Atmos. Chem. Phys. Discuss., 11, 9769–9795, 2011
www.atmos-chem-phys-discuss.net/11/9769/2011/
doi:10.5194/acpd-11-9769-2011

Atmospheric Chemistry and Physics Discussions

A novel calibration approach of MODIS AOD data to predict $PM_{2.5}$ concentrations

H. J. Lee[1], Y. Liu[2], B. A. Coull[3], J. Schwartz[1], and P. Koutrakis[1]

[1]Department of Environmental Health, Harvard School of Public Health, Boston, MA 02215, USA
[2]Department of Environmental and Occupational Health, Rollins School of Public Health, Emory University, Atlanta, GA 30322, USA
[3]Department of Biostatistics, Harvard School of Public Health, Boston, MA 02115, USA

Received: 7 March 2011 – Accepted: 20 March 2011 – Published: 23 March 2011

Correspondence to: H. J. Lee (hlee@hsph.harvard.edu)

Published by Copernicus Publications on behalf of the European Geosciences Union.

Abstract

Epidemiological studies investigating the human health effects of $PM_{2.5}$ are susceptible to exposure measurement errors, a form of bias in exposure estimates, since they rely on data from a limited number of $PM_{2.5}$ monitors within their study area. Satellite data can be used to expand spatial coverage, potentially enhancing our ability to estimate location- or subject-specific exposures to $PM_{2.5}$, but some have reported poor predictive power. A new methodology was developed to calibrate aerosol optical depth (AOD) data obtained from the Moderate Resolution Imaging Spectroradiometer (MODIS). Subsequently, this method was used to predict ground daily $PM_{2.5}$ concentrations in the New England region. 2003 MODIS AOD data corresponding to the New England region were retrieved, and $PM_{2.5}$ concentrations measured at 26 US Environmental Protection Agency (EPA) $PM_{2.5}$ monitoring sites were used to calibrate the AOD data. A mixed effects model which allows day-to-day variability in daily $PM_{2.5}$-AOD relationships was used to predict location-specific $PM_{2.5}$ levels. $PM_{2.5}$ concentrations measured at the monitoring sites were compared to those predicted for the corresponding grid cells. Both cross-sectional and longitudinal comparisons between the observed and predicted concentrations suggested that the proposed new calibration approach renders MODIS AOD data a potentially useful predictor of $PM_{2.5}$ concentrations. Furthermore, the estimated $PM_{2.5}$ levels within the study domain were examined in relation to air pollution sources. Our approach made it possible to investigate the spatial patterns of $PM_{2.5}$ concentrations within the study domain.

1 Introduction

Atmospheric aerosols originate from natural and anthropogenic emission sources. Particularly, anthropogenic aerosols are considered to have major human health implications, and numerous studies have reported associations between mortality and morbidity and particulate matter with aerodynamic diameter $\leq 2.5\,\mu m$ ($PM_{2.5}$) (Bell et al., 2007; Dominici et al., 2006; Franklin et al., 2007; Gent et al., 2003, 2009; Schwartz et al.,

1996; Slama et al., 2007). The $PM_{2.5}$ health effect studies generally use $PM_{2.5}$ measurements from ground monitoring sites, but there are many regions with no ground $PM_{2.5}$ measurements available due to their sparse monitoring networks. This limits the ability of estimating human exposures to $PM_{2.5}$, which is likely to cause less reliable health effect assessments.

Satellite remote sensing can be used to assess $PM_{2.5}$ air quality for areas where surface $PM_{2.5}$ monitors are not available (Di Nicolantonio et al., 2009; Engel-Cox et al., 2004; Gupta and Christopher, 2008; Gupta et al., 2006; Koelemeijer et al., 2006; Liu et al., 2004; Schaap et al., 2009; van Donkelaar et al., 2010). The most applicable satellite-retrieved product for estimating $PM_{2.5}$ concentrations is aerosol optical depth (AOD), which measures the light extinction by aerosol scattering and absorption in the atmospheric column. Since the AOD reflects the integrated amount of particles in the vertical column, it has been used as an input parameter in statistical models predicting $PM_{2.5}$ levels. In addition to AOD values, several studies have also included other predictor parameters such as local meteorology and land use information (e.g., population density). As reported by previous studies, these parameters influence the relationship between AOD and ground-level $PM_{2.5}$ concentrations, thus can be used as additional predictors (Liu et al., 2005, 2007a, b, c, 2009). However, these models, developed by us and others, generally predict <60% of the variability in daily $PM_{2.5}$ concentrations (Paciorek et al., 2008). Additional time-varying parameters influence the $PM_{2.5}$-AOD relationship, including $PM_{2.5}$ vertical and diurnal concentration profiles, PM optical properties, and others. Therefore, it is reasonable to expect that the relationship between $PM_{2.5}$ and AOD varies by day. In this paper we introduce a new approach to calibrate Moderate Resolution Imaging Spectroradiometer (MODIS) AOD data to accurately predict $PM_{2.5}$ ground concentrations.

Our method is unique because it establishes day-specific $PM_{2.5}$-AOD relationships using a mixed effects model to fully exploit satellite data. To the best of our knowledge, no previous studies have suggested a statistical approach establishing the $PM_{2.5}$-AOD relations on a daily basis.

2 Methods

2.1 Ground-level $PM_{2.5}$ data

Our study region includes the States of Massachusetts (MA), Connecticut (CT), and Rhode Island (RI) in the Northeastern US. To calibrate satellite data, daily $PM_{2.5}$ concentrations measured at 26 US Environmental Protection Agency (EPA) $PM_{2.5}$ monitoring sites were used (Fig. 1). For collocated monitors, we calculated the daily averages of the $PM_{2.5}$ concentrations. Samples were collected at 15 Connecticut sites and 11 Massachusetts sites during the period 1 January through 31 December 2003. Sampling frequency differed by site including collecting samples every day, every third day, and every sixth day.

2.2 AOD retrieval

MODIS aboard the National Aeronautics and Space Administration (NASA)'s Earth Observing System (EOS) satellites, Terra and Aqua, was used to retrieve AOD (Collection 5; Level 2 aerosol product) for the year 2003. The Terra and Aqua satellites were launched in December 1999 and in May 2002, respectively. These polar-orbiting satellites, operating at an altitude of approximately 700 km, provide data every one to two days under cloud-free conditions. Their sensors scan the swath of 2330 km (cross-track) by 10 km (along-track at nadir) and gather information on particle abundance once from each satellite: approximately 10:30 a.m. and 1:30 p.m. local times for Terra and Aqua, respectively. In the Collection 5 retrieval algorithm, three different channels of 0.47, 0.66, and 2.12 μm are primarily employed for over-land retrievals. The channels of 0.47 and 0.66 μm are used to retrieve AOD values which are interpolated to report AOD values at the wavelength of 0.55 μm, and the uncertainty of the MODIS AOD is expected to be $\Delta AOD = \pm 0.05 \pm 0.15 \times AOD$ over land. Furthermore, the maximum AOD value is constrained to be 5.0, and negative AOD values down to −0.05 were retained in order to avoid bias that can occur when truncating or omitting low

exposure values. More details about MODIS satellite data are reported in Remer et al. (2005) and Levy et al. (2007). Following the nominal resolution of MODIS (10 km), we created 387 grid cells of $10 \times 10\,km^2$ covering our study region in ArcGIS (Version 9.3; ESRI). Subsequent analyses were based on these grid cells.

Since Terra and Aqua satellites retrieve AOD data at two different times each day, the average of these two measurements should be used to predict daily $PM_{2.5}$ levels (Kaufman et al., 2000). However, there are many days where only one of the two retrievals is available. To fully exploit the measurements of both satellites we primarily used Terra AOD data for our predictions, and for days with no Terra data, Aqua AOD measurement values were used to estimate the missing Terra values. This was accomplished by multiplying Aqua AOD measurements by an adjustment factor, which was necessary to account for diurnal variations (Green et al., 2009) and potential calibration differences in two satellite sensors. This factor was equal to the average Terra AOD/Aqua AOD ratio which was calculated for days where both Terra and Aqua data were available.

2.3 Statistical model

Since time-varying parameters such as relative humidity, $PM_{2.5}$ vertical and diurnal concentration profiles, and $PM_{2.5}$ optical properties influence the $PM_{2.5}$-AOD relationship, our statistical model allows for day-to-day variability in this relationship. Furthermore, we hypothesize that these time-varying parameters exhibit little spatial variability and consequently the $PM_{2.5}$-AOD relationship varies minimally spatially on a given day over the spatial scale of our study. Therefore, a quantitative relationship between $PM_{2.5}$ concentrations measured at 26 $PM_{2.5}$ monitoring sites and AOD values in their corresponding grid cells can be determined on a daily basis. A simple approach would be to calculate such $PM_{2.5}$-AOD slopes separately for each day in the study. However, this simplistic approach can yield highly variable slope estimates, since some days might have a small to moderate amount of monitoring data. An alternative approach that pools daily slope estimates but uses data from all days to stabilize the estimates

is to use a mixed effects model with random intercepts and slopes (Fitzmaurice et al., 2004), shown by the following equations:

$$PM_{ij} = (\alpha + u_j) + (\beta_1 + v_j) \times AOD_{ij} + s_i + \varepsilon_{ij} \quad (1)$$

$$(u_j v_j) \sim N[(00), \Sigma_\beta]$$

where PM_{ij} is the $PM_{2.5}$ concentration at a spatial site i on a day j; AOD_{ij} is the AOD value in the grid cell corresponding to site i on a day j; α and u_j are the fixed and random intercepts, respectively; β_1 and v_j are the fixed and random slopes, respectively; $s_i \sim N(0, \sigma_s^2)$ is the random intercept of site i; ε_{ij} is the error term at site i on a day j; and Σ_β is the variance-covariance matrix for the day-specific random effects. In the statistical model, the AOD fixed effect represents the average effect of AOD on $PM_{2.5}$ for all study days. The AOD random effects explain the daily variability in the $PM_{2.5}$-AOD relationship. The site bias may arise since an AOD value in a $10 \times 10\,km^2$ grid cell is an average optical depth in the given grid cell, while the $PM_{2.5}$ concentrations measured at a given site may not be representative of the whole grid cell. Specifically, the bias can indicate spatial sites presenting high $PM_{2.5}$ levels due to their locations near high traffic areas. To control for the site bias, we added a site term as a random effect into the mixed effects model. It should be noted that the random estimates for the site term were omitted when estimating grid-specific $PM_{2.5}$ concentrations from AOD values, since AOD values are unbiased representatives of the corresponding grid cells. Because a slope cannot be estimated from a single data point, we excluded all the pairs of measured $PM_{2.5}$ concentrations and their corresponding AOD values when there was only one pair on a given day before running the mixed effects model. This resulted in the exclusion of 29 days. Furthermore, the model prediction was examined using the root mean squared error (RMSE) between the measured and predicted $PM_{2.5}$ concentrations on each day. Four sample days with RMSE $>5\,\mu g\,m^{-3}$ were excluded from the analysis, since the daily $PM_{2.5}$-AOD relationships were not considered reliable

enough to calibrate AOD data. Finally, $PM_{2.5}$ estimates covering the whole study area were produced using the AOD calibration model described above.

To demonstrate whether the mixed effects model improved the ability of AOD to predict $PM_{2.5}$ concentrations we compared our model performance to that of a previously used model which assumes that the $PM_{2.5}$-AOD relationship remains constant over time (Wang and Christopher, 2003). In the previous model, measured $PM_{2.5}$ concentrations were regressed on AOD values retrieved in the corresponding grid cells as a fixed effect, establishing a single linear $PM_{2.5}$-AOD relation applied to all sampling days. It is noted that the comparison of those two models was based on identical sampling days. As measures of accuracy and precision of the two models, we used coefficient of determination (R^2) and precision (% Precision) between the measured and predicted $PM_{2.5}$ concentrations. The precision was estimated as the square root of the mean of the squared errors, and % Precision was calculated as follows:

$$\%\ \text{Precision} = 100 \times (\text{precision}/\text{measured mean } PM_{2.5}) \quad (2)$$

2.4 Model validation

To test this new approach we analyzed the 2003 MODIS data for MA, CT, and RI. We utilized a cross-validation (CV) method to examine whether the model is generalizable to any grid cell in the study domain. Toward this end the data of one site (test site) were separated from those of the other 25 sites (calibration sites). Subsequently, a model was developed using the data from the calibration sites. Finally, the model was used to predict $PM_{2.5}$ concentrations for the test site. This process was repeated until each of the 26 spatial sites was tested, and the measured $PM_{2.5}$ concentrations were compared to those predicted at each site. Furthermore, Pearson correlation coefficients were used to examine the relationship between the measured and predicted concentrations in each site. Since time-series studies examine longitudinal associations between exposures and health outcomes, high correlation coefficients would imply that satellite AOD data can be used to assess exposures for these health investigations. In addition,

we examined the agreement between the measured and predicted annual mean $PM_{2.5}$ concentration levels for each of the 26 sites, which was assessed by the correlation between the measured and predicted mean $PM_{2.5}$ concentrations. This comparison is important for determining whether model predictions are reliable for cross-sectional studies, which require accurate assessment of spatial patterns in exposure.

2.5 $PM_{2.5}$ levels in the study region

$PM_{2.5}$ levels were estimated for each of the 387 grid cells. Since the AOD retrieval rate varies by location, the number of $PM_{2.5}$ concentration predictions varied by grid cell. Therefore, a direct comparison among cell means would not be adequate for the investigation of the $PM_{2.5}$ spatial patterns within the study domain. To minimize the potential impact of varying predictions per grid cell we estimated the mean differences between the predicted grid cell and regional $PM_{2.5}$ concentrations for the days where grid cell predictions were available. Note that daily regional $PM_{2.5}$ concentrations were calculated by averaging the predicted $PM_{2.5}$ concentrations for each of the grid cells. Since the number of AOD retrievals varied by day, the number of available $PM_{2.5}$ concentrations used to estimate the daily regional average levels varied by day as well. To obtain reliable and representative regional $PM_{2.5}$ concentrations we limited our estimations to days with 50 or more grid cell predictions. Finally, the grid cell-specific $PM_{2.5}$ mean differences between the grid cell and the regional $PM_{2.5}$ concentrations were presented using septiles, which split the distribution of the mean differences into seven equally-sized bins, in ArcGIS. Positive mean differences, expressed in $\mu g\,m^{-3}$, indicate that on average levels at a given grid cell are higher relative to the regional $PM_{2.5}$ levels, while the opposite is true for negative values.

3 Results and discussion

3.1 Descriptive statistics

The mean $PM_{2.5}$ concentrations measured at the 26 EPA $PM_{2.5}$ monitoring sites in 2003 are summarized in Table 1. The mean (SE) $PM_{2.5}$ concentrations ranged from 9.0 (0.7) $\mu g\,m^{-3}$ in Haverhill, MA (Site ID: 25-009-5005) to 17.0 (0.5) $\mu g\,m^{-3}$ in New Haven, CT (Site ID: 09-009-0018). The mean $PM_{2.5}$ concentration at the New Haven site was exceptionally high as compared to those monitored at other sites, possibly because the site was located on a ramp connecting to interstate I-95. Many of the monitoring sites showed similar mean $PM_{2.5}$ concentrations. However, it should be noted that the number of samples used to estimate these means varied by site due to differences in sampling frequencies among sites and missing data. Furthermore, mean (SE) daily AOD values observed for the 387 grid cells varied from 0.08 (0.02) to 0.36 (0.04). On average 67 AOD values were retrieved per grid cell which corresponds to 18% of the entire study period of 365 days.

3.2 $PM_{2.5}$ prediction

In the mixed effects model, 99 different daily $PM_{2.5}$-AOD relations were generated in 2003. The fixed effects of intercept and slope (AOD) were statistically significant ($\alpha = 11.9$, $p < 0.0001$; $\beta_1 = 4.4$, $p = 0.0049$), respectively. The random effects of intercept and slope (AOD) varied considerably by day, with standard deviations of the daily intercepts and slopes of 8.0 and 2.3, respectively. This supports our hypothesis that parameters influencing the relationship between $PM_{2.5}$ and AOD vary daily but not spatially. Therefore, it is possible to perform daily calibrations using data from the multiple $PM_{2.5}$ monitoring sites in the study domain. It is noted that the daily intercepts and slopes were independent of the number of $PM_{2.5}$-AOD pairs on a given day. In addition, the averages of the daily intercepts and slopes were found to be 12.7 (SD = 8.7)

and 4.6 (SD = 2.5) in warm season (15 April–14 October) and 10.1 (SD = 5.4) and 3.8 (SD = 1.3) in cold season (15 October–14 April), respectively. The random effect estimates of the site term for densely populated and high traffic areas were positive as presented in Table 2. Therefore, inclusion of the site term was necessary to adjust for the site bias in our model. As shown by Table 3 and Fig. 2a, the mixed effects model performed quite well. Table 3 presents the site-specific comparisons between the measured and predicted $PM_{2.5}$ concentrations in the mixed effects model, and the model prediction was reliable for most spatial sites (mean % Precision = 13.16%, Range = 7.38 to 25.45%). Moreover, Fig. 2a depicts the results of the linear regression model which was used to compare the measured and predicted daily concentrations for all 26 monitoring sites ($R^2 = 0.97$, slope = 0.96 (SE = 0.01), and intercept = 0.44 (SE = 0.11)). In addition, Fig. 2b shows the results of the linear regression model used to compare the measured concentrations to those obtained from the CV procedure ($R^2 = 0.92$, slope = 0.92 (SE = 0.01), and intercept = 0.88 (SE = 0.18)). It is noted that the predicted $PM_{2.5}$ concentrations from the CV procedure were not adjusted for the site bias, due to the fact that this term would not be available for location-specific predictions in an epidemiological health effects study. The more pronounced difference between the measured and predicted concentrations in Fig. 2b as compared to Fig. 2a is likely to reflect the bias. As it can be seen, both the model fit and CV test resulted in high R^2, slopes close to 1, and intercepts close to 0, indicating a good agreement between the measured and predicted concentrations.

In the mixed effects model, the differences between measured and predicted $PM_{2.5}$ levels can be attributed to a combination of monitoring site-specific characteristics as well as $PM_{2.5}$ measurement and AOD retrieval errors. For instance, the monitoring site location may not be representative of a given $10 \times 10\,km^2$ grid cell for an average optical depth retrieved value. For example, concentrations measured at the New Haven site (Site ID: 09-009-0018), which was located on a ramp to interstate I-95, were significantly higher than those observed at the other sites, including the site (Site ID: 09-009-0026) located nearby (0.7 km). Therefore, the difference (4.65 $\mu g\,m^{-3}$) between

the measured and predicted mean $PM_{2.5}$ concentrations before taking the site bias into account at the New Haven site can be explained by the fact that this site is not representative of the corresponding grid cell $10 \times 10\,km^2$ area, and it indicates that the approach of controlling for the site bias in the mixed effects model is reasonable for the comparisons between the measured and predicted $PM_{2.5}$ concentrations. However, considering that AOD-derived $PM_{2.5}$ concentrations reflect the overall $PM_{2.5}$ levels in the grid cell, the unadjusted predicted $PM_{2.5}$ levels may be more representative of the average population exposures to $PM_{2.5}$.

AOD retrieval errors due to unscreened clouds could introduce positive bias. The current cloud screening algorithm in AOD retrievals (Collection 5) effectively masks clouds, but it is still possible to have AOD values affected by clouds, particularly for isolated and residual clouds (Levy et al., 2007). The comparison between MODIS AOD and the Aerosol Robotic Network (AERONET) AOD (Level 2.0; within ± 30 min of Terra measurements) in Billerica could indicate days with positive bias potentially from isolated and residual clouds in the area (correlation $r = 0.92$; slope = 1.20; intercept = −0.002 in a linear regression model between the MODIS AOD and the AERONET AOD data) (Holben et al., 1998). Consequently, the AOD values overestimated by the clouds may cause positive bias in predicted $PM_{2.5}$ concentrations. In part, $PM_{2.5}$ measurement errors might cause positive or negative bias in measured $PM_{2.5}$ levels.

The ability of the mixed effects and linear regression models to predict $PM_{2.5}$ concentrations was compared. For each model the predicted concentrations were regressed on the observed ones for each site separately (Table 4 and Fig. 3). It should be noted that the CV method produces less biased estimates than those obtained from the model fit (shown in Tables 3 and 4). The two models were compared using results from CV analyses to avoid over-fitting thus to produce more robust results. Note that the predicted $PM_{2.5}$ concentrations in the mixed effects model were not adjusted for the site bias (Table 4 and Fig. 3). The mixed effects model explained 95% of the variability in the measured $PM_{2.5}$ concentrations on average, ranged from 82% in Boston, MA (Site ID: 25-025-0027) to 100% in Bridgeport, CT (Site ID: 09-001-0010). On the

other hand, in the linear regression model, the mean variability of the measured $PM_{2.5}$ explained by the predicted $PM_{2.5}$ was 51%, ranging from 12% in Boston, MA (Site ID: 25-025-0027) to 88% in Stamford, CT (Site ID: 09-001-2124). While the regression model yielded modest and considerably varying predictability by site, our model demonstrated consistently high predictability for most of the sites. These findings suggest that predicting $PM_{2.5}$ within a domain requires the use of daily calibrations. This explains why previous investigations have not demonstrated that AOD can be a robust predictor of $PM_{2.5}$ (Paciorek and Liu, 2009; Paciorek et al., 2008).

The predictive ability of our model was also compared to that of the regression model in terms of percent precision (% Precision) (Table 4 and Fig. 3). Note that this comparison was performed using the CV results as well. Since the R^2 does not reflect systematic differences between the measured and predicted $PM_{2.5}$ levels, the measure of precision (% Precision) is necessary to better assess model performance. In the mixed effects model, the % CV precision ranged from 8.8% ($1.08\,\mu g\,m^{-3}$) in New Haven, CT (Site ID: 09-009-0026) to 38.6% ($4.08\,\mu g\,m^{-3}$) in Lynn, MA (Site ID: 25-009-2006) with the mean value of 20.0% ($2.45\,\mu g\,m^{-3}$). For the regression model the estimated mean % CV precision was 59.5% ($7.40\,\mu g\,m^{-3}$), varying from 41.1% ($4.59\,\mu g\,m^{-3}$) in Fall River, MA (Site ID: 25-005-1004) to 89.8 % ($12.73\,\mu g\,m^{-3}$) in Boston, MA (Site ID: 25-025-0027).

With regard to the measures of CV R^2 and precision values, our model presented considerably higher CV R^2 (0.95) and lower CV precision (20.0%, $2.45\,\mu g\,m^{-3}$) than those estimated for the regression model (CV $R^2 = 0.51$, % CV precision = 59.5% ($7.40\,\mu g\,m^{-3}$)). Also, the cross-sectional comparison between the measured and predicted site mean $PM_{2.5}$ concentrations was performed for both models. As shown in Fig. 4, a higher correlation coefficient ($R^2 = 0.62$; Pearson $r = 0.79$) was determined for the mixed effects model as compared to that estimated for the linear regression model ($R^2 = 0.26$; Pearson $r = 0.51$). Overall, the performance of the mixed effects model to predict surface-level $PM_{2.5}$ concentrations was superior as compared to that of the regression model. Collectively, these performance tests suggest that the mixed effects

model can be used to produce concentration data sets reliable for both time-series and cross-sectional health effect studies.

3.3 Spatial variability in $PM_{2.5}$ levels

The spatial patterns of $PM_{2.5}$ levels within the study domain are shown in Fig. 5. To highlight the spatial patterns, we used the mean differences between grid-specific $PM_{2.5}$ and regional $PM_{2.5}$ levels, as mentioned above. Mean concentration differences varied from -0.36 to $0.87\,\mu g\,m^{-3}$ (mean = $0.01\,\mu g\,m^{-3}$, SD = $0.17\,\mu g\,m^{-3}$), and were log-normally distributed, which led us to use septiles for characterizing the spatial variability of $PM_{2.5}$ levels in our study region. The relatively small difference between the lowest and highest values ($1.23\,\mu g\,m^{-3}$) compared to the one presented in Table 4 can be explained by the fact that the result of Fig. 5 represented average cell concentrations which were based on the large number of overlapping days, while the large variability in average $PM_{2.5}$ concentrations between sites in Table 4 was derived from the limited number of samples, used to calculate the means, which did not correspond to the same time period. As expected, highly populated areas such as Bridgeport, New Haven, Hartford, Boston, Springfield, and Providence exhibited higher $PM_{2.5}$ levels. Also, higher $PM_{2.5}$ levels were predicted along the major interstate highways (e.g., I-91/95) and areas with high point emission sources (e.g., power plants located in coastal cities) (US EPA, 2008). The concentration spatial patterns observed in eastern Massachusetts were similar to those found by our previous studies (Gryparis et al., 2007). Furthermore, the estimated $PM_{2.5}$ levels in western Massachusetts were generally lower, which is due to the lower population density and traffic density in the area. However, it must be noted that the reported $PM_{2.5}$ spatial patterns may not be representative of the entire year, since AOD values are less likely to be collected during the cold season due to more frequent cloud conditions during this period. The average number of the predicted $PM_{2.5}$ concentrations in each grid cell was 39 (SD = 6) days in warm season and 14 (SD = 5) days in cold season.

4 Conclusions

Satellite AOD data have been increasingly used for $PM_{2.5}$ air pollution studies. Remote sensing technologies have a great potential to expand current ground-level $PM_{2.5}$ monitoring networks. To date, the application of satellite data to health effect studies has been limited mostly due to the insufficient power of AOD to predict $PM_{2.5}$ and the high frequency of non-retrieval days. We have introduced an AOD calibration method which made it possible to determine the temporal and spatial patterns of $PM_{2.5}$ in a large study domain comprising the States of Massachusetts, Connecticut, and Rhode Island. An approach to $PM_{2.5}$ prediction for non-retrieval days will be presented in our forthcoming paper.

Finally, it is anticipated that future satellite technologies will provide data with finer spatial and temporal resolutions and more accurate data retrievals. In addition, the advanced capability of discriminating by aerosol species in satellite technologies will further contribute to health effect studies investigating species-specific health implications. Since satellite data are readily available, $PM_{2.5}$ concentrations can be predicted in a cost-effective way. Considering the sparse ground-level $PM_{2.5}$ monitoring networks, our method will help to investigate the associations between subject-specific exposures to $PM_{2.5}$ and their health effects.

Acknowledgements. This study was funded by the Harvard EPA PM Center (R-832416) and the Yale Center for Perinatal, Pediatric and Environmental Epidemiology (NIH-NIEHS R01-ES-016317). Also support was provided by NIEHS grants (ES009825 and ES00002).

References

Bell, M. L., Ebisu, K., and Belanger, K.: Ambient air pollution and low birth weight in Connecticut and Massachusetts, Environ. Health Perspect., 115, 1118–1125, 2007.

Di Nicolantonio, W., Cacciari, A., and Tomasi, C.: Particulate matter at surface: Northern Italy monitoring based on satellite remote sensing, meteorological fields, and in-situ samplings, IEEE J. Selected Topics Appl. Earth Observ. Rem. Sens., 2, 284–292, 2009.

Dominici, F., Peng, R. D., Bell, M. L., Pham, L., McDermott, A., Zeger, S. L., and Samet, J. M.: Fine particulate air pollution and hospital admission for cardiovascular and respiratory diseases, Jama.-J. Am. Med. Assoc., 295, 1127–1134, 2006.

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

Fitzmaurice, G. M., Laird, N. M., and Ware, J. H.: Applied longitudinal analysis, New York, Wiley & Sons, 2004.

Franklin, M., Zeka, A., and Schwartz, J.: Association between $PM_{2.5}$ and all-cause and specific-cause mortality in 27 US communities, J. Expo. Sci. Environ. Epidemiol., 17, 279–287, 2007.

Gent, J. F., Triche, E. W., Holford, T. R., Belanger, K., Bracken, M. B., Beckett, W. S., and Leaderer, B. P.: Association of low-level ozone and fine particles with respiratory symptoms in children with asthma, Jama.-J. Am. Med. Assoc., 290, 1859–1867, 2003.

Gent, J. F., Koutrakis, P., Belanger, K., Triche, E., Holford, T. R., Bracken, M. B., and Leaderer, B. P.: Symptoms and medication use in children with asthma and traffic-related sources of fine particle pollution, Environ. Health Perspect., 117, 1168–1174, 2009.

Green, M., Kondragunta, S., Ciren, P., and Xu, C.: Comparison of GOES and MODIS aerosol optical depth (AOD) to aerosol robotic network (AERONET) AOD and IMPROVE $PM_{2.5}$ mass at Bondville, Illinois, J. Air Waste Manag. Assoc., 59, 1082–1091, 2009.

Gryparis, A., Coull, B. A., Schwartz, J., and Suh, H. H.: Semiparametric latent variable regression models for spatiotemporal modeling of mobile source particles in the greater Boston area, J. R. Stat. Soc. Ser. C Appl. Stat., 56, 183–209, 2007.

Gupta, P. and Christopher, S. A.: Seven year particulate matter air quality assessment from surface and satellite measurements, Atmos. Chem. Phys., 8, 3311–3324, doi:10.5194/acp-8-3311-2008, 2008.

Gupta, P., Christopher, S. A., Wang, J., Gehrig, R., Lee, Y., and Kumar, N.: Satellite remote

sensing of particulate matter and air quality assessment over global cities, Atmos. Environ., 40, 5880–5892, 2006.

Holben, B. N., Eck, T. F., Slutsker, I., Tanre, D., Buis, J. P., Setzer, A., Vermote, E., Reagan, J. A., Kaufman, Y. J., Nakajima, T., Lavenu, F., Jankowiak, I., and Smirnov, A.: AERONET – A federated instrument network and data archive for aerosol characterization, Remote Sens. Environ., 66, 1–16, 1998.

Kaufman, Y. J., Holben, B. N., Tanre, D., Slutsker, I., Smirnov, A., and Eck, T. F.: Will aerosol measurements from Terra and Aqua polar orbiting satellites represent the daily aerosol abundance and properties?, Geophys. Res. Lett., 27, 3861–3864, 2000.

Koelemeijer, R. B. A., Homan, C. D., and Matthijsen, J.: Comparison of spatial and temporal variations of aerosol optical thickness and particulate matter over Europe, Atmos. Environ., 40, 5304–5315, 2006.

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

Liu, Y., Park, R. J., Jacob, D. J., Li, Q., Kilaru, V., and Sarnat, J. A.: Mapping annual mean ground-level $PM_{2.5}$ concentrations using Multiangle Imaging Spectroradiometer aerosol optical thickness over the contiguous United States, J. Geophys. Res., 109, D22206, doi:10.1029/2004JD005025, 2004.

Liu, Y., Sarnat, J. A., Kilaru, V., Jacob, D. J., and Koutrakis, P.: Estimating ground-level $PM_{2.5}$ in the eastern United States using satellite remote sensing, Environ. Sci. Technol., 39, 3269–3278, 2005.

Liu, Y., Franklin, M., Kahn, R., and Koutrakis, P.: Using aerosol optical thickness to predict ground-level $PM_{2.5}$ concentrations in the St. Louis area: A comparison between MISR and MODIS, Remote Sens. Environ., 107, 33–44, 2007a.

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

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

Liu, Y., Paciorek, C. J., and Koutrakis, P.: Estimating regional spatial and temporal variability of $PM_{2.5}$ concentrations using satellite data, meteorology, and land use information, Environ. Health Perspect., 117, 886–892, 2009.

Paciorek, C. J. and Liu, Y.: Limitations of remotely sensed aerosol as a spatial proxy for fine particulate matter, Environ. Health Perspect., 117, 904–909, 2009.

Paciorek, C. J., Liu, Y., Moreno-Macias, H., and Kondragunta, S.: Spatiotemporal associations between GOES aerosol optical depth retrievals and ground-level $PM_{2.5}$, Environ. Sci. Technol., 42, 5800–5806, 2008.

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., Eck, T. F., Vermote, E., and Holben, B. N.: The MODIS aerosol algorithm, products, and validation, J. Atmos. Sci., 62, 947–973, 2005.

Schaap, M., Apituley, A., Timmermans, R. M. A., Koelemeijer, R. B. A., and de Leeuw, G.: Exploring the relation between aerosol optical depth and $PM_{2.5}$ at Cabauw, the Netherlands, Atmos. Chem. Phys., 9, 909–925, doi:10.5194/acp-9-909-2009, 2009.

Schwartz, J., Dockery, D. W., and Neas, L. M.: Is daily mortality associated specifically with fine particles?, Air Waste Manag. Assoc., 46, 927–939, 1996.

Slama, R., Morgenstern, V., Cyrys, J., Zutavern, A., Herbarth, O., Wichmann, H. E., and Heinrich, J.: Traffic-related atmospheric pollutants levels during pregnancy and offspring's term birth weight: A study relying on a land-use regression exposure model, Environ. Health Perspect., 115, 1283–1292, 2007.

US Environmental Protection Agency (US EPA): National Emissions Inventory (NEI), available at: http://www.epa.gov/ttn/chief/eiinformation.html (last access: 14 January 2010), 2008.

van Donkelaar, A., Martin, R. V., Brauer, M., Kahn, R., Levy, R., Verduzco, C., and Villeneuve, P. J.: Global estimates of ambient fine particulate matter concentrations from satellite-based aerosol optical depth: Development and application, Environ. Health Perspect., 118, 847–855, 2010.

Wang, J. and Christopher, S. A.: Intercomparison between satellite-derived aerosol optical thickness and $PM_{2.5}$ mass: Implications for air quality studies, Geophys. Res. Lett., 30, 2095, doi:10.1029/2003GL018174, 2003.

Table 1. $PM_{2.5}$ concentrations ($\mu g\,m^{-3}$) observed at the 26 EPA monitoring sites in 2003.

Site ID	City	*N*	Mean	SE
09-001-0010	Bridgeport, CT	97	12.2	0.8
09-001-0113	Bridgeport, CT	94	11.7	0.8
09-001-1123	Danbury, CT	101	13.0	0.9
09-001-2124	Stamford, CT	100	13.3	0.9
09-001-3005	Norwalk, CT	99	12.0	0.8
09-001-9003	Westport, CT	108	11.0	0.7
09-003-1003	E. Hartford, CT	310	11.4	0.4
09-003-1018	Hartford, CT	92	12.3	0.9
09-009-0018	New Haven, CT	307	17.0	0.5
09-009-0026	New Haven, CT	70	11.5	1.1
09-009-1123	New Haven, CT	108	13.4	0.8
09-009-2008	New Haven, CT	79	12.0	1.1
09-009-2123	Waterbury, CT	110	12.4	0.8
09-009-8003	W. Haven, CT	77	12.6	1.1
09-011-3002	Norwich, CT	79	10.7	0.7
25-005-1004	Fall River, MA	90	10.2	0.8
25-009-2006	Lynn, MA	78	10.3	1.2
25-009-5005	Haverhill, MA	87	9.0	0.7
25-013-0008	Chicopee, MA	237	9.7	0.4
25-013-0016	Springfield, MA	265	12.8	0.5
25-013-2009	Springfield, MA	75	11.3	1.0
25-023-0004	Brockton, MA	97	10.0	0.8
25-025-0027	Boston, MA	198	11.7	0.5
25-025-0042	Boston, MA	246	11.5	0.5
25-025-0043	Boston, MA	96	13.1	0.8
25-027-0020	Worcester, MA	231	11.7	0.5

Table 2. Site bias ($\mu g\,m^{-3}$) estimates for 26 EPA $PM_{2.5}$ monitoring sites.

Site ID	City	Bias[a]	p-value
09-001-0010	Bridgeport, CT	0.77	0.18
09-001-0113	Bridgeport, CT	0.57	0.28
09-001-1123	Danbury, CT	0.47	0.40
09-001-2124	Stamford, CT	1.22	0.03
09-001-3005	Norwalk, CT	1.18	0.03
09-001-9003	Westport, CT	0.24	0.68
09-003-1003	E. Hartford, CT	−0.57	0.17
09-003-1018	Hartford, CT	−0.09	0.86
09-009-0018	New Haven, CT	4.49	< .0001
09-009-0026	New Haven, CT	0.30	0.58
09-009-1123	New Haven, CT	1.35	0.006
09-009-2008	New Haven, CT	0.03	0.96
09-009-2123	Waterbury, CT	0.46	0.34
09-009-8003	W. Haven, CT	1.70	0.002
09-011-3002	Norwich, CT	−0.08	0.89
25-005-1004	Fall River, MA	−0.27	0.66
25-009-2006	Lynn, MA	−2.64	< .0001
25-009-5005	Haverhill, MA	−1.64	0.003
25-013-0008	Chicopee, MA	−1.92	< .0001
25-013-0016	Springfield, MA	−0.001	0.998
25-013-2009	Springfield, MA	−0.55	0.31
25-023-0004	Brockton, MA	−1.71	0.002
25-025-0027	Boston, MA	−1.37	0.04
25-025-0042	Boston, MA	−0.43	0.45
25-025-0043	Boston, MA	0.004	0.996
25-027-0020	Worcester, MA	−1.48	0.002

[a] Bias represents the random effect estimates of the site term in the mixed effects model.

Table 3. Mixed effects model performance by site[a].

Site ID	City	*N*	$PM_{2.5\ measured}$	$PM_{2.5\ predicted}$	Bias[b]	R^2	Precision[c]	% Precision[d]
09-001-0010	Bridgeport, CT	15	11.59	11.50	−0.08	1.00	0.96	8.31
09-001-0113	Bridgeport, CT	19	9.64	9.59	−0.05	0.97	1.15	11.89
09-001-1123	Danbury, CT	16	13.96	13.91	−0.05	0.98	1.85	13.29
09-001-2124	Stamford, CT	14	12.63	12.48	−0.14	0.98	1.42	11.21
09-001-3005	Norwalk, CT	18	13.49	13.38	−0.11	0.99	1.32	9.81
09-001-9003	Westport, CT	15	11.07	11.05	−0.03	0.99	1.12	10.16
09-003-1003	E. Hartford, CT	56	13.99	14.01	0.02	0.98	1.41	10.04
09-003-1018	Hartford, CT	18	8.98	8.99	0.01	0.97	0.76	8.44
09-009-0018	New Haven, CT	45	19.46	19.30	−0.16	0.97	2.17	11.17
09-009-0026	New Haven, CT	18	12.32	12.30	−0.03	0.99	0.91	7.38
09-009-1123	New Haven, CT	25	12.54	12.45	−0.09	0.99	1.01	8.02
09-009-2008	New Haven, CT	25	14.36	14.35	0.00	0.99	1.40	9.72
09-009-2123	Waterbury, CT	25	11.44	11.41	−0.03	0.99	1.07	9.38
09-009-8003	W. Haven, CT	16	17.04	16.87	−0.18	0.98	2.77	16.28
09-011-3002	Norwich, CT	14	8.21	8.22	0.01	0.97	0.81	9.83
25-005-1004	Fall River, MA	12	11.16	11.20	0.04	0.95	2.37	21.21
25-009-2006	Lynn, MA	13	10.57	10.90	0.34	0.97	2.20	20.81
25-009-5005	Haverhill, MA	17	11.44	11.60	0.16	0.98	1.47	12.88
25-013-0008	Chicopee, MA	34	9.33	9.42	0.09	0.93	2.14	22.98
25-013-0016	Springfield, MA	44	12.73	12.73	0.00	0.97	1.70	13.39
25-013-2009	Springfield, MA	18	10.30	10.35	0.05	0.98	1.30	12.59
25-023-0004	Brockton, MA	18	8.99	9.15	0.16	0.95	2.29	25.45
25-025-0027	Boston, MA	15	14.17	14.32	0.15	0.92	2.90	20.50
25-025-0042	Boston, MA	22	15.24	15.27	0.03	0.98	1.81	11.85
25-025-0043	Boston, MA	6	15.45	15.45	0.00	0.99	1.53	9.89
25-027-0020	Worcester, MA	38	10.25	10.32	0.06	0.93	1.60	15.62

[a] The measured and predicted $PM_{2.5}$ concentrations, bias, and precision are in the unit of $\mu g\,m^{-3}$.

[b] Bias is defined as ($PM_{2.5\ predicted} - PM_{2.5\ measured}$).

[c] Precision is estimated as the square root of the mean of the squared errors.

[d] % Precision is defined as ($100 \times (precision/PM_{2.5\ measured})$).

Table 4. Comparisons of CV R^2 and % CV Precision ($\mu g\,m^{-3}$ for CV precision) between the measured and predicted $PM_{2.5}$ concentrations using mixed effects model and regression model[a].

Site ID	City	N	$PM_{2.5\ measured}$	$PM_{2.5\ predicted}$	Bias[b]	R^2	Precision[c]	% Precision[d]
Mixed effects model								
09-001-0010	Bridgeport, CT	15	11.59	10.66	−0.93	1.00	1.45	12.54
09-001-0113	Bridgeport, CT	19	9.64	8.84	−0.80	0.95	1.77	18.33
09-001-1123	Danbury, CT	16	13.96	13.43	−0.53	0.96	2.39	17.11
09-001-2124	Stamford, CT	14	12.63	11.23	−1.40	0.98	2.10	16.64
09-001-3005	Norwalk, CT	18	13.49	12.16	−1.33	0.99	2.03	15.06
09-001-9003	Westport, CT	15	11.07	10.72	−0.36	0.98	1.44	13.00
09-003-1003	E. Hartford, CT	56	13.99	14.64	0.65	0.95	2.29	16.34
09-003-1018	Hartford, CT	18	8.98	9.05	0.06	0.95	0.98	10.91
09-009-0018	New Haven, CT	45	19.46	14.56	−4.90	0.95	5.66	29.11
09-009-0026	New Haven, CT	18	12.32	12.00	−0.32	0.99	1.08	8.78
09-009-1123	New Haven, CT	25	12.54	11.01	−1.53	0.99	1.92	15.34
09-009-2008	New Haven, CT	25	14.36	14.31	−0.05	0.99	1.54	10.72
09-009-2123	Waterbury, CT	25	11.44	10.94	−0.50	0.99	1.33	11.62
09-009-8003	W. Haven, CT	16	17.04	15.09	−1.95	0.97	3.67	21.55
09-011-3002	Norwich, CT	14	8.21	8.30	0.09	0.96	1.03	12.49
25-005-1004	Fall River, MA	12	11.16	11.25	0.09	0.92	3.03	27.13
25-009-2006	Lynn, MA	13	10.57	13.77	3.20	0.96	4.08	38.58
25-009-5005	Haverhill, MA	17	11.44	13.52	2.08	0.97	2.88	25.20
25-013-0008	Chicopee, MA	34	9.33	11.55	2.22	0.90	3.43	36.78
25-013-0016	Springfield, MA	44	12.73	12.81	0.08	0.94	2.37	18.63
25-013-2009	Springfield, MA	18	10.30	10.89	0.59	0.97	1.60	15.51
25-023-0004	Brockton, MA	18	8.99	10.94	1.95	0.94	3.30	36.69
25-025-0027	Boston, MA	15	14.17	16.04	1.86	0.82	4.64	32.72
25-025-0042	Boston, MA	22	15.24	15.70	0.46	0.95	2.90	19.02
25-025-0043	Boston, MA	6	15.45	15.48	0.03	0.99	1.82	11.75
25-027-0020	Worcester, MA	38	10.25	12.10	1.84	0.87	2.94	28.66

Table 4. Continued.

Regression model								
09-001-0010	Bridgeport, CT	15	11.59	13.52	1.93	0.40	8.14	70.29
09-001-0113	Bridgeport, CT	19	9.64	11.95	2.30	0.67	4.43	45.95
09-001-1123	Danbury, CT	16	13.96	10.26	−3.69	0.71	8.21	58.80
09-001-2124	Stamford, CT	14	12.63	9.80	−2.83	0.88	6.96	55.10
09-001-3005	Norwalk, CT	18	13.49	12.24	−1.25	0.38	8.53	63.24
09-001-9003	Westport, CT	15	11.07	12.09	1.01	0.48	5.94	53.64
09-003-1003	E. Hartford, CT	56	13.99	13.04	−0.96	0.39	7.78	55.64
09-003-1018	Hartford, CT	18	8.98	10.76	1.78	0.41	3.84	42.79
09-009-0018	New Haven, CT	45	19.46	13.47	−5.99	0.44	11.00	56.52
09-009-0026	New Haven, CT	18	12.32	13.74	1.42	0.70	6.03	48.96
09-009-1123	New Haven, CT	25	12.54	11.84	−0.70	0.62	6.74	53.73
09-009-2008	New Haven, CT	25	14.36	14.05	−0.31	0.66	8.18	57.00
09-009-2123	Waterbury, CT	25	11.44	10.16	−1.28	0.63	5.75	50.25
09-009-8003	W. Haven, CT	16	17.04	13.28	−3.76	0.58	11.57	67.90
09-011-3002	Norwich, CT	14	8.21	9.32	1.11	0.43	4.04	49.25
25-005-1004	Fall River, MA	12	11.16	12.46	1.31	0.79	4.59	41.10
25-009-2006	Lynn, MA	13	10.57	13.80	3.23	0.72	8.73	82.56
25-009-5005	Haverhill, MA	17	11.44	12.48	1.04	0.73	6.36	55.61
25-013-0008	Chicopee, MA	34	9.33	10.19	0.86	0.25	6.19	66.39
25-013-0016	Springfield, MA	44	12.73	11.44	−1.28	0.30	8.13	63.86
25-013-2009	Springfield, MA	18	10.30	11.66	1.36	0.36	6.24	60.62
25-023-0004	Brockton, MA	18	8.99	11.02	2.03	0.44	5.76	64.03
25-025-0027	Boston, MA	15	14.17	21.08	6.90	0.12	12.73	89.83
25-025-0042	Boston, MA	22	15.24	18.24	3.00	0.40	10.25	67.24
25-025-0043	Boston, MA	6	15.45	19.44	3.99	0.68	9.54	61.74
25-027-0020	Worcester, MA	38	10.25	12.54	2.28	0.17	6.63	64.69

[a] The measured and predicted $PM_{2.5}$ concentrations, bias, and precision are in the unit of $\mu g\,m^{-3}$.

[b] Bias is defined as ($PM_{2.5\ predicted} - PM_{2.5\ measured}$).

[c] Precision is estimated as the square root of the mean of the squared errors.

[d] % Precision is defined as ($100 \times (precision/PM_{2.5\ measured})$).

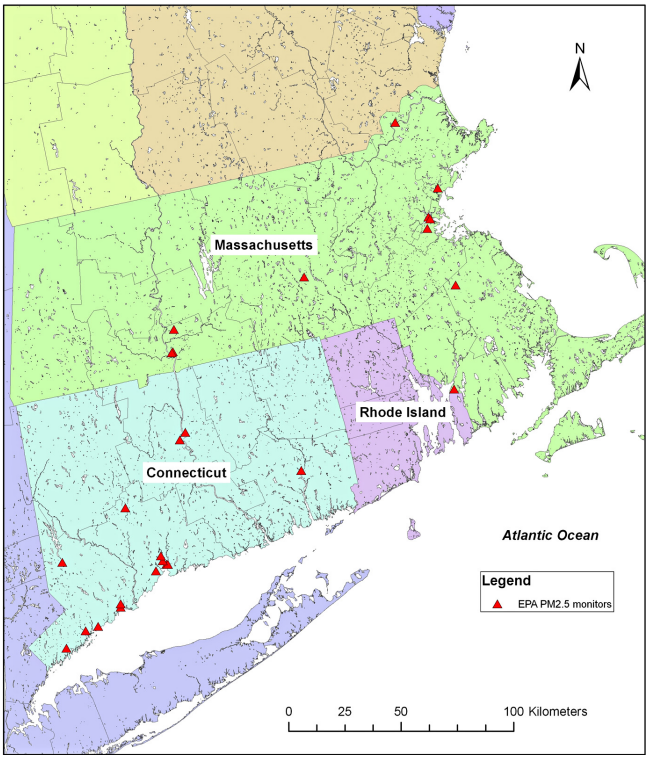

Fig. 1. $PM_{2.5}$ monitoring site locations in 2003.

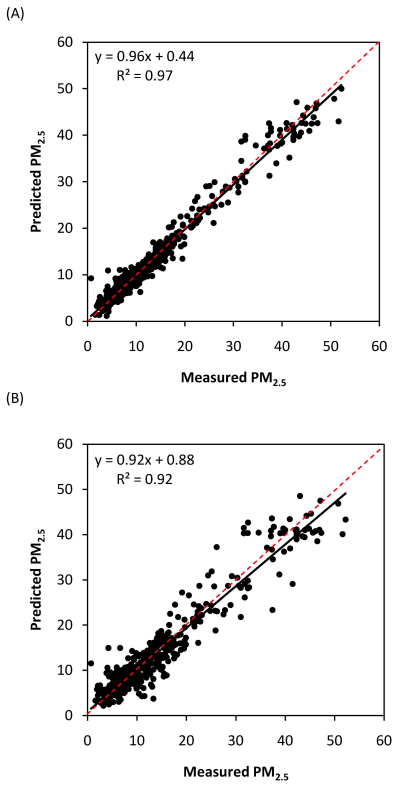

Fig. 2. Mixed effects model performance assessed by 576 measured and predicted daily $PM_{2.5}$ concentrations ($\mu g\,m^{-3}$) from: **(A)** mixed effects model and **(B)** CV mixed effects model. The solid line represents the regression line, and the dashed line displays the 1:1 line.

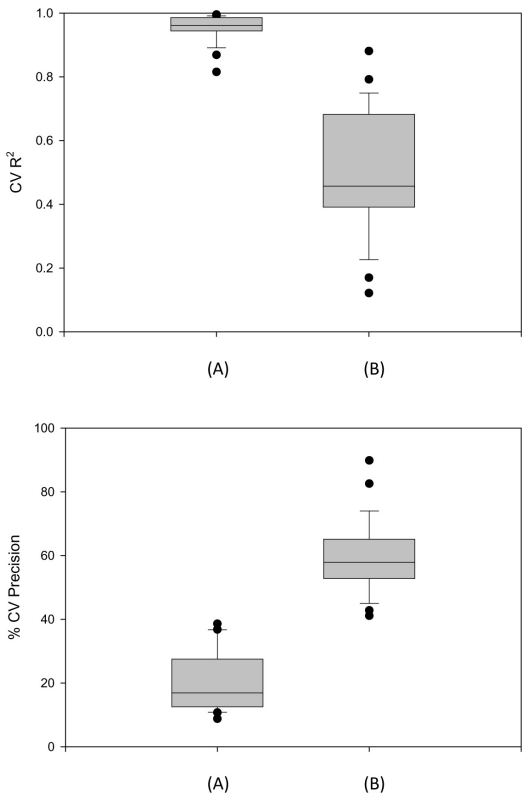

Fig. 3. Cross-validation correlation coefficients and % Precision between the measured and predicted $PM_{2.5}$ concentrations for the: **(A)** mixed effects model and **(B)** regression model.

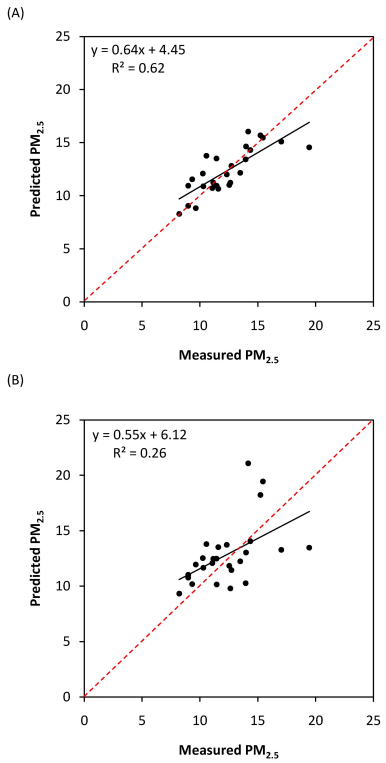

Fig. 4. Cross-sectional comparisons between the measured and predicted site mean $PM_{2.5}$ concentrations ($\mu g\,m^{-3}$) for the: **(A)** mixed effects model and **(B)** regression model (both from CV analyses). The solid line represents the regression line, and the dashed line displays the 1:1 line.

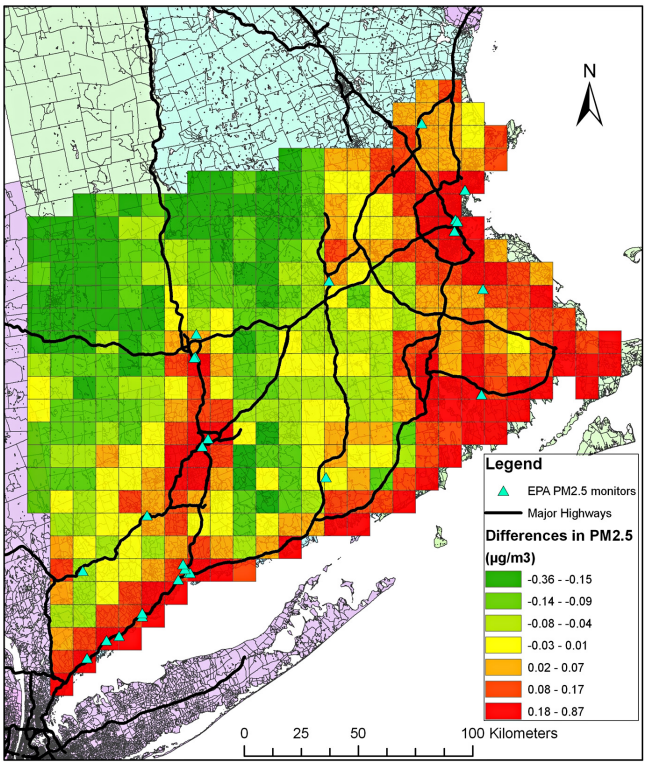

Fig. 5. Spatial variability in $PM_{2.5}$ levels in the study region. $PM_{2.5}$ levels are expressed as differences between grid-specific predicted and regional $PM_{2.5}$ concentrations ($\mu g\,m^{-3}$).