Atmospheric Environment 43 (2009) 4303–4310

Contents lists available at ScienceDirect

Atmospheric Environment

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

# Optimal temporal scale for the correlation of AOD and ground measurements of $PM_{2.5}$ in a real-time air quality estimation system

Hui Li [a,*], Fazlay Faruque [a], Worth Williams [a], Mohammad Al-Hamdan [b], Jeffrey Luvall [c], William Crosson [b], Douglas Rickman [c], Ashutosh Limaye [b]

[a] University of Mississippi Medical Center, 2500 North State Street, Jackson, MS 39216-4505, USA
[b] Universities Space Research Association, NASA Marshall Space Flight Center, Huntsville, AL 35805, USA
[c] Earth Science Office, NASA Marshall Space Flight Center, Huntsville, AL 35805, USA

ARTICLE INFO

*Article history:*
Received 19 February 2009
Received in revised form
29 May 2009
Accepted 3 June 2009

*Keywords:*
Aerosol optical depth (AOD)
$PM_{2.5}$
Air quality
Temporal scale
Remote sensing
Real-time system

ABSTRACT

Aerosol optical depth (AOD), an indirect estimate of particulate matter using satellite observations, has shown great promise in improving estimates of $PM_{2.5}$ (particulate matter with aerodynamic diameter less than or equal to 2.5 μm) surface. Currently, few studies have been conducted to explore the optimal way to apply AOD data to improve the model accuracy of $PM_{2.5}$ in a real-time air quality system. We believe that two major aspects may be worthy of consideration in that area: 1) an approach that integrates satellite measurements with ground measurements in the estimates of pollutants and 2) identification of an optimal temporal scale to calculate the correlation of AOD and ground measurements. This paper is focused on the second aspect, identifying the optimal temporal scale to correlate AOD with $PM_{2.5}$. Five following different temporal scales were chosen to evaluate their impact on the model performance: 1) within the last 3 days, 2) within the last 10 days, 3) within the last 30 days, 4) within the last 90 days, and 5) the time period with the highest correlation in a year. The model performance is evaluated for its accuracy, bias, and errors based on the following selected statistics: the Mean Bias, the Normalized Mean Bias, the Root Mean Square Error, Normalized Mean Error, and the Index of Agreement. This research shows that the model with the temporal scale: within the last 30 days, displays the best model performance in a southern multi-state area centered in Mississippi using 2004 and 2005 data sets.


## 1. Introduction

Aerosol optical depth (AOD), derived from satellite measurements, is a measure of atmospheric extinction of radiance, due to scattering and absorption by aerosols, through a vertical column in the atmosphere. AOD offers indirect estimates of particle matter. Previous research showed a significant positive correlation between satellite-based measurements of AOD and ground-based measurements of particulate matter with aerodynamic diameter less than or equal to 2.5 μm ($PM_{2.5}$) and aerodynamic diameter less than or equal to 10 μm ($PM_{10}$) (Chu, 2006; Gupta et al., 2006; Li et al., 2003; Engel-Cox et al., 2004; Wang and Christopher, 2003; Liu et al., 2007). In addition, satellite observations have shown great promise in enhancing air quality event monitoring (Engel-Cox et al., 2004, 2005; Hutchison, 2003; Kacenelenbogen et al., 2006; Chu et al., 2003), improving estimates of $PM_{2.5}$ air quality spatial surface (Al-Hamdan et al., 2009; Gupta et al., 2006; Kumar et al., 2007), and enhancing national air quality forecasts (Al-Saadi et al., 2005). Satellite observations have also been used as a direct input in models (data assimilation) to increase the model performance (Collins et al., 2001; Koelemeijer et al., 2006). Research shows that correlations between AOD and ground $PM_{2.5}$ are affected by a combination of many factors, including inherent characteristics of satellite observations, aerosol optical depth algorithms, errors of estimate of regression models, terrain, cloud cover, height of the mixing layer, relative humidity, wind velocity, temperature, aerosol chemical composition, and sea-level atmospheric pressure conditions (Kumar et al., 2007; Gupta et al., 2006; Schaap et al., 2008; Liu et al., 2007). Therefore, the corrections may vary widely in different regions, different seasons, and even on different days in one location. For example, Engel-Cox et al. (2004) finds that the correlations are stronger in the eastern half of the United States, while they are weaker in the western United States. They believe that some of the general variation between AOD and PM measurements is caused by the artifact of linear analysis, different terrain conditions, and inherent differences in the data sets. Gupta et al. (2006) found that

* Corresponding author. Tel.: +1 601 984 6288; fax: +1 601 984 6214.
*E-mail address:* hli@dor.umsmed.edu (H. Li).

1352-2310/$ – see front matter © 2009 Elsevier Ltd. All rights reserved.
doi:10.1016/j.atmosenv.2009.06.004

the relationship between AOD and $PM_{2.5}$ was higher for cloud-free conditions, low boundary layer heights, and low relative humidity. Pelletier et al. (2007) found that meteorological variables improved the relationship between AOD and $PM_{10}$.

The temporal scale of the correlation between AOD and $PM_{2.5}$ in this study is defined as the number of preceding days that cumulate AOD data for a modeled day (see Fig. 1). Analysis of correlating AOD with ground measured $PM_{2.5}$ on a day-to-day basis in this research suggests that the temporal scale to be used in determining their correlations needs to be considered to improve estimates of air quality surface. In addition, correlation coefficients between AOD and ground measured $PM_{2.5}$ cannot be predetermined in real-time daily air quality estimation and need to be calculated for each day's run, because the coefficients can vary in different seasons and even different days. Few studies have been conducted to explore the optimal way to apply AOD data to improve model accuracies of $PM_{2.5}$ spatial surface estimation in a real-time air quality system. We believe that two major aspects may be worth considering when applying satellite data to improve the performance of pollution spatial surface models: 1) an approach that integrates satellite observations with ground measurements for the pollution estimates of pollutants and 2) identification of an optimal temporal scale to calculate the correlation of AOD and ground measurements. This paper will focus on the second aspect and discuss the best temporal scale to calculate the correlation of AOD and ground particulate matter to improve the performance of air quality models in a real-time system.

## 2. Real-time $PM_{2.5}$ estimation system

The near real-time $PM_{2.5}$ estimation system used in this research is built from a $PM_{2.5}$ spatial surface model, originally developed by NASA Marshall Space Flight Center (NASA/MSFC) and described in Al-Hamdan et al. (2009). This surface model was improved and integrated into a real-time geo-spatial health surveillance system (GeoMedStat) developed at the University of Mississippi Medical Center. The system estimates daily concentration of $PM_{2.5}$ for Mississippi and its neighboring states of Arkansas, Tennessee, Alabama, Florida, and Texas, using NASA MODIS AOD data on board Terra and Aqua satellites and EPA air quality ground measurements from the AirNow gateway system. The system calculates daily average ground-level $PM_{2.5}$ in a batch mode on a daily-basis with a 2-day delay due to the lag in satellite data received in the system. The model adopts the same spatial resolution as that of satellite data for its outputs (about $10 \times 10$ km). Ground measurements of $PM_{2.5}$ from EPA monitoring stations and satellite-derived AOD from MODIS instruments on aboard NASA's Earth Observing System Terra and Aqua satellites are two major

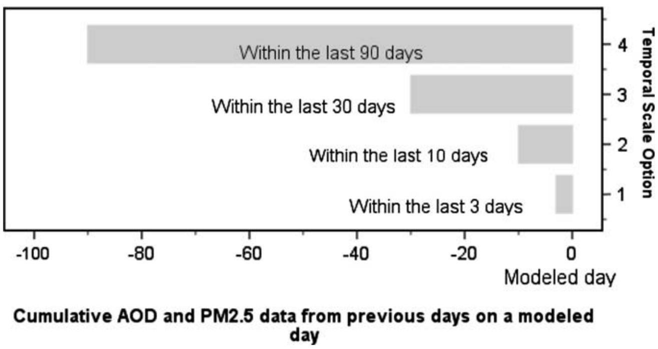


**Fig. 1.** Schematic diagram of the first four temporal scales in calculating the correlation of AOD and ground measurements in this study.

data sources used to estimate daily ground-level air quality surface in the system. The ground measurements of daily average $PM_{2.5}$ data are downloaded automatically from the AirNow gateway system using the File Transfer Protocol (FTP) on a daily-basis. MODIS-derived AOD data from Terra and Aqua, stored in a Hierarchical Data Format (HDF), are also obtained automatically from the NASA-Goddard Earth Sciences (GES) FTP site on a daily-basis. Satellite-derived AOD is a Level 2 atmospheric product from MODIS instruments on board Terra and Aqua platforms at a spatial resolution of about 10*10 km nearly covering the entire surface of the Earth everyday. Two NASA MODIS sensors are currently in orbit on the Terra and Aqua satellites, which, in sun-synchronous orbit, observe any location on the Earth's surface at about 10:30 AM and 1:30 PM local standard time, respectively, each day. Because MODIS AOD data from Terra are found to have better relationship with ground measurements of $PM_{2.5}$ than MODIS AOD from Aqua in the study area, AOD data from Terra are used in the model by default. However, whenever the relationship is found not statistically significant, the system will automatically switch to using AOD data generated on board Aqua. If neither relationship is statistically significant, then only ground measurements are used in the model.

The system includes the following three main components: 1) AOD–$PM_{2.5}$ linear regression models for AOD-derived $PM_{2.5}$; 2) a surface model to interpolate AOD-derived $PM_{2.5}$ and ground measurements of $PM_{2.5}$ to a continuous grid surface respectively using B-spline algorithms; and 3) an approach that integrates the two interpolated surfaces above into a final surface output if a significant relationship is found between them on each calculated day; otherwise, only ground measurements are used for the model output. The model domain is shown in Fig. 2, which also illustrates the distribution of the monitoring stations used in the air quality model.

## 3. Methodology

To identify the optimal temporal scale for the AOD–$PM_{2.5}$ correlations, we chose the following five different temporal scales to evaluate their impact on the performance of the daily-basis air quality surface model in both 2004 and 2005: 1) within the last 3 days, 2) within the last 10 days, 3) within the last 30 days, 4) within the last 90 days, and 5) the time period with the highest correlation in a year (August–October in 2004 and June–September in 2005). For the first four temporal scales, the regression analysis was conducted on the fly to determine the significant relationship between AOD and $PM_{2.5}$ on each modeled day, based on the $p$-value at a significant level of 5%. First, each EPA monitoring station was identified in the study area by its longitude and latitude. Second, all corresponding pixels of satellite observations within the 0.1° distance range of each station were identified in the AOD data set on each accumulated day, which was inside the evaluated temporal scale range. Third, for each involved AOD daily data, only the first three identified pixels, closest to their paired station, were kept for further process. Fourth, the pairing AOD value of each station was estimated by averaging the AOD values from all identified pixels in the above process on each accumulated day. Once the satellite measurements were paired with all stations on a modeled day, a linear regression model was fitted to the identified paired data on a day-by-day basis. When their relationship was considered statistically significant ($p$-value less or equal to 5%), AOD data were determined to be used in the model. As for the last temporal scale, a predetermined regression model was used for the model estimation in the defined time period with the highest correction in each evaluating year (August–October, 2004 and June–September, 2005).

To make the accuracy assessment subjective, a station site with the ID 280810005 (seen in Fig. 1) was left out in the air quality estimation and was only used for the performance evaluation. The