A neural-network-based forward model to improve air quality estimation from spaceborne polarimeters

***UVA Team**: Stephen Whetzel, Abhi Dommalapati, Jack Peele, Sebastian Ranasinghe*

***Client**: NASA Langley Research Center*

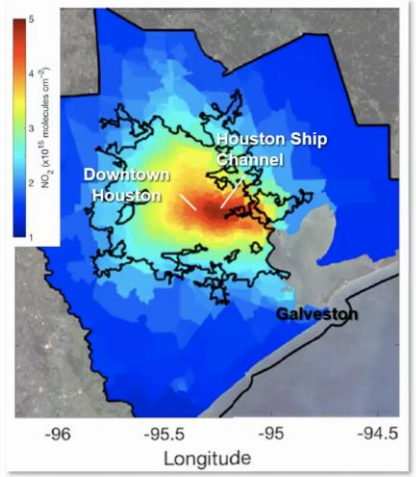
Introduction and Background

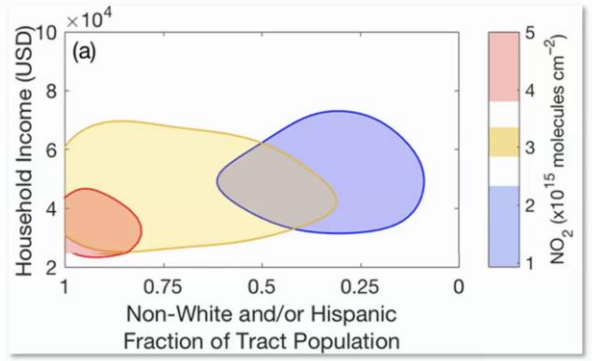
- Project goal: accurately and efficiently retrieve surface-level PM2.5 from satellite polarimeter and lidar measurements
  - PM2.5 is the mass-concentration of any particles in the air 2.5 microns or less in diameter (pollution & natural sources)

- The polarimeter satellite remote sensing retrieval problem is an *inverse problem*:
  - Optimal estimation inverts satellite polarimeter data (the output, or the measurements) into aerosol and ocean properties (the input, or the state vector)
  - Optimal estimation requires iteratively calling NASA's vector radiative transfer forward model (accurate but has a significant computational burden)

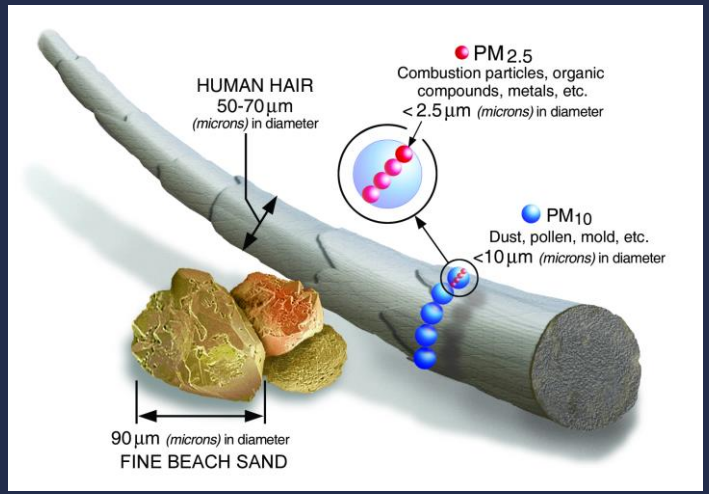
- Solution: Train an accurate and fast neural network to simulate PARASOL polarimeter measurements to replace the forward model and operationalize retrievals of aerosol properties and PM2.5

Pollution Disproportionately Impacts Low-Income, Non-White, and Hispanic Neighborhoods in Houston, TX





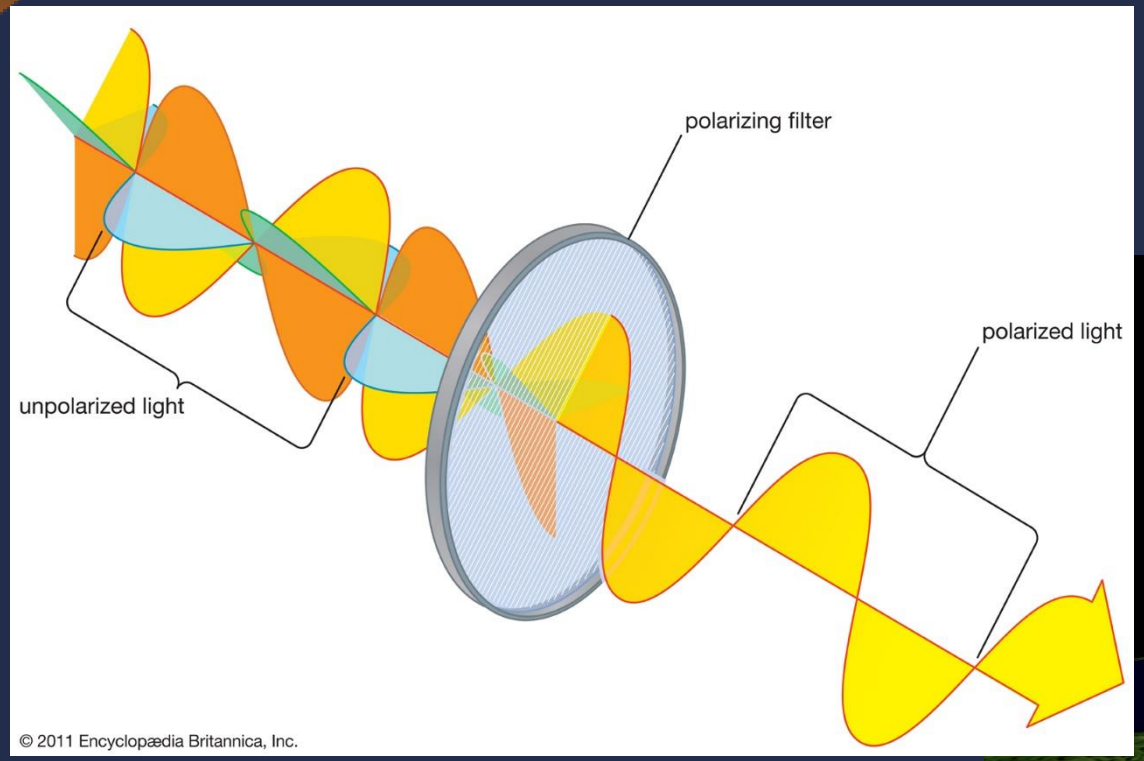
Demetillo**, M. A. G., Navarro, A., Knowles, K. K., Fields, K. P., Geddes, J. A., et al (2020). Observing nitrogen dioxide air pollution inequality using high-spatial-resolution remote sensing measurements in Houston, Texas. *Environmental Science & Technology, 54*(16), 9882–9895. https://doi.org/10.1021/acs.est.0c01864 **FINESST proposal

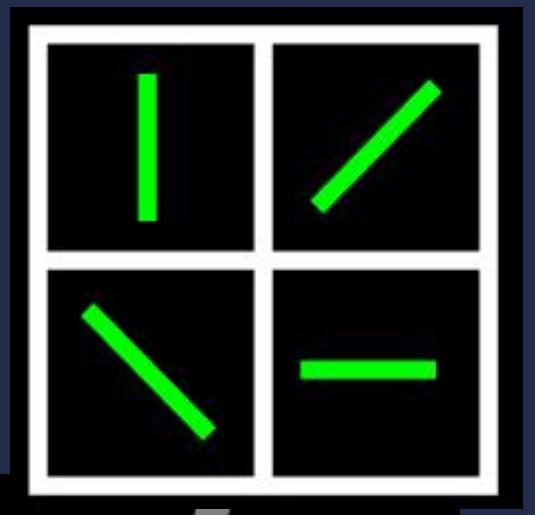


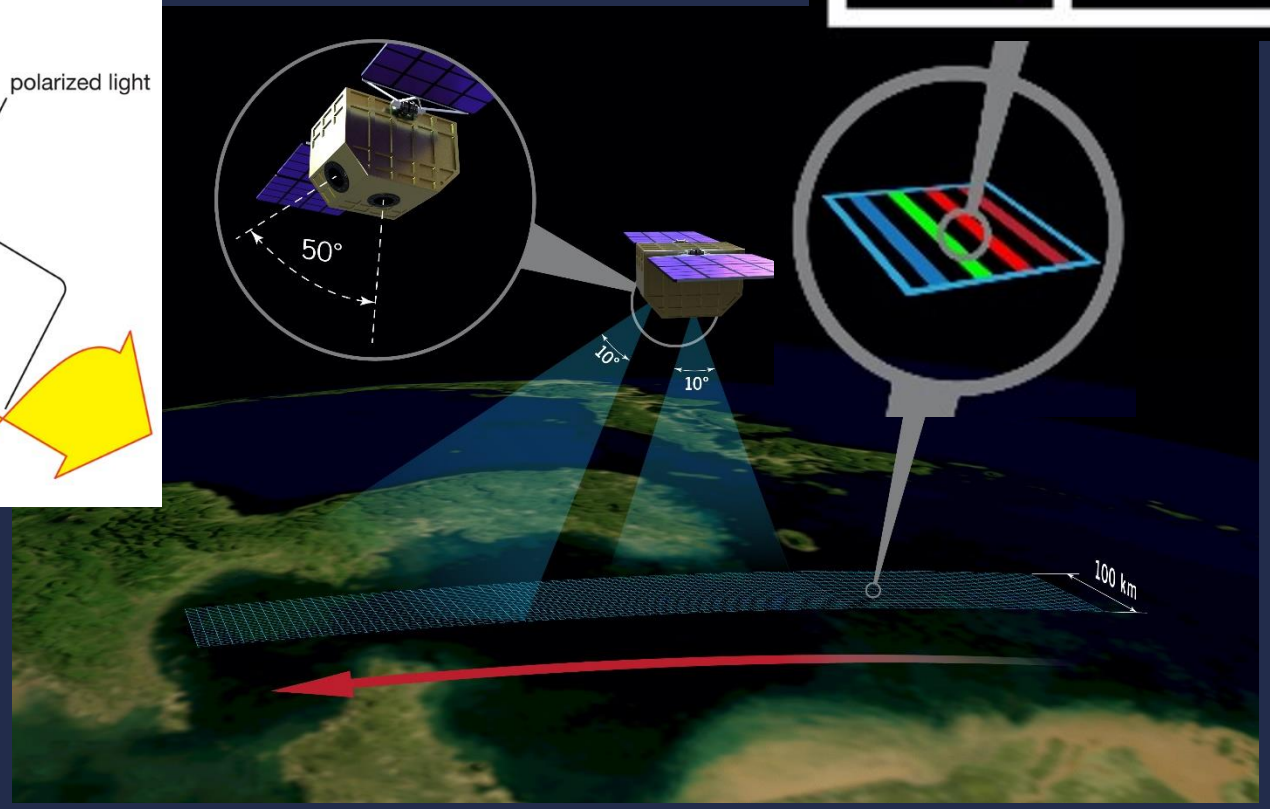
UNIVERSITY of VIRGINIA

NASA

Satellite Polarimeters

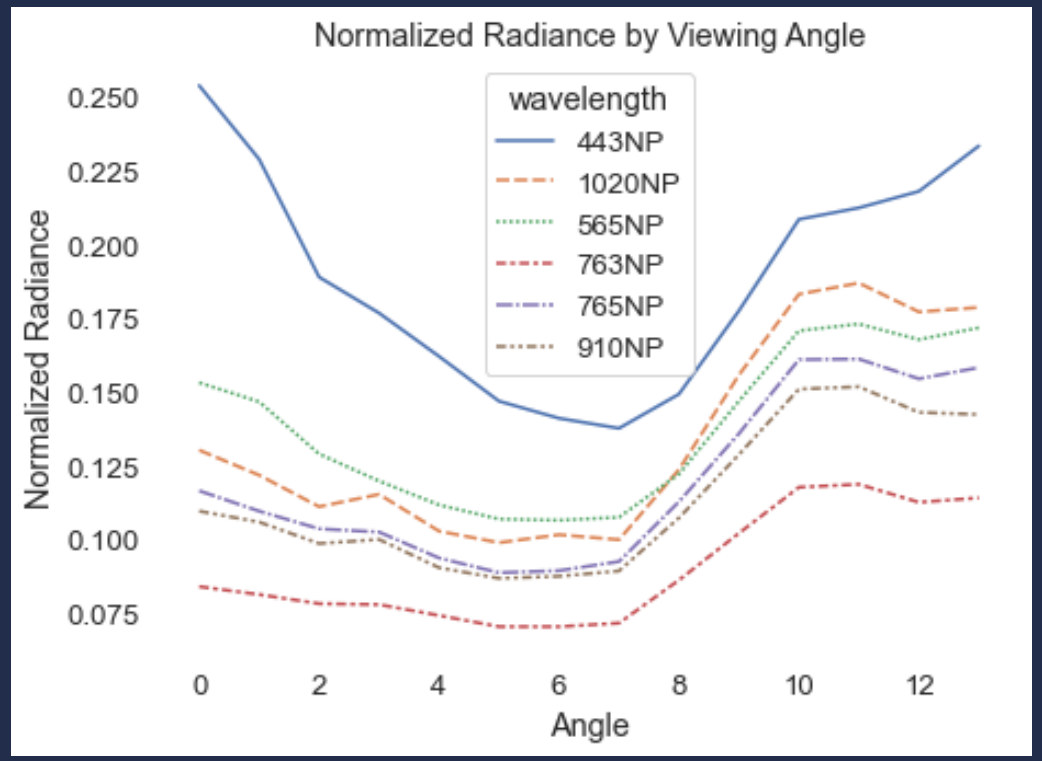


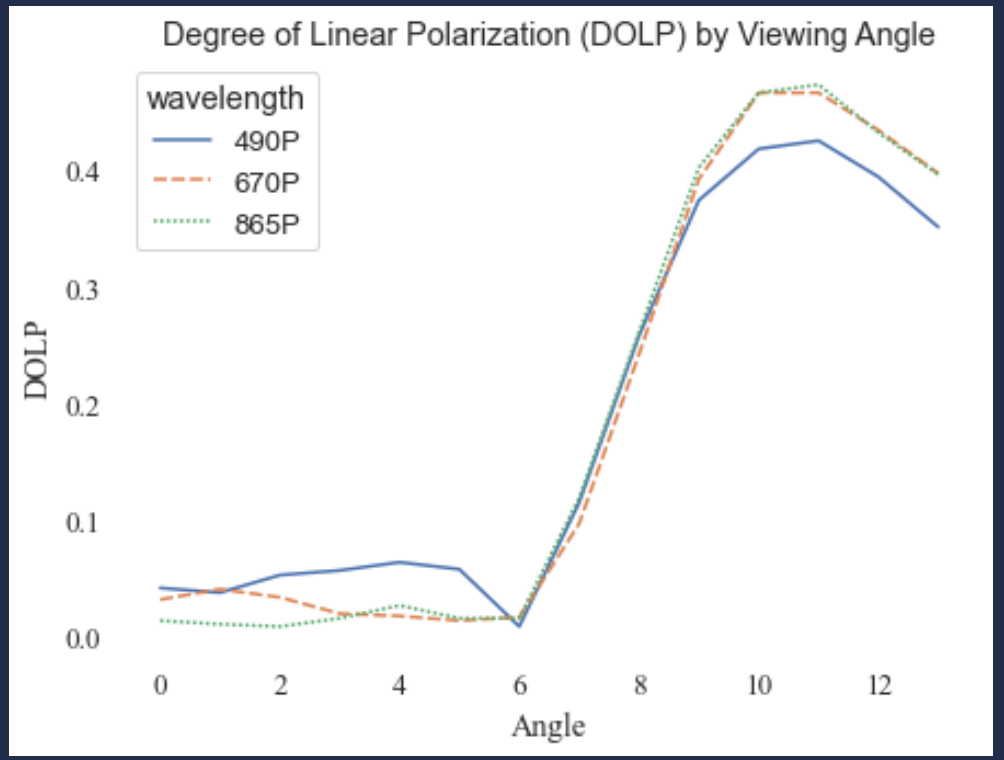




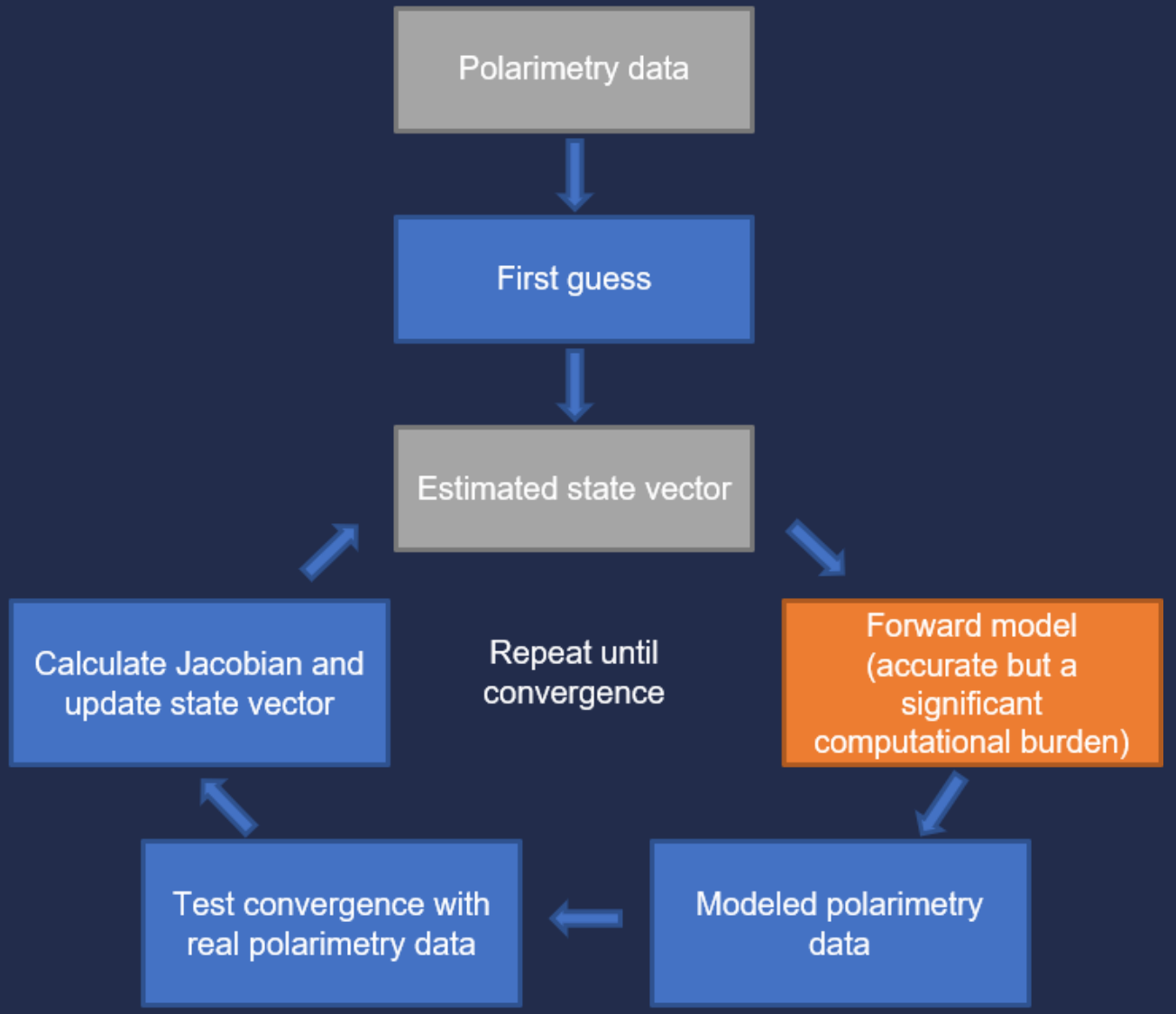
Data Description

- The University of Lille stores large volumes of satellite polarimeter data through ICARE
  - The data include spaceborne measurements at multiple wavelengths for a wide range latitude/longitude values and are stored in HDF files.

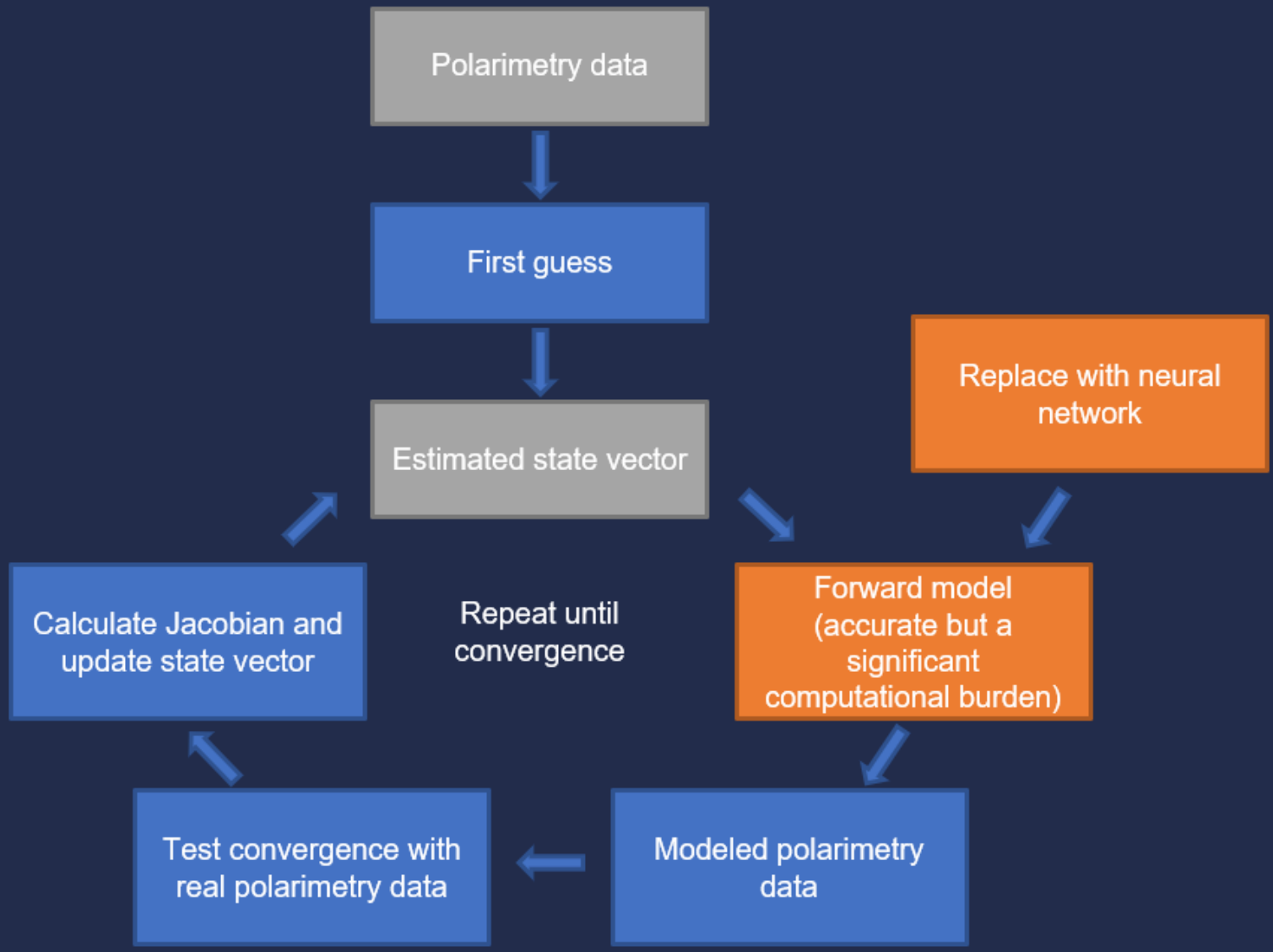




Retrieval Algorithm



Retrieval Algorithm



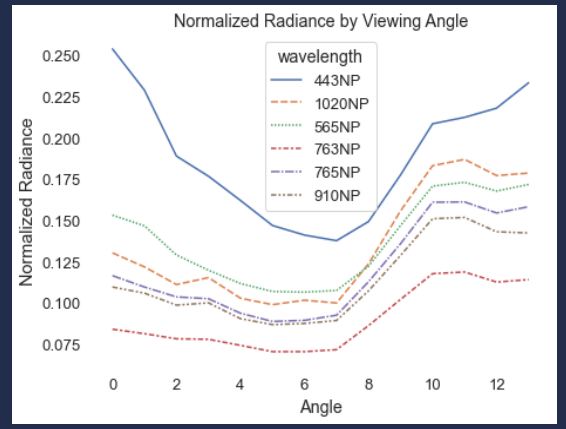
Methodology

- Simulating training dataset
  - Problem: Real data contain only measured polarized radiance, not state vectors
    - Retrieval of sufficient amount of state vectors with real data is prohibitively inefficient
    - Difficult to ensure sufficient state vector diversity
  - Solution: simulate data by sampling state vector values from physically representative distributions and then run the forward model to estimate spaceborne polarimetric measurements
- Neural network design and training
  - Sufficiently complex for accuracy but drastically reduces the computational resources required

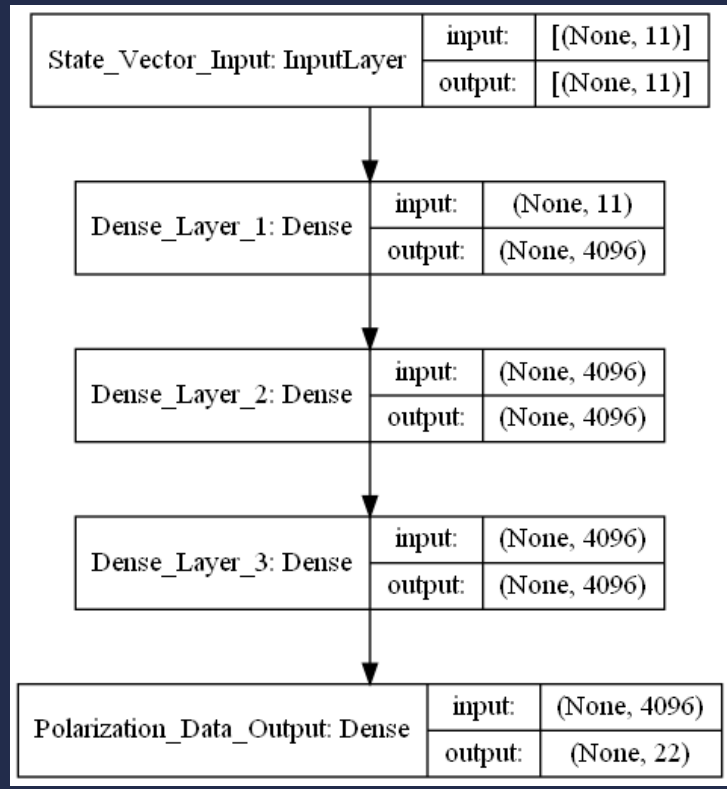
Input: aerosol and ocean input parameters (the state vector*)

$$\mathbf{x} = \langle\ \tau_{555f}\quad r_{nf}\quad \sigma_{gf}\quad n_{rf}\quad n_{if}\quad \tau_{555c}\quad r_{nc}\quad \sigma_{gc}\quad n_{rc}\quad n_{ic}\quad v\quad \mathrm{CHL}\quad z_c\quad z_f\ \rangle$$

Forward Model



Output: polarimeter measurements



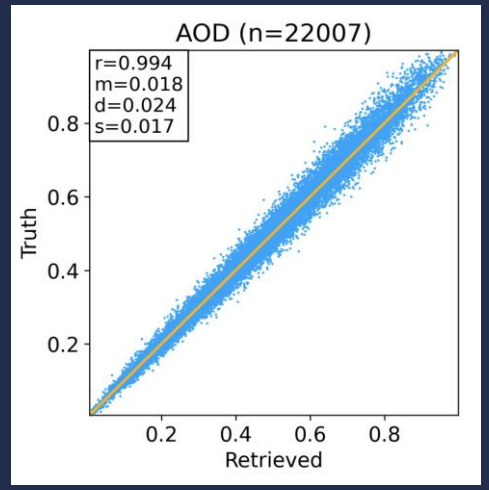
*State Vector Parameters
Bimodal (fine- and coarse-mode) aerosol optical depth (AOD) at 555 nm
Bimodal aerosol effective radius and variance
Aerosol complex refractive index
Ocean surface roughness
Ocean chlorophyll concentration

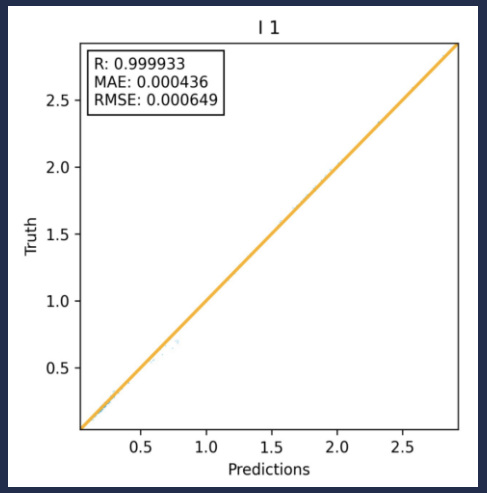
Results

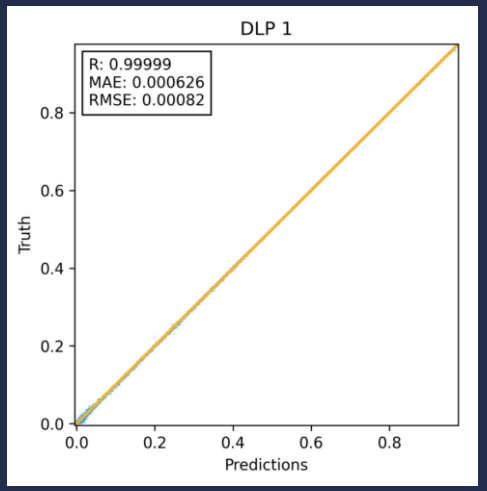
- The neural network model improves the speed of the forward model by a factor of 1000x without sacrificing appreciable accuracy
  - Overall satellite retrieval algorithm is likewise sped up

- Demonstrates that neural networks can effectively replace the vector radiative transfer forward model; this physics-based approach is an example of scientific machine learning, and thus preserves the depth and nuance of these data

- The speed enhancement of neural networks makes low-latency aerosol retrievals and PM2.5 products computationally feasible, and will allow for the development of pollution forecast models that require highly accurate/low latency data acquisition







UNIVERSITY of VIRGINIA

NASA

Future Work

- After this successful test using simulated data, the next step is to retrieve aerosol properties and PM2.5 from real satellite polarimeter data from PARASOL
  - Scripts have been created to interface with online data archives (ICARE) - download data files and process to L1C file format
- Collocation with Lidar measurements to model surface-level aerosol concentration and PM2.5
- Compare PM2.5 retrievals with those from GRASP, the algorithm for PARASOL polarimeter aerosol retrievals and PM2.5
- Create polarimeter-derived PM2.5 products for the future NASA Plankton, Aerosols, Clouds, Ecosystems (PACE) mission

References

S. Feng, D. Gao, F. Liao, F. Zhou, and X. Wang, “The health effects of ambient PM2.5 and potential mechanisms,” Ecotoxicology and Environmental Safety, vol. 128, pp. 67–74, 2016.

V. A. Southerland, M. Brauer, A. Mohegh, M. S. Hammer, A. van Donkelaar, R. V. Martin, J. S. Apte, and S. C. Anenberg, “Global Urban Temporal Trends in fine particulate matter (PM2·5) and attributable health burdens: Estimates from Global Datasets,” *The Lancet Planetary Health*, vol. 6, no. 2, Feb. 2022.

R. V. Martin, “Satellite remote sensing of Surface Air Quality,” Atmospheric Environment, vol. 42, no. 34, pp. 7823–7843, 2008.

S. Stamnes, C. Hostetler, R. Ferrare, S. Burton, X. Liu, J. Hair, Y. Hu, A. Wasilewski, W. Martin, B. van Diedenhoven, J. Chowdhary, I. Cetinić, L. K. Berg, K. Stamnes, and B. Cairns, “Simultaneous polarimeter retrievals of Microphysical Aerosol and ocean color parameters from the ‘MAPP’ algorithm with comparison to high-spectral-resolution lidar aerosol and Ocean Products,” Applied Optics, vol. 57, no. 10, p. 2394, 2018.

S. Shi, T. Cheng, X. Gu, H. Guo, Y. Wu, Y. Wang, F. Bao, and X. Zuo, “Probing the dynamic characteristics of aerosol originated from South Asia biomass burning using polder/grasp satellite data with relevant accessory technique design,” Environment International, vol. 145, p. 106097, 2020.

P. Lier and M. Bach, “Parasol a microsatellite in the A-train for earth atmospheric observations,” Acta Astronautica, vol. 62, no. 2-3, pp. 257–263, 2008.

Parasol, 24-Apr-2015. [Online]. Available: https://parasol.cnes.fr/en/PARASOL/index.htm. [Accessed: 08-Apr-2022].

G. Shaddick, M. L. Thomas, A. Jobling, M. Brauer, A. van Donkelaar, R. Burnett, H. Chang, A. Cohen, R. Van Dingenen, C. Dora, S. Gumy, Y. Liu, R. Martin, L. A. Waller, J. West, J. V. Zidek, and A. Prüss-Ustün, “Data Integration Model for air quality: A hierarchical approach to the global estimation of exposures to ambient air pollution,” *Applied Statistics*, 2017.

E. Gorman, D. A. Kubalak, P. Deepak, A. Dress, D. B. Mott, G. Meister, and J. Werdell, “The NASA plankton, aerosol, cloud, ocean ecosystem (PACE) mission: An emerging era of global, Hyperspectral Earth System Remote Sensing,” *Sensors, Systems, and Next-Generation Satellites XXIII*, 2019.

“Parasol”, distributed by ICARE On-line Data Archive, https://www.icare.univ-lille.fr/asd-content/archive/?dir=PARASOL/

D. M. Winker, M. A. Vaughan, A. Omar, Y. Hu, K. A. Powell, Z. Liu, W. H. Hunt, and S. A. Young, “Overview of the CALIPSO mission and CALIOP data processing algorithms,” Journal of Atmospheric and Oceanic Technology, vol. 26, no. 11, pp. 2310–2323, 2009.

Pollution Disproportionately Impacts Low-Income, Non-White, and Hispanic Neighborhoods in Houston, TX