ASSESSMENT OF PM2.5 RETRIEVALS USING A COMBINATION OF SATELLITE AOD AND WRF PBL HEIGHTS IN COMPARISON TO WRF/CMAQ BIAS CORRECTED OUTPUTS

Lina Cordero*, Nabin Malakar, Yonghua Wu, Barry Gross, Fred Moshary
Optical Remote Sensing Lab, City College of New York, New York, NY, USA 10031

1. INTRODUCTION

Fine particulate matter with particle diameters less than 2.5 microns (PM2.5) has been linked to respiratory and pulmonary difficulties and for this reason, strong concentration guidelines have been developed by the U.S. Environmental Protection Agency (EPA) to limit exposure [Pope et al. (2002); US EPA (2004); Hoff and Christopher (2009)]. To assess compliance, the EPA generally rely on air quality measurements made by expensive ground-based monitors which unfortunately limits the spatial extent of air quality networks. In order to overcome this limitation, satellite remote sensing of aerosol path integrated properties has become a major tool.

In particular, significant efforts have been made to connect Aerosol Optical Depth (AOD), which is a measure of the path integrated aerosol extinction, to estimate ground-level PM2.5; however, a wide range of factors such as aerosols variability, meteorology and the vertical structure of aerosols, which is often (but not always) constrained by the Planetary Boundary Layer (PBL) height, affect the relationship between AOD and PM2.5 [Tsai et al. (2009); Zhang et al. (2009); Boyouk et al. (2010)].

In another approach, a global model (GEOS-CHEM) is being used to estimate the spatial relationship between PM2.5 forecast and column path AOD on a daily basis. This approach is currently being used by IDEA (Infusing satellite Data into Environmental air quality Applications) providing real time spatial maps of PM2.5 [van Donkelaar (2012]. Unfortunately, the low spatial resolution (0.5 deg) reduce the effectiveness of this approach in dealing with urban-suburban domains. To improve on this, the WFR/CMAQ model is used. This is a high-resolution algorithm that accounts for physical meteorological factors and surface boundary conditions including emission inventories to estimate particulate concentrations and vertical distributions [Hu et al. (2010); Hogrefe et al. (2011)].

In this paper, we first focus on local ground measurements from a CIMEL sun/sky radiometer, LIDAR and TEOM instruments at The City College of New York (CCNY) to explore a neural network (NN) approach at one urban location. In particular, we assess the importance of ingesting the lidar derived planetary boundary layer height into the NN fine particulate matter estimator. Later, we discuss a NN method over the entire NY state region ingesting WRF meteorological information and compare its performance to existing GEOS-CHEM and CMAQ products.

2. METHODOLOGY

We used a total of 41 stations from the New York state area, and perform regional seasonal comparisons between ground PM2.5 and satellite/model PM2.5. Figure 1 illustrates the NY state area and the locations of the stations as well as the urban classification, which is primarily based on site location; mostly, stations located in the New York City metro area are depicted as urban.

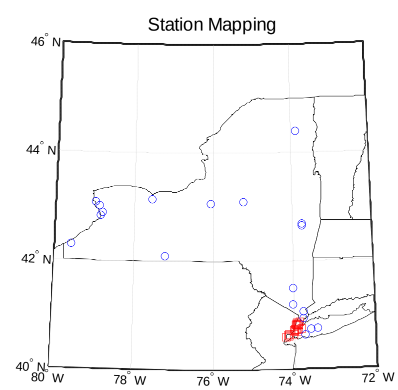

Figure 1. New York State station mapping: red squares are urban stations and blue circles are non-urban stations.

In order to assess the PM2.5-AOD relationship, we have collected hourly satellite AOD data from the Moderate Resolution Imaging Spectroradiometer (MODIS) on board of TERRA

*Corresponding author: Lina Cordero, The City College of New York, New York, NY, USA 10031; e-mail: lina.cordero@gmail.com

and AQUA for the period corresponding from January 1, 2006 to December 31, 2007. Also, hourly PM2.5 data from the New York State Department of Environmental Conservation (NYSDEC) have been collected along with Community Multiscale Air Quality - Weather Research Forecast (CMAQ-WRF) planetary boundary layer height. Datasets are converted to daily averages following EPA regulations for further comparisons with GEOS-CHEM and CMAQ products.

Additionally, a dataset corresponding to measurements acquired from an AERONET (AErosol RObotic NETwork) Cimel sun/sky radiometer (CE-318), a Lidar (Light Detection and Ranging) and a TEOM (Tapered Element Oscillating Microbalance) instruments plus meteorological information (relative humidity, wind speed, etc.) have been used to perform initial comparisons between ground measurements versus satellite and model products. These instruments are located at the City College of New York (CCNY, 40.821 °N/73.949 °W) and therefore this analysis is limited to one urban station.

3. RESULTS

3.1 PBL height and seasonality assessment on a local neural network

While strong correlations can be found connecting AOD to PM2.5 in the northeast, PBL height information can provide significant improvements. This may be due to the fact that the PBL typically contains greater aerosol concentration than the overlying troposphere and hence has larger backscatter. The PBL height is higher during late afternoon in the summer (HPBL>1.5km) and smaller (HPBL<1.2km) during most of the winter and fall.

In addition to PBL height information, we explored the potential effect of other variables such as temperature and relative humidity in combination with total AOD as inputs to the neural network as well as the seasonal effect. Figure 2 illustrates the outcome from this experiment and as expected, due to the well-mixed PBL development in an urban environment, summer returned the highest performance (R~0.94 and RMSE~0.27) in comparison to other seasons and yearly estimations (R~0.77 for fall and spring, R~0.87 for winter and R~0.75 for year).

3.2 Regional neural network

From our previous results, we've seen the importance of including PBL information as well as the seasonal factor in our NN approach. Then, our regional experiment uses MODIS AOD and WRF PBL daily data to estimate 24-hour in-situ PM2.5 measurements which we compare against the GEOS-CHEM estimated fine PM and the CMAQ PM2.5 product. Figure 3 illustrates the results for summer comparisons.

As we can observe, the NN estimated output returns the highest correlation when compared to in-situ measurements against the other two approaches (R_{NN}~0.73, $R_{GEOS\text{-}CHEM}$~0.61, R_{CMAQ}~0.69). GEOS-CHEM estimations seem to return the largest errors which mainly come from fine PM overestimations.

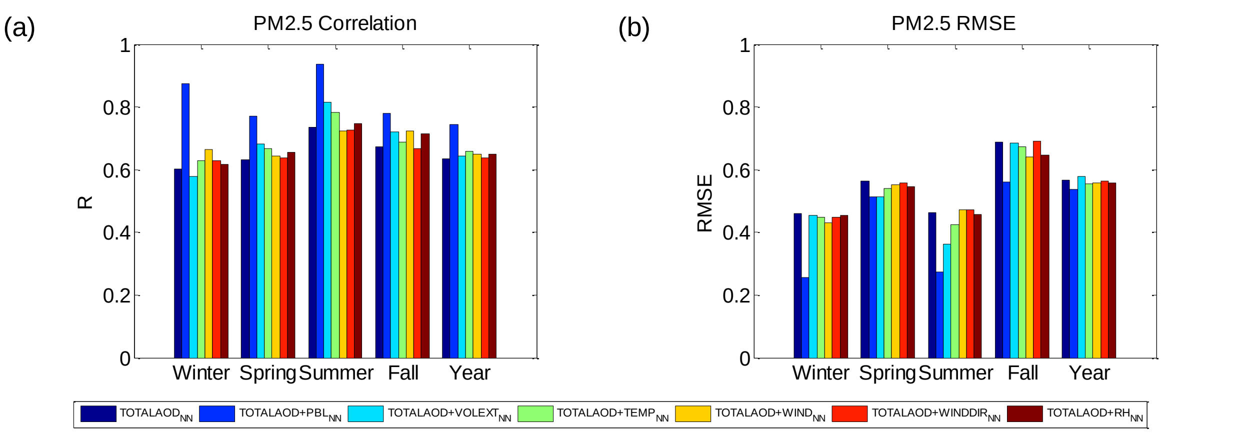

Figure 2. NN estimated PM2.5 vs. ground PM2.5 (a) seasonal correlation and (b) root-mean-square-error.

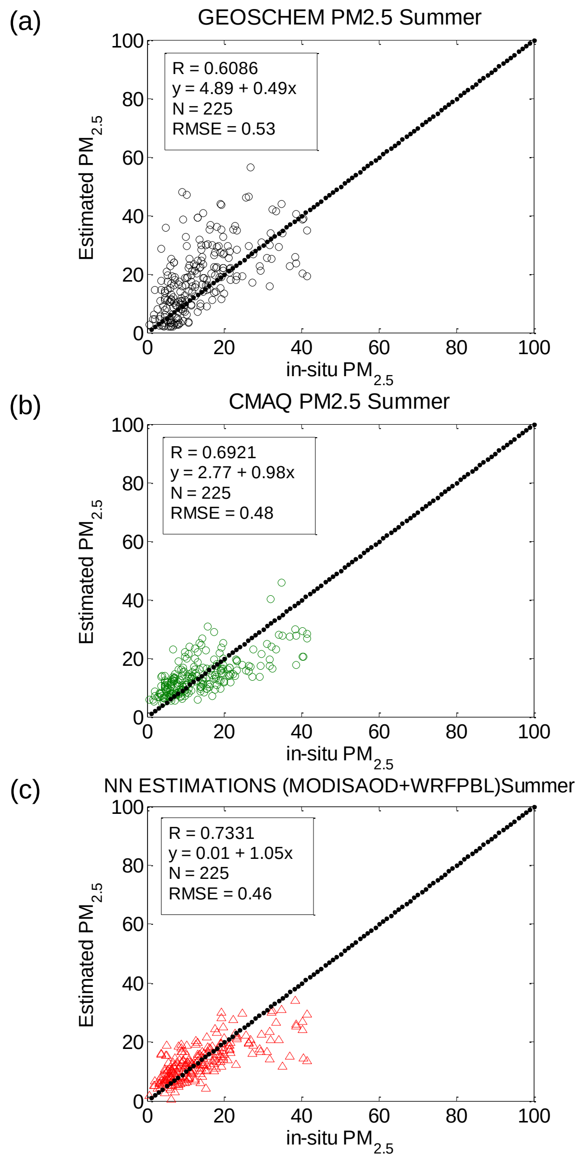

Figure 3. PM2.5 summer comparisons between ground PM2.5 and estimated PM2.5 from (a) GEOS-CHEM, (b) CMAQ and (c) neural network (AOD+PBL).

Moreover, we performed other seasons and yearly comparisons using the three methods and illustrate their correlation values in Figure 4. We see a higher performance of the neural network in fall (R~0.77) as well as the lowest correlation for GEOS-CHEM and CMAQ PM2.5 products ($R_{GEOS\text{-}CHEM}$~0.56, R_{CMAQ}~0.55). Spring comparisons using the neural network and CMAQ PM2.5 resulted in similar values (R~0.64). Additionally, because of the lack of satellite measurements during winter, there are no comparisons during this season. In general, the NN estimated fine PM performs better than GEOSCHEM and CMAQ outputs when taking into consideration PBL height and seasonality into our NN scheme.

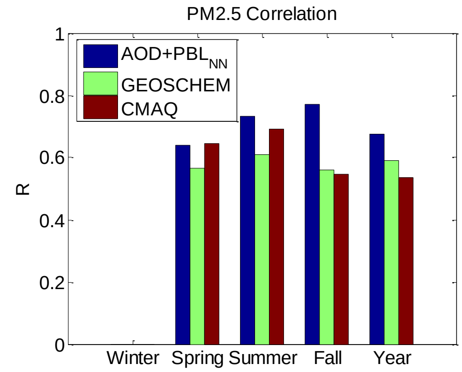

Figure 4. PM2.5 correlation comparisons using NN, GEOS-CHEM and CMAQ outputs.

3.3 Implementing NN for PM2.5 mapping

In order to estimate fine PM to conform with EPA standards, we use daily AOD retrievals together with WRF PBL measurements to illustrate our NN performance as a spatial map. However, even in a "good" case, cloud cover can significantly reduce spatial coverage. Therefore, we iteratively apply inverse distance weighted (IDW) averages of the data as defined in (1) to improve spatial coverage while using a 0.1 degree radial domain.

$$AOD_{IDW} = \frac{\sum_{i=1}^{n}(\frac{AOD}{d^2})_i}{\sum_{i=1}^{n}(\frac{1}{d^2})_i} \qquad (1)$$

Figure 5 shows MODIS AOD, WRF PBL and estimated fine PM maps for July 18, 2006 and the scatter plot between stations measurements and NN estimations. Figure 5a-b shows the improved-spatial-coverage for both AOD and PBL. Figure 5c displays the NY regional PM2.5 map obtained from our neural network together plus the station readings in that particulate date. We see a good agreement between station and estimations data with low PM2.5 values observed in the non-urban region while high fine PM values are observed in the metropolitan area.

Figure 5d shows the scatter plot between ground measurements and NN estimations at the corresponding stations available in this particulate date. In general, neural network estimations are underestimated in comparison to the site data. However, good statistics are obtained for the specific date illustrated (July 18, 2006) (R~0.90, RMSE~0.20).

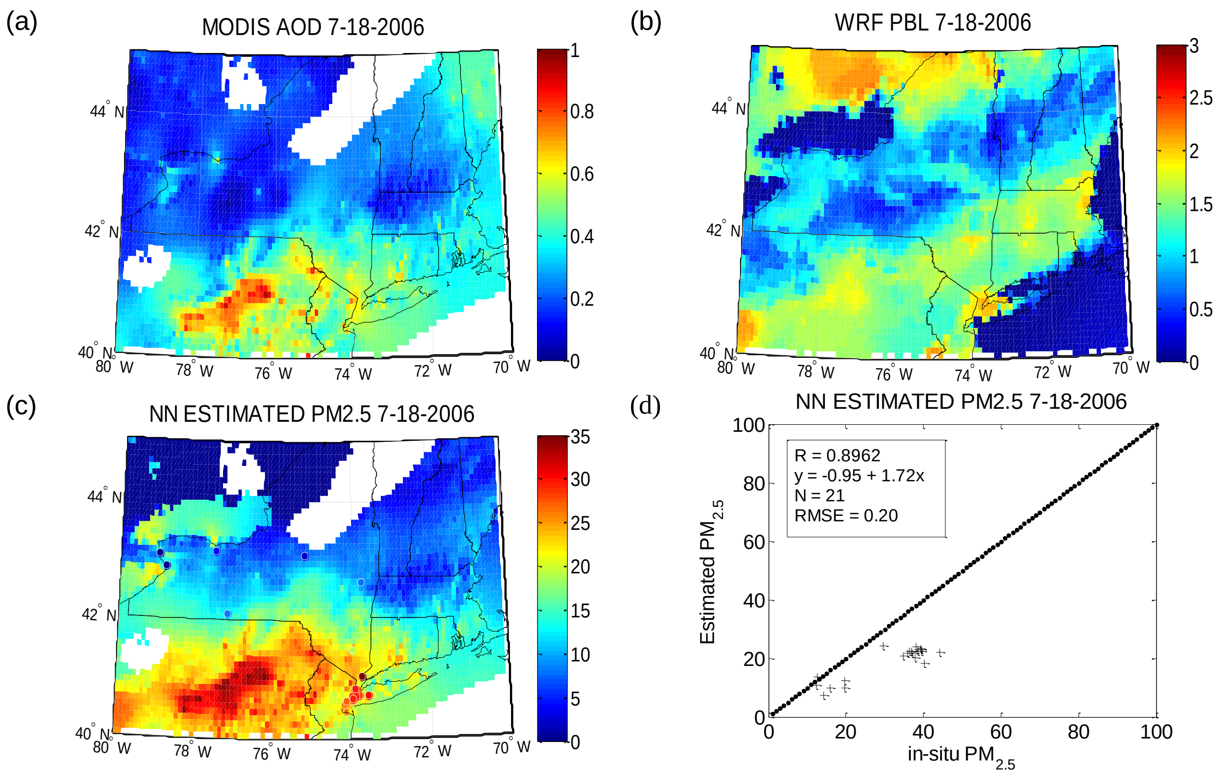

Figure 5. Improved spatial coverage based on IDW for (a) MODIS AOD daily average, (b) WRF PBL daily average, (c) PM2.5 daily average estimations based on the NN approach and (d) fine PM comparisons between ground measurements and NN estimations on July 18, 2006.

3.4 Assessment using satellite remote sensing variables

In this section we focus on satellite observations, and study their influence on PM2.5 estimations. However, before training the NN, a careful choice of variables should be made since the presence of redundant information may result in less robust solutions.

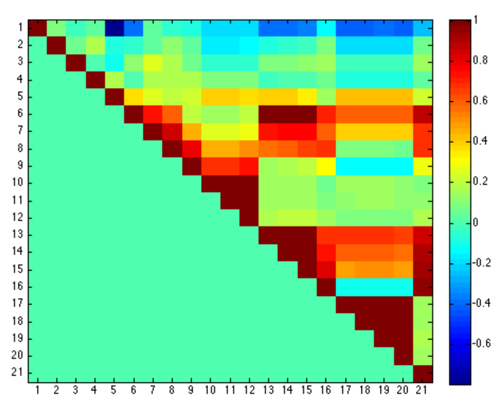

Figure 6. Correlation map between 21 satellite variables.

We reduced the number of inputs by removing the higher correlated variables (R>0.5). Figure 6 shows the correlation coefficient between the 21 candidate satellite variables. After removing the highly correlated variables (for example, "Corrected Optical Depth Land" at different wavelengths), our input set to NN included the seven input variables listed in Table 1.

Table 1 - Satellite input variables for NN.

Variable	Wavelength (μm)
Solar_Zenith	--
Solar_Azimuth	--
Sensor_Zenith	--
Sensor_Azimuth	--
Scattering_Angle	--
Optical_Depth_Land_And_Ocean	0.55
Mean_Reflectance_Land_All	2.1

Figure 7 shows the result of the multivariate NN training with the 7 satellite variables taken as the input and the ground station PM data as the target. In particular, we find that the multivariate satellite model using only the satellite remote

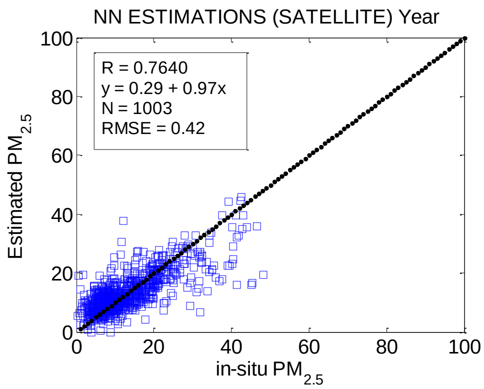

Figure 7. Regression plot of satellite multivariate NN.

sensing input variables are found to have the highest correlation with the target in comparison to the other methods. This clearly motivates further investigation of infusing surface as well as PBL height meteorological inputs.

6. CONCLUSIONS

Our main focus in this work was to demonstrate that regionally trained NN approaches are more accurate with less bias that operational approaches which do not directly input regional information. Before exploring the regional satellite based approach, we used combined active/passive radiometer and lidar measurements to assess PM2.5 estimations. First, we established that adding lidar derived PBL significantly improved the correlation between the estimated and observed PM2.5. Therefore, it is the most important factor that must be accounted for, which is particularly reasonable for urban conditions when convective mixing is expected to be magnified. In addition, we also investigated the importance of seasonality into the PM2.5-AOD relationship because PBL and other factors, which affect physical and chemical properties of aerosol, also depend on seasons. We found that separate seasonal training can provide considerable improvements too.

The success of the local experiments motivated the study of combining satellite AOD and WRF PBL height where seasonal factors were integrated into the training. These results were compared against GEOS-CHEM estimated PM2.5 and modeled PM2.5 product from CMAQ. We found that the regionally trained NN performed significantly better with much less over-bias at low PM2.5 values.

As an application, we developed daily PM2.5 maps based on the NN approach using high resolution AOD and PBL grids for the NY state region. Since the spatial coverage was still sparse, we applied an inverse distance weighting (IDW) method to both AOD and PBL products to obtain better spatial coverage. The resulting maps were in good agreement with station data.

Finally, we also explored the potential of adding geometric and land surface satellite variables as additional regressors to the neural network. The results showed better correlation and even lower biases at high PM25. Additional research is in progress to include other meteorological information as well as neighboring states ground PM2.5 information in order to improve the robustness of the NN approach.

ACKNOWLEDGMENTS

This project was made possible by the National Oceanic and Atmospheric Administration, Office of Education Educational Partnership Program award NA11SEC4810004 as well as NYSERDA under grant # 22885. Its contents are solely the responsibility of the award recipient and do not necessarily represent the official views of the U.S. Department of Commerce, National Oceanic and Atmospheric Administration (NOAA) or New York State Energy Research and Development Authority (NYSERDA).

8. REFERENCES

[1] Boyouk, N., J. F. Leon, H. Delbarre, T. Podvin and C. Deroo, 2010: Impact of the mixing boundary layer on the relationship between PM2.5 and aerosol optical thickness. *J Atmos. Environ*., **44**, 271-277.

[2] Hoff, R.M. and S. A. Christopher, 2009: Remote sensing of particulate pollution from space: have we reached the promised land? *J. Air. & Waste Manage. Assoc*., **59**, 645-675.

[3] Hogrefe. C., P. Doraiswamy, B. Colle, K. Demerjian, W. Hao, M. Beauharnois, M. Erickson, M. Souders and J.-Y. Ku, 2011: Effects of Grid Resolution and Perturbations in Meteorology and Emissions on Air Quality Simulations Over the Greater New York City Region. 10th Annual CMAS Conference, Chapel Hill, NC.

[4] Hu, X., J. W. Nielsen-Gammon and F. Zhang, 2010: Evaluation of Three Planetary Boundary Layer Schemes in the WRF Model. *J. Appl. Meteor. Climatol.*, **49**, 1831–1844.

[5] Pope, C. A., III, R. T. Burnett, M. J. Thun, E. E. Calle, D. Krewski, K. Ito, et al. 2002: Lung cancer, cardiopulmonary mortality, and long-term exposure to fine particulate air pollution.

J. of the American Medical Association, **287**, 1132−1141.

[6] T.-C. Tsai, et al. 2009: Analysis of the relationship between MODIS aerosol optical depth and particulate matter from 2006 to 2008. *Atmos. Environ*. **45**, 4777-4788.

[7] U. S. Environmental Protection Agency. Air quality criteria for particulate matter. 2004. EPA/600/P-99/002aF, Research Triangle Park, N. C.

[8] van Donkelaar, A., R. V. Martin, A. N. Pasch, J. J. Szykman, L. Zhang, Y. X. Wang and D. Chen, 2012: Improving the accuracy of daily satellite-derived ground-level fine aerosol concentration estimates for North America. *Environ. Science & Technology*, **46**, 11971-11978**.**

[9] Zhang, H., R. M. Hoff and J. A. Engel-Cox, 2009: The relation between Moderate Resolution Imaging Spectroradiometer (MODIS) aerosol optical depth and PM2.5 over the United States: a geographical comparison by EPA regions. *J. Air & Waste Manage. Assoc.* **59**, 1358-1369.