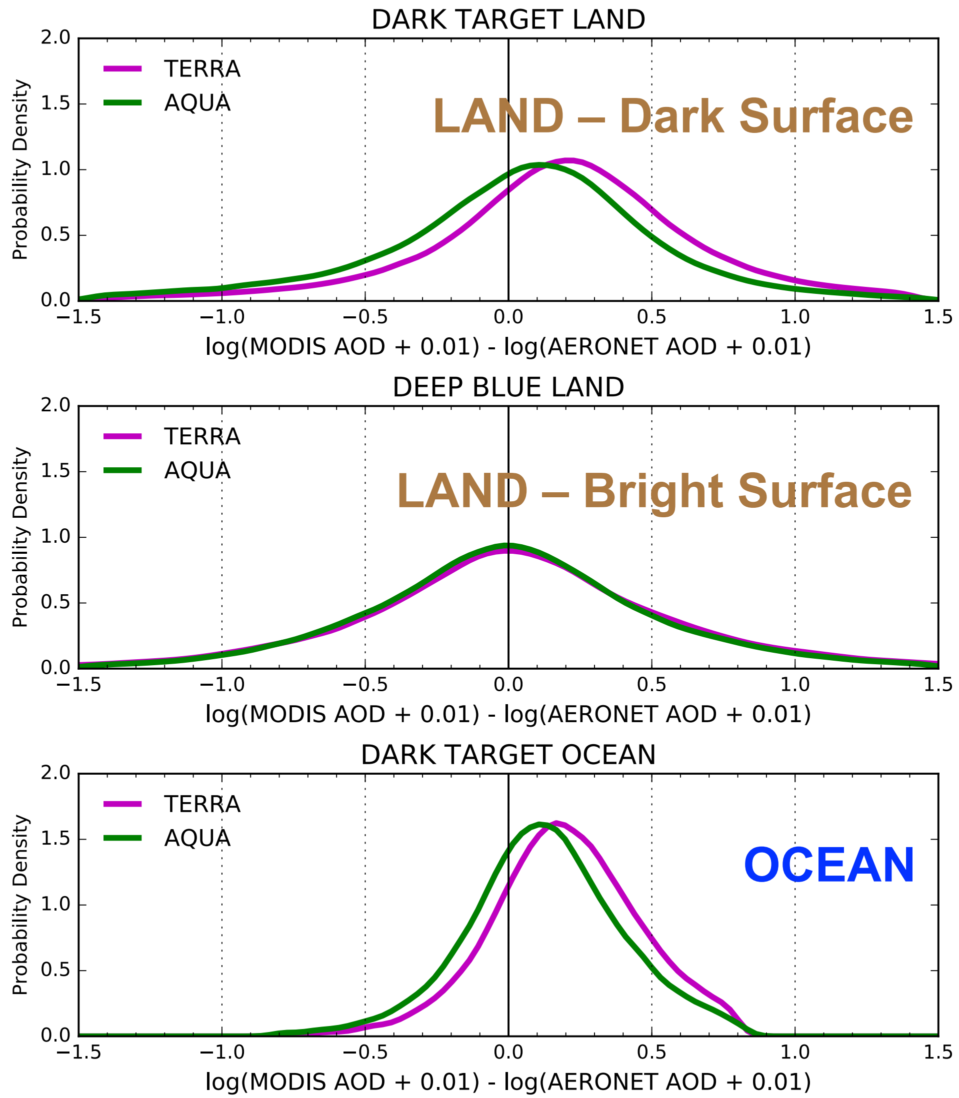
GOES-5 AOD Data Assimilation: Plans for Geostationary Observations

Patricia Castellanos & Arlindo da Silva

CEOS Atmospheric Composition Virtual Constellation AC-VC-14
Friday May 4, 2018

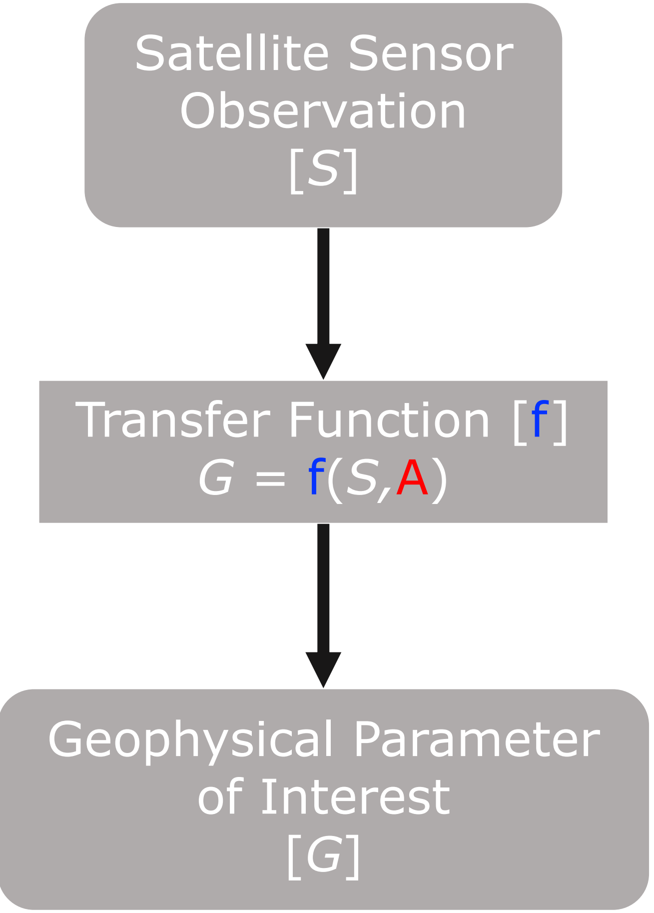
Comparison of MODIS Standard AOD Retrievals



- The aerosol data assimilation problem requires a homogenized AOD observing system across different platforms
- Biases between datasets can propagate in the model forecast and lead to artificial time variability.

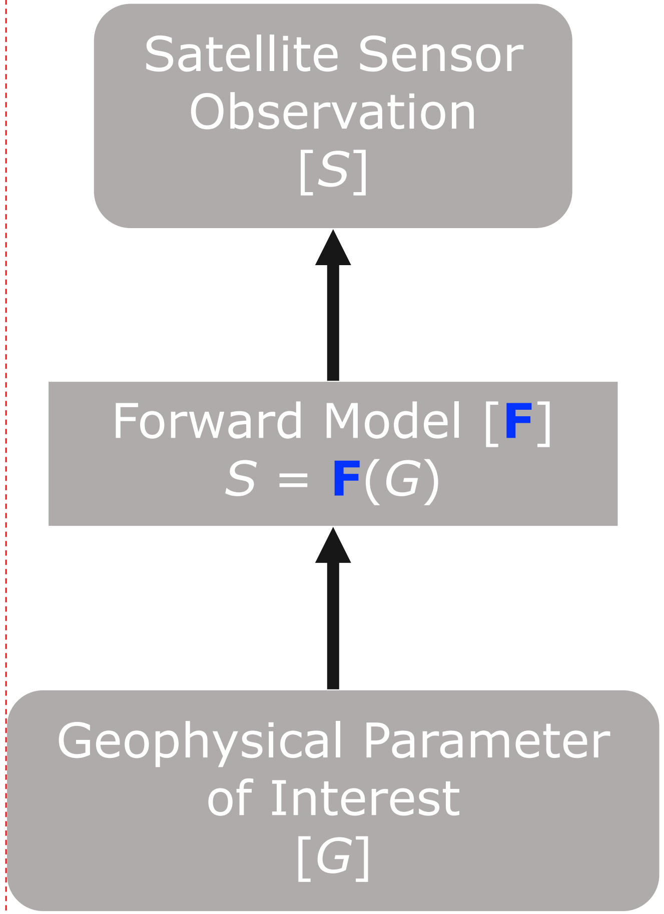
patricia.castellanos@nasa.gov

Empirical Retrievals

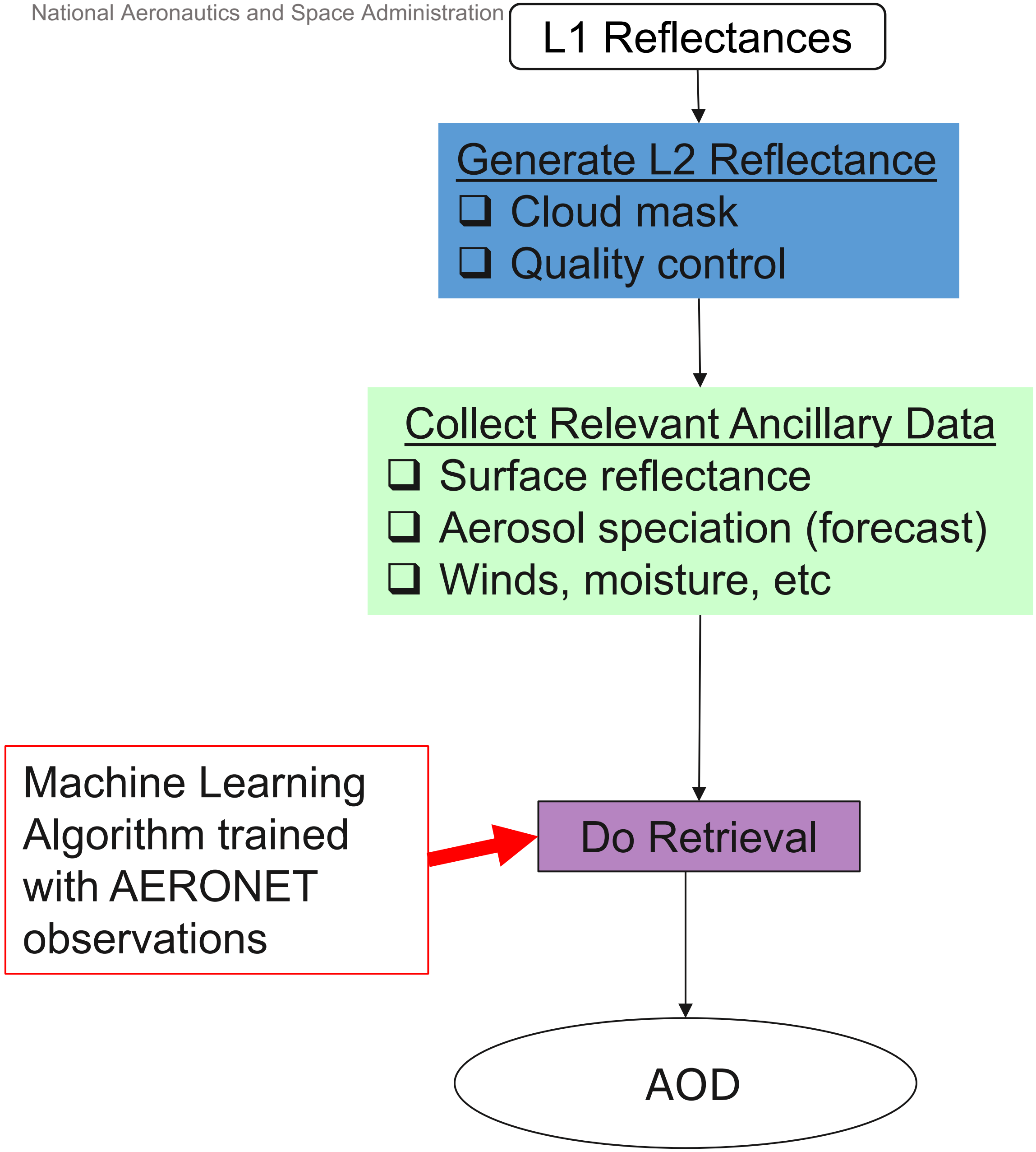


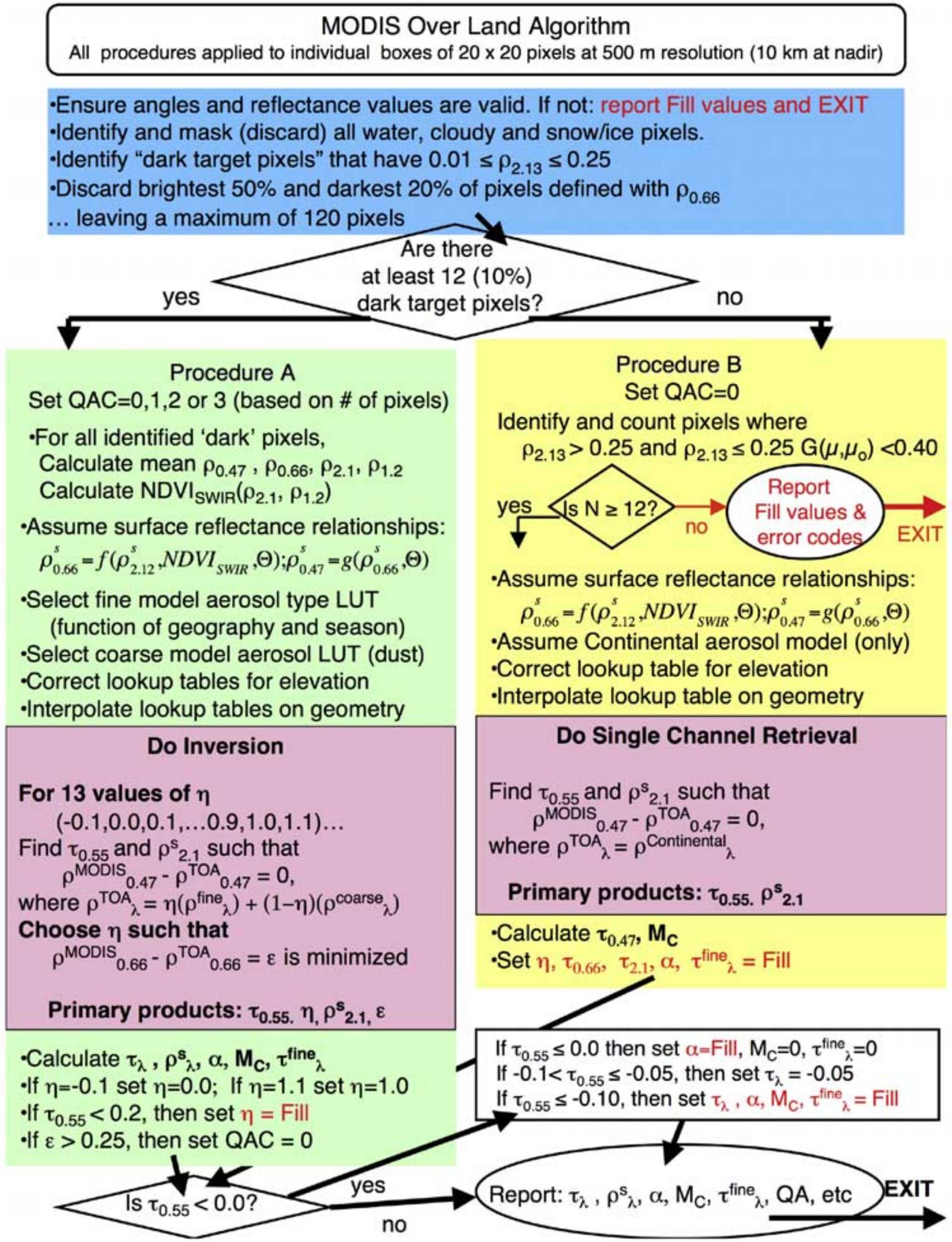
- f is a continuous function that maps S to G
- Represent f with a <u>mathematical function</u> that contains a set of empirical parameters, A
- A are determined from a training dataset of pairs of G and S <u>observations</u>.
- Training empirically captures physical relationships and tunes away calibration issues.

Physically Based Retrievals



- **F** is a <u>physical model</u> derived from first principles (e.g. radiative transfer model)
- **F** is not easily inverted
- The objective of the retrieval algorithm is to search for a G* that minimizes $||S - \mathbf{F}(G)||$
- Product quality affected by calibration issues.

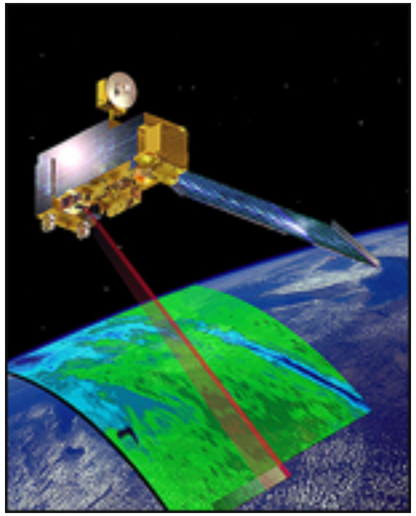

L1 Reflectances
Generate L2 Reflectance
Cloud mask
Quality control
Collect Relevant Ancillary Data
Surface reflectance
Aerosol speciation (forecast)
Winds, moisture, etc
Machine Learning Algorithm trained with AERONET observations
Do Retrieval
AOD


MODIS Over Land Algorithm
All procedures applied to individual boxes of 20 x 20 pixels at 500 m resolution (10 km at nadir)
•Ensure angles and reflectance values are valid. If not: report Fill values and EXIT
•Identify and mask (discard) all water, cloudy and snow/ice pixels.
•Identify "dark target pixels" that have 0.01 ≤ ρ2.13 ≤ 0.25
•Discard brightest 50% and darkest 20% of pixels defined with ρ0.66
… leaving a maximum of 120 pixels
Are there at least 12 (10%) dark target pixels?
yes
no
Procedure A
Set QAC=0,1,2 or 3 (based on # of pixels)
•For all identified 'dark' pixels,
Calculate mean ρ0.47, ρ0.66, ρ2.1, ρ1.2
Calculate NDVISWIR(ρ2.1, ρ1.2)
•Assume surface reflectance relationships:
•Select fine model aerosol type LUT (function of geography and season)
•Select coarse model aerosol LUT (dust)
•Correct lookup tables for elevation
•Interpolate lookup tables on geometry
Procedure B
Set QAC=0
Identify and count pixels where
ρ2.13 > 0.25 and ρ2.13 ≤ 0.25 G(μ,μo) <0.40
yes
Is N ≥ 12?
no
Report Fill values & error codes
EXIT
•Assume surface reflectance relationships:
•Assume Continental aerosol model (only)
•Correct lookup table for elevation
•Interpolate lookup table on geometry
Do Inversion
For 13 values of η
(-0.1,0.0,0.1,…0.9,1.0,1.1)…
Find τ0.55 and ρs2.1 such that
ρMODIS0.47 - ρTOA0.47 = 0,
where ρTOAλ = η(ρfineλ) + (1−η)(ρcoarseλ)
Choose η such that
ρMODIS0.66 - ρTOA0.66 = ε is minimized
Primary products: τ0.55, η, ρs2.1, ε
Do Single Channel Retrieval
Find τ0.55 and ρs2.1 such that
ρMODIS0.47 - ρTOA0.47 = 0,
where ρTOAλ = ρContinentalλ
Primary products: τ0.55, ρs2.1
•Calculate τ0.47, MC
•Set η, τ0.66, τ2.1, α, τfineλ = Fill
•Calculate τλ, ρsλ, α, MC, τfineλ
•If η=-0.1 set η=0.0; If η=1.1 set η=1.0
•If τ0.55 < 0.2, then set η = Fill
•If ε > 0.25, then set QAC = 0
If τ0.55 ≤ 0.0 then set α=Fill, MC=0, τfineλ=0
If -0.1< τ0.55 ≤ -0.05, then set τλ = -0.05
If τ0.55 ≤ -0.10, then set τλ, α, MC, τfineλ = Fill
Is τ0.55 < 0.0?
yes
no
Report: τλ, ρsλ, α, MC, τfineλ, QA, etc
EXIT

Observations

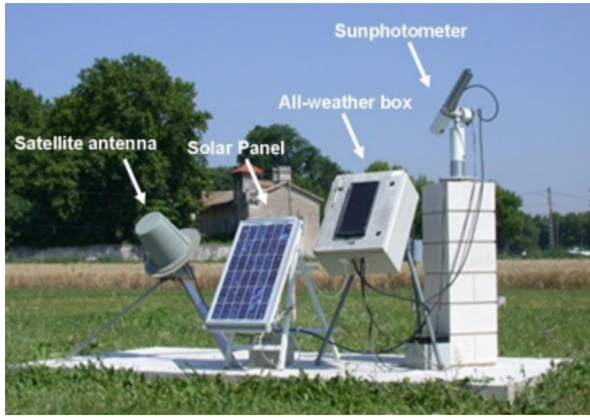
Satellite Sensor Observation [*S*]:
MODIS MOD04 /MYD04 Level 2 Reflectance

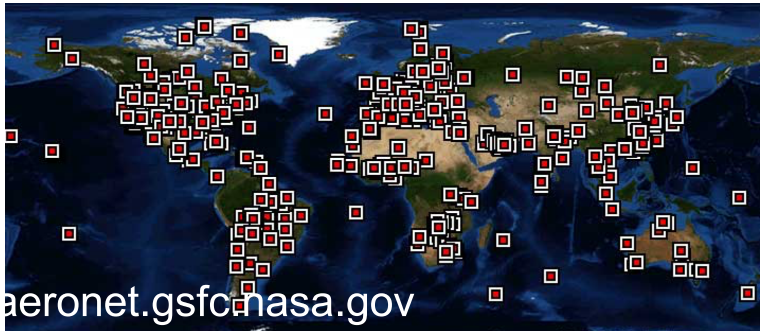
- Cloud masked, quality controlled, 10 km data
- Deep Blue Land
  - 3 channels over bright surfaces
  - 412 nm, 470 nm, and 670 nm
- Dark Target Land
  - 9 channels over dark surfaces
  - 412-2100 nm
- Dark Target Ocean
  - 7 channels over ocean
  - 470-2100 nm



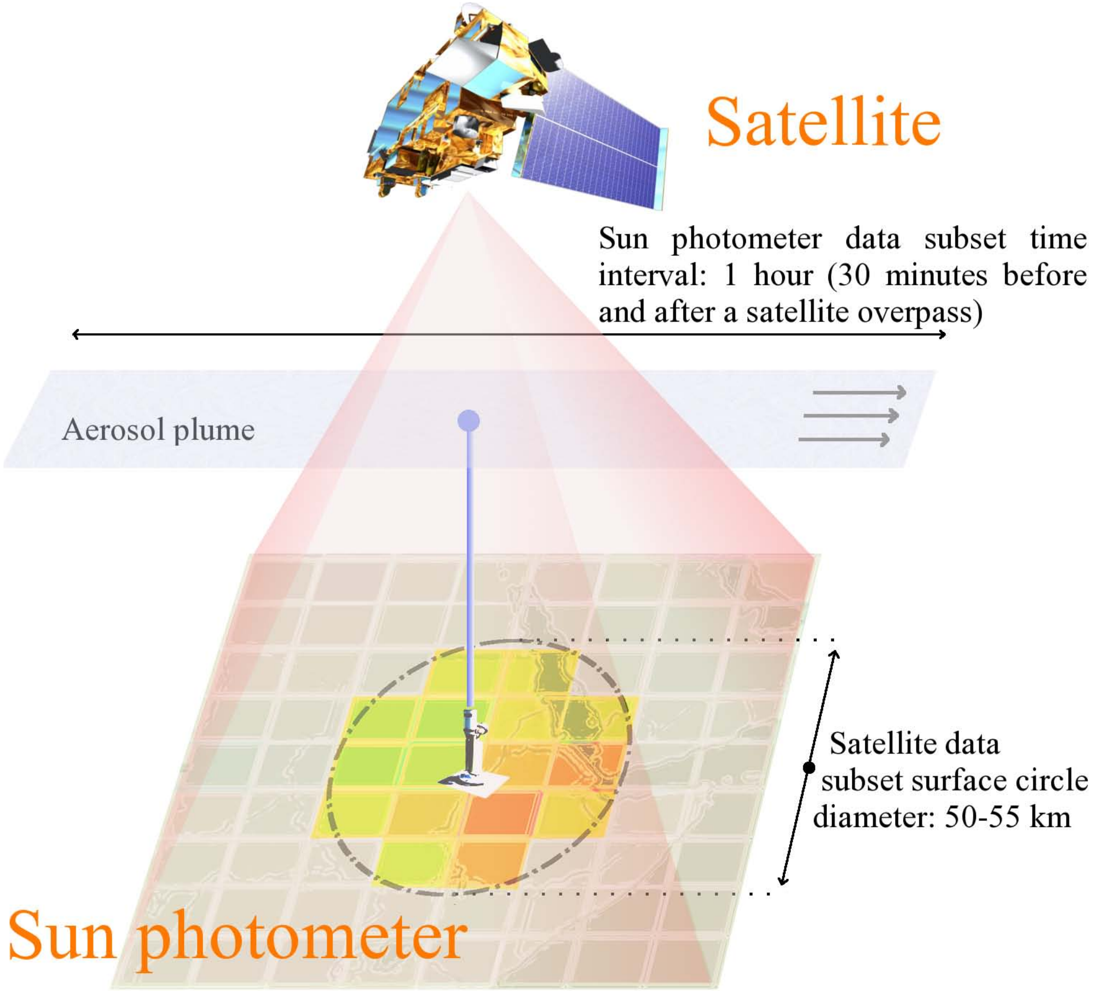
Geophysical Parameter of Interest [*G*]:
470, 550, 670, 870 nm AOD

- Aerosol Robotic Network (AERONET) observations of AOD
  - Global network of sunphotometers
  - 15 minute sampling
  - Low uncertainty (±0.01)





MODIS-AERONET Data Pairs



Petrenko, et al. AMT (2012)

- 15 years of data (2000-2015)
- ~150K Deep Blue data pairs
- ~100K DT-Land data pairs
- ~40K DT-Ocean data pairs

Additional Data Screening

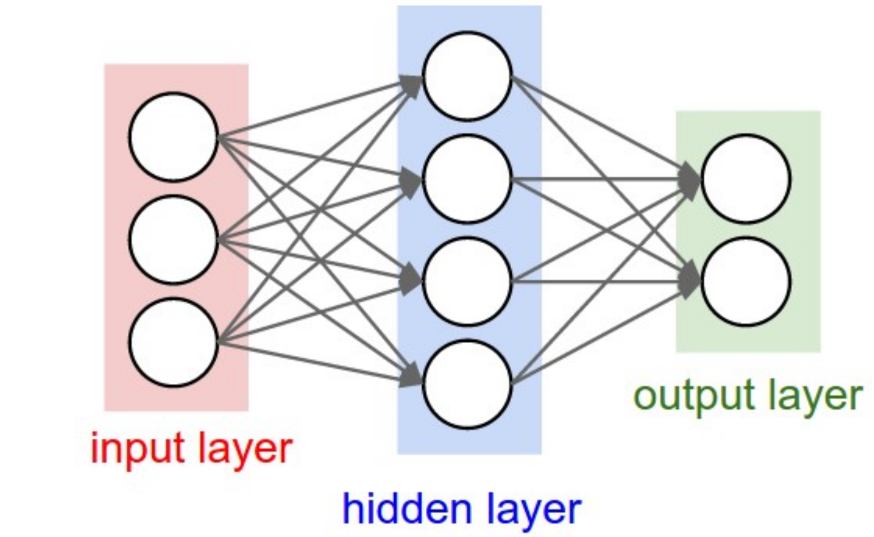
- Outlier removal
- Cloud Fraction < 0.7
- Used MERRA-2 to "balance" the dataset by aerosol type
  - Dust
  - Smoke (Black Carbon + Organic Carbon)
  - Sea Salt
  - Sulfate
- Reduces number of data pairs by half

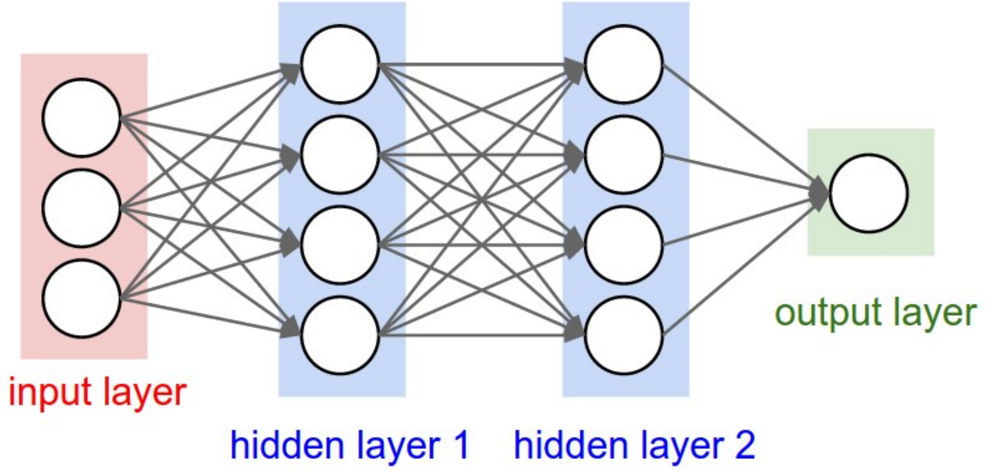
patricia.castellanos@nasa.gov

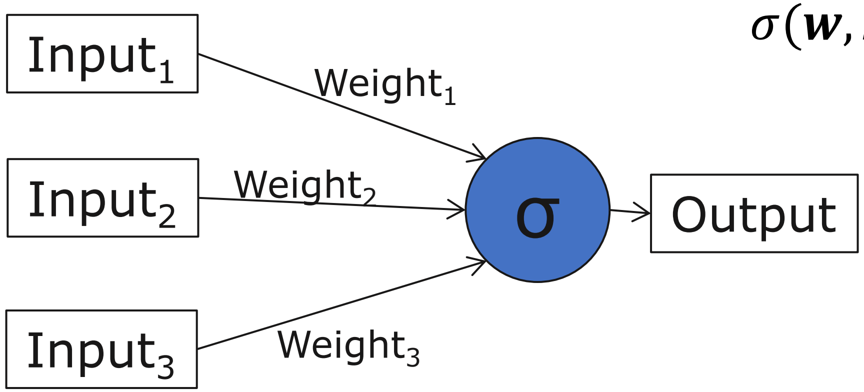
Mathematical Function: Neural Network

- A set of nodes connected by numerical weights.
- The weights are tuned based on a specific training dataset containing input-output data pairs.
- The superposition of many nonlinear transfer functions (nodes) allow NN's to approximate extremely nonlinear functions

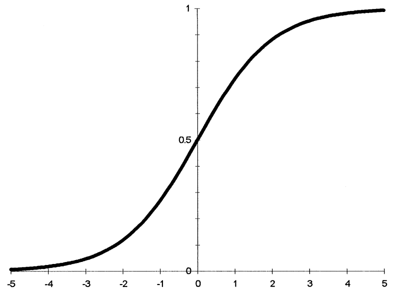
Feed-forward Neural Network



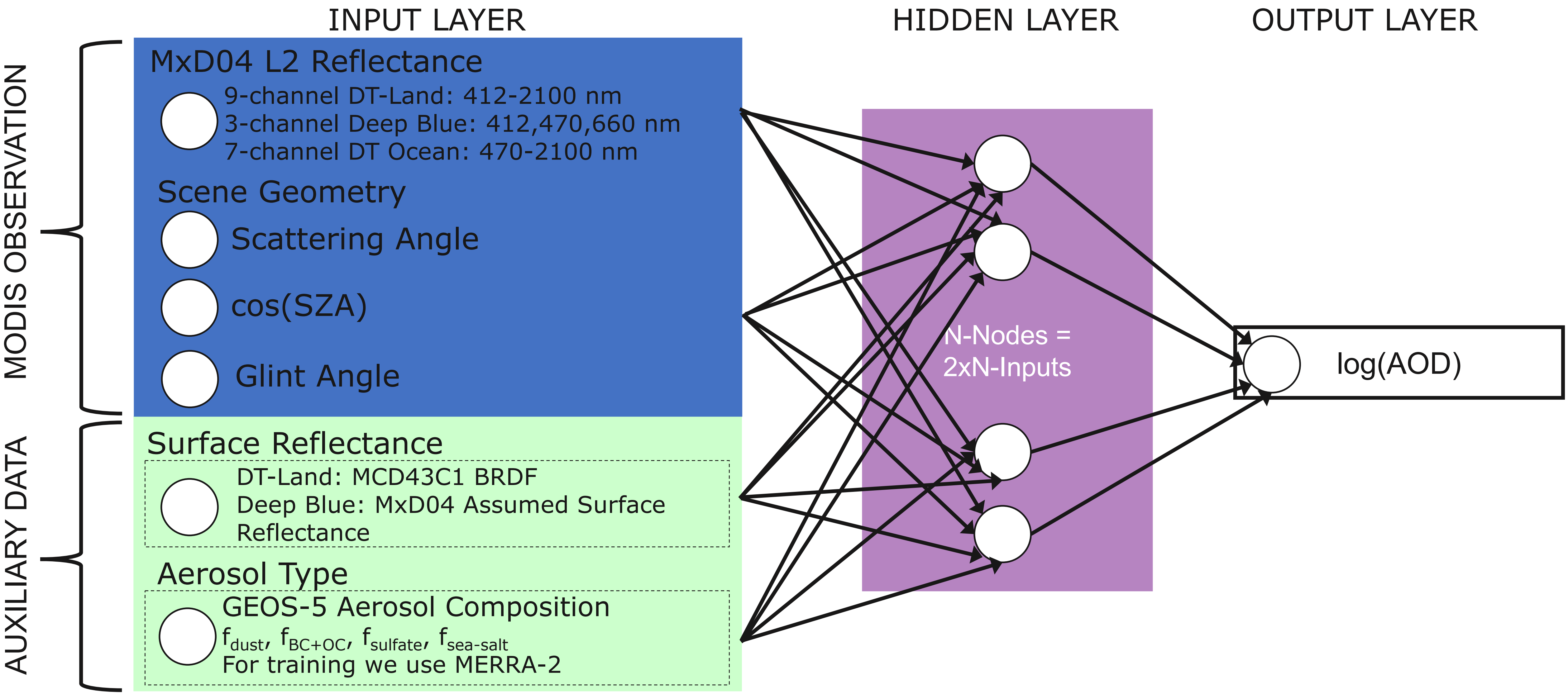




$$\sigma(\boldsymbol{w}, \boldsymbol{x}) = \frac{1}{1 + \exp\left(-\boldsymbol{w} \cdot \boldsymbol{x} - c\right)}$$

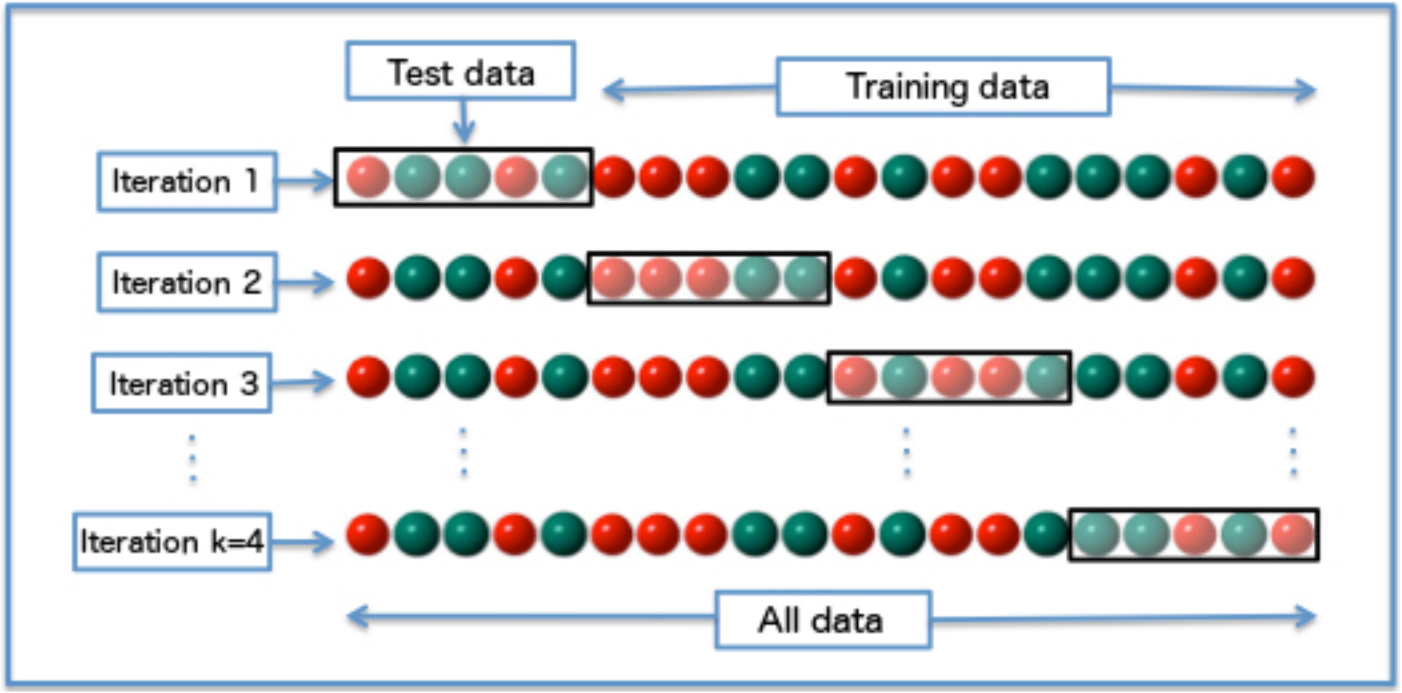


GEOS-5 NNR for AOD

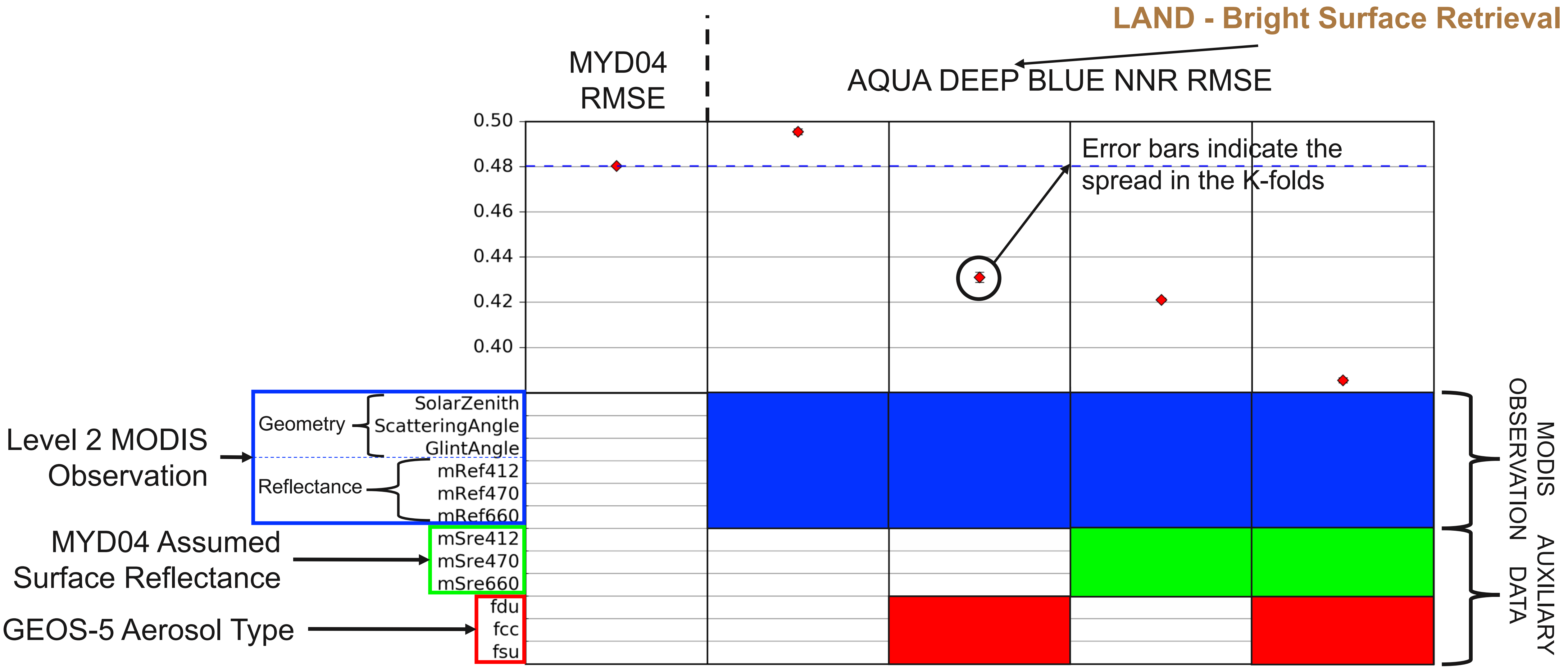


NNR Training, Testing, Validation

- Train & Test
  - Iteratively train and test adjusting input variables, architecture, etc. to optimize neural network
  - Cross Validation
    - K-folding: create K subsets of data, using K-1 for training, and 1 for testing. Iterate K times.
- Validate
  - Use a separate dataset, not used to train/test
  - Observations after 2015



NNR Training & Testing

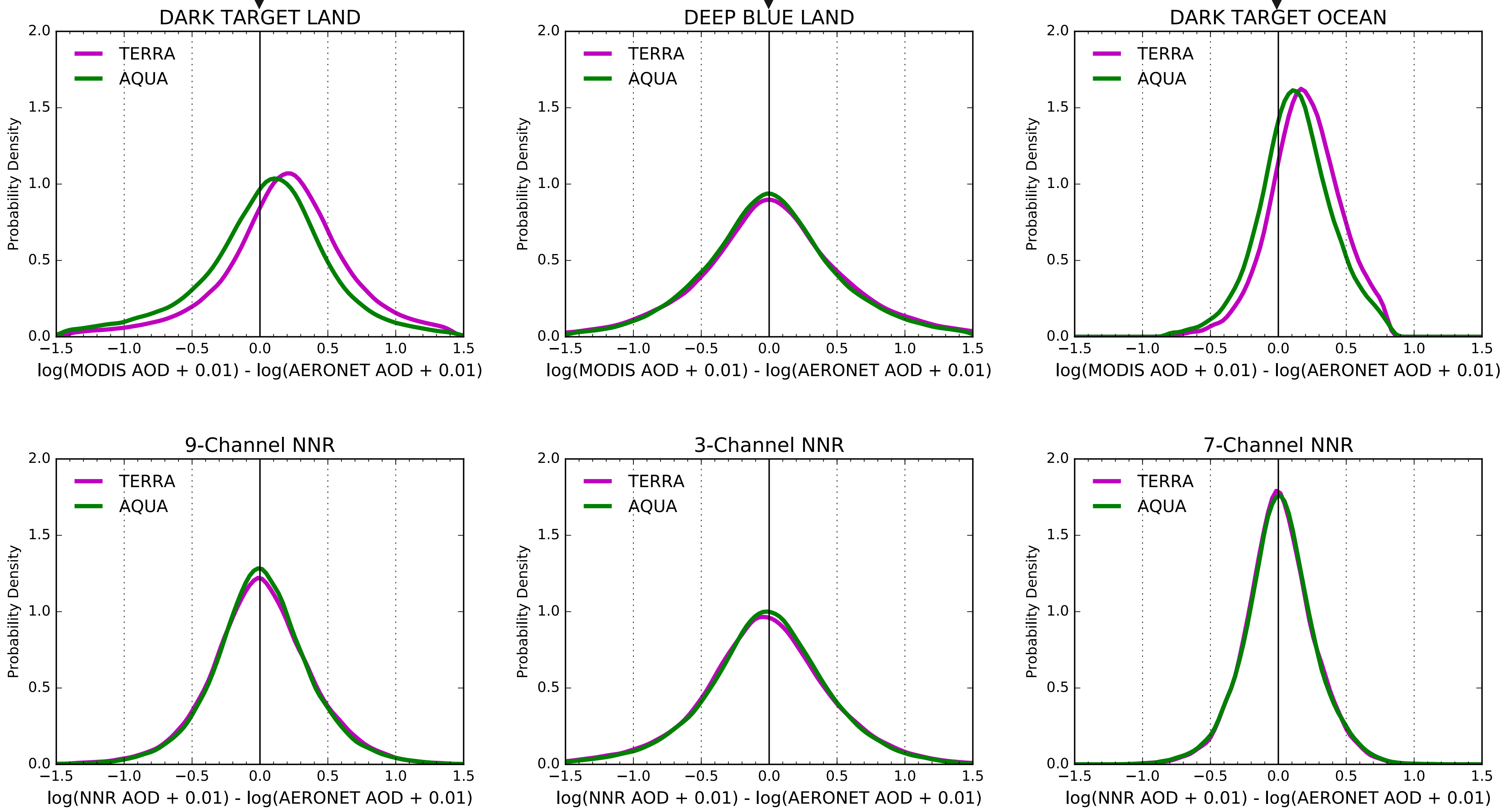


NNR Testing

LAND - Dark Surface | LAND - Bright Surface | OCEAN

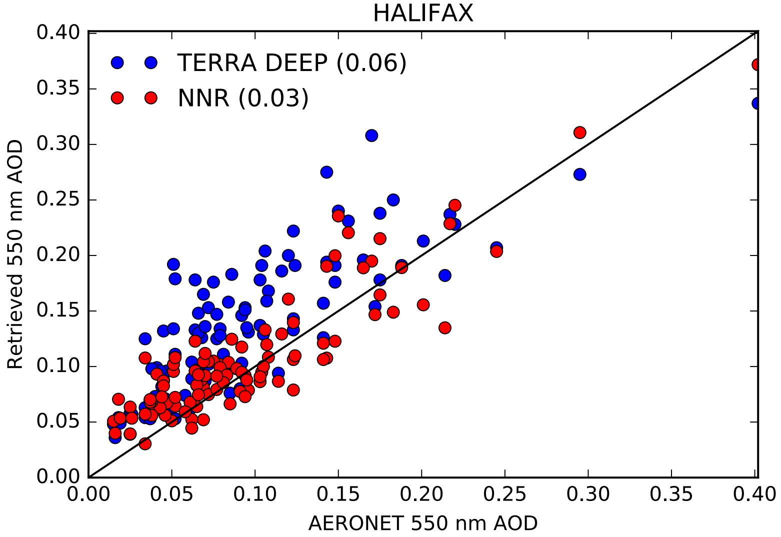
MODIS Standard Retrievals

Neural Network Retrievals

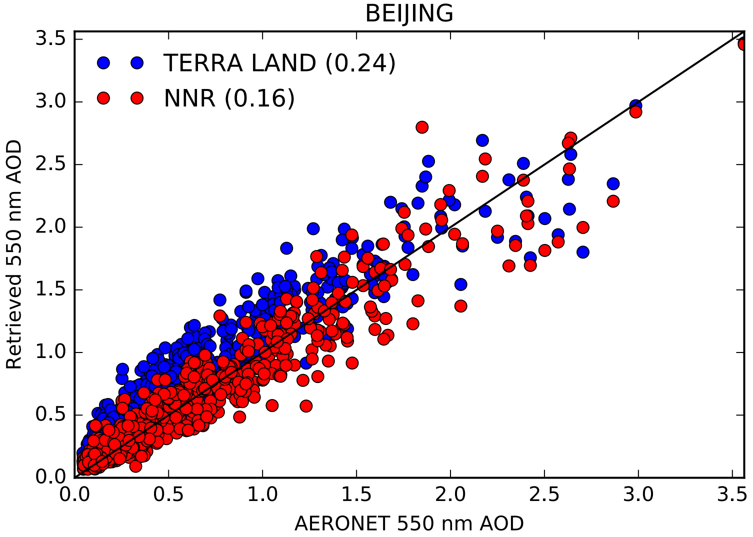


NNR Testing at Some Individual Sites

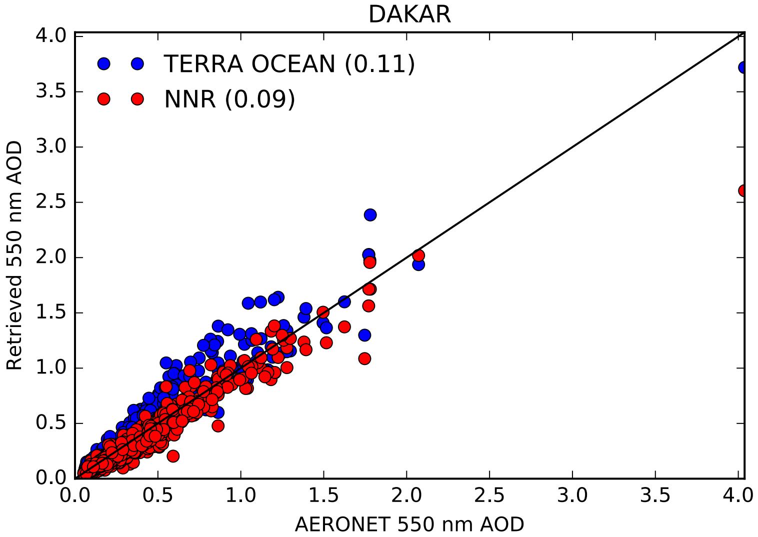
LAND - Bright Surface



LAND – Dark Surface



OCEAN

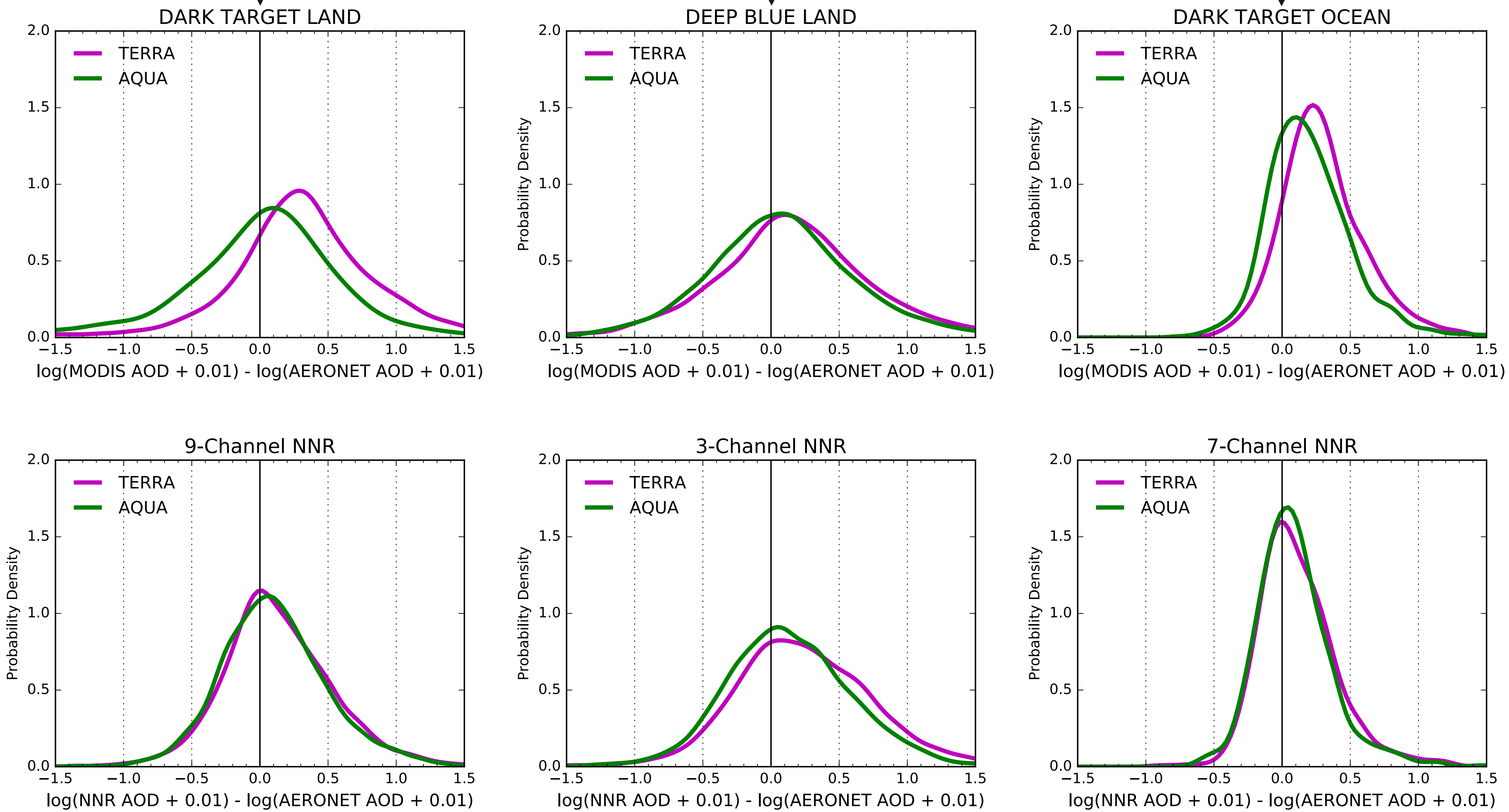


NNR Validation (Observations after 2015)

LAND - Dark Surface | **LAND - Bright Surface** | **OCEAN**

MODIS Standard Retrievals

Neural Network Retrievals



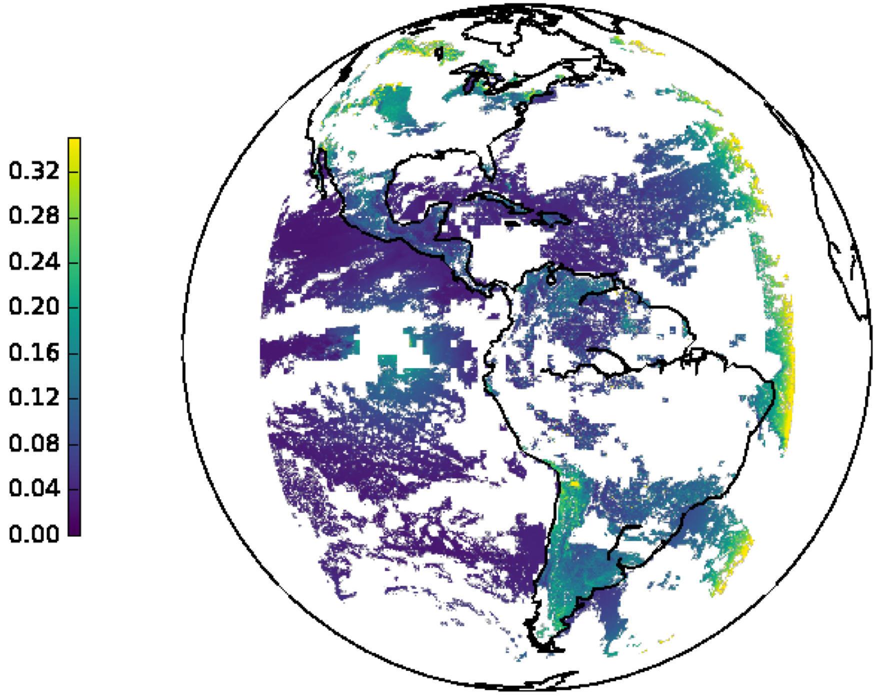
patricia.castellanos@nasa.gov

From LEO to GEO: Calibration Transfer

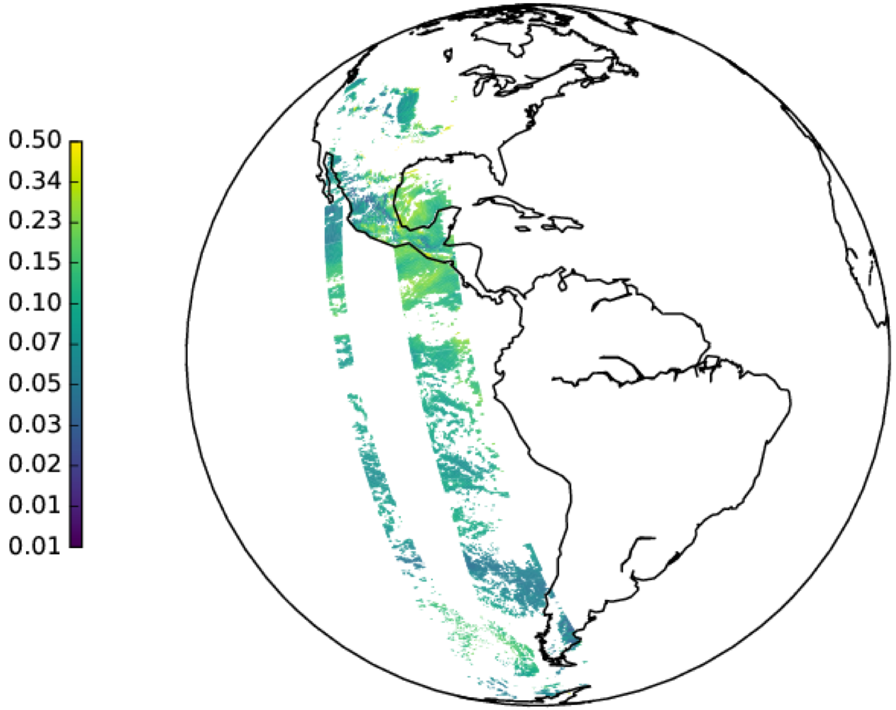
- NRT SatCORPS cloud cleared GOES-16 and AHI-8 TOA reflectances
  - ➢ Provided by NASA Langley SatCORPS Group (R. Palikonda, W. Smith)
  - ➢ 1 full disk scan per hour
- Use MODIS-NNR AOD as targets for training GEO Reflectance observations

INPUTS → TARGETS

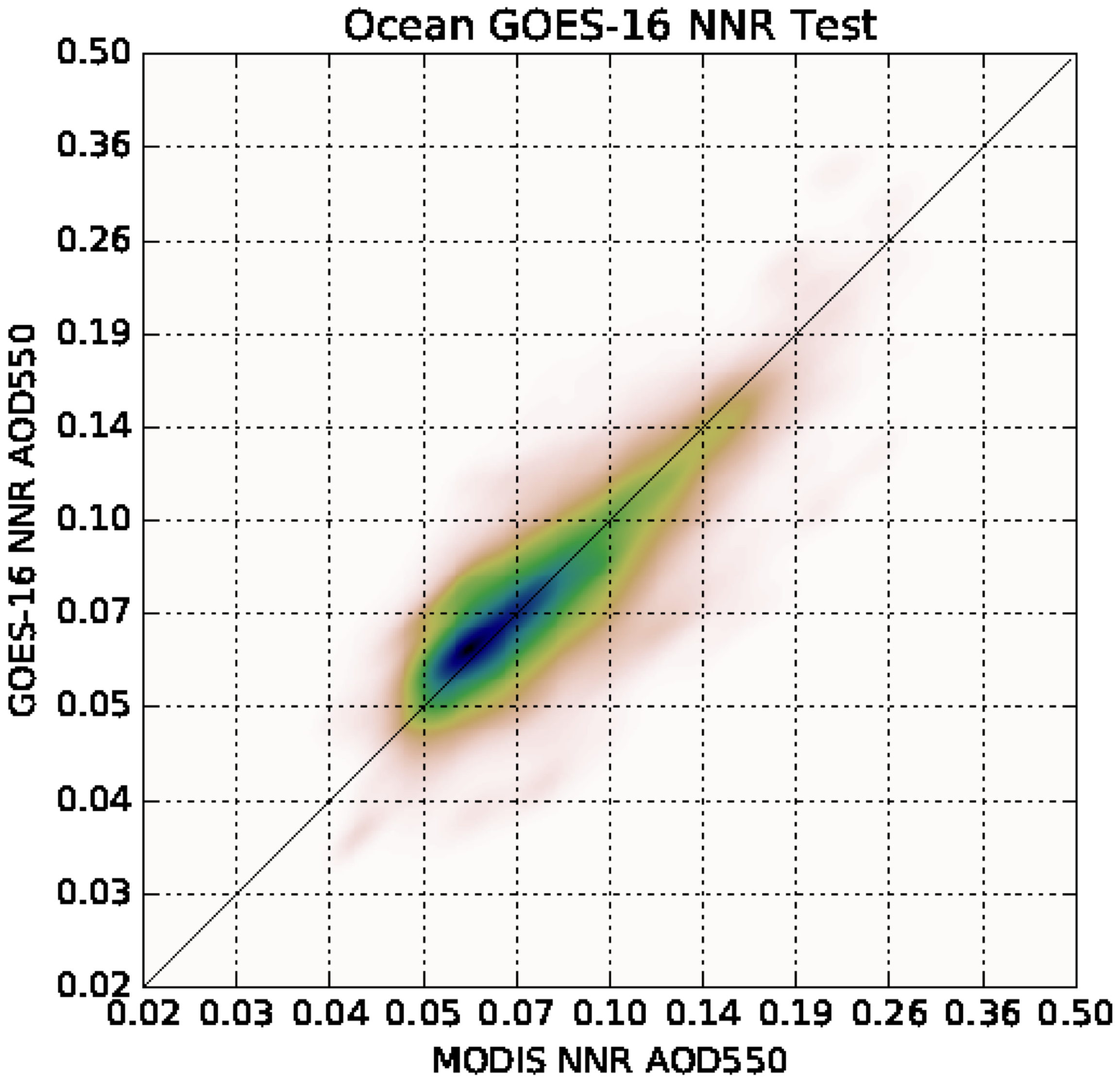
GOES-16 TOA 640 Reflectance 2018-03-11 20Z



MODIS Aqua NNR AOD 2018-03-11 20Z



patricia.castellanos@nasa.gov



Preliminary Test with Ocean Data

- Trained NNR with 2-months of MODIS/GOES-16 collocated observations
- ~100K data points
- Currently only 640, 860, and 1600 nm channels provided
- No water vapor correction

patricia.castellanos@nasa.gov

Current Aerosol Analysis: Splitting

2D AOD ANALYSIS

- Observable 550 nm AOD is 2D
  - MODIS Aqua/Terra: Ocean, Land, Deep Blue channels
  - Constrains column averaged optics
  - Cannot constrain speciation or vertical distribution
- Analysis in observation space:

$$\tau^a \equiv Hq^a = H\left(q^b + \delta q^a\right)$$
$$= \tau^b + \delta\tau^a$$

GOING TO 3D CONCENTRATIONS

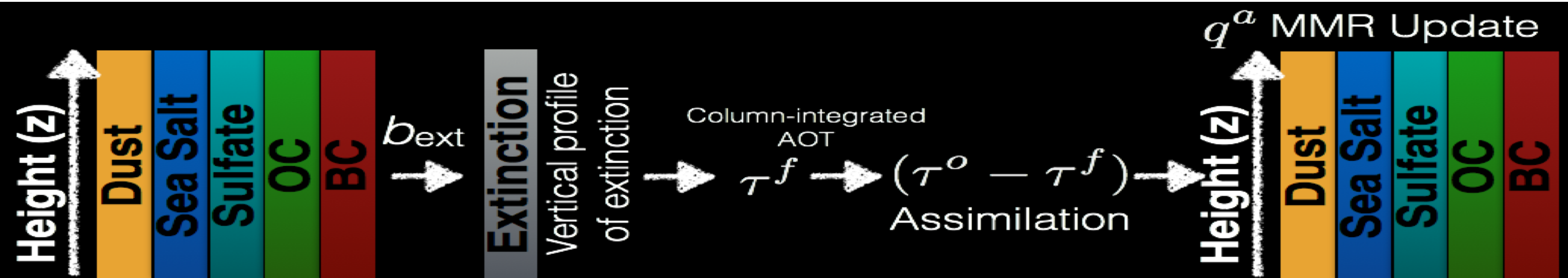
- Based on error covariances:

$$\delta q^a = BH^T\left(HBH^T\right)^{-1}\delta\tau^a$$

- Using ensemble perturbations,

$$\delta q^a = XY^T\left(YY^T\right)^{-1}\delta\tau^a$$

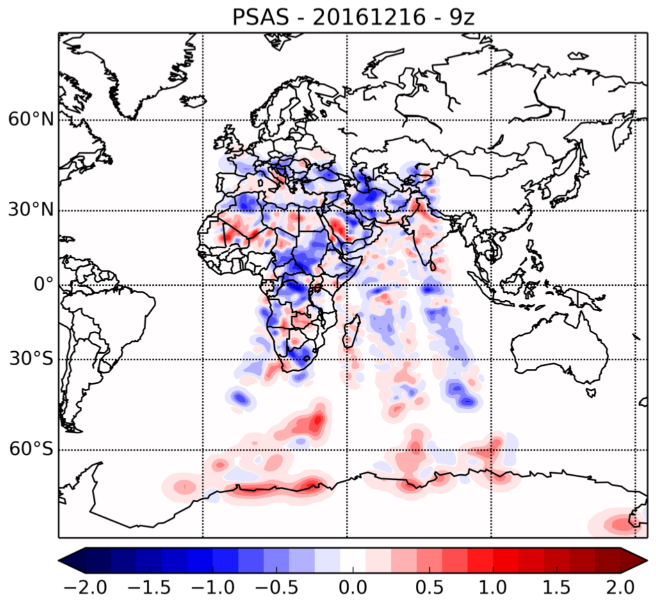
- NRT GEOS-5 uses Local Displacement Ensembles (LDE), in 1D
- Developing EnKF for Aerosols



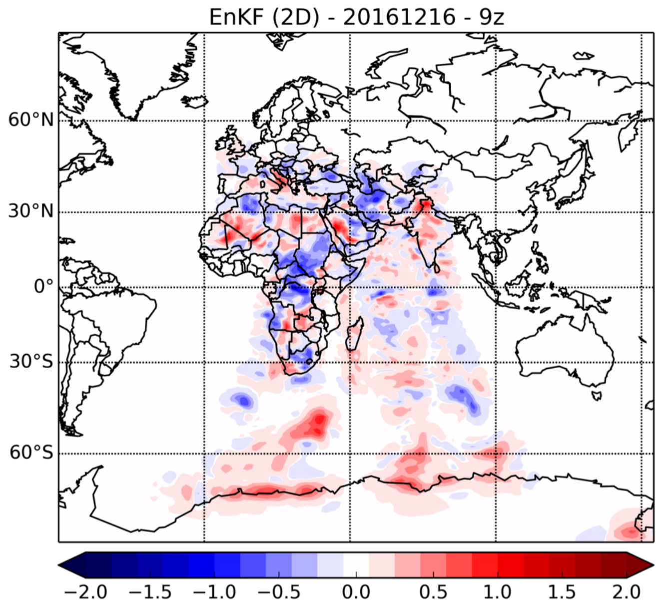
In Development: Aerosol EnKF

- **As part of GMAO's hybrid system, aerosol ensemble members are produced as a matter of routine**
- **The same Whitaker-Hamill EnKF used the hybrid Meteorological assimilation has been adapted for aerosols**
- **Target observation systems**
  - Multi-spectral AOD: 470, 550 and 870 nm
  - Lidar attenuated backscatter
  - Sensors: MODIS, VIIRS, **GEO**, CATS/CALIOP, TropOMI

PSAS AOD inc (current method)



EnKF AOD inc Ensmean



Summary & Outlook

- The NNR provides a way to homogenize the AOD observing system for data assimilation
- A NNR for GOES-16 and AHI-8 will be developed using the MODIS-NNR AOD data as targets
  - Demonstrates the need for continuity in the observing system
- Innovate the training methodology
  - Multi-channel AOD training allows for physical constraints on cost function
  - Use p.d.f. type of targets – spread in the solution indicates the uncertainty in the NNR AOD.