Environmental Research 159 (2017) 9–15

Contents lists available at ScienceDirect

Environmental Research

journal homepage: www.elsevier.com/locate/envres

Development of a model for particulate matter pollution in Australia with implications for other satellite-based models

Gavin Pereira[a,*], Hyung Joo Lee[b], Michelle Bell[c], Annette Regan[a], Eva Malacova[a], Ben Mullins[a], Luke D. Knibbs[d]

[a] School of Public Health, Curtin University, WA, Australia
[b] NASA Ames Research Center, CA, USA
[c] School of Forestry and Environmental Studies, Yale University, CT, USA
[d] School of Population Health, University of Queensland, Qld, Australia

ARTICLE INFO

Keywords:
Particulate matter
PM_{10}
Aerosol optical depth (AOD)
Air pollution
Land-use regression
Satellite remote sensing

ABSTRACT

Estimating exposure to particulate matter (PM_{10}) air pollution concentrations in Australia is challenging due to relatively few monitoring sites relative to the geographic distribution of the population. We modelled daily ground-level PM_{10} concentrations for the period 2006–2011 for Australia using linear mixed models with satellite remote-sensed AOD, land-use and geographical variables as predictors. The variation in daily PM_{10} explained by the model was 51% for Australia overall, and ranged from 51% for Tasmania to 78% for South Australia. Cross-validation indicated that the models were most suitable for prediction in New South Wales and Victoria and least suitable for prediction in Western Australia, the Australian Capital Territory and Tasmania. Most of the variation in PM_{10} concentrations was explained by temporal rather than spatial variation. The inclusion of AOD and other predictors did not substantially improve model performance. Temporal models were sufficient to account for daily PM_{10} variation recorded by statutory monitors.

1. Introduction

Particulate matter (PM) air pollution is a major air quality issue (WHO, 2013). In Australia, it is estimated that more than 3000 people die prematurely each year as a result of air pollution (AIHW et al., 2007). Studies have linked both PM_{10} (< 10 μm aerodynamic diameter) and $PM_{2.5}$ (< 2.5 μm aerodynamic diameter) with a range of health problems including respiratory and cardiovascular morbidity (Barnett et al., 2005; Crabbe, 2012; Pereira et al., 2010a), adverse perinatal outcomes (Bell et al., 2007; Simpson, 2006) and lung cancer (Hamra et al., 2014). Major sources of PM_{10} in Australia are bushfires (Dennekamp and Abramson, 2011), dust storms (Merrifield et al., 2013), and anthropogenic combustion emissions (Pereira et al., 2010b), which vary both geographically and seasonally. To ascertain exposure, health studies generally require wide geographical coverage at sufficient temporal resolution. This underscores the need for a spatio-temporal model with a daily resolution to estimate exposure to PM_{10}.

A challenge in estimating exposure to ambient particulate matter is that in some countries, such as Australia, there are few regulatory ground monitoring sites relative to the geographic distribution of the population (Knibbs et al., 2014), which introduces considerable sample loss when populations who do not live close to a monitor are excluded to minimize exposure misclassification (Ebisu et al., 2014). This challenge can been addressed by use of land-use regression (LUR) (Hoek et al., 2008) that first uses geographically varying predictors (e.g., proximity to major roads) to fit the model with measured pollutant concentrations, and next applies that model at unmonitored locations (Ryan and LeMasters, 2007). The relatively recent addition of satellite remote sensing measurements of Aerosol Optical Depth (AOD) as a predictor to these models (Liu et al., 2005) has led to a putative improvement in their geographic accuracy. Consequently, models for PM air pollution covering large geographic regions have been developed for countries including the United States (Li et al., 2015), Canada (Hystad et al., 2011) and China (Ma et al., 2015).

However, substantial uncertainties remain. Firstly, it is unclear as to whether equally reliable estimates can be obtained in settings with relatively lower concentrations due to the lower signal-to-noise ratio. Moreover, the improvement in models for daily PM_{10} attributable to satellite remote-sensed AOD has not been quantified. It is also unknown as to the extent to which transient air pollution events (e.g., bushfire or dust storm events) affect the validity of these models. In this situation it is plausible that such events might inflate the proportion of variation

* Corresponding author.
E-mail address: gavin.f.pereira@curtin.edu.au (G. Pereira).

http://dx.doi.org/10.1016/j.envres.2017.07.044
Received 7 April 2017; Received in revised form 28 June 2017; Accepted 25 July 2017
0013-9351/

explained by model (i.e. high R^2) yet the model might not fit the non-event periods. Finally, there is no such model for daily PM_{10} in Australia.

In this study, we developed state-specific models for daily ground-level PM_{10} concentrations using satellite remote-sensed AOD and other geographic predictors for the period 2006–2011 in Australia, a country with relatively lower pollution levels. We investigated the influence of major air pollution events on model performance. We also quantified the value of including satellite remote-sensed AOD relative to more parsimonious models.

2. Methods

2.1. Data sources

2.1.1. Ground-monitored PM_{10}

PM_{10} in Australia is measured daily (average 24 h concentrations), in contrast with the United States and many other countries where it is measured every three to six days (Lee et al., 2012). We obtained, from each state's Environmental Protection Authority (EPA), daily ground-level PM_{10} measurements from 1st January 2006 to 31st December 2011 (2191 days) from 75 monitoring sites across Australia (Table 1 and Supplementary Fig. 1a–b). The PM_{10} monitoring sites were concentrated in and around capital cities, which are located near the coast. Continuous measurements of PM_{10} were obtained using Tapered Elemental Oscillating Microbalance (TEOM) monitors (PRC). Since the TEOM heats air samples (to 50 °C), it can underestimate PM10 levels when particles contain semi-volatile or volatile material (Allen et al., 1997). To account for this, PM_{10} measurements have an internal correction factor applied by the TEOM. While this method does not fully address sample loss in areas dominated by volatile particles (AQEG, 2005), it is widely-used in Australia and elsewhere for assessing compliance with regulations and in health studies. The EPA in each state runs their own quality assurance algorithms over the data before it is released. We also performed additional checks to ensure that the data values were reasonable and that there were not too many missing values, before proceeding with the analysis.

For the 75 sites over the six year period, there were 143,129PM_{10} measurements available for analysis. The highest PM_{10} concentrations were recorded during the Australian Dust Storm in New South Wales and Queensland between 22nd and 24th September 2009. The highest daily PM_{10} concentration was over 2400 μg/m^3 in Newcastle, New South Wales, on 23rd September 2009, with an average concentration of 1080 μg/m^3 across New South Wales and Queensland.

2.1.2. Satellite remote-sensed Aerosol Optical Depth

Collection 6 MODIS AOD (Level 2; 10 km resolution) was obtained from the National Aeronautics and Space Administration (NASA)'s Earth Observing System (EOS) satellites, Terra (launched in 2000) and Aqua (launched in 2002), over Australia for the period 2006–2011. AOD is a measure of light extinction (i.e., scattering and absorption) by aerosols in the atmospheric column, which makes the AOD data useful for particle concentration prediction. MODIS AOD data are retrieved every one or two days at a global scale but only in cloud-free conditions. The Terra and Aqua satellites cross the equator at 10.30 a.m. (descending orbit) and 1.30 p.m. (ascending orbit) local sun times respectively with a scanning swath of 2330 km (cross track) by 10 km (along-track at nadir) (Lee et al., 2012). Therefore, these two satellites provide the information of particle abundance at two different times, morning (Terra) and early afternoon (Aqua), indicating part of the diurnal variability in aerosol levels. Despite the difference in overpass time, same retrieval algorithms are applied to both Aqua and Terra AOD data. To have the best spatial coverage of AOD retrievals, we used AOD data products, which merged Dark Target (DT) and Deep Blue (DB) algorithms (Levy et al., 2013; Hsu et al., 2013; Sayer et al., 2013). The merged AOD data are useful for a country consisting of mixed land cover (i.e. vegetation, semi-arid, and desert areas) such as Australia. Only AOD data with the quality assurance flag of 2 and 3 (scale of 0 – 3) were selected for high data quality. More details about the DT/DB AOD data can be found in Levy et al. (2013), Hsu et al. (2013) and Sayer et al. (2013).

For each PM_{10} site, the AOD for each day was calculated as the average of the AOD values within a 10 km radius of the site (based on the 10 km nominal resolution of MODIS). AOD values were calculated separately for Aqua and Terra for each site and day. The Aqua and Terra values had non-equivalent distributions, most likely due to diurnal patterns of aerosols and calibration issues, particularly for Terra AOD (Lee et al., 2012).

Table 1
Measured PM_{10} concentrations (μg/m^3) by state/territory for 2006–2011.

State	Number of sites	Number of observations	Mean	SD	25th centile	50th centile	75th centile	99th centile
All	75	143,129	18	8	11	16	21	53
NSW	26	49,633	18	8	11	16	21	51
QLD	18	33,362	18	7	12	15	20	50
VIC	11	22,279	19	8	13	17	23	58
SA	9	17,728	18	9	11	15	22	63
WA	6	11,799	18	7	12	16	21	47
ACT	2	2543	14	8	7	11	17	50
TAS	2	4045	16	7	10	14	19	41
NT	1	1740	15	9	8	13	20	46

NSW = New South Wales, QLD = Queensland, VIC = Victoria, SA = South Australia, WA = Western Australia, ACT = Australian Capital Territory, TAS = Tasmania, NT = Northern Territory, SD = Standard Deviation (excludes values > 99th centile).

2.1.3. Explanatory (X) variables

For each PM_{10} measurement site, we obtained data on geographical and land-use variables that are potentially associated with PM_{10} concentrations. There were 14 area-level explanatory variables calculated at 25 different circular areas (buffers), four area-level variables calculated at five buffers, and 25 point-level variables, resulting in 395 explanatory variables (Table 2).

The variables related to bushfires (annual and monthly active fire and burnt area) had five buffers with radii from 10 km to 250 km (Supplementary Table 1) to give a total of 20 buffer variables (four variables calculated at five buffers each). The land-use and geographical variables had 25 buffers from 100 m to 100 km (Supplementary Table 1) giving a total of 350 buffer variables (14 variables calculated at 25 buffers each). Buffer variables were calculated using either the sum or the average of the variable within the buffer.

The 25 point variables included meteorological, elevation and distance variables and were calculated at each monitoring point. We also included a continuous variable *day* to account for longer term trend (Supplementary Fig. 2) and a categorical variable *season* (Supplementary Fig. 3) with *autumn (fall)* as the reference group. Detailed information about the variables is contained in Supplementary Table 1.

2.2. Analyses

2.2.1. Aqua versus Terra AOD values

Although Terra values can be used when Aqua is not available, in order to reduce missing values, this might introduce additional uncertainty due to the difference between the two satellite observations. Terra and Aqua AOD data reflect aerosol levels at two different time points (i.e., Terra in the morning and Aqua in the afternoon). Therefore, it may not be reasonable to use Terra AOD when Aqua AOD is not available because of the diurnal variability in aerosol levels influenced by emissions and local meteorology. The calibration issue particularly for Terra AOD can also make such an approach less appropriate.