

# QUANTITATIVE EVALUATION OF SATELLITE AEROSOL DATA FOR MAPPING FINE PARTICULATE AIR POLLUTION IN THE CONTERMINOUS UNITED STATES

Zhiyong Hu  
Department of Environmental Studies  
University of West Florida  
11000 University Parkway  
Pensacola, FL 32514 USA  
Email: [zhu@uwf.edu](mailto:zhu@uwf.edu)

## Abstract

**Introduction** Existing studies have shown that fine aerosol particles have more adverse health effects than course particles and gases. Air quality monitoring has long relied on ground measurement at point monitoring sites. Ground monitoring data often lacks spatially and temporally complete coverage. In recently years, repetitive and broad coverage capabilities of satellites allow atmospheric remote sensing to offer a unique opportunity to monitor air quality at continental, national and regional scales. The NASA Earth Observation System (EOS)'s MODIS (Moderate Resolution Imaging Spectrometer) aerosol optical depth (AOD) data could be used to improve ground measurement of fine particulate matter.

**Objectives** This paper quantitatively examined the relationship between  $PM_{2.5}$  concentration and MODIS AOD data in the conterminous U.S. and assessed the potential of using AOD data as an indicator of particulate air pollution. If an association should be found, AOD data would be used to map particulate matter concentration.

**Methodology** Relationship between AOD and ground-based  $PM_{2.5}$  observations was examined at two levels. One used MODIS Level 2 data, and the other used Level 3 annual mean data. Daily MODIS Level 2 images covering the conterminous U.S. were compared with hourly  $PM_{2.5}$  data for the year 2004 with imaging time collated with  $PM_{2.5}$  measurement time within one hour and monitoring site locations matched to conjugate pixels. The Pearson's correlation value was calculated for each point containing at least 8 pairs of data values. For the year 2004, a geographically weighted regression (GWR) model was also fitted to examine the spatial variation of the relationship between  $PM_{2.5}$  and AOD using annual mean  $PM_{2.5}$  and MODIS Level 3 AOD. The possible association between AOD and  $PM_{2.5}$  was then used to predict annual mean  $PM_{2.5}$  values using AOD

Level 3 annual mean values and the prediction accuracy was assessed against ground measurement.

**Results** Significant positive correlations between PM<sub>2.5</sub> and AOD were found to be to the east of the -100° longitude line. Eighteen of the twenty sites with highest correlation values ( $r > 0.8$ ) are in the east. GWR predicts well in the eastern U.S. and poorly in the west, as indicated by the map of local R square. The coefficient raster surface (AOD) exhibits regional variation. The relationship between PM<sub>2.5</sub> and AOD is not spatially consistent (stationary) across the conterminous states. Eastern U.S. shows higher AOD coefficient values, while values in the west are lower. PM<sub>2.5</sub> calculation for the east and the year 2003 using the linear regression equation fitted on the year 2004 data achieved an accuracy of 2.24  $\mu\text{m}/\text{m}^3$ .

**Conclusions** EPA PM<sub>2.5</sub> ground data is positively related to MODIS AOD data in the eastern U.S. where the aerosol retrieval algorithm used the urban/industrial aerosol model. Urban-industrial aerosols are mainly from fossil fuel combustion in populated industrial regions and dominated by fine particles. The positive relationship in the east could exist in other regions in the world where the aerosol retrieval algorithm uses the urban/industrial aerosol model. Satellite aerosol optical depth data can be used as an alternative indicator of air quality for regions dominated by anthropogenic fine model aerosol particles.

## **Introduction**

Fine particulate matter (PM) causes more visibility degradation and possibly more health problems than do gases. Air quality monitoring has long relied on ground measurement at individual monitoring sites. Ground monitoring data often lacks spatially complete coverage. In recently years, repetitive and broad coverage capabilities of satellites allow atmospheric remote sensing to offer a unique opportunity to monitor air quality at continental, national and regional scales. This project examined the relationship between PM<sub>2.5</sub> concentration and aerosol optical depth (AOD) measured by the NASA MODIS (Moderate Resolution Imaging Spectrometer) sensor and assessed the potential of using AOD data as an indicator of particulate air pollution.

## **MODIS**

The MODIS instrument flies on polar-orbiting and sun-synchronous Terra and Aqua satellites of the Earth Observation System (EOS). MODIS performs measurements in the solar to thermal infrared spectrum region from 0.41 to 14.235  $\mu\text{m}$  (Salomonson et al., 1989). The wide spectral range and high spatial resolution with daily global coverage (e.g., 500 m at 0.47 to 2.12  $\mu\text{m}$  with 250 m at 0.66 and 0.86  $\mu\text{m}$  and 1 km at 1.38  $\mu\text{m}$ ) allow excellent cloud rejection while maintaining high statistics of cloud free pixels. The

MODIS sensor was expected to be the key for monitoring global aerosol properties. Not only have MODIS aerosol products been used to answer scientific questions about radiation and climate e.g., IPCC, 2001; Yu, et al., 2006), they are being used for applications not previously intended. Some examples include monitoring surface air quality for health (e.g., Chu et al., 2003; Hutchison, 2003; Engel-Cox et al., 2004; Al-Saadi et al., 2005; Hutchison et al., 2005; Gupta et al., 2006).

One of the fundamental aerosol products from MODIS is spectral AOD (or  $\tau$ ). The latest version of the MODIS aerosol retrieval algorithm is Collection 5 (C005) (Remer et al., 2006). The aerosol retrieval makes use of seven wavelength bands (channels 1 to 7), and a number of other bands to help with cloud and other screening procedures. Only data from MODIS daytime orbits are considered for the retrieval. The aerosol algorithm relies on calibrated, geolocated reflectance data (known as 'L1B'). In addition, the MODIS algorithm uses two processed products known as 'Level 2' (L2). These include cloud mask product (Ackerman et al., 1998) and atmospheric profile product. Finally, the algorithm expects ancillary data from NCEP (National Center for Environmental Prediction) analyses, including hourly meteorological analysis and daily ozone analysis. Daily optical depth and mass concentration are stored on a resolution of 10x10 pixels (of 1 km nadir resolution). To retrieve aerosol data, an aerosol model is selected that describes the aerosol size distribution, assigns values for refractive index and single scattering albedo, and describes the effect of nonsphericity on the phase function. The models are derived from studies at different places around the world in which *in situ* measurements were taken and ground instruments used to measure the ambient column aerosol size distribution. Four aerosol models are utilized in the algorithm over the land, including a continental model, a biomass burning model, an industrial/urban aerosol model, and a dust model.

### **Data**

MODIS daily Level 2 (2003-2004) and monthly Level 3 (2000-2004) data were obtained from the NASA Level 1 and Atmosphere Archive and Distribution System (LAADS Web) at (<http://ladsweb.nascom.nasa.gov/>). MODIS Level 2 data are produced at the spatial resolution of a 10x10 1-km (at nadir)-pixel array. The study used both two MODIS aerosol data product files: MOD04\_L2, containing data collected from the Terra platform; and MYD04\_L2, containing data collected from the Aqua platform. The Scientific Data Set (SDS) 'Optical Depth Land And Ocean' was used for analysis. The 'Optical Depth Land And Ocean' includes only those retrievals with quality assurance confidence (QAC) greater than zero. This is the recommended product for quantitative studies.

MODIS Level-3 aerosol monthly data was derived from daily Level-3 data, then yearly average images (2003, 2004) were calculated using the monthly data. Data from both

satellite platforms (MOD08\_M3 for Terra and MYD08\_M3 for Aqua) were used. The Level-3 MODIS Atmosphere Daily Product was derived from approximately 80 scientific parameters from four Level-2 MODIS Atmosphere Products: Aerosol, Water Vapor, Cloud, and Atmosphere Profile. Statistics are sorted into 1x1 degree cells on an equal-angle grid that spans a (calendar) monthly interval and then summarized over the globe. Land-based measures of fine-mode fraction of AOD can only be used as a qualitative indicator of whether AOD values are dominated by natural or anthropogenic emissions. This project did not use fine-mode AOD for statistical analyses. Over the land of the conterminous U.S., the aerosol retrieval algorithm used the urban/industrial aerosol model for the east and the moderate absorption model (based on biomass burning) for the west. Industrial/urban aerosol and biomass burning are considered as the main anthropogenic sources of aerosol. The line splitting between non-absorbing and moderately absorbing aerosol types is at approximately -100° longitude.

PM<sub>2.5</sub> ground data was obtained from the EPA Air Quality System (AQS) online Data Mart at <http://www.epa.gov/ttn/airs/aqsdatamart/index.htm>. PM<sub>2.5</sub> values measured within 1 hour of the MODIS imaging time were retrieved for the year 2004. Annual statistical summary (mean) PM<sub>2.5</sub> data for 2003 and 2004 were also obtained for monitoring sites covering the conterminous land. Calculation of annual statistics included exceptional air pollution events.

### **Pearson's correlation**

At Level 2, collation was made between AOD and PM<sub>2.5</sub> data using GIS with satellite imaging time collated with ground measurement time within one hour. A monitor was assigned the temporally collated AOD value of the pixel containing the site. This resulted in 3,378 pairs of ground PM<sub>2.5</sub> and satellite AOD observations with point location information. Pearson's correlation value was calculated for each point containing at least 8 pairs of data values. Eighteen of the twenty sites with highest correlation values ( $r > 0.8$ ) are located east of the -100° longitude line. The highest correlation ( $r = 0.96$ ) is at Nashville, Tennessee. A monitoring site at Cedar Rapids, Iowa was selected for comparison of patterns of temporal changes in PM<sub>2.5</sub> and AOD in 2004. The patterns of PM<sub>2.5</sub> and AOD at this site match very well ( $r = 0.806$ ). A correlation surface (Figure 1) was created using the Kriging spatial interpolation method. The surface indicates that the eastern part of the U.S. has higher positive correlation between PM<sub>2.5</sub> and AOD.

### **Geographically weighted regression**

For annual mean data at Level 3 for the year 2004, a geographically weighted regression (GWR) model was fitted to examine the relationship between PM<sub>2.5</sub> and AOD. GWR is a local form of linear regression used to model spatially varying relationships (Fotheringham et al., 2002). GWR is one of several spatial regression techniques, increasingly used in geography and other disciplines. GWR provides a local model of the

variable or process we are trying to understand/predict by fitting a regression equation to every feature in the dataset. GWR constructs these separate equations by incorporating

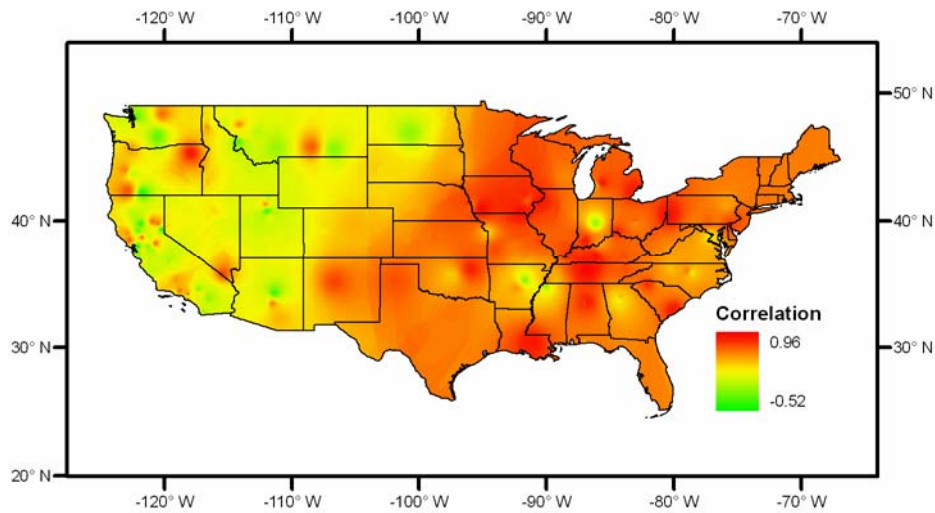


Figure 1. Surface of Pearson's correlation between  $PM_{2.5}$  and AOD

the dependent and explanatory variables of features falling within a kernel of each target feature. GWR accounts for the effect of local multicollinearity which occurs when the values for a particular explanatory variable cluster spatially. GWR creates coefficient raster surfaces for the model intercept and each explanatory variable. At Level 3, there are 3,264 pairs of data covering 867 monitoring sites across the conterminous U.S. for different time points. A fixed kernel was used to fit the regression. The bandwidth value was chosen by using the corrected Akaike Information Criterion (AICc). This option tries to identify an optimal fixed distance.

Results from the GWR show that all the monitoring sites have a condition number less than 30. The condition number evaluates local collinearity. In the presence of strong local collinearity, results become unstable. Results associated with condition numbers larger than 30, may be unreliable. Figure 2 shows a map of local  $R^2$  values.  $R^2$  values range between 0.0 and 1.0 and indicate how well the local regression model fits observed  $PM_{2.5}$  values. Very low values indicate the local model is performing poorly. It can be seen that GWR predicts well in the eastern U.S. and poorly in the west. Figure 3 shows the coefficient raster surface for the explanatory variable AOD created by the GWR. The map exhibits regional variation in the model explanatory variable. The relationship between  $PM_{2.5}$  and AOD is not spatially consistent (stationary) across the conterminous states. Like Pearson's correlation surface, eastern U.S. shows higher AOD coefficient

values, while values in the west are lower. Even negative values are found in part of the western region.

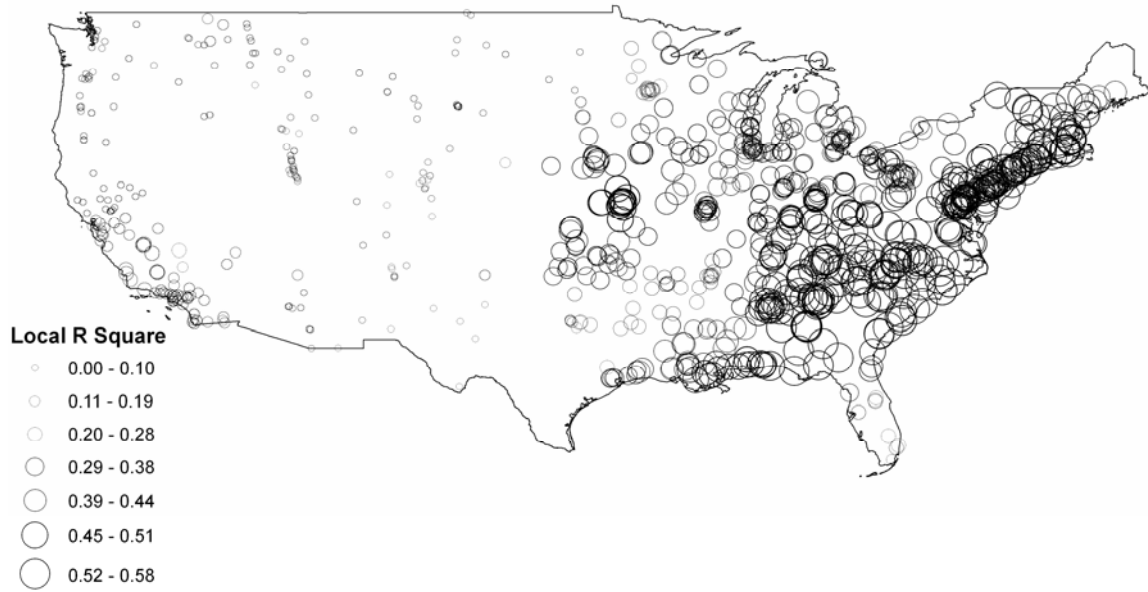


Figure 2. Local R square of geographically weighted regression

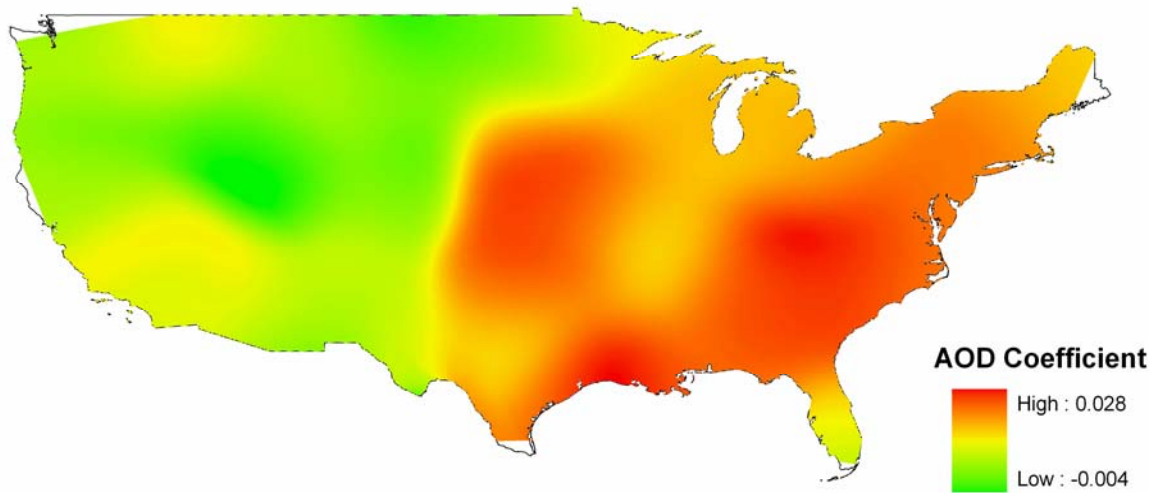


Figure 3. Coefficient raster surface for AOD from geographically weighted regression

### Mapping PM<sub>2.5</sub> concentration using MODIS AOD data

The strong correlation between PM<sub>2.5</sub> and AOD in the east suggests that AOD could be used as an alternative air quality indicator and a predictor of PM<sub>2.5</sub> for the region. To test the ability of AOD to predict PM<sub>2.5</sub>, a linear regression equation was fitted using the 2004 level 3 data for the east region. Figure 4 shows the spatial collation of the 2004 annual mean MODIS Level 3 AOD with annual mean PM<sub>2.5</sub>. Like Level 2 data collation, the AOD value of the pixel containing a monitoring site was assigned to the site. If there are multiple sites within a pixel, the average PM<sub>2.5</sub> was calculated for the pixel. The regression result ( $PM_{2.5} = 8.14 + 27.57 \times AOD$ ,  $R^2 = 0.53$ ,  $p < 0.001$  for both coefficients) indicates that that Level 3 data exhibits similar correlation as Level 2 data. The equation was applied to the 2003 Level 3 annual mean AOD data to calculate a PM<sub>2.5</sub> surface (Figure 5). The calculated values were compared to the ground measures to calculate the root mean square error (RMSE) of the backward prediction. The RMSE value of  $2.24 \mu g/m^3$  indicates good predictability.

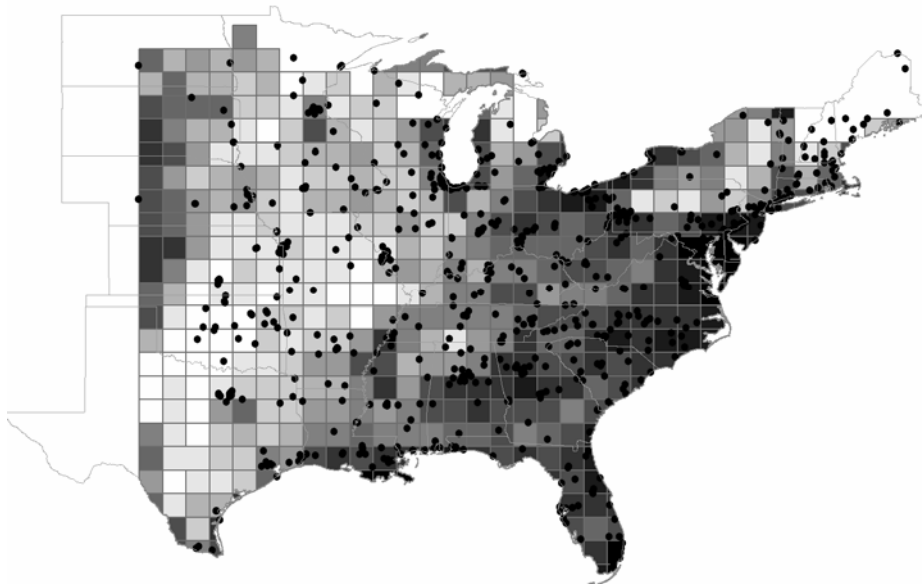


Figure 4. Collation of 2004 annual mean MODIS Level 3 AOD with annual mean PM<sub>2.5</sub>

### Conclusion and discussion

In the west of the U.S., MODIS AOD does not correlate with PM<sub>2.5</sub>. PM<sub>2.5</sub> ground data is positively related to MODIS AOD data in the eastern U.S. where the aerosol retrieval algorithm uses the urban/industrial aerosol model. Urban–industrial aerosols are mainly from fossil fuel combustion in populated industrial regions and dominated by fine particles. The positive relationship in the east could exist in other regions of the world

where the aerosol retrieval algorithm uses the urban/industrial aerosol model. While spatial coverage for  $PM_{2.5}$  is adequate in the U.S., many developing countries do not

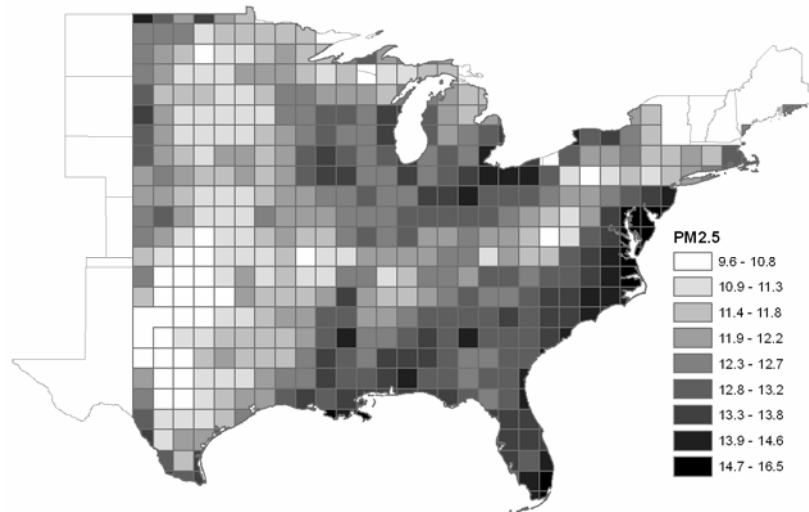


Figure 5. 2003 annual mean  $PM_{2.5}$  calculated using the regression equation of 2004 collated level 3 data.

have an extensive network of air quality monitors. Satellite aerosol optical depth data can be used as an alternative indicator of air quality for regions where aerosols are dominated by anthropogenic fine model particles.

It must be noted that there are limitations of AOD data. Since AOD does not differentiate between fine-mode and coarse-mode particles, the relationship does not necessarily hold for all regions. AOD does not specify the location of aerosols within a column and the AOD values do not necessarily represent ground conditions. Emerging Light Detection and Ranging (LIDAR) systems could provide vertical resolution for AOD and quantify pollutant levels on the ground.

## References

- Ackerman, SA, Strabala, KI, Menzel, WP, Frey, RA, Moelle, CC, Bumley, LE, 1998. Discriminating clear sky from clouds with MODIS. *Journal of Geophysical Research* 103(32), 139-140.
- Al-Saadi J, Szykman J, Pierce RB, Kittaka C, Neil D, Chu DA, Remer L, Gumley L, Prins E, Weinstock L, MacDonald C, Wayland R, Dimmick F, Fishman J, 2005. Improving national air quality forecasts with satellite aerosol observations. *Bulletin of American Meteorological Society* 86(9), 1249-1261.



- Chu, DA, Kaufman YJ, Zibordi G, Chern JD, Mao J, Li C, Holben BN, 2003. Global monitoring of air pollution over land from EOS-Terra MODIS. *Journal of Geophysical Research* 108 (D21), 4661, doi: 10.1029/2002JD003179.
- Engel-Cox JA, Holloman CH, Coutant BW, Hoff RM, 2004. Qualitative and quantitative evaluation of MODIS satellite sensor data for regional and urban scale air quality. *Atmospheric Environment* 38, 2495-2509.
- Fotheringham SA, Brunson, Chris, Charlton, Martin, 2002. *Geographically Weighted Regression: the analysis of spatially varying relationships*. John Wiley & Sons.
- Gupta P, Christopher SA, Wang J, Gehrig R, Lee Y, Kumar N, 2006. Satellite remote sensing of particulate matter and air quality assessment over global cities. *Atmospheric Environment* 40,5880-5892.
- Hutchison KD, 2003. Applications of MODIS satellite data and products for monitoring air quality in the state of Texas. *Atmospheric Environment* 37, 2403-2412.
- Hutchison KD, Smith S, Faruqui SJ, 2005. Correlating MODIS aerosol optical thickness data with ground-based PM<sub>2.5</sub> observations across Texas for use in a real-time air quality prediction system. *Atmospheric Environment* 39, 7190-7203.
- Intergovernmental Panel on Climate Change (IPCC), 2001. *Climate Change 2001: The Scientific Basis*, JT Houghton, Y Ding, DJ Griggs, M Noguer, PJ van der Linden and D Xiaosu (Eds.), Cambridge University Press, UK. pp 944.
- Kaufman, YJ, Wald AE et al., 1997. The MODIS 2.1- $\mu$ m channel - Correlation with visible reflectance for use in remote sensing of aerosol. *IEEE Transactions On Geoscience and Remote Sensing* 35(5), 1286-1298.
- Kaufman, YJ, Boucher O et al., 2005. Aerosol anthropogenic component estimated from satellite data. *Geophysical Research Letters* 32(17).
- Remer LA, Tanré D, Kaufman YJ, Levy R, Mattoo S, 2006. Algorithm for remote sensing of tropospheric aerosol from MODIS: Collection 005. [http://modis-atmos.gsfc.nasa.gov/docs/MOD04:MYD04\\_ATBD\\_C005\\_rev1.pdf](http://modis-atmos.gsfc.nasa.gov/docs/MOD04:MYD04_ATBD_C005_rev1.pdf)
- Salomonson, VV, Barnes WL et al., 1989. MODIS: Advanced Facility Instrument for Studies of the Earth as a System. *IEEE Transactions On Geoscience and Remote Sensing* 27, 145-153.
- Yu H, Kaufman YJ, Chin M, Feingold G, Remer L, Anderson T, Balkanski Y, Bellouin N, Boucher O, Christopher S, DeCola P, Kahn R, Koch D, Loeb N, Reddy MS, Schulz M, Takemura T, Zhou M, 2006. A review of measurement-based assessments of aerosol direct radiative effect and forcing. *Atmospheric Chemical Physics* 6, 613-666.