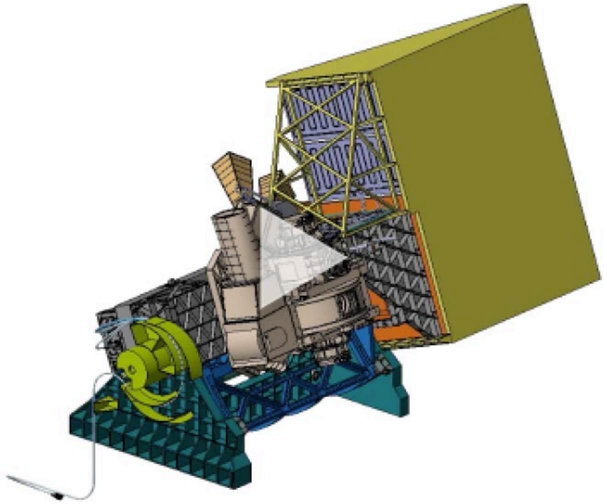
Can we retrieve NO_2 accurately with lower spectral resolution hyperspectral *imagers* such as those designed for ocean color?

Joanna Joiner, Sergey Marchenko, Zachary Fasnacht, Lok Lamsal, Can Li, Alexander Vasilkov, and Nickolay Krotkov

Joiner, J., Marchenko, S., Fasnacht, Z., Lamsal, L., Li, C., Vasilkov, A., and Krotkov, N.: Use of machine learning and principal component analysis to retrieve nitrogen dioxide (NO_2) with hyperspectral imagers and reduce noise in spectral fitting, Atmos. Meas. Tech., 16, 481–500, https://doi.org/10.5194/amt-16-481-2023, 2023

Coming soon: NASA Plankton, Aerosol, Cloud, and ocean Ecosystem (PACE) Ocean Color Instrument (OCI) and Geosynchronous Littoral Imaging and Monitoring Radiometer (GLIMR)

- OCI is a hyperspectral imaging radiometer whose continuous coverage extends from 340 to 890 nm in the ultraviolet (UV) to near infrared (NIR) spectrum at 5 nm resolution with spectral steps of 2.5 nm.
  - It acquires Earth view data at a 1 km x 1 km ground sample footprint at nadir and an angular range of ±56.5° for a ground swath width of 2663 km.
  - Launch into low Earth (polar) orbit in early 2024.
- GLIMR is a geostationary hyperspectral imager (NASA Earth Ventures) for coastal ocean observations with variable spectral resolution and 300 m spatial resolution.



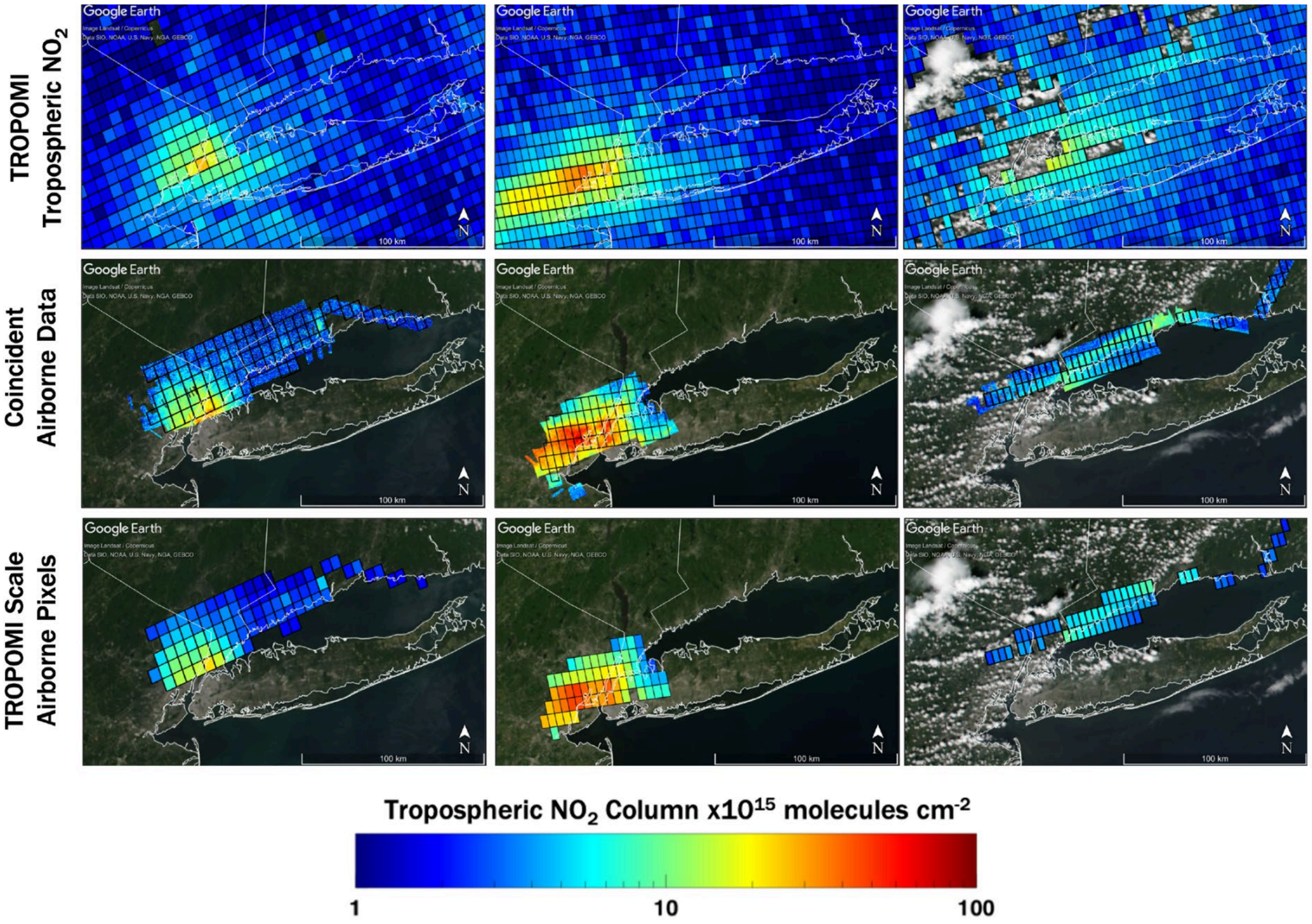
Motivation

- NO_2 is a pollutant and climate agent (contributes to the formation of aerosols and tropospheric ozone)
- NO_2 impacts ocean color algorithms. Errors under high loading (1×10^{16} mol cm^{-2}) of 10-20% in water leaving radiance (Ahmad et al., 2007).

- Can we retrieve NO_2 accurately with new GEO and LEO hyper-spectral sensors?
- If so, these sensors could provide value added data for AQ and atmospheric chemistry communities (for downscaling emissions, etc). If not, need to prepare to collocate TROPOMI or TEMPO data for ocean color products.

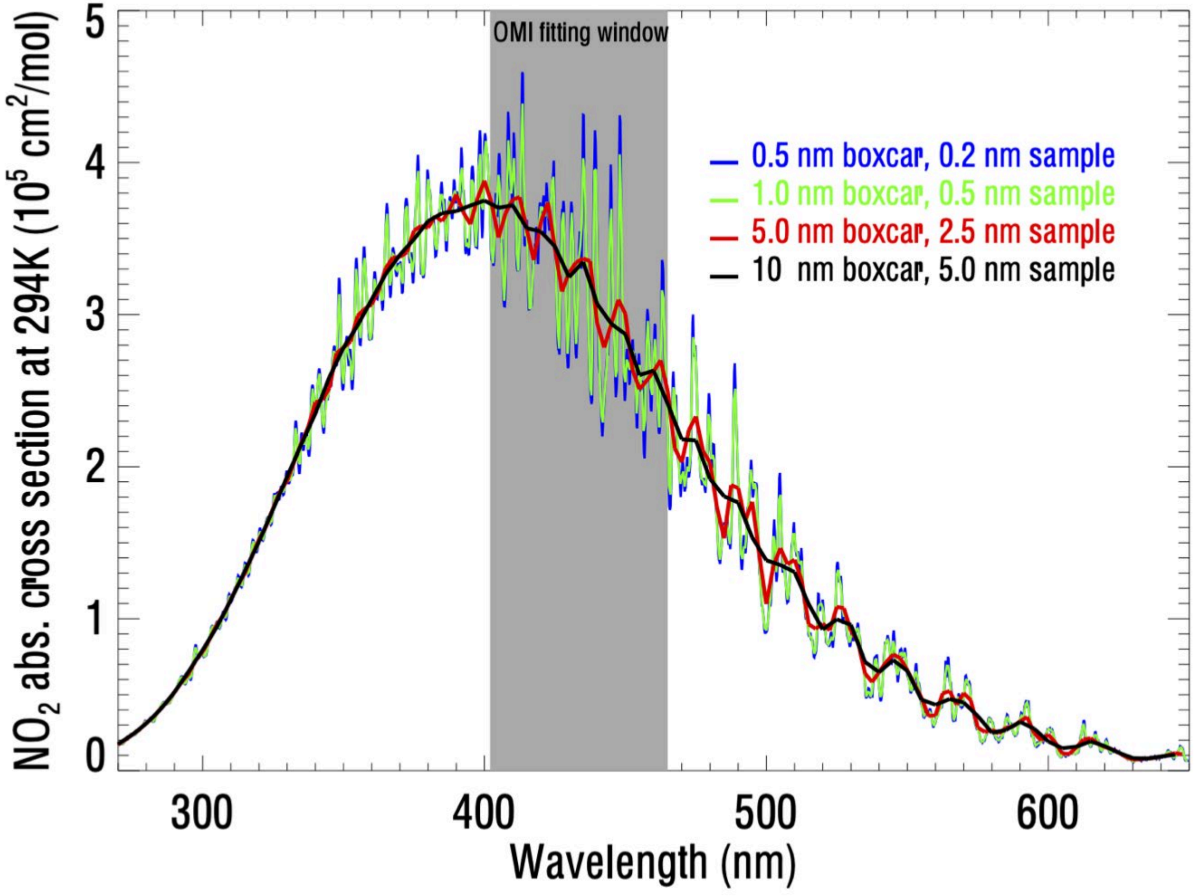
Can NO_2 from ocean color instruments be beneficial for atmospheric science?

- PACE OCI will have native pixels ~1 km x 1 km. TROPOMI resolution is ~ 3.5 x 5.5 km at nadir, so PACE OCI resolution is ~20 higher.
- Aircraft data can give an indication of benefits of higher spatial resolution



From Judd et al. (2020), Atmos. Meas. Tech.

NO_2 Absorption Cross Sections



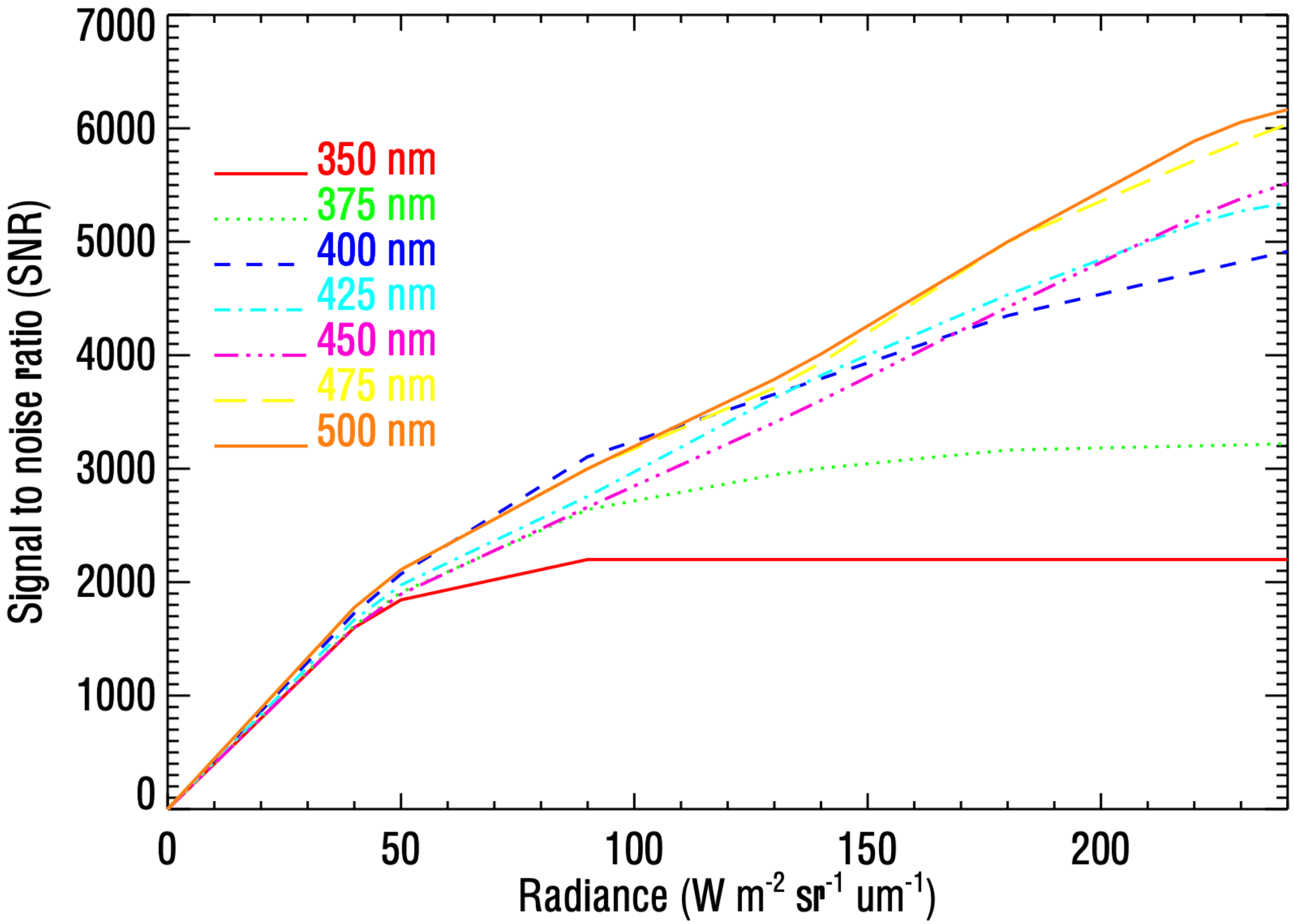
UV (more Rayleigh scattering, less surface sensitivity)

Blue/green (less Rayleigh scattering, more surface sensitivity)

- With OMI-like sensors, NO_2 slant columns densities (SCD) typically retrieved using high frequency structure in the range 400-465 nm
- With spectral resolution of PACE OCI (red), there is still some high frequency structure in range 400-500 nm
- Can we utilize a wider spectral range to retrieve accurate NO_2 (SCD) given instrument specifications for signal-to-noise?
- Machine learning trained on OMI standard retrievals may be an effective strategy

Signal-to-noise ratios (SNR) used in simulations

- Average OMI data over PACE OCI bands and resample to simulate PACE OCI data (over available OMI wavelengths)
- Add noise according to PACE SNR's from measurements (right)
- Run experiments using different wavelength ranges and predictors
  - Adding UV wavelengths to basic fitting window of 400-465 nm didn't help much
  - Adding observing geometry (angles) as predictors didn't help much
  - Adding stratospheric column as a predictor helped a little



Train a neural network to retrieve NO_2 slant columns

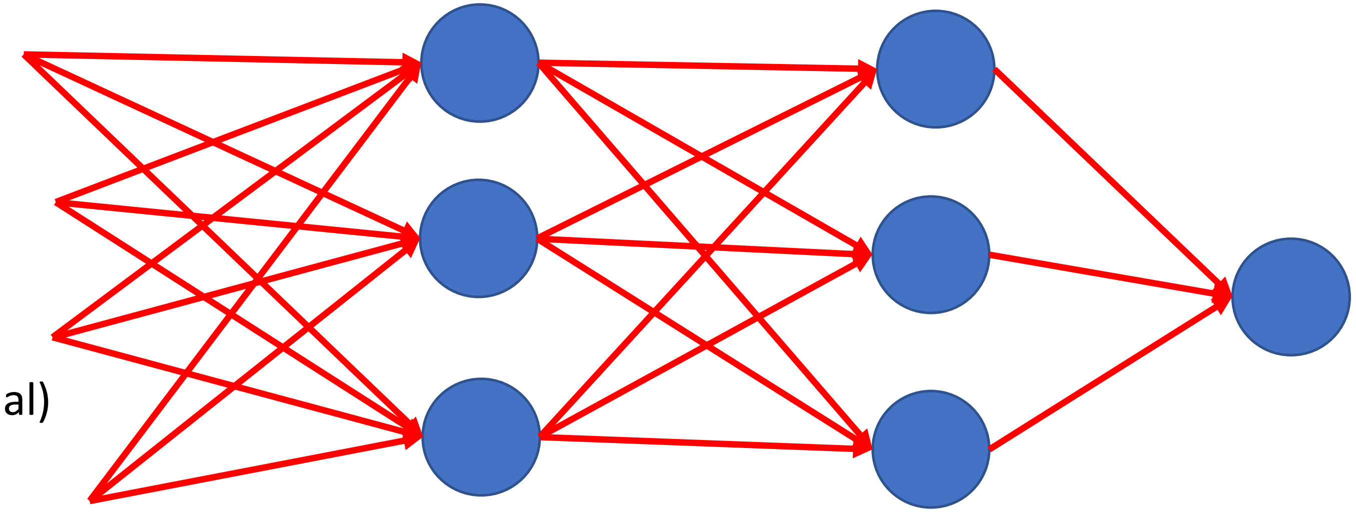
Inputs

Coefficients of leading spectral modes (half of total) of simulated radiances from a satellite instrument

Angles that define sun-satellite geometry (optional)

Stratospheric column amount (optional, from GMI, *a priori*)

Neural Network (2 hidden layers)



Output

OMI NO_2 slant columns derived from high frequency spectral structure (standard NASA product, destriped)

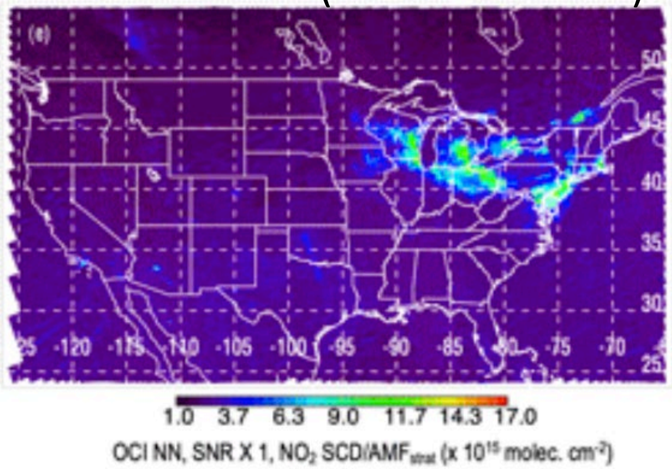
- Train each OMI row separately (each row can be considered a unique instrument)
- Train using data from 1 day every other month (in 2005) + additional high pollution days in winter

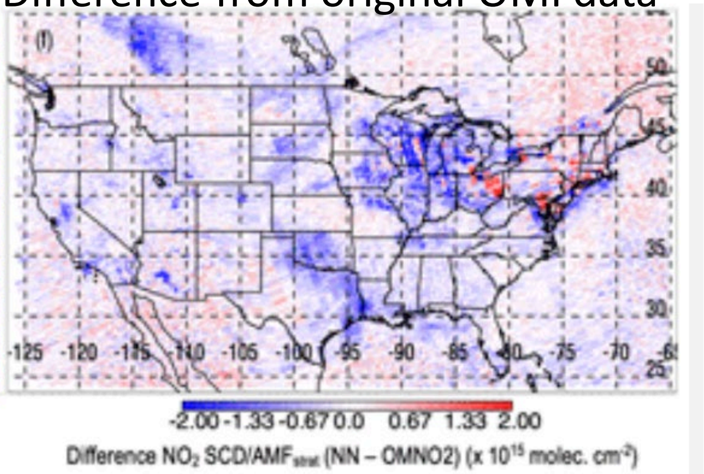
Results shown at OMI resolution for convenience 28 Jan 2005 (day independent of training)

- Results apply to PACE OCI native resolution (1 km, use your imagination)
- Results with nominal OCI noise (top panels) appear biased low over polluted regions
- Averaging 16 observations reduces errors as expected, so that high spatial resolution observations averaged monthly could yield useful information for spatial downscaling of TROPOMI data.

Nominal PACE OCI noise

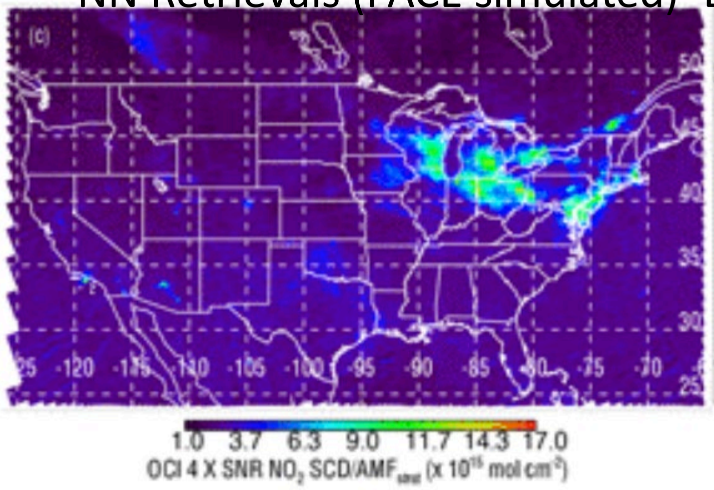
NN Retrievals (PACE simulated) Difference from original OMI data

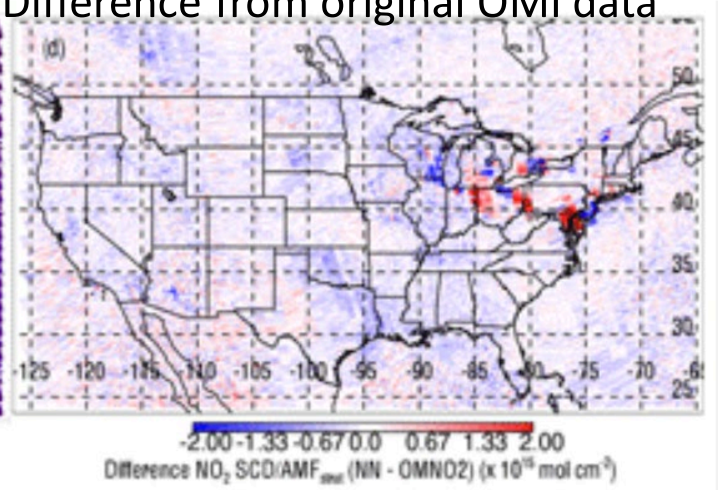




PACE OCI SNR increased by factor of 4 (average 16 observations)

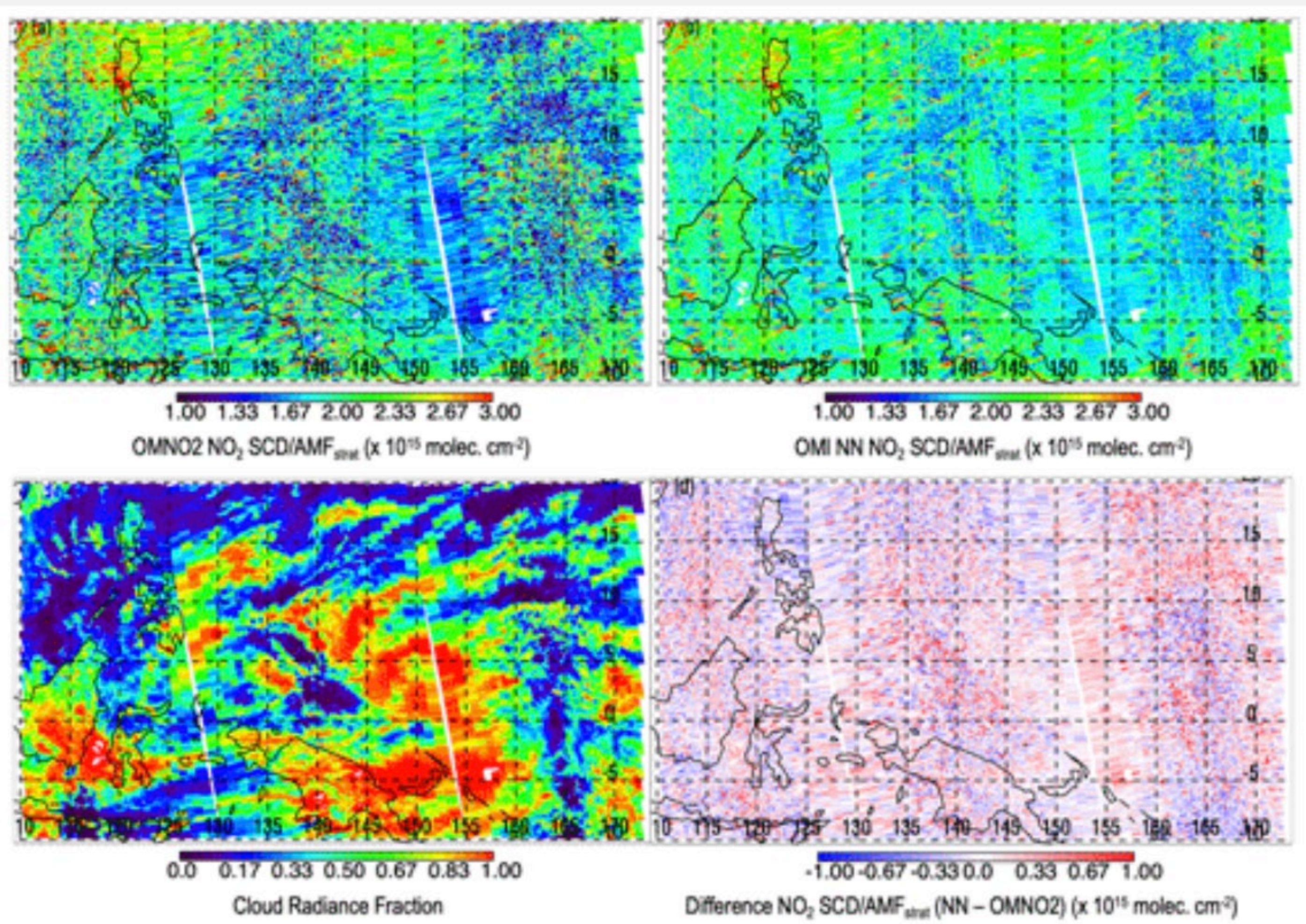
NN Retrievals (PACE simulated) Difference from original OMI data





Can machine learning with principal component analysis benefit existing NO_2 retrievals?

Results over clear Pacific with OMI original data



- Applied same approach to OMI at original spectral resolution using wavelengths 402-465 nm for potential noise reduction
- Use of leading principal components (half of total) with machine learning enables retrievals with apparent lower noise (inputs do not allow perfect reconstruction of noisy retrievals); this is currently a work in progress

Summary

- PACE OCI NO_2 algorithm could be trained up fairly quickly after launch with TROPOMI by collocating and adjusting air mass factors appropriately to the observed geometry.
- Can average PACE OCI NO_2 retrievals over ~1 month or more to provide results at 1 km resolution, may be useful for downscaling TROPOMI averages (environmental justice).
- Machine learning algorithms may be useful for reducing noise in retrievals.
  - Can be applied as a post processing step or as a stand alone fast spectral fit
  - Full traditional spectral fitting algorithms still needed, and periodic retraining may be necessary