes including, but not limited to, mortality [1], respiratory and cardiovascular morbidity [34, 23], negative birth outcomes [15], and lung cancer [10]. Although $PM_{2.5}$ concentrations have been declining in many parts of the United States due to policies to limit emissions of air pollutants [7], $PM_{2.5}$ levels have been increasing in parts of the northwestern US [21]. This increase has been shown to be associated with wildfire smoke [21, 22], which can cause $PM_{2.5}$ concentrations that are several times higher than the Environmental Protection Agency's (EPA's) daily $PM_{2.5}$ National Ambient Air Quality Standard (NAAQS) in areas downwind of the wildfires for several days at a time [25].

Estimates of $PM_{2.5}$ concentrations for health studies have traditionally been derived from data from stationary air quality monitors placed in and around populated areas for regulatory purposes. In the US, the EPA's Federal Reference Method (FRM) monitors often only measure every third or sixth day and do not provide enough spatial coverage to obtain a good estimate of the air pollution exposures where every person lives. In fact, most US counties do not contain a regulatory air pollution monitor [3]. Using solely monitoring data in health studies leads to exposure misclassification, which often, but not always, drives effect estimates of the association between air pollution and health towards the null [35].

To improve population exposure assessment of $PM_{2.5}$, epidemiological researchers have increasingly been using methods to estimate $PM_{2.5}$ exposures in the temporal and spatial gaps between regulatory monitors using data from satellites (such as aerosol optical depth (AOD) or polygons of smoke plumes) or air pollution models [3, 19]) over the past two decades. Each of these data sources has its own benefits and limitations, and researchers are increasingly statistically "blending" information from a combination of data sources to better estimate $PM_{2.5}$ in space and time. Various methods of blending have been used including spatiotemporal regression kriging (e.g., [13]), geographically-weighted regression (e.g., [16]), and machine learning methods (e.g., [24, 14, 5]).

Machine learning methods train large auxiliary datasets, often including satellite AOD, meteorological data, chemical transport model output, and land cover and land use data to provide optimal estimates of $PM_{2.5}$ where people breathe. These models have been implemented in various locations around the world at city, regional, and national scales [2]. Some epidemiological questions can only be addressed in longitudinal studies with large sample sizes. Exposure models with large spatial and temporal domains will help enable such studies. Within the US, Di et al. [5, 6] and Hu et al. [14] have separately used machine learning algorithms to create fine-resolution daily $PM_{2.5}$ estimates for the continental US. These models, however, have performed poorly in the western US [5, 14] and particularly the mountain west [6] compared to the rest of the country. Given the increasing trends in $PM_{2.5}$ concentrations in parts of the western US and the importance of wildfires as a source of $PM_{2.5}$ there, it is important to have a model that is tailored to this region to capture the variability in space and time in this region.

The dataset we describe here improves upon previous daily estimates of $PM_{2.5}$ concentrations from machine learning models in the following ways: (1)

use of a more extensive monitoring station network than used in previous models that captures more spatial locations and also proximity to wildfires, a key driver of $PM_{2.5}$ in the western US, (2) use of an ensemble of machine learning algorithms which have been shown to improve model performance [6], (3) better temporal prediction through the use of a nonlinear function (cosine) on day of year, (4) allowance for different prediction models for fire-affected and non-fire affected days to better capture and predict high $PM_{2.5}$ levels during wildfires, and (5) incorporation of errors in prediction back into daily estimates through spatial interpolation. We are making these data available as daily estimates of $PM_{2.5}$ exposures at census tract, ZIP-code, county scales in a public repository, which the above cited papers have not done, to be used in future studies of the societal impacts of air pollution exposure in the western US, where wildfires are a significant contributor to $PM_{2.5}$ concentrations.

[insert Figure 1: monitor locations (points) and state boundaries]

[insert Table 1: list variables]

Methods

Study Area

Our study area includes 11 western US states: Arizona, California, Colorado, Idaho, Montana, Nevada, New Mexico, Oregon, Utah, Washington, and Wyoming. Our temporal domain were all days between January 1, 2008 and ***, 2018. We predicted daily estimates of $PM_{2.5}$ at the ZIP code and county levels from machine learning ensembles trained on observed daily $PM_{2.5}$ values from monitoring stations from a variety of sources (**put in all $PM_{2.5}$ data sources). The predictor variables for the machine learning ensemble included (**put in all variables here) More information on the sources of these data can be found in Table 1.

$PM_{2.5}$ Measurements

To get a more comprehensive set of locations and time points of $PM_{2.5}$ measurement throughout the western US, we did an extensive search for as many $PM_{2.5}$ monitoring data within our spatial and temporal study area as we could find. We downloaded $PM_{2.5}$ data from the US EPA AQS Air Data Query Tool [28] for the 11-state region (Figure **??**) including any of the following parameter codes: 88101, 88500, 88502, 81104 [27, 29, 31]. These data include the IMPROVE monitors that capture air quality information in more rural areas [30]. We also retrieved all available $PM_{2.5}$ data in the Fire Cache Smoke Monitor Archive (https://wrcc.dri.edu/cgi-bin/smoke.pl), which includes U.S. Forest Service monitors that were deployed to capture air quality impacts during wildfire events.

Some states have additional $PM_{2.5}$ monitors beyond those required by the U.S. EPA. We reached out to the department charged with air quality in every state within our study domain and obtained additional $PM_{2.5}$ data from California Air Resources Board and the Utah Department of Environmental Quality. We only included data that was in addition to the monitors in those states that was part of the U.S. EPA's AQS and IMPROVE data.

We also reached out to researchers who may have had their own monitoring networks of $PM_{2.5}$ throughout the region. We were able to obtain data from the Uintah Basin, Utah from Seth Lyman at Utah State University, and $PM_{2.5}$ measurements from the Persistent Cold Air Pool Study (PCAPS) [26] conducted in the Salt Lake Valley, Utah in January–February, 2011 from Dr. Geoff Silcox at the University of Utah.

All of this yielded a total of XX daily $PM_{2.5}$ observations, which represent XX locations.

Predictor Variables

[Write short description of each predictor data set and refer to Table 1]

Satellite Aerosol Optical Depth (AOD) is a measure of particle loading in the atmosphere from the ground to the satellite, which can be used as a proxy measure of $PM_{2.5}$ when combined with other variables such as ... We obtained daily estimates of AOD from the MODIS Terra and Aqua combined Multi-angle Implementation of Atmospheric Correction (MAIAC) dataset https://ladsweb.modaps.eosdis.nasa.gov/archive/allData/6/MCD19A2/. This is the finest resolution (1 km) AOD dataset currently available and was available for our whole time period and spatial domain. After downloading each Hierarchical Data Format (HDF) file from the online repository, we calculated the average daily AOD values at each location, and took the nearest neighbor value at each $PM_{2.5}$ monitoring location. MAIAC AOD has been shown to better predict $PM_{2.5}$ than coarser resolution AOD [4] and has been used in many studies in various geographic regions in blended models to predict daily $PM_{2.5}$ [17, 8, 18].

We obtained meteorological data from the North American Mesoscale (NAM), Analysis meteorological model https://www.ncdc.noaa.gov/data-access/model-data/model-datasets/north-american-mesoscale-forecast-system-nam because it includes all of the standard meteorological variables, including planetary boundary layer height, which are important for converting AOD to $PM_{2.5}$ [20]. We calculated 24-hour averages from 6-hourly data for temperature, relative humidity, sea level pressure, surface pressure, planetary boundary layer height, dew point temperature, precipitation, snow coverage, and the U and V components of wind speed. NAM has 12 km resolution and is available from 2004 onward.

Because one of the reasons that $PM_{2.5}$ concentrations have been increasing in the western US is the increasing number and magnitude of wildfires, we wanted to have variables about the proximity of a location to an active fire. We collected daily data about fire detection locations and size from the MODIS Thermal Anomalies/Fire Daily L3 Global 1km product (MOD14 and MYD14) [9, 11]. As fires in closer proximity are likely to influence $PM_{2.5}$ more than fires further away, we calculated the number of active fires in radial buffers of 25, 50, 100, and 500 km radii around each monitoring location.

Elevation can influence $PM_{2.5}$ concentrations. For example, $PM_{2.5}$ can accumulate in mountain valleys during persistent cold air pools (commonly referred to as inversions) during winter [33]. We obtained elevation data from the 3D Elevation Program, which has a resolution of 1 arc-second, which is approximately 30 m north/south and varies east/west with latitude [32].

Surrounding land cover can be a proxy for air pollution emissions. We used the land cover class information from the Landsat-derived National Land Cover Dataset (NLCD) [12] to calculate the percentage of urban development (codes 22, 23, and 24), agriculture (codes 81 and 82), and vegetated area other than agricultural land (codes 21, 41, 42, 43, 52, and 71) within buffer radii of 1 km, 5 km, and 10 km around each monitor. NLCD 2011 has a spatial resolution of 30 m and uses circa 2011 Landsat satellite data. We obtained the Normalized Difference Vegetation Index (NDVI) from the MODIS satellite product MOD13A3 https://lpdaac.usgs.gov/products/mod13a3v006/ at 1 km resolution by month as another measure of vegetation that was not just a measure of agricultural vegetation but all vegetation.

As a proxy indicator of emissions from vehicles, we calculated the sum of all road lengths of type Arterial and Collector within 100, 250, 500, 1000 m buffers of each monitoring location. Arterial roads are high-capacity urban roads. Collector roads are low-to-moderate capacity roads. The road data came from the National Highways Planning Network https://www.fhwa.dot.gov/planning/processes/tools/nhpn/index.cfm which contains spatial information on over 450,000 miles of highways in the United States.

We included population density as an additional proxy for emisions as areas with higher population have more sources of air pollution emissions. Population density was obtained from the American Community Survey at the XXX (spatial resolution) for each year or five year averages?

To account for seasonality in $PM_{2.5}$ data, we created the following predictor variables: cosine of day-of-year and cosine of month. We also created dummy variables for each state and month in our study domain to allow for spatial and temporal variation in the data that could not be explained by any of the other spatial, temporal, or spatiotemporal variables.

Data merging

We created three datasets: one dataset to train the model and two prediction datasets. The training dataset merged all predictor variables to each 24-hour average $PM_{2.5}$ monitoring observation by linking the data temporally (using date) and spatially (by selecting the nearest observation for each predictor variable). Similarly, the prediction datasets were created by spatially and temporally linking all predictor variables to the population-weighted centroid of each ZIP code and county for each day in the study domain.

Machine learning modeling and mapping

Finally, we created an indicator variable for whether there were one or more fires within 500 km of a monitor in the last week.

Code availability

[Insert brief description of how to access code on GitHub.] The code was written and annotated in R [version number] and Python [version number] and is available from GitHub [doi citation link]. The key package for implementing the ML model was [caretEnsemble?].

Data Records

All data are freely available from [repository name, data doi citation]. We provide ... [reference Figure 2]

[insert Figure 2: choropleths at zip code level - 4-panel: a) highest year $PM_{2.5}$, Aug or Sept, b) highest year $PM_{2.5}$, Jan/Feb, c) lowest year $PM_{2.5}$, Aug or Sept, d) lowest year $PM_{2.5}$, Jan/Feb.]

[insert Figure 3: Time series of select cities]

[Insert Table 3: list of files]

Technical Validation

[Write description of goodness of fit methods/metrics - out-of-bag data, RMSE, R^2, models run on subsets of data, etc.]

[Insert Figure 4: a) out-of bag observed $PM_{2.5}$ vs predicted, b) full model observed $PM_{2.5}$ vs predicted, c-j) various subsets of data - oob and full model plots (see figure 5 of example paper)]

[Write discussion about variable importance, possibly referring to the suggested figure of variable importance panel figure. Could make an observation or two about the complexity of the variables, e.g., $PM_{2.5}$ can be highest at highest and lowest temperatures (summer fire season and winter inversions), etc.]

[Thoughts - insert figure of predicted $PM_{2.5}$ vs predictor variable for the 8 (or so) most important variables (panel figure)]

Thoughts: compare to $PM_{2.5}$. Concerned comparing to HMS will take too long?

Usage Notes

[Write brief description of things the provided code can be adapted to do, such as making plots of specific years, use in health/pollution studies.]

Acknowledgements

[Write acknowledgements text here.]

Author contributions

[Write brief description of contribution from each author.]

Competing interests

The authors declare not competing interests.

Figures and figures legends

[All figures go here and are referred to in the text]

Tables

Table 1: Variables used in the machine learning models.

Variable	**Type**	**Source**
Date		
Coordinates (Latitude and Longitude)	Spatial	
Active Fire Points Count (25 km, 50 km, 100 km, and 500 km buffer radii; 0-7 day lags)	Spatial and Temporal	
Binary Fire indicator (0 for no active fire points in any buffer radii or lag for given point; 1 otherwise)	Spatial and Temporal	
Summed length of arterial (A) and collector (C) roads within 100, 250, 500, and 1000 m buffer radii, A and C separately and together	Spatial	National Highways Planning Network
Population Density	Spatial	
MAIAC AOD	Spatial and Temporal	NAM
HPBL.surface	Spatial and Temporal	NAM
TMP.2.m.above.ground	Spatial and Temporal	NAM
RH.2.m.above.ground	Spatial and Temporal	NAM
DPT.2.m.above.ground	Spatial and Temporal	NAM

APCP.surface	Spatial and Temporal	NAM
WEASD.surface	Spatial and Temporal	NAM
SNOWC.surface	Spatial and Temporal	NAM
UGRD.10.m.above.ground	Spatial and Temporal	NAM
VGRD.10.m.above.ground	Spatial and Temporal	NAM
PRMSL.mean.sea.level	Spatial and Temporal	NAM
PRES.surface	Spatial and Temporal	NAM
DZDT.850.mb	Spatial and Temporal	NAM
DZDT.700.mb	Spatial and Temporal	NAM
TimeZone	Spatial	
National Land Cover Database (NLCD) (1 km, 5 km, and 10 km)	Spatial and Temporal	
NDVI	Spatial and Temporal	
Season	Temporal	
State	Spatial	
Cosine of Day of Year	Temporal	

References

[1] Souzana Achilleos, Marianthi-Anna Kioumourtzoglou, Chih-Da Wu, Joel D. Schwartz, Petros Koutrakis, and Stefania I. Papatheodorou. Acute effects of fine particulate matter constituents on mortality: A systematic review and meta-regression analysis. *Environment International*, 109:89–100, 2017.

[2] Colin Bellinger, Mohomed Shazan Mohomed Jabbar, Osmar Zaiane, and Alvaro Osornio-Vargas. A systematic review of data mining and machine learning for air pollution epidemiology. *Bmc Public Health*, 17:907, 2017. WOS:000416433100002.

[3] Cole Brokamp, Eric B. Brandt, and Patrick H. Ryan. Assessing Exposure to Outdoor Air Pollution for Epidemiological Studies: Model-based and Personal Sampling Strategies. *The Journal of Allergy and Clinical Immunology*, May 2019.

[4] A. A. Chudnovsky, A. Kostinski, A. Lyapustin, and P. Koutrakis. Spatial scales of pollution from variable resolution satellite imaging. *Environ Pollut*, 172C:131–138, September 2012.

[5] Q. Di, I. Kloog, P. Koutrakis, A. Lyapustin, Y. Wang, and J. Schwartz. Assessing PM2.5 Exposures with High Spatiotemporal Resolution across the Continental United States. *Environ Sci Technol*, 50(9):4712–21, May 2016.

[6] Qian Di, Heresh Amini, Liuhua Shi, Itai Kloog, Rachel Silvern, James Kelly, M. Benjamin Sabath, Christine Choirat, Petros Koutrakis, Alexei

Lyapustin, Yujie Wang, Loretta J. Mickley, and Joel Schwartz. An ensemble-based model of PM2.5 concentration across the contiguous United States with high spatiotemporal resolution. *Environment International*, 130:104909, July 2019.

[7] Neal Fann, Sun-Young Kim, Casey Olives, and Lianne Sheppard. Estimated Changes in Life Expectancy and Adult Mortality Resulting from Declining PM2.5 Exposures in the Contiguous United States: 1980-2010. *Environmental Health Perspectives*, 125(9):097003, 2017.

[8] Guannan Geng, Nancy L. Murray, Daniel Tong, Joshua S. Fu, Xuefei Hu, Pius Lee, Xia Meng, Howard H. Chang, and Yang Liu. Satellite-Based Daily PM2.5 Estimates During Fire Seasons in Colorado. *Journal of geophysical research. Atmospheres: JGR*, 123(15):8159–8171, August 2018.

[9] Louis Giglio, Ivan Csiszar, and Christopher O. Justice. Global distribution and seasonality of active fires as observed with the Terra and Aqua Moderate Resolution Imaging Spectroradiometer (MODIS) sensors. *Journal of Geophysical Research: Biogeosciences*, 111(G2), 2006. G02016; https://modis.gsfc.nasa.gov/data/dataprod/mod14.php.

[10] Ghassan B. Hamra, Neela Guha, Aaron Cohen, Francine Laden, Ole Raaschou-Nielsen, Jonathan M. Samet, Paolo Vineis, Francesco Forastiere, Paulo Saldiva, Takashi Yorifuji, and Dana Loomis. Outdoor particulate matter exposure and lung cancer: a systematic review and meta-analysis. *Environmental Health Perspectives*, 122(9):906–911, September 2014.

[11] Todd J. Hawbaker, Melanie K. Vanderhoof, Yen-Ju Beal, Joshua D. Takacs, Gail L. Schmidt, Jeff T. Falgout, Brad Williams, Nicole M. Fairaux, Megan K. Caldwell, Joshua J. Picotte, Stephen M. Howard, Susan Stitt, and John L. Dwyer. Mapping burned areas using dense time-series of Landsat data. *Remote Sensing of Environment*, 198(Supplement C):504 – 522, 2017.

[12] Collin Homer, Jon Dewitz, Limin Yang, Suming Jin, Patrick Danielson, George Xian, John Coulston, Nathaniel Herold, James Wickham, and Kevin Megown. Completion of the 2011 National Land Cover Database for the Conterminous United States – Representing a Decade of Land Cover Change Information. *Photogrammetric Engineering & Remote Sensing*, 81(5):345 – 354, 2017. https://www.mrlc.gov/nlcd2011.php.

[13] Hongda Hu, Zhiyong Hu, Kaiwen Zhong, Jianhui Xu, Feifei Zhang, Yi Zhao, and Pinghao Wu. Satellite-based high-resolution mapping of ground-level PM2.5 concentrations over East China using a spatiotemporal regression kriging model. *The Science of the Total Environment*, 672:479–490, April 2019.

[14] Xuefei Hu, Jessica H. Belle, Xia Meng, Avani Wildani, Lance A. Waller, Matthew J. Strickland, and Yang Liu. Estimating PM2.5 Concentrations

in the Conterminous United States Using the Random Forest Approach. *Environmental Science & Technology*, 51(12):6936–6944, June 2017.

[15] Petra Klepac, Igor Locatelli, Sara Korošec, Nino Künzli, and Andreja Kukec. Ambient air pollution and pregnancy outcomes: A comprehensive review and identification of environmental public health challenges. *Environmental Research*, 167:144–159, 2018.

[16] William Lassman, Bonne Ford, Ryan W. Gan, Gabriele Pfister, Sheryl Magzamen, Emily V. Fischer, and Jeffrey R. Pierce. Spatial and temporal estimates of population exposure to wildfire smoke during the washington state 2012 wildfire season using blended model, satellite, and in situ data. *GeoHealth*, 1(3):106–121, 2017.

[17] Hyung Joo Lee. Benefits of High Resolution PM2.5 Prediction using Satellite MAIAC AOD and Land Use Regression for Exposure Assessment: California Examples. *Environmental Science & Technology*, September 2019.

[18] Runkui Li, Tianxiao Ma, Qun Xu, and Xianfeng Song. Using MAIAC AOD to verify the PM2.5 spatial patterns of a land use regression model. *Environmental Pollution (Barking, Essex: 1987)*, 243(Pt A):501–509, September 2018.

[19] Jia C. Liu, Gavin Pereira, Sarah A. Uhl, Mercedes A. Bravo, and Michelle L. Bell. A systematic review of the physical health impacts from non-occupational exposure to wildfire smoke. *Environ Res*, 136:120–132, 2015.

[20] Y. Liu, J. A. Sarnat, V. Kilaru, D. J. Jacob, and P. Koutrakis. Estimating ground-level PM2.5 in the eastern United States using satellite remote sensing. *Environ Sci Technol*, 39(9):3269–78, May 2005.

[21] Crystal D. McClure and Daniel A. Jaffe. Us particulate matter air quality improves except in wildfire-prone areas. *Proc Natl Acad Sci U S A*, pages 1–6, 2018.

[22] Katelyn O'Dell, Bonne Ford, Emily V. Fischer, and Jeffrey R. Pierce. The contribution of wildland-fire smoke to US PM2.5 and its influence on recent trends. *Environmental Science & Technology*, January 2019.

[23] Sanjay Rajagopalan, Sadeer G. Al-Kindi, and Robert D. Brook. Air Pollution and Cardiovascular Disease: JACC State-of-the-Art Review. *Journal of the American College of Cardiology*, 72(17):2054–2070, October 2018.

[24] C. E. Reid, M. Jerrett, M. L. Petersen, G. G. Pfister, P. E. Morefield, I. B. Tager, S. M. Raffuse, and J. R. Balmes. Spatiotemporal prediction of fine particulate matter during the 2008 northern California wildfires using machine learning. *Environ Sci Technol*, 49(6):3887–96, March 2015.

[25] Colleen E. Reid, Ellen M. Considine, Gregory L. Watson, Donatello Telesca, Gabriele G. Pfister, and Michael Jerrett. Associations between respiratory health and ozone and fine particulate matter during a wildfire event. *Environment International*, 129:291–298, August 2019.

[26] Geoffrey D. Silcox, Kerry E. Kelly, Erik T. Crosman, C. David Whiteman, and Bruce L. Allen. Wintertime pm2.5 concentrations during persistent, multi-day cold-air pools in a mountain valley. *Atmospheric Environment*, 46:17 – 24, 2012.

[27] US EPA. *AQS Memos - Technical Note on Reporting PM2.5 Continuous Monitoring and Speciation Data to the Air Quality System (AQS)*, 2017, accessed November 2, 2017. https://www.epa.gov /aqs/aqs-memos-technical-note-reporting-pm25-continuous-monitoring- and-speciation-data-air-quality.

[28] US EPA. *Outdoor Air Quality Data Download Daily Data*, 2017, accessed November 2, 2017. https://www.epa.gov/outdoor-air-quality-data/download-daily-data.

[29] US EPA. *Parameters*, 2017, accessed November 2, 2017. https://aqs.epa.gov/aqsweb/docu- ments/codetables/parameters.html.

[30] US EPA. *PM 2.5 - Visibility (IMPROVE)*, 2017, accessed November 2, 2017. https://www3.epa.gov/ttnamti1/visdata.html.

[31] US EPA. *Sampling Methods for All Parameters*, 2017, accessed November 2, 2017. https://aqs.epa.gov/aqsweb/documents/codetables/methods_all.html.

[32] USGS. *About 3DEP Products and Services*, 2017, accessed November 6, 2017. https://national map.gov/3DEP/3dep_prodserv.html.

[33] C. David Whiteman, Sebastian W. Hoch, John D. Horel, and Allison Charland. Relationship between particulate air pollution and meteorological variables in Utah's Salt Lake Valley. *Atmospheric Environment*, 94(Supplement C):742 – 753, 2014.

[34] Yu-Fei Xing, Yue-Hua Xu, Min-Hua Shi, and Yi-Xin Lian. The impact of PM2.5 on the human respiratory system. *Journal of Thoracic Disease*, 8(1):E69–74, January 2016.

[35] S.L. Zeger, D. Thomas, F. Dominici, J.M. Samet, J. Schwartz, D. Dockery, and A. Cohen. Exposure measurement error in time-series studies of air pollution: concepts and consequences. *Environ Health Perspect*, 108(5):419–426, 2000.