Atmospheric Environment 184 (2018) 156–165

Contents lists available at ScienceDirect

Atmospheric Environment

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

Development of non-linear models predicting daily fine particle concentrations using aerosol optical depth retrievals and ground-based measurements at a municipality in the Brazilian Amazon region

Karen dos Santos Gonçalves[a,b,c,*], Mirko S. Winkler[a,b], Paulo Roberto Benchimol-Barbosa[d], Kees de Hoogh[a,b], Paulo Eduardo Artaxo[e], Sandra de Souza Hacon[c], Christian Schindler[a,b], Nino Künzli[a,b]

[a] *Swiss Tropical and Public Health Institute, Basel, Switzerland*
[b] *University of Basel, Basel, Switzerland*
[c] *National School of Public Health Sergio Arouca, Oswaldo Cruz Foundation – ENSP/FIOCRUZ, Rio de Janeiro, Brazil*
[d] *Clinical Coordination of Pedro Ernesto University Hospital, Rio de Janeiro State University – HUPE/UERJ, Rio de Janeiro, Brazil*
[e] *Physics Institute, University of São Paulo – IFUSP/USP, São Paulo, Brazil*

A R T I C L E I N F O

Keywords:
Aerosol optical depth
Particulate matter
Air pollution
Forest fire
Validation approach
Brazilian Amazon Region

A B S T R A C T

Epidemiological studies generally use particulate matter measurements with diameter less 2.5 µm ($PM_{2.5}$) from monitoring networks. Satellite aerosol optical depth (AOD) data has considerable potential in predicting $PM_{2.5}$ concentrations, and thus provides an alternative method for producing knowledge regarding the level of pollution and its health impact in areas where no ground $PM_{2.5}$ measurements are available. This is the case in the Brazilian Amazon rainforest region where forest fires are frequent sources of high pollution. In this study, we applied a non-linear model for predicting $PM_{2.5}$ concentration from AOD retrievals using interaction terms between average temperature, relative humidity, sine, cosine of date in a period of 365,25 days and the square of the lagged relative residual. Regression performance statistics were tested comparing the goodness of fit and R^2 based on results from linear regression and non-linear regression for six different models. The regression results for non-linear prediction showed the best performance, explaining on average 82% of the daily $PM_{2.5}$ concentrations when considering the whole period studied. In the context of Amazonia, it was the first study predicting $PM_{2.5}$ concentrations using the latest high-resolution AOD products also in combination with the testing of a non-linear model performance. Our results permitted a reliable prediction considering the AOD-$PM_{2.5}$ relationship and set the basis for further investigations on air pollution impacts in the complex context of Brazilian Amazon Region.

1. Introduction

In spite of the efforts to improve air quality during the past decades, levels of air pollution experienced by human populations continue to cause a large burden of disease (Cohen et al., 2005; Brauer et al., 2015; Global Burden of Diseases (GBD), 2010). Atmospheric aerosols and particulate matter that are breathable (< 2.5 µm diameter = $PM_{2.5}$) and inhalable (< 10 µm = PM_{10}), generated from natural and anthropogenic emission sources present known effects for a number of causes of death, particularly the increase in cardio-respiratory diseases in areas with high concentrations (Brook et al., 2010; World Health Organization (WHO), 2014).

Intensive and indiscriminate occurrence of forest fire has become a serious environmental problem in Brazil, affecting ecosystems' balance and human health with consequences at the local, regional and global level (Gonçalves et al., 2012; Becker, 2005). Brazilian Amazon region has geographic and environmental circumstances that are distinct from other world regions. For this reason, the occurrence of fire and emissions of $PM_{2.5}$ exposes every year increasingly large portions of vulnerable populations (Fearnside, 2005; Goncalves et al., 2014).

To understand the association between $PM_{2.5}$ and effects on human health, epidemiological studies have employed $PM_{2.5}$ measurements from monitoring sites. However, due to cost and lack of appropriate infrastructure, especially in rural and remote areas, no fixed site $PM_{2.5}$ measurements are available in many regions of Brazil. This is a major limitation for estimating exposure to $PM_{2.5}$ and assessing health

* Corresponding author. ENSP/FIOCRUZ: Rua Leopoldo Bulhões, 1480, CEP: 21041-210, Manguinhos, Rio de Janeiro, Brazil.
E-mail address: karengoncalves@ensp.fiocruz.br (K.d.S. Gonçalves).

https://doi.org/10.1016/j.atmosenv.2018.03.057
Received 19 June 2017; Received in revised form 19 March 2018; Accepted 28 March 2018
Available online 31 March 2018
1352-2310/

impacts associated with forest fires as one of its major source (Lee et al., 2011; Ruckerl et al., 2011; Ye et al., 2011; Yi et al., 2010; Arbex et al., 2009; McMichael et al., 2008).

An alternative approach to estimate the air quality in areas without direct $PM_{2.5}$ measurements is by means of satellite remote sensing using aerosols optical depth (AOD). AOD is an electromagnetic radiation measure and reflects the integrated number of particles at a given wavelength. It is an important satellite-retrieved property for predicting the $PM_{2.5}$ concentrations due repeated observations of the atmosphere and its extensive spatial coverage (Kloog et al., 2014). The AOD has been successfully used in statistical models for estimating $PM_{2.5}$ levels. As shown by previous studies, parameters such as local meteorology and land use information influence the relationship between AOD and daily $PM_{2.5}$ concentrations, which need to be considered as additional predictors (Lee et al., 2011; Liu et al., 2004, 2005, 2007a, 2007b, 2007c, 2009; Hoff and Christopher, 2009; Xie et al., 2015).

Traditionally, the health exposure studies have used the standard MODIS (Moderate Resolution Imaging Spectroradiometer) AOD product of the "Dark Target" algorithm published by Levy et al. (2007, 2010), which has a resolution of $10 \times 10\,km^2$. Later, Remer et al. (2013, 2005) described AOD algorithm applying a higher resolution of $3 \times 3\ Km^2$. (Remer et al., 2005, 2013; Levy et al., 2007, 2010).

Concerning the applicability of the statistical methods for predicting $PM_{2.5}$ concentration using AOD retrievals, de Hoogh et al. (2017) used a higher spatial resolution for modelling daily $PM_{2.5}$ concentrations across Switzerland during the period between 2003 and 2013. Their models result explained on average 73% of the total,71% of the spatial and 75% of the temporal variation (all cross validated) in measured $PM_{2.5}$ concentrations. Kloog et al. (2012) described a new hybrid spatio-temporal model for estimating daily $PM_{2.5}$ concentrations across northeastern USA using high resolution AOD data. Their results showed a high predictive accuracy at high spatial resolutions using a mixed model regressing $PM_{2.5}$ measurements with an excellent model performance (R2 = 0.88).

These recent studies still have the challenge of reducing exposure error, although shows better fits than previous models. In spite our model showed a good performance, it is important to reproduce it in another region with different meteorological and geographical patterns. Our model can be applied to other sites if site-specific AOD and meteorological data are available, to be inserted into the prediction equation. The lagged relative residual added as a further predictor variable it is cautious strategy to remove the serial autocorrelation and to further improve the model. As another important challenge is that AOD data availability is much greater in the dry seasons compared to the rainy period. This is mostly due to heavily clouded days which results in missing AOD data. This non-random lack of AOD readings could negatively affect predictive performance. Also, treating large areas, such as Brazilian Amazon region, can add additional selection bias since there may be meteorological variations in the daily calibration between PM2.5 and AOD (Kloog et al., 2012).

In this paper we developed a non-linear model predicting daily fine particle concentrations using AOD retrievals at 3×3 km resolution and ground-based measurements at a municipality of Porto Velho, Brazil during the period between 2009 and 2011. For Brazilian Amazon region, it is the first study to develop this approach considering a non-linear model predicting PM2.5 concentrations. This study assessment is part of an investigation that aims at analysing the impact of PM2.5 exposure on cardiovascular disease in Porto Velho.

2. Material and methods

2.1. Ground-level $PM_{2.5}$ data

Daily averages were derived during the period from 25 September 2009 to 21 October 2011 with a total of 757 days. Over the study period, $PM_{2.5}$ concentrations were measured for 24 h at one air quality monitoring station in Porto Velho municipality, which was implanted in partnership between Institute of Physics at University of São Paulo (USP), University of Rondônia (UNIR), Environmental Biogeochemistry Laboratory Wolfgang H. Pfeiffer and National School of Public Health-Oswaldo Cruz Foundation (FIOCRUZ) in Brazil. The $PM_{2.5}$ monitor is located at 15 km north of to the centre of urban area (Fig. 1). Porto Velho municipality is the third capital in the Brazilian Amazon region with 67 districts within the urban area. With an area of 34,096 km^2 Porto Velho has a population of 503,000 inhabitants according to *Brazilian Institute of Geography and Statistics* (IBGE, Census 2010). $PM_{2.5}$ measurements were collected by means of a stacked filter unit (SFU) and were analysed gravimetrically according to the World Health Organization Air Quality Guidelines for particulate matter, ozone, nitrogen dioxide and sulfur dioxide (World Health Organization (WHO), 2006).

This methodology involves the sampling site (8.69° S, 63.87° W) located in a region with large land use changes and associated regional biomass burning. The SFU (Stacked Filter Unit) type samplers and the analysis follows routine gravimetric techniques[33]. In addition, trace elements and ionic compounds are collected, allowing for future analyses.

There is an AFG sampler, which collects aerosols for elemental PIXE and black carbon analyses on the roof of the shelter, 24 h sampling. The collection of aerosol particles using filters is a simple and very common method for sampling aerosol particles. Filters allow elemental and ionic analysis through a series of measurement techniques. The sampler collects fine and coarse particles and contains an inlet that allows the entry of particles in the range of $2 < Dp < 10\,\mu m$. The filters are polycarbonate, having a diameter of 47 mm and are arranged in series. In the first step the particles of the coarse fraction are retained using Nucleopore filters with pores of 8 μm in diameter, in the second stage, they are the fine particles that are retained using the filter Nucleopore with pores of 0.4 μm. The samples collected with the AFG sampler was used to determine the mass of the aerosols by means of gravimetric analysis, the concentration of black carbon and to quantify the elemental concentration of the material deposited in the filters.

2.2. MODIS 3 km AOD retrieval

The Moderate Resolution Imaging Spectroradiometer (MODIS) is a key instrument aboard the Terra and Aqua satellites of the National Aeronautics and Space Administration (NASA) and has been in operation since 1999 and 2002, respectively. While Terra passes the equator in the morning, from north to south, Aqua passes the equator from south to north in the afternoon. These satellites were used to retrieve AOD aerosol products with a 3 km resolution (MOD04_3K and MYD04_3K), operating at an altitude of approximately 700 km (http://modis-atmos.gsfc.nasa.gov/). In the Collection 6, Level 2 aerosol products, the most recent 3 km AOD dark target retrieval algorithm is similar to the 10 km standard product (Collection 5, Level 2) and has three different wavelength channels of 0.47, 0.66 and 2.12 μm employed for AOD retrieval over land. The other channels are used for screening procedures (e.g., coverage of cloud, snow and ice) (Remer et al., 2013; Levy et al., 2007; Munchak et al., 2013). More details on the retrieval of MODIS satellite aerosol data have previously been published by Remer et al. (2013, 2005) and Levy et al. (2007, 2010). For the AOD daily averages, we used the algorithm retrieval in MATLAB (version 2015a, MathWorks) and the software ArcGIS (version 10, ESRI) to create 820 grid cells of 3 x 3-km covering the study area for spatial analyses.

2.3. Statistical model and validation

In this study we considered five different types of prediction models of $PM_{2.5}$ concentrations from AOD retrievals. They were all of the form $PM_{2.5}$ = exp(linear predictor), with the linear predictor involving terms

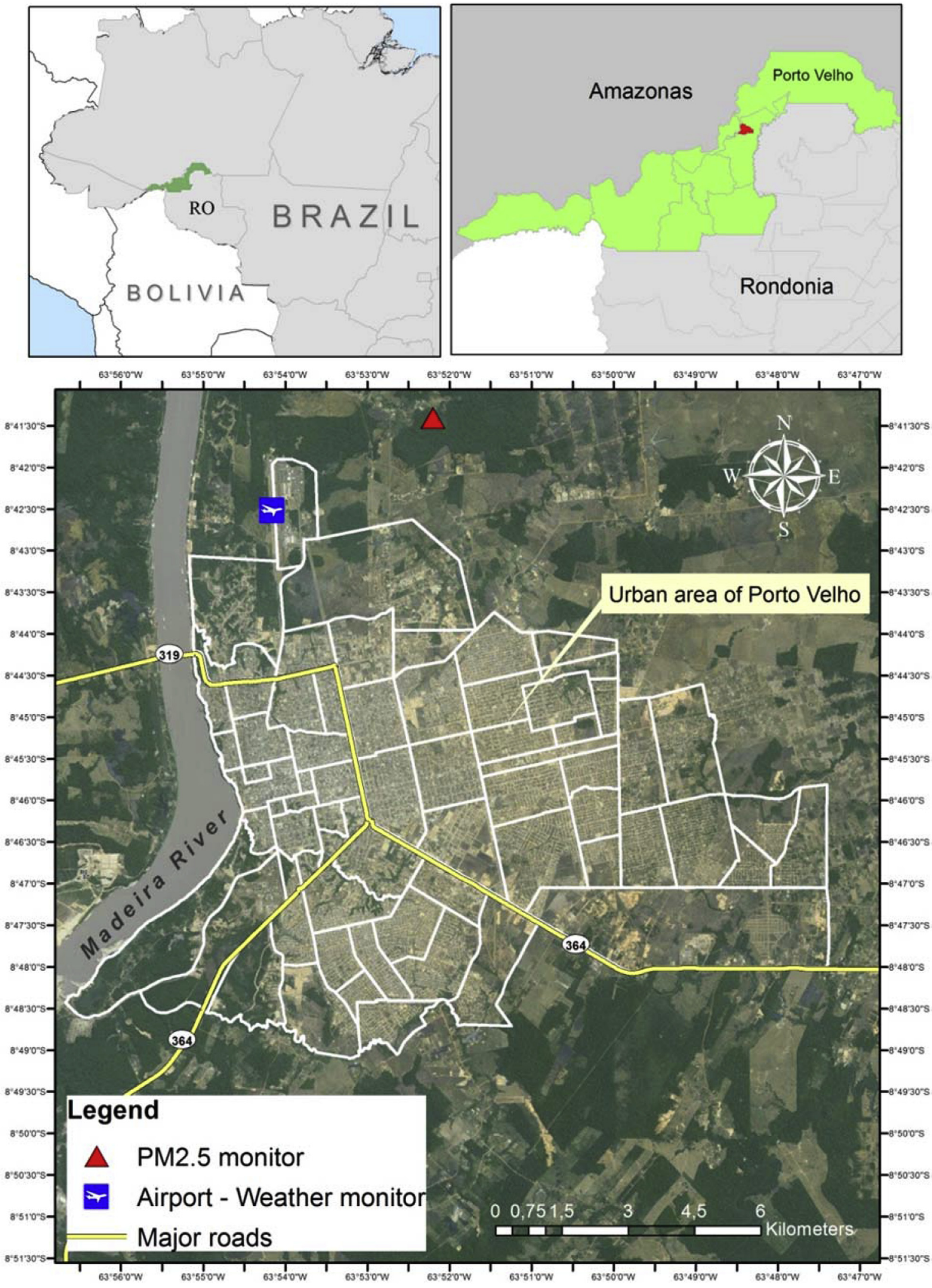

Fig. 1. Study area with the locations of $PM_{2.5}$ and weather monitors. Municipality of Porto Velho, Rondônia state, Brazilian Amazon region.

composed of AOD and other influencing factors. The advantage of such a model over one for log-transformed outcome data is that it provides estimates of mean exposure levels while estimates of geometric mean levels are obtained when modelling log-transformed outcome data and then exponentiating the resulting predictions (which are on the log-scale). In Model 1, we took the linear predictor to be a cubic polynomial in AOD with time-independent coefficients:

(Model 1)

$$PM_{2.5} = \exp(\alpha + \beta_1(AOD) + \beta_2 (AOD)^2 + \beta_3 (AOD)^3)$$

In a next step, we considered the coefficients of Model 1 to be polynomials of second degree in average temperature (TEMP) and relative humidity (RH). Thus, each of the coefficients was assumed to be of the form $\gamma_0 + \gamma_1 \cdot TEMP + \gamma_2 \cdot RH + \gamma_3 \cdot TEMP^2 + \gamma_4 \cdot RH^2 + \gamma_5 \cdot TEMP \cdot RH$. Multiplying these polynomials with AOD, AOD^2 and AOD^3 and the intercept, respectively, provided:

(Model 2) $$PM_{2.5} = \text{Model 1} \cdot \exp(\beta_4(AOD \cdot TEMP) + \beta_5(AOD \cdot TEMP^2) + \beta_6(AOD \cdot RH) + \beta_7(AOD \cdot (RH^2) + \beta_8(AOD \cdot TEMP \cdot RH) + \beta_9(AOD^2 \cdot TEMP) + \beta_{10}(AOD^2 \cdot TEMP^2) + \beta_{11}(AOD^2 \cdot RH) + \beta_{12}(AOD^2 \cdot RH^2) + \beta_{13}(AOD^2 \cdot TEMP * RH) + \beta_{14}(AOD^3 \cdot TEMP) + \beta_{15}(AOD^3 \cdot TEMP^2) + \beta_{16}(AOD^3 \cdot RH) + \beta_{17}(AOD^3 \cdot RH^2) + \beta_{18}(AOD^3 \cdot TEMP \cdot RH))$$

In an attempt to further improve the model, we added interaction terms between AOD, AOD^2 and AOD^3 with rainy season (Model 3) and tested interactions of the three AOD-terms with sine and cosine functions of date with a period of 365.25 days in the Model 4.

In the last step, we included the lagged relative residual and its square as additional predictor variables (Model 5). This was to reduce serial autocorrelation. The final model obtained after some backward elimination steps was of the form:

(Final model)

$$PM_{2.5} = \exp(\alpha + \beta_1 \cdot AOD + \beta_2 \cdot AOD^2) + \beta_3 \cdot AOD^3 + \beta_4 \cdot (AOD \cdot TEMP) + \beta_5(AOD \cdot RH) + \beta_6(AOD \cdot TEMP^2) + \beta_7(AOD \cdot RH^2) + \beta_8(AOD \cdot cos\_days) + \beta_9(AOD \cdot sin\_days) + \beta_{10}(AOD^2 \cdot cos\_days) + \beta_{11}(AOD^2 \cdot sin\_days) + \beta_{12}(AOD^3 \cdot cos\_days) + \beta_{13}(AOD^3 \cdot sin\_days) + \beta_{14}(residual) + \beta_{15}(residual)^2$$

In these equations, $PM_{2.5}$ denotes the predicted concentrations, where *exp* is the exponential function; *cos_days* and *sin_days* denote the cosine and sine terms of date with a period of 365,25 days; *residual* denotes the lagged relative residual. Relative residuals were defined as ratio between residuals and predicted values. Model performance was evaluated by comparing the predictions with the ground measurements using the adjusted coefficient of determination (R^2 adj), residual standard deviation (RMSE), Akaike's information criterion (AIC), and partial autocorrelation of residuals by lags. High values of adjusted R squared suggest that MODIS AOD data can be used to estimate ambient concentrations. Furthermore, we calculated mean, standard deviation and maximum/minimum values to summarize the descriptive statistics of our sample for the whole period, the dry season (months June to November, when forest fires occur in Brazilian Amazon region) and the rainy season (months from December to May).

Table 1
Descriptive statistics of the parameters analysed during the study period (September 25th, 2009 to October 21th, 2011).

Variable	Entire period (25 September 2009 to 21 October 2011)				
	N[a]	Mean	SD[b]	Min	Max
MODIS AOD 3 km (unitless)	649	0,29	0,36	0,03	2,19
$PM_{2.5}$ (μg/m^3)	649	11,36	20,06	1.68	164,41
Average temperature (°C)	649	26,82	1,41	16,24	31,26
Relative humidity (%)	649	84,93	5,80	61,50	98,75
Preciptation (mm)	649	5,06	11,37	0	71,40
	Dry season (June to November)				
Variable	**N**	**Mean**	**SD**	**Min**	**Max**
MODIS AOD 3 km (unitless)	323	0,44	0,46	0,03	2,19
$PM_{2.5}$ (μg/m^3)	323	20,51	25,19	1.68	164,41
Average temperature (°C)	323	27,00	1,68	16,24	31,26
Relative humidity (%)	323	82,18	5,90	61,50	98,75
Preciptation (mm)	323	0	0	0	0
	Rainy season (December to May)				
Variable	**N**	**Mean**	**SD**	**Min**	**Max**
MODIS AOD 3 km (unitless)	326	0,14	0,07	0.03	0,76
$PM_{2.5}$ (μg/m^3)	326	2,28	2,87	1,68	26,62
Average temperature (°C)	326	26,66	1,07	22,07	29,05
Relative humidity (%)	326	87,65	4,21	73,80	98,00
Preciptation (mm)	326	7,76	14,09	0,10	71,40
	Pearson correlation				
Variable	**AOD**	**PM2.5**	**TEMP**	**RH**	**PRECIP**
MODIS AOD 3 km (unitless)	1				
$PM_{2.5}$ (μg/m^3)	0.5811	1			
Average temperature (°C)	0.245	0.1007	1		
Relative humidity (%)	−0.3957	−0.4455	−0.4411	1	
Preciptation (mm)	−0.101	−0.1454	−0.1638	0.2413	1

Note: a = Number of values observed in days; b = Standard deviation (SD).

Table 2
Comparison between prediction models.

Prediction models	N	df[a]	R^2	R^2 adj[b]	RMSE[c]
Model 1	649	3	0.544	0.541	13.59
Model 2	649	18	0.683	0.674	11.47
Model 3	649	21	0.707	0.697	11.04
Model 4	649	27	0.782	0.772	9.58
Model 5	649	15	0.823	0.816	8.60

Note: a = degrees of freedom; b = R-squared adjusted; c = Residual standard deviation (RMSE).
Model 1 = simple model with linear, quadratic and cubic term of AOD.
Model 2 = Model 1 + interactions of AOD, AOD^2 and AOD^3 with linear and quadratic terms in temperature and relative humidity; Model 3 = Model 2 + interactions between AOD, AOD^2 and AOD^3 and rain.
Model 4 = Model 3 + interactions of AOD, AOD^2 and AOD^3 with sine and cosines of date with a period of 365.25 days (without the term for rainy season).
Model 5 = Model 4 + lagged relative residual and its square as additional predictor variables.

Daily meteorological data on average temperature and relative humidity were obtained for the period 25 September 2009 to 21 October 2011 from the monitoring station of INMET (*National Meteorology Institute*) in Porto Velho. The information is publicly available on the website of the institute (www.inmet.gov.br).

The spatial distribution of the 3 × 3km resolution MODIS AOD average during the study period was derived by spatial interpolation using the inverse distance weighting (IDW). We present the results for all states in Brazil and for our study area. It is important to highlight that all regression results were presented with the original AOD and $PM_{2.5}$ datasets. The software R (version 3.1.3) was used for statistical analyses.

3. Results

3.1. Descriptive statistics

Table 1 shows the descriptive statistics of daily measured $PM_{2.5}$, values of AOD, relative humidity, average temperature and precipitation from 25 September 2009 to 21 October 2011, as well as for the dry and rainy seasons. Average daily level of $PM_{2.5}$ from the ground-level monitor was 11 μg/m³ with a standard deviation (SD) of ± 20 μg/m³ over the three years studied. Of note, the analysis of the data for 2010, the year when one of the most extreme dry seasons in Brazilian Amazon region occurred, revealed an annual-mean of 36 μg/m³ (± 46 μg/m³ SD).

Considering the differences between seasons, the maximum daily value was exceptionally high (164 μg/m³) during the dry seasons of the period studied compared to 27 μg/m³ in the rainy seasons.

Over the entire study period, the daily AOD values observed varied from 0.03 to 2.19. On average, 649 AOD values were retrieved per grid cell which corresponds to 86% of the entire study period of 757 days.

All the meteorological variables such as relative humidity, average temperature and precipitation were consistent with the climatic patterns expected for the Brazilian Amazon region and thus support the analysis in the regression models.

3.2. Non-linear prediction models

To test the performance of the five regressions models we use a total of 649 valid days for the model fitting. The comparisons between the models analysed and parameters estimated are shown in Tables 2 and 3

Model 1 shows an adj R^2 of 0.54, RMSE of 13.59 μg/m³ and AIC of 5 234.3 for the whole period. Model 2 including interactions between AOD, AOD^2 and AOD^3 and linear and quadratic terms in temperature and relative humidity provided a better fit (R^2 = 0.67). After adding interactions between the three AOD-terms and rain the fit only slightly improved (R^2 = 0.70). In Model 4 we excluded the rain term and added interactions of the three AOD-terms with sine and cosine of date with a period of 365.25 days. This model performed considerably better (R^2 = 0.77; RMSE = 9.59 μg/m³; AIC = 4 803.1).

After adding the lagged relative residual and its square as additional predictors (Model 5) the adjusted R2 further increased to 0.82 (RMSE = 8.60 μg/m³; AIC = 4 797.6). This means that this non-linear prediction model explains 82% of the variance of daily $PM_{2.5}$ concentrations in combination with meteorological and seasonal variables. The introduction of these two terms also led to a drastic reduction in residual autocorrelation, in that lag1-autocorrelation of residuals was no longer significant (Fig. 2). As visualized in Fig. 2, the time series of predicted $PM_{2.5}$ concentrations follows a very similar pattern as the measured $PM_{2.5}$ confirming the high performance of the prediction models. The period from mid-July to end of October 2010 – a dry period with plumes of biomass burning – is characterized by very high AOD values, reaching peaks 50–100 times above the typical values observed before and after this period. The comparisons between the measured and predicted $PM_{2.5}$ concentrations for Model 1 and Model 5 are illustrated in Fig. 3.

The Spatial distribution of $PM_{2.5}$ predicted concentration over the basin during different seasons for all Brazilian states are shown in Fig. 4. The highest predicted $PM_{2.5}$ concentrations were observed in the Brazilian Amazon region during the forest fires season (months between September, October and November). In our study area, $PM_{2.5}$ averages reached 44 μg/m³ in the urban area of Porto Velho, and 54 μg/m³ across the Rondônia state during the forest fires between 2009 and 2011. Information about the distribution of $PM_{2.5}$ within the urban districts of Porto Velho and the relation with the health data will be presented in a separate manuscript about the impacts of $PM_{2.5}$ on human health in Brazilian Amazon region.

4. Discussion

The results of our final non-linear prediction model for $PM_{2.5}$ showed a good performance, explaining on average 82% of the variance in measured $PM_{2.5}$ concentrations during the period studied. This result is similar and in accordance with the findings presented by Lee et al. (2011) and Xie et al. (2015), who showed prediction models that explained 92% and 82% of the variance in $PM_{2.5}$ concentrations in the North-eastern, US and in Beijing, China, respectively. Our model has the advantage that it does not produce negative predictions and fits the mean of the data as a function of the predictor variables. Moreover, by including the lagged relative residual and its square as additional predictor variables it was possible to remove the significant lag1-autocorrelation, and to further improve the model fit.

Observing the temporal distribution of predicted and measured $PM_{2.5}$ concentrations it is important to highlight the enormous peak of $PM_{2.5}$ observed between days 300 and 400 during the dry periods in our study area with the maximum daily value of 164 μg/m³. This value more than 6.5 times higher than the daily mean guideline value proposed by WHO to protect public health (25 μg/m³). During the dry season, this value was exceeded on X% of all days. As a consequence, the long-term mean concentration during the dry period was 2 times above the WHO annual mean guideline value, set at 10 μg/m³. Currently, these values were adopted in only a few countries as legally binding targets, thus, policy makers accept major impacts on morbidity and mortality.[34] On the other hand, during the rainy seasons concentrations were low (2 μg/m³; ± 3 μg/m³ SD) and fully in line with both, the daily and annual targets proposed by WHO (Hopke et al., 1997; Künzli et al., 2015). This confirms the dominant role of fires as source of ambient air pollution in the Amazon region. This result highlight the importance to set limits for $PM_{2.5}$ in the Brazil Air quality Standards defined by the National Environmental Agency (CONAMA) that currently set limits only for PM_{10}.[34] Our data provide unique input

Table 3
Description of parameters, standard error and p-value for each prediction models.

Model 1	Estimate	Std. Error	t value	Pr(> |t|)
β0	113.565	0.5333	21.295	< 2e-16 ***
β1	3.733.880	135.860	27.483	< 2e-16 ***
β2	369.936	135.860	2.723	0.00665 **
β3	319.761	135.860	2.354	0.01889 *

Model 2	**Estimate**	**Std. Error**	**t value**	**Pr(> |t|)**
β0	1,10E+06	1,89E+05	5.803	1.03e-08 ***
β1	8,13E+07	9,91E+06	8.202	1.33e-15 ***
β2	3,19E+07	8,50E+06	3.751	0.000193 ***
β3	2,32E+06	5,36E+06	0.432	0.665709
β4	8,85E+04	7,41E+04	1.195	0.232703
β5	−1,62E+04	1,26E+04	−1.286	0.198907
β6	−2,44E+03	9,76E+02	−2.503	0.012569 *
β7	4,02E+02	4,75E+02	0.847	0.397177
β8	−4,14E+05	2,03E+05	−2.046	0.041196 *
β9	2,82E+04	3,92E+04	0.720	0.472077
β10	9,39E+03	3,01E+03	3.119	0.001895 **
β11	5,05E+01	1,17E+02	0.431	0.666751
β12	−1,19E+03	1,17E+03	−1.017	0.309673
β13	9,79E+04	9,91E+04	0.988	0.323655
β14	−5,25E+04	2,15E+04	−2.445	0.014744 *
β15	−3,73E+03	1,53E+03	−2.437	0.015070 *
β16	9,53E+01	7,64E+01	1.247	0.212722
β17	1,34E+03	5,54E+02	2.420	0.015795 *

Model 3	**Estimate**	**Std. Error**	**t value**	**Pr(> |t|)**
β0	1,09E+06	1,85E+05	5.870	7.05e-09 ***
β1	8,22E+07	9,58E+06	8.584	< 2e-16 ***
β2	3,36E+07	8,29E+06	4.057	5.60e-05 ***
β3	2,95E+06	5,21E+06	0.566	0.571754
β4	8,48E+04	7,17E+04	1.183	0.237140
β5	−1,68E+04	1,22E+04	−1.379	0.168533
β6	−2,47E+03	9,43E+02	−2.616	0.009109 **
β7	5,13E+02	4,60E+02	1.116	0.264968
β8	−4,16E+05	1,96E+05	−2.123	0.034177 *
β9	4,21E+04	3,87E+04	1.088	0.276841
β10	9,67E+03	2,91E+03	3.325	0.000935 ***
β11	−1,90E+01	1,19E+02	−0.159	0.873546
β12	−1,46E+03	1,14E+03	−1.285	0.199287
β13	8,94E+04	9,59E+04	0.932	0.351465
β14	−5,79E+04	2,12E+04	−2.734	0.006433 **
β15	−3,69E+03	1,48E+03	−2.498	0.012757 *
β16	1,20E+02	7,65E+01	1.572	0.116357
β17	1,46E+03	5,39E+02	2.708	0.006950 **
β18	−1,01E+05	2,00E+04	−5.063	5.44e-07 ***
β19	3,62E+05	1,18E+05	3.063	0.002283 **
β20	−2,58E+05	1,33E+05	−1.935	0.053453.

Model 4	**Estimate**	**Std. Error**	**t value**	**Pr(> |t|)**
β0	5,55E+05	1,68E+05	3.305	0.001003 **
β1	3,85E+07	1,04E+07	3.692	0.000242 ***
β2	4,62E+06	9,84E+06	0.469	0.638882
β3	−6,61E+06	5,32E+06	−1.245	0.213777
β4	1,04E+05	6,25E+04	1.659	0.097687.
β5	1,28E+02	1,08E+04	0.012	0.990548
β6	−1,88E+03	8,21E+02	−2.291	0.022301 *
β7	−1,21E+02	4,08E+02	−0.296	0.767175
β8	−3,37E+05	1,72E+05	−1.956	0.050898 .
β9	−5,94E+04	4,05E+04	−1.467	0.142785
β10	4,84E+03	2,56E+03	1.893	0.058828 .
β11	2,10E+02	1,19E+02	1.773	0.076650 .
β12	9,88E+02	1,11E+03	0.894	0.371821
β13	9,74E+04	8,57E+04	1.136	0.256252
β14	2,08E+04	2,51E+04	0.832	0.405926
β15	−1,46E+03	1,30E+03	−1.129	0.259177
β16	−8,98E+01	8,23E+01	−1.091	0.275639
β17	−2,22E+02	5,79E+02	−0.383	0.701887
β18	−1,24E+04	2,54E+04	−0.488	0.625602
β19	7,89E+04	1,19E+05	0.664	0.506904
β20	−9,50E+04	1,24E+05	−0.767	0.443244

Table 3 (*continued*)

Model 1	Estimate	Std. Error	t value	Pr(> |t|)
β21	6,99E+04	1,49E+04	4.681	3.51e-06 ***
β22	−3,71E+04	7,47E+03	−4.966	8.83e-07 ***
β23	−2,64E+05	5,20E+04	−5.068	5.31e-07 ***
β24	4,90E+04	2,17E+04	2.257	0.024378 *
β25	1,08E+05	3,06E+04	3.545	0.000423 ***
β26	−3,07E+04	1,37E+04	−2.244	0.025199 *

Model 5	**Estimate**	**Std. Error**	**t value**	**Pr(> |t|)**
β0	7,46E+03	1,19E+03	6.266	6.92e-10 ***
β1	1,18E+06	1,62E+06	0.732	0.464505
β2	4,13E+06	3,97E+06	1.041	0.298197
β3	−1,48E+06	1,97E+06	−0.754	0.451224
β4	5,84E+04	6,22E+04	0.938	0.348512
β5	−4,32E+04	2,33E+04	−1.858	0.063575 .
β6	−1,53E+03	7,45E+02	−2.049	0.040855 *
β7	2,09E+02	8,55E+01	2.443	0.014849 *
β8	2,10E+02	4,51E+02	0.465	0.641789
β9	−3,03E+05	1,66E+05	−1.833	0.067247 .
β10	5,14E+01	5,64E+04	0.001	0.999274
β11	4,05E+03	2,31E+03	1.750	0.080635 .
β12	−1,70E+02	2,04E+02	−0.833	0.404931
β13	1,11E+03	1,13E+03	0.981	0.327124
β14	9,58E+04	8,07E+04	1.187	0.235520
β15	2,21E+03	2,94E+04	0.075	0.939958
β16	−1,14E+03	1,17E+03	−0.975	0.329859
β17	5,66E+01	1,06E+02	0.533	0.594386
β18	−4,24E+02	5,72E+02	−0.741	0.459242
β19	7,01E+04	7,96E+03	8.812	< 2e-16 ***
β20	−3,67E+04	5,74E+03	−6.389	3.27e-10 ***
β21	−2,55E+05	3,53E+04	−7.223	1.49e-12 ***
β22	5,66E+04	1,77E+04	3.195	0.001470 **
β23	1,18E+05	2,26E+04	5.234	2.27e-07 ***
β24	−4,10E+04	1,17E+04	−3.504	0.000491 ***
β25	4,55E+02	3,75E+01	12.129	< 2e-16 ***

Model 6	**Estimate**	**Std. Error**	**t value**	**Pr(> |t|)**
β0	2,35E+03	9,58E+01	24.568	< 2e-16 ***
β1	−1,12E+04	5,55E+03	−2.019	0.0439 *
β2	2,02E+04	2,16E+03	9.358	< 2e-16 ***
β3	−9,94E+03	9,56E+02	−10.405	< 2e-16 ***
β4	1,71E+02	3,39E+02	0.503	0.6154
β5	5,82E+01	7,62E+01	0.764	0.4454
β6	−3,13E+00	6,32E+00	−0.495	0.6211
β7	−4,23E-01	4,83E-01	−0.875	0.3817
β8	8,42E+03	9,67E+02	8.708	< 2e-16 ***
β9	−6,59E+03	4,36E+02	−15.117	< 2e-16 ***
β10	−2,08E+04	1,97E+03	−10.561	< 2e-16 ***
β11	1,01E+04	8,26E+02	12.217	< 2e-16 ***
β12	1,00E+04	8,98E+02	11.130	< 2e-16 ***
β13	−4,31E+03	3,51E+02	−12.275	< 2e-16 ***
β14	5,57E+02	4,34E+01	12.841	< 2e-16 ***
β15	−8,85E+01	1,11E+01	−8.012	5.45e-15 ***

Note: Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1.

to evaluate whether the CONAMA may add to the Brazil Air Q S limits for $PM_{2.5}$ rather than PM_{10} alone. This will be particularly worth if sources and spatio-temporal patterns of the two markers of air pollution largely vary across Brazil.

The non-linear prediction model demonstrated a high performance in predicting the daily $PM_{2.5}$ concentrations. However, some limitations, such as cloud properties and uncertainties need to be mentioned. The use of only one air monitoring station for the development and evaluation of the model is a major limitation of this study. However, although this limits our ability to draw firm conclusions about the applicability of the model across Brazil, it does provide a valid approach to predict $PM_{2.5}$ across our main study area. $PM_{2.5}$ is spatially rather homogenously distributed, thus extrapolation of the model from the measurement site to the adjacent urban area of Porto Velho is expected to be reliable. The model gives also good indications of possible hot

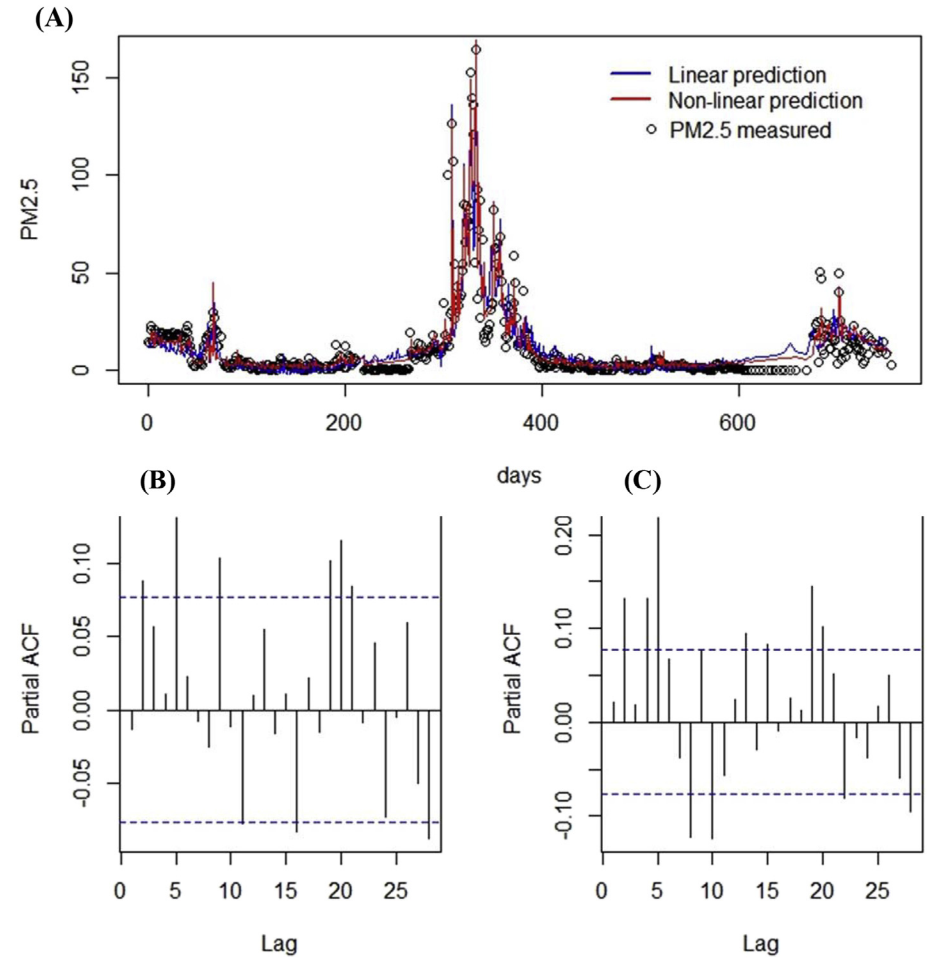

Fig. 2. (A) Comparisons between $PM_{2.5}$ measured and $PM_{2.5}$ predictions (μg/m³) across time (25 September 2009 to 21 October 2011). (B) Partial autocorrelation plot of residuals before introducing the lagged relative residual and its square as additional predictor variables into the model (C) Partial residual autocorrelation plot after adding the two variables to the model.

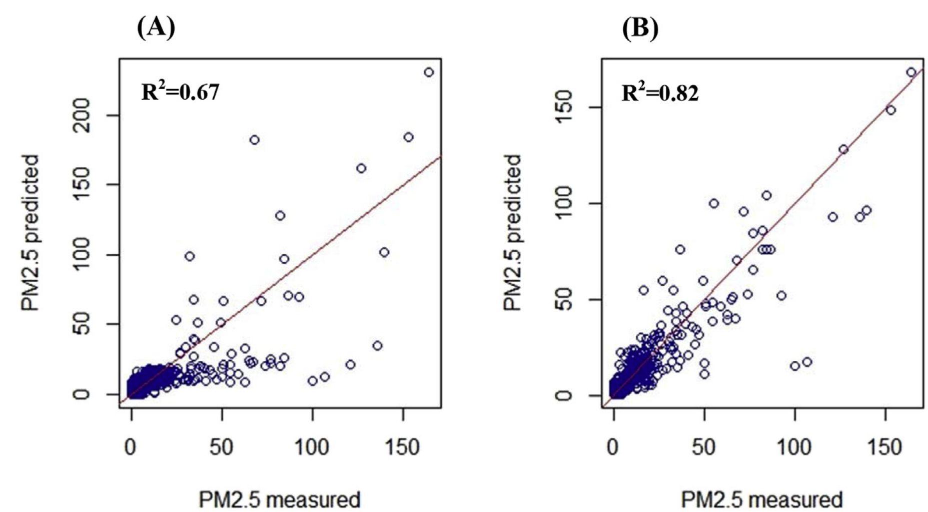

Fig. 3. Comparisons between the measured and predicted $PM_{2.5}$(μg/m³) for (A) Model 2 and (B) non-linear prediction (Final model). The red line represents the regression line. (For interpretation of the references to colour in this figure legend, the reader is referred to the Web version of this article.)

spots of pollution where it may be worth installing additional monitors operating continuously or at least during dry seasons. With the use of mobile stations, one could characterize the spatial pattern of air pollution across a larger area with only one or a few monitors while using the current central monitoring station as a reference point to understand the temporal variation. The installation of air quality monitoring networks is an important step for the future evaluation of progress in clean air management and the assessment of its health impact.

The predictions were based on different spatial scales which may be

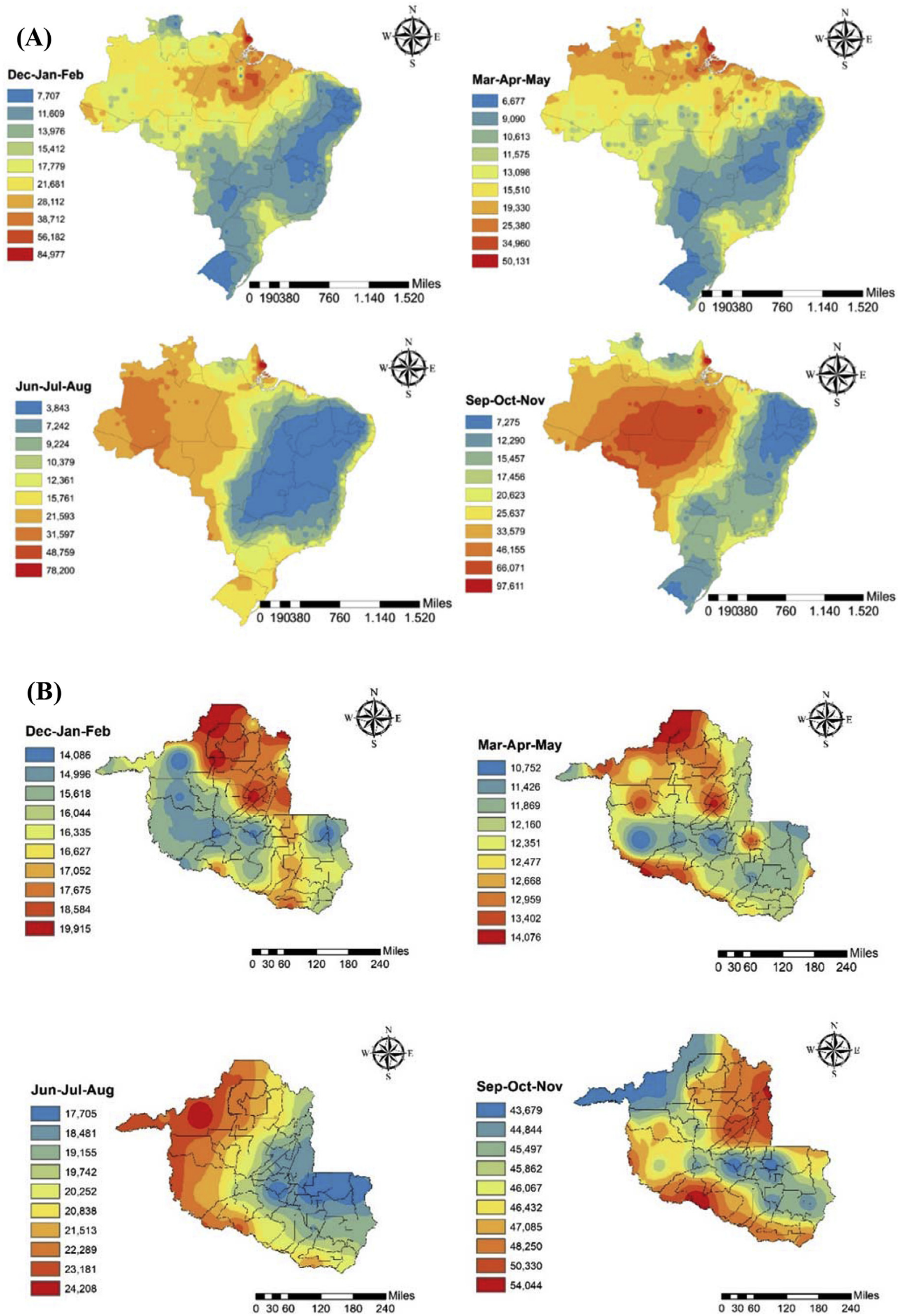

Fig. 4. Spatial distribution of $PM_{2.5}$ predicted concentration over the basin during different seasons. Forest fires occurs between the months of September, October and November. (A): $PM_{2.5}$ averages predicted concentration interpolated to all Brazilian states; (B): Rondônia State including the study area of Porto Velho.

a source of uncertainties. The AOD satellite data is based on a grid cell with 3 km resolution while $PM_{2.5}$ ground-level is measured at a fixed point. It is also important to highlight the complex relationship between AOD and $PM_{2.5}$ due the nature of the forest aerosol in Brazilian Amazon. Other important predicting factors such as wind speed, atmosphere or physical-chemical components were not analysed in this study.

Despite the uncertainties, this study is the first to predict $PM_{2.5}$ concentrations using a non-linear prediction model and the higher-resolution MODIS AOD products in Porto Velho. By modelling the AOD-$PM_{2.5}$ relationship in a time-dependent manner reflecting seasonal fluctuations and influences of temperature and relative humidity we were able to develop a prediction model for PM25 with good fitting properties.

Our model also can be applied to other sites of the region if site-specific AOD- and meteorological data are available, to be inserted into the prediction equation. On the other hand, as measured $PM_{2.5}$ data are only available for the reference station, the relative lag1-residuals of $PM_{2.5}$ at the reference site would have to be used for all other sites. About the applicability of the model to other sites, the Model 5 (not involving lagged residuals) could be applied to any other site of the Amazon region, as AOD-values and estimates of meteorological parameters can be obtained for any such site based on satellite data. On the other hand, model 6 involves an input variable which is only available at the reference site and whose values would therefore have to be used at all other sites. A cautious strategy might therefore consist in using both model 5 and 6 when conducting time series analyses of deaths and hospital admissions.

5. Conclusions

Satellite data has an important potential for the spatio-temporal prediction of $PM_{2.5}$ concentrations. It offers an alternative method to describe the impacts of forest fires on air quality and to assess the related health effects in the Brazilian Amazon region. Our method provides valuable inputs on how to strengthen and optimize the $PM_{2.5}$ monitoring networks with well-placed complementary measurement sites.

This is needed to understand the impact of air pollution in Brazil and to demonstrate the improvements of air quality and the related health benefits due to the adoption of clean air policies.

Conflicts of interest

Note: The authors declare no conflict of interest.

Acknowledgement

This work was supported by the CAPES foundation [grant number 88881.068027/2014–01] from Brazil. Special thanks to Glauber G.C. Silva (INPE/Brazil), who was always kind and available to help with MATLAB scripting and technical solutions. This manuscript is part of a larger project on *Cardiovascular diseases and the exposure to forest fires in Porto Velho municipality, Rondônia state, Brazil* that was submitted and accepted by the Research Ethics Committee of the Sergio Arouca National Public Health School (*Comitê de Ética em Pesquisa da Escola Nacional de Saúde Pública Sergio Arouca* – ENSP/FIOCRUZ) according to Resolution number 466/2012 from National Research Ethics Council (*Conselho Nacional de Pesquisa* – CONEP) under CAAE number 41732615.4.0000.5 240.

References

Arbex, M.A., de Souza Conceição, G.M., Cendon, S.P., Arbex, F.F., Lopes, A.C., Moysés, E.P., et al., 2009 Oct. Urban air pollution and chronic obstructive pulmonary disease-related emergency department visits. J. Epidemiol. Community Health 63 (10), 777–783. http://dx.doi.org/10.1136/jech.2008.078360.

Becker, B.K., 2005. Geopolítica da Amazônia. Estud. Avançados 1 (53), 71–86.

Brauer, M., Freedman, G., Frostad, J., van Donkelaar, A., Martin, R.V., Dentener, F., Van Dingenen, R., Estep, K., Amini, H., Apte, J.S., Balakrishnan, K., Barregard, L., Broday, D.M., Feigin, V., Ghosh, S., Hopke, P.K., Knibbs, L.D., Kokubo, Y., Liu, Y., Ma, S., Morawska, L., Sangrador, J.L.T., Shaddick, G., Anderson, H.R., Vos, T., Forouzanfar, M.H., Burnett, R.T., Cohen, A., 2015 Nov 23. Ambient air pollution exposure estimation for the Global Burden of Disease 2013. Environ. Sci. Tech. http://dx.doi.org/10.1021/acs.est.5b03709.

Brook, R.D., Rajagopalan, S., Pope III, C.A., Brook, J.R., Bhatnagar, A., Diez-Roux, A.V., Holguin, F., Hong, Y., Luepker, R.V., Mittleman, M.A., Peters, A., Siscovick, D., Smith, S.C., Whitsel, L., Kaufman, J.D., 2010. Particulate matter air pollution and cardiovascular disease: an update to the scientific statement from the American Heart Association. J. AHA. Circ. 2331–2378. http://dx.doi.org/10.1161/CIR.0b013e3181dbece1.

Cohen, A.J., Anderson, H.R., Ostro, B., Pandey, K.D., Krzyzanowski, M., Künzli, N., Gutschmidt, K., Pope, A., Romieu, I., Samet, J.M., Smith, K., 2005. The global burden of disease due to outdoor air pollution. J. Toxicol. Environ. Health, Part A 68, 1301–1307. http://dx.doi.org/10.1080/15287390590936166.

Fearnside, P.M., Julho 2005. Desmatamento na Amazônia brasileira: história, índices e conseqüências. MEGADIVERSIDADE 1 (1) Available: http://philip.inpa.gov.br/publ_livres/2005/Desmatamento%20historia-Megadiversidade.pdf Accessed: May/2016.

Global Burden of Diseases (GBD), 2010. Institute for Health Metrics and Evaluation 2013. http://vizhub.healthdata.org/irank/heat.php.

Goncalves, K.S., Siqueira, A.S.P., Castro, H.A., Hacon, S.S., 2014. Indicator of socio-environmental vulnerability in the western Amazon. The case of the city of Porto Velho, state of Rondônia, Brazil. Ciência Saúde Coletiva 19 (9), 3809–3817. http://dx.doi.org/10.1590/1413-81232014199.14272013.

Gonçalves, K.S., Castro, H.A., Hacon, S.S., 2012. As queimadas na região amazônica e o adoecimento respiratório. Rev Ciência e Saúde Coletiva 17 (6), 1523–1532. https://doi.org/10.1590/S1413-81232012000600016.

Hoff, R.M., Christopher, S.A., 2009. Remote sensing of particulate pollution from space: have we reached the promised land? J. Air Waste Manag. Assoc. 59, 645–675. http://dx.doi.org/10.3155/1047-3289.59.6.645.

de Hoogh, Kees, Héritier, Harris, Stafoggia, Massimo, Künzli, Nino, Kloog, Itai, 2017. Modelling daily PM2.5 concentrations at high spatio-temporal resolution across Switzerland. Environ. Pollut ISSN 0269–7491. https://doi.org/10.1016/j.envpol.2017.10.025.

Hopke, Philip K., Xie, Ying, Raunemaa, Taisto, Biegalski, Steven, Landsberger, Sheldon, Maenhaut, Willy, Artaxo, Paulo, Cohen, David, 1997. Characterization of the gent stacked filter unit PM10 sampler. Aerosol. Sci. Technol. 27 (6), 726–735. http://dx.doi.org/10.1080/02786829708965507.

Kloog, I., Nordio, F., Coull, B.A., Schwartz, J., 2012. Incorporating local land use regression and satellite aerosol optical depth in a hybrid model of spatiotemporal PM2.5 exposures in the mid-Atlantic states. Environ. Sci. Tech. 46 (21), 11913–11921. http://dx.doi.org/10.1021/es302673e.

Kloog, I., Chudnovsky, A.A., Just, A.C., Nordio, F., Koutrakis, P., Coull, B.A., Lyapustin, A., Wang, Y., Schwartz, J., 2014. A new hybrid spatio-temporal model for estimating daily multi-year PM 2.5 concentrations across northeastern USA using high resolution aerosol optical depth data. Atmos. Environ. 95, 581–590. http://dx.doi.org/10.1016/j.atmosenv.2014.07.014.

Künzli, N., Kutlar Joss, M.K., Gintowt, E., 2015. Global standards for global health in a globalized economy! [Editorial]. Int. J. Publ. Health 60, 757–759. http://dx.doi.org/10.1007/s00038-015-0729-0.

Lee, H.J., Liu, Y., Coull, B.A., Schwartz, J., Koutrakis, P., 2011. A calibration method of MODIS AOD data to predict PM2.5. Atmos. Chem. Phys. 11, 7991–8002. http://dx.doi.org/10.5194/acp-11-7991-2011.

Levy, R.C., Remer, L.A., Mattoo, S., Vermote, E.F., Kaufman, Y.J., 2007. Second-generation operational algorithm: retrieval of aerosol properties over land from inversion of Moderate Resolution Imaging Spectroradiometer spectral reflectance. J. Geophys. Res. 112, D13211. http://dx.doi.org/10.1029/2006JD007811.

Levy, R.C., Remer, L.A., Kleidman, R.G., Mattoo, S., Ichoku, C., Kahn, R., Eck, T.F., 2010. Global evaluation of the Collection 5 MODIS dark-target aerosol products over land. Atmos. Chem. Phys. 10, 10399–10420. http://dx.doi.org/10.5194/acp-10-10399-2010.

Liu, Y., Park, R.J., Jacob, D.J., Li, Q.B., Kilaru, V., Sarnat, J.A., 2004. Mapping annual mean ground-level PM2.5 concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States. J. Geophys. Res. 109, D22206. http://dx.doi.org/10.1029/2004JD005025.

Liu, Y., Sarnat, J.A., Kilaru, V., Jacob, D.J., Koutrakis, P., 2005. Estimating ground-level PM2.5 in the eastern United States using satellite remote sensing. Environ. Sci. Technol. 39, 3269–3278.

Liu, Y., Franklin, M., Kahn, R., Koutrakis, P., 2007a. Using aerosol optical thickness to predict ground-level PM2.5 concentrations in the St. Louis area: a comparison between MISR and MODIS, Remote Sens. Environ. Times 107, 33–44.

Liu, Y., Koutrakis, P., Kahn, R., 2007b. Estimating fine particulate matter component concentrations and size distributions using satellite-retrieved fractional aerosol optical depth: Part 1 – method development. J. Air Waste Manag. Assoc. 57, 1351–1359.

Liu, Y., Koutrakis, P., Kahn, R., Turquety, S., Yantosca, R.M., 2007c. Estimating fine particulate matter component concentrations and size distributions using satellite-retrieved fractional aerosol optical depth: Part 2 – a case study. J. AirWaste Manag. Assoc. 57, 1360–1369.

Liu, Y., Paciorek, C.J., Koutrakis, P., 2009. Estimating regional spatial and temporal variability of PM2.5 concentrations using satellite data, meteorology, and land use information. Environ. Health Perspect. 117, 886–892.

McMichael, A.J., Wilkinson, P., Kovats, R.S., Patternden, S., Hajat, S., Armstrong, B., et al., 2008 Oct. International study of temperature, heat and urban mortality: the 'ISOTHURM' project. Int. J. Epidemiol. 37 (5), 1121–1131. http://dx.doi.org/10.1093/ije/dyn086.

Munchak, L.A., Levy, R.C., Mattoo, S., Remer, L.A., Holben, B.N., Schafer, J.S., Hostetler, C.A., Ferrare, R.A., 2013. MODIS 3km aerosol product: applications over land in an urban/suburban region. Atmos. Meas. Tech 6 (7), 1747–1759. https://doi.org/10.5194/amt-6-1747-2013.

Remer, L.A., Kaufman, Y.J., Tanre, D., Mattoo, S., Chu, D.A., Martins, J.V., Li, R.R., Ichoku, C., Levy, R.C., Kleidman, R.G., et al., 2005. Environmental Science & Technology. The MODIS aerosol algorithm, products, and validation. J. Atmos. Sci. 62 (4), 947–973. http://dx.doi.org/10.1021/acs.est.5b01413.

Remer, L.A., Mattoo, S., Levy, R.C., Munchak, L.A., 2013. MODIS 3km aerosol product: algorithm and global perspective. Atmos. Meas. Tech. 6, 1829–1844. http://dx.doi.org/10.5194/amt-6-1829-2013. www.atmos-meas-tech.net/6/1829/2013/.

Ruckerl, R., Schneider, A., Breitner, S., Cyrys, J., Peters, A., 2011 Aug. Health effects of particulate air pollution: a review of epidemiological evidence. Inhal. Toxicol. 23 (10), 555–592. http://dx.doi.org/10.3109/08958378.2011.593587.

World Health Organization (WHO), 2006. Air Quality Guidelines for Particulate Matter, Ozone, Nitrogen Dioxide and Sulfur Dioxide—global Update 2005 Geneva. WHO Office for Europe. http://www.euro.who.int/_data/assets/pdf_file/0005/78638/E90038.pdf.

World Health Organization (WHO), 2014. 7 million Premature Deaths Annually Linked to Air Pollution. http://www.who.int/mediacentre/news/releases/2014/air-pollution/en/.

Xie, Yuanyu, Wang, Yuxuan, Zhang, Kai, Dong, Wenhao, Baolei, Lv, Bai, Yuqi, 2015. Daily estimation of ground-level PM2.5 concentrations over Beijing using 3km resolution MODIS AOD. Environ. Sci. Technol. 49 (20), 12280–12288. http://dx.doi.org/10.1021/acs.est.5b01413.

Ye, X., Wolff, R., Yu, W., Vaneckova, P., Pan, X., Tong, S., 2011 Aug. Ambient temperature and morbidity: a review of epidemiological evidence. Environ. Health Perspect. http://dx.doi.org/10.1289/ehp.1003198.

Yi, O., Hong, Y.C., Kim, H., 2010 Jan. Seasonal effect of PM(10) concentrations on mortality and morbidity in Seoul, Korea: a temperature-matched case-crossover analysis. Environ. Res. 110 (1), 89–95. http://dx.doi.org/10.1016/j.envres.2009.09.009.