MODIS Aerosol Optical Depth Bias Adjustment Using Machine Learning Algorithms

AGU 2011 – Poster A53C-0371

Arif Albayrak[1], Jennifer Wei[2], Maksym Petrenko[3], David Lary[4], and Gregory Leptoukh[5]

1,2 - NASA GES DISC, ADNET , 3 – NASA, UMCP, 4 - University of Texas at Dallas, 5 – NASA GES DISC

Goal and Application

Goal : The goal of this work is to adjust the biases and the systematic errors from MODIS (both Terra and Aqua) aerosol product, using Machine Learning algorithms (Neural Network).

Application : The results of bias adjustment for MODIS Terra and Aqua are incorporated into the **AeroStat Giovanni as part of the NASA ACCESS funded AeroStat project** (see Poster IN51C-1604 for AeroStat Giovanni).

Background

To monitor the earth atmosphere and its surface changes, satellite based instruments collect continuous data. While some of the data is directly used, some others such as aerosol properties are indirectly retrieved from the observation data. While retrieved variables (RV) form very powerful products, they don't come without obstacles. Different satellite viewing geometries, calibration issues, dynamically changing atmospheric and earth surface conditions, together with complex interactions between observed entities and their environment affect them greatly. This results in random and systematic errors in the final products.

Methodology

Neural Networks (NN) is a mathematical modeling technique used to mimic the performance of a system. It consists of a set of elements that start out connected in a random pattern, and, based upon operational feedback, are molded into the patterns to generate the required results. (PCMagazineEncyclopedia)

- Network Architecture: Feed-Forward (no cycles)
- Optimization Algorithm: Back Propagation
- Hidden Layers: One
- Nodes: 40
- Number of Regressors
  - Land: 15
  - Ocean: 14

Selection of Regressors: Based on Expert opinions. During the regressors analysis the Aerosol Robotic Network of sun-photometers (AERONET) is used as a baseline for evaluating the MODIS aerosol products (Dark Target for land and ocean retrieval algorithms). Detailed study on the topic continues by Dr. Lary and his team.

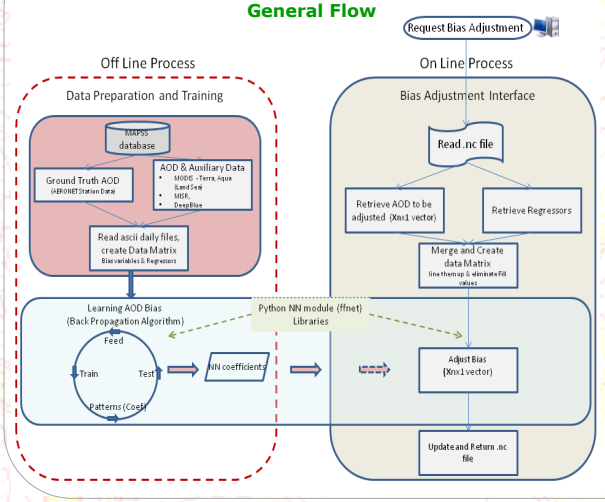
Code: Python and ffnet module by Merek Wojciechowski (Fortran + f2py)

Statistical Methods and Visualization Tools: Because of high number of data with different regimes, scatter plot with 1:1 line is not enough to draw conclusions. Results farther supported with

- Standard and Non-standard tests for analysis
- Correlation coefficient analysis
- Taylor Diagrams

Neural Network Model and Results

General Flow



Data Set and Regressors For Training

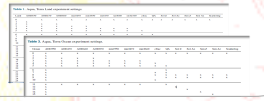
Data set for Terra-Aqua MODIS (ocean land), is prepared by using Multisensor Aerosol Products Sampling System (MAPSS) database.

Training Data Set is prepared over 10 year period from 2001 to 2011. The number of data is used for training is summarized below:

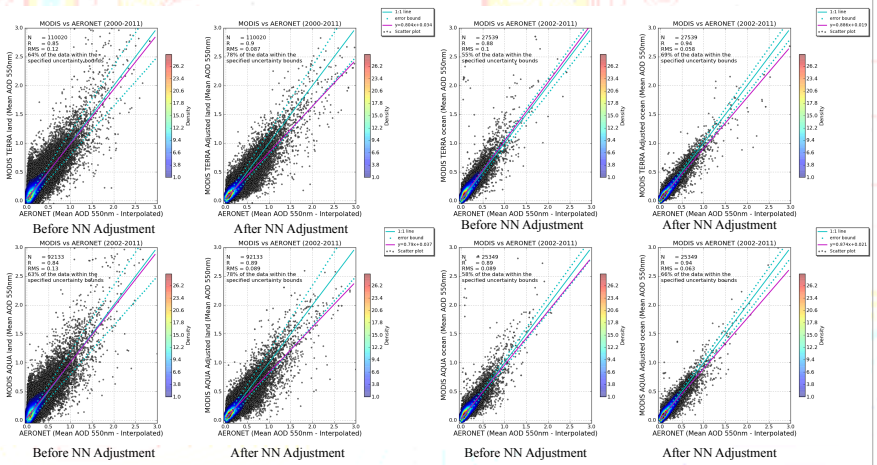
	Land	Ocean
Aqua	92133	25349
Terra	110020	27539

MODIS Common Regressors:
Target: mean AOD0550intrp (for training only)
Geometry: SolarZenith, SolarAzimuth, SensorZenith, SensorAzimuth, ScatteringAngle
Ocean: mean AOD0550, mean AOD0470, mean AOD0660, meanAOD0870, mean mref0550, meanmref0470, mean mref0660, mean cfrac, QAavg
Land: mean AOD0550, mean AOD0470, mean AOD0660, mean mref0470, mean mref0550, meansurfre0660, mean surfre0470, mean surfre2100, mean cfrac, QA

Preliminary Tests for Regressors



Bias Adjustment Results For AQUA and TERRA (Global)



Error Cones are the confidence intervals defined below:

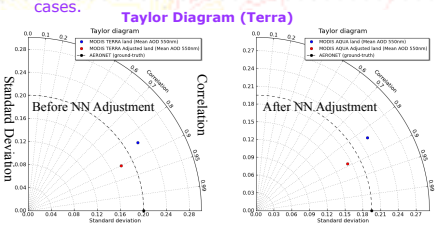
- Land = 0.15*xtime +/- 0.05
- Ocean = 0.05*txime +/- 0.03
- The more the number of data in the cone, the better the results

Linear Fit and 1:1 line:
Even though NN estimations brought high improvement in the Data Set, the linear fit line gets further from 1:1 line. This indicates multi regimes on the global data sets which will be analyzed in a future article

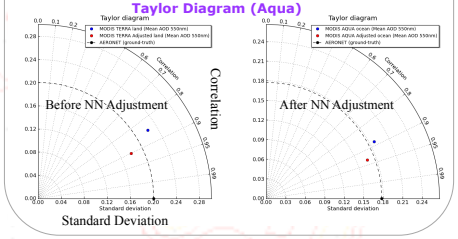
Summary

- Both correlation coefficient and std are improved after NN is applied for both, training and testing.
- Independent analysis showed that there are 2 possible regimes in data set causing the slop to move away from the 1:1 line. As a result multiple statistics need to be used for realistic evaluation.
- Best possible regressor combination currently studied by Dr. David Larry and his group.
- NN approach gives better Neural Network approximations with Terra land and ocean in all cases.

Taylor Diagram (Terra)



Taylor Diagram (Aqua)



Current and Future Work

- Currently we are extending the work to Deep-Blue and MISR. In the future we plan to:
- Consider multi regimes in the data sets: Running same NN system on the global clusters that are obtained from multi-dimensional data sets.
- Combine regressors that are obtained from MISR and MODIS.
- Perform Data fusion using unbiased data sets and compare other similar studies.

Acknowledgement

Support for the development of this on-line platform for statistical intercomparison of aerosols has been provided by NASA HQ (PM: Stephen Berrick) through the ROSES 2009 ACCESS Program (PI:Gregory Leptoukh). The AERONET data is a contribution by the International AERONET Federation (PI: Brent Holben).