High Spatiotemporal Resolution Estimates of $PM_{2.5}$ in West Africa Using Well-Calibrated Air Sensors, Satellite Data, and Machine Learning

Prof. Dan Westervelt
Columbia University, Lamont-Doherty Earth Observatory, New York, USA (primary appointment)
University of Mohammed VI Polytechnic, Marrakech, Morocco

Lamont-Doherty Earth Observatory
COLUMBIA UNIVERSITY | EARTH INSTITUTE

UM6P
University Mohammed VI Polytechnic

Westervelt Aerosol Group

Garima Raheja
(PhD student)

Benjamin Yang
(PhD student)

Joe Adabouk Amooli
(PhD student)

Yanda Zhang (Postdoc)

Celeste McFarlane
(fmr Research Staff Assistant)

Victoria Owusu Tawiah
(visiting PhD Student)

Josephine Kanyeria
(MS student)

Paulson Kasereka
(PhD student, UM6P)

Andrea Belvis Aquino
(Summer intern, UPRM)

Moses Njeru
(MS student, UoN)

CLEAN AIR FUND

COLUMBIA UNIVERSITY
DATA SCIENCE INSTITUTE

Satellite data: what role for air sensors?

- **The problem**: satellite estimates provide excellent spatial coverage of pollutants at decent resolution (1 km^2 daily), but measure the entire atmospheric column, rather than "breathing level"

- **One solution**: geophysical model to convert AOD to surface PM2.5 (van Donkelaar, Martin, etc)

$$PM_{2.5} = \eta\,(x, y, t) \times AOD_{BE}$$

GEOS-Chem

- **Another solution**: use statistical models to develop an empirical relationship between surface PM2.5 and column AOD

$$PM2.5 = f(AOD, T, Pr, NO2, CO)$$

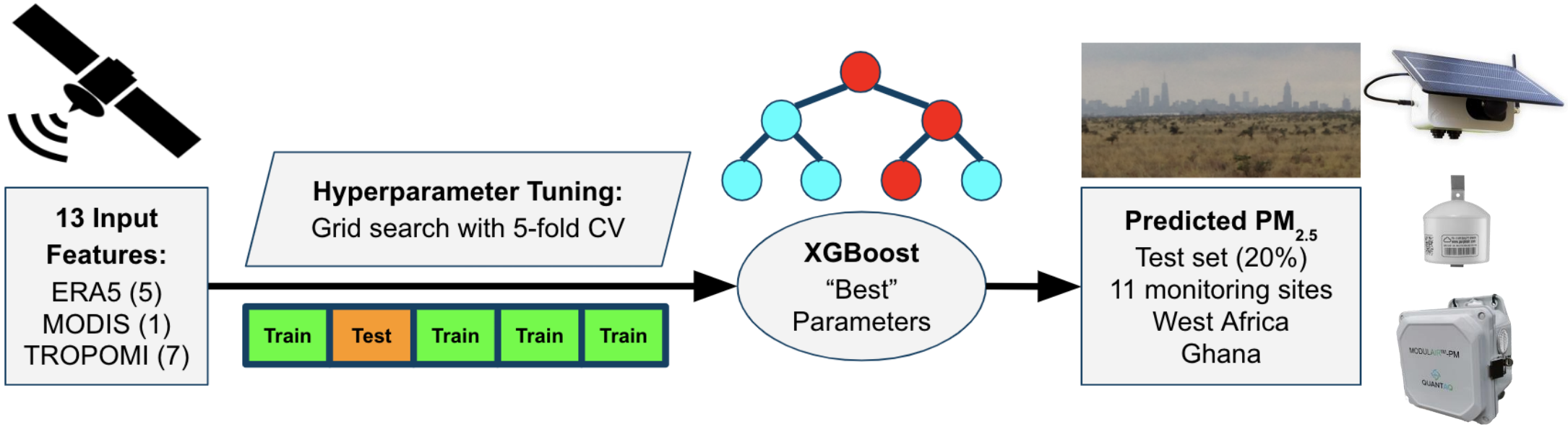
High Spatiotemporal Resolution Modeling of $PM_{2.5}$ in West Africa Using Satellite Data and Machine Learning

Goal: Estimate $PM_{2.5}$ at high spatiotemporal resolution (1 km^2, daily) in West Africa over the past 2 decades

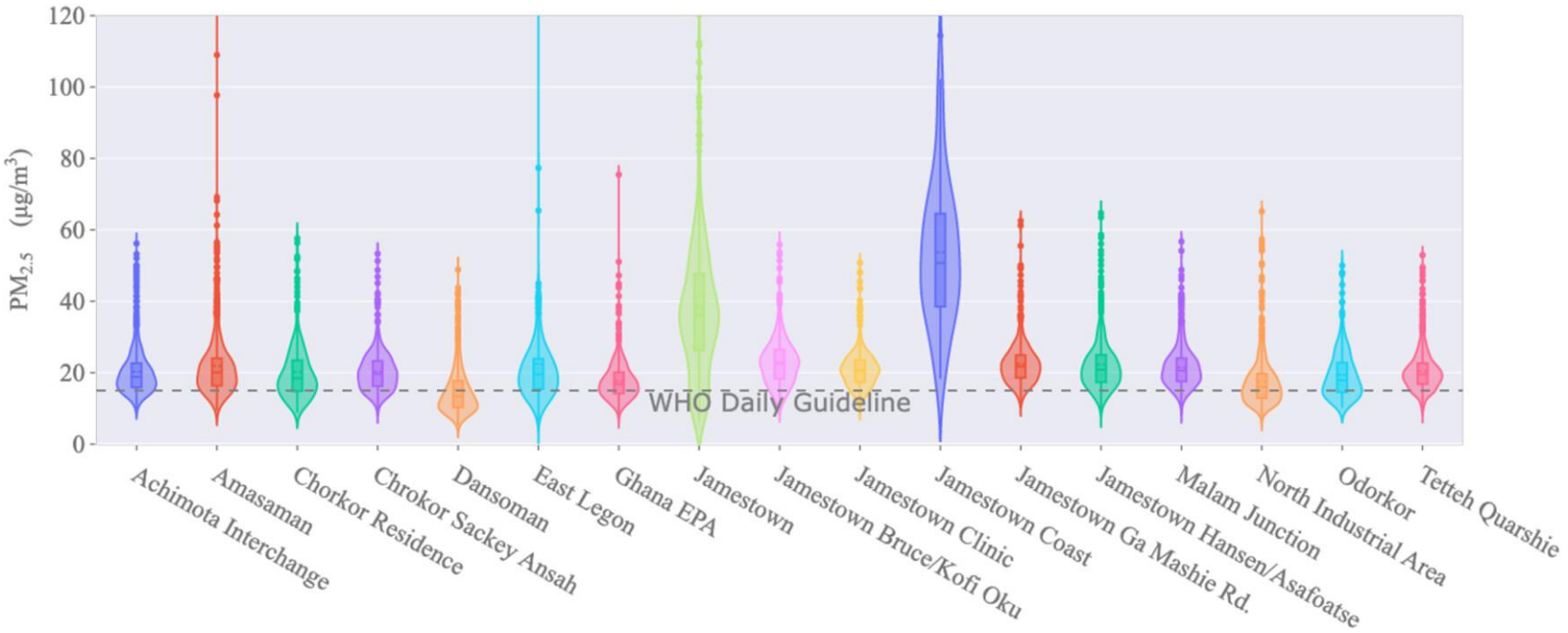
Data & Methods: Trained, tested, and fine-tuned a machine learning (XGBoost) model:

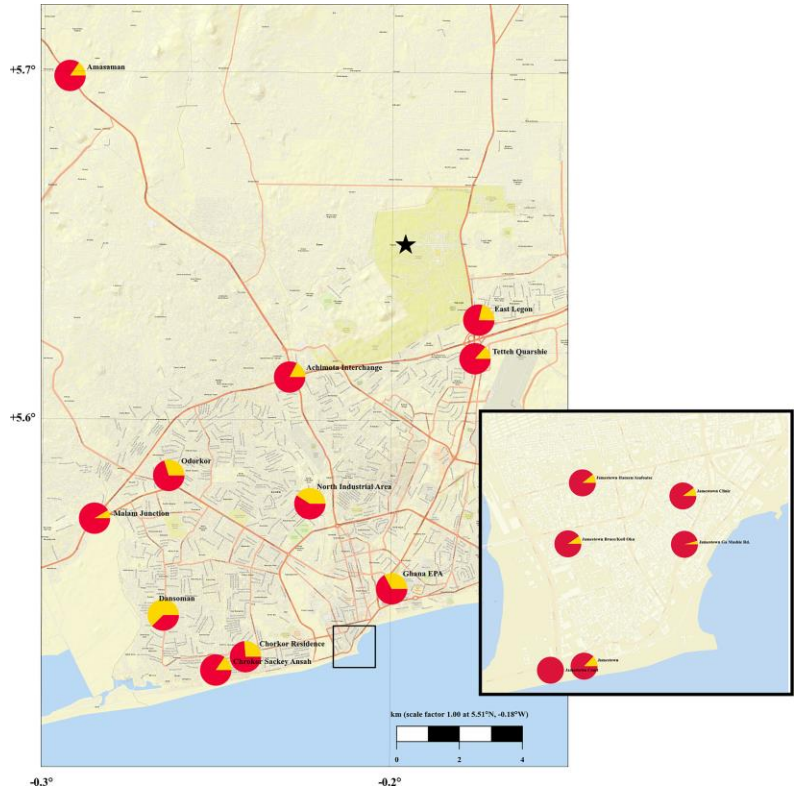
- ✓ $PM_{2.5}$ from air sensor and reference grade monitoring sites
- ✓ AOD from MODIS MAIAC satellite retrievals
- ✓ 5 meteorological features from ERA5
- ✓ 7 trace gas or aerosol features from TROPOMI satellite retrievals
  - ✓ Value of trace gases demonstrated in Zheng et al. (2023)

Applications: Epidemiological, policy, and environmental justice studies (Ghana)



Example of air sensor networks that are leveraged – Accra, Ghana

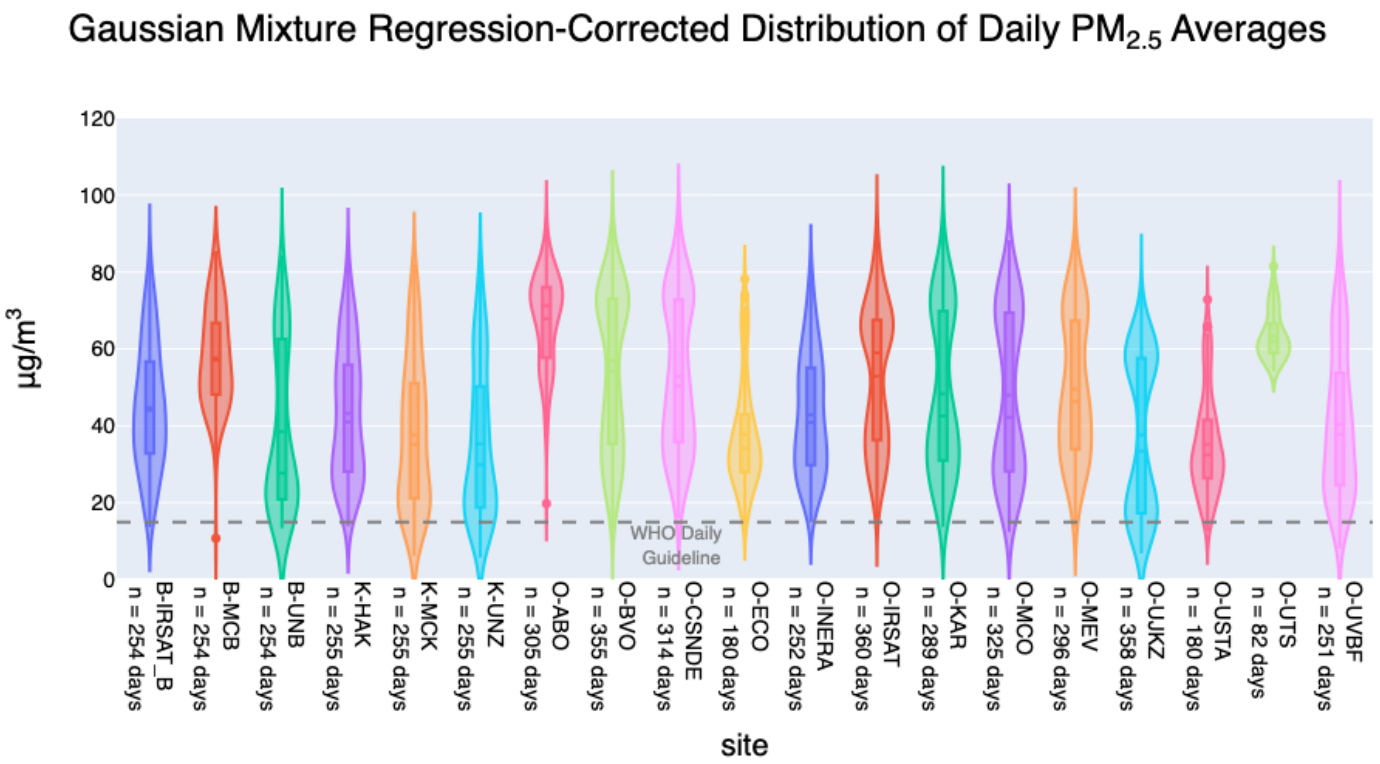


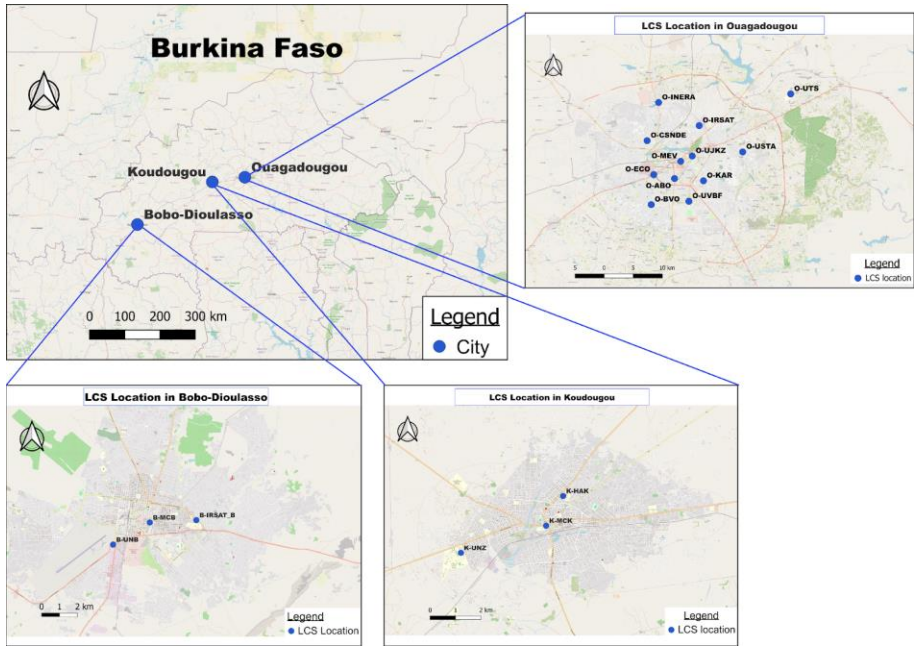


- 3 years of locally-calibrated Clarity Node-S sensors in Accra, Ghana

Raheja et al., 2023, Environmental Science and Technology

Example of air sensor networks that are leveraged – Burkina Faso

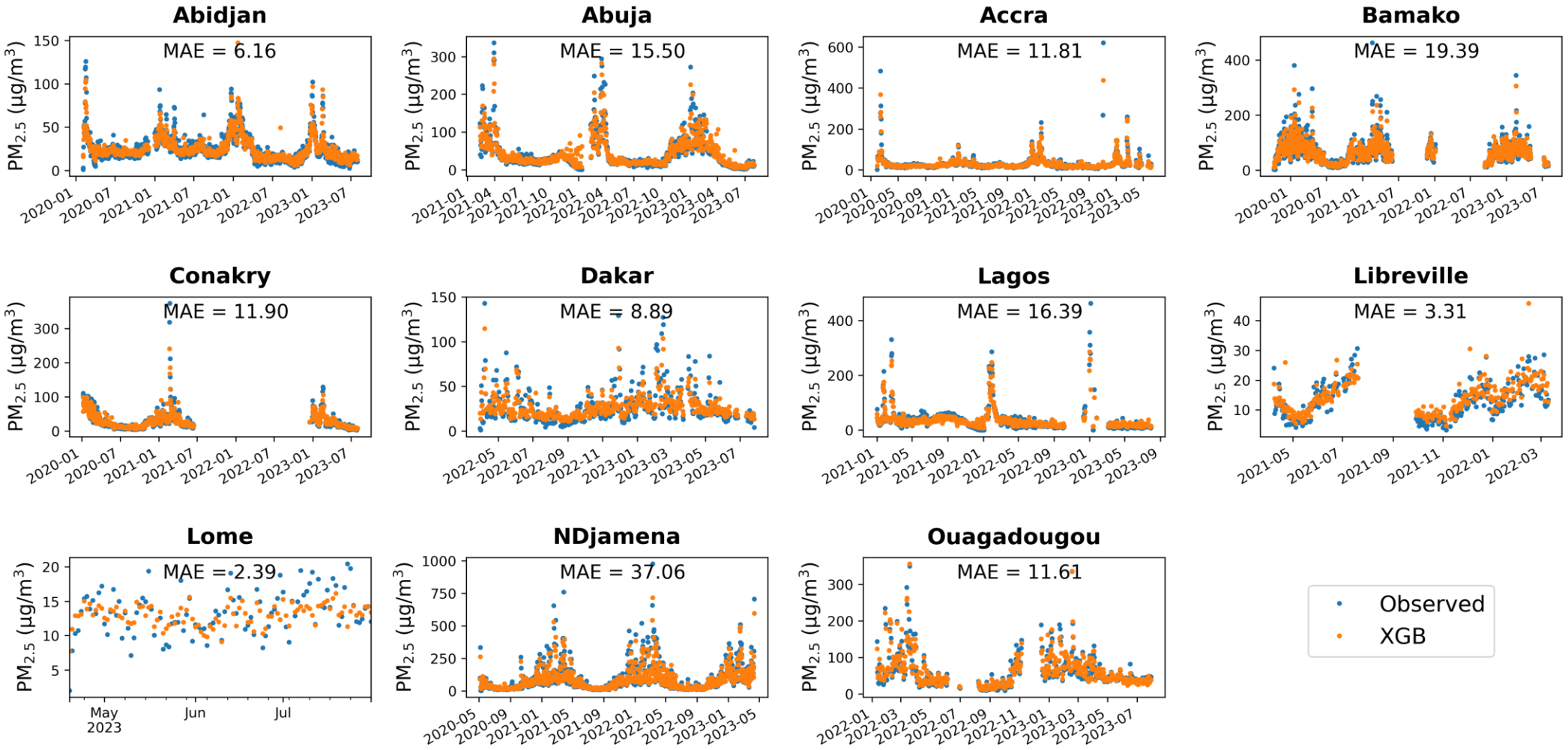




- 1 year of locally-calibrated Clarity Node-S sensors in 3 cities in Burkina Faso

Nana et al., 2024, submitted to ES:Atmosphere

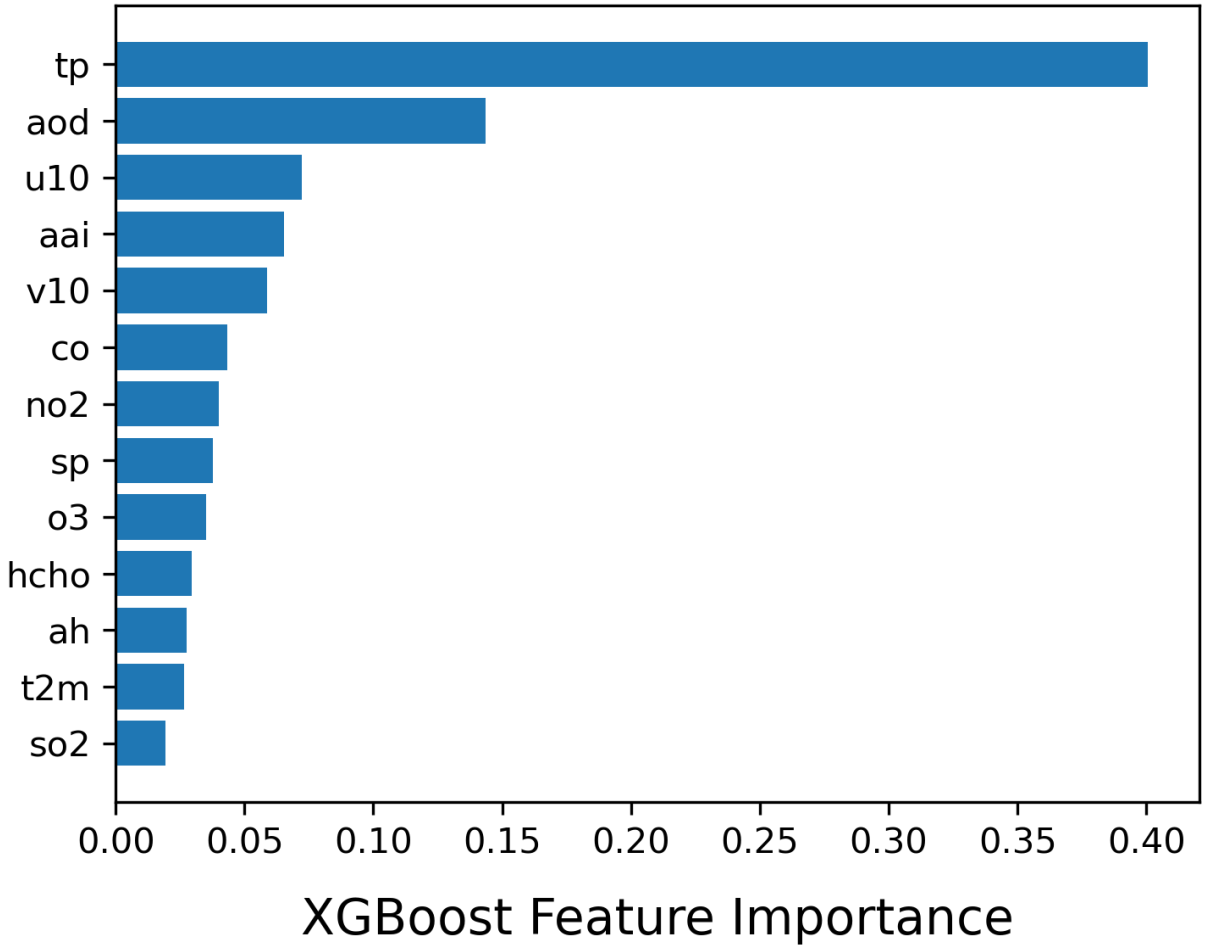
Predicted ML-satellite PM2.5 versus observations

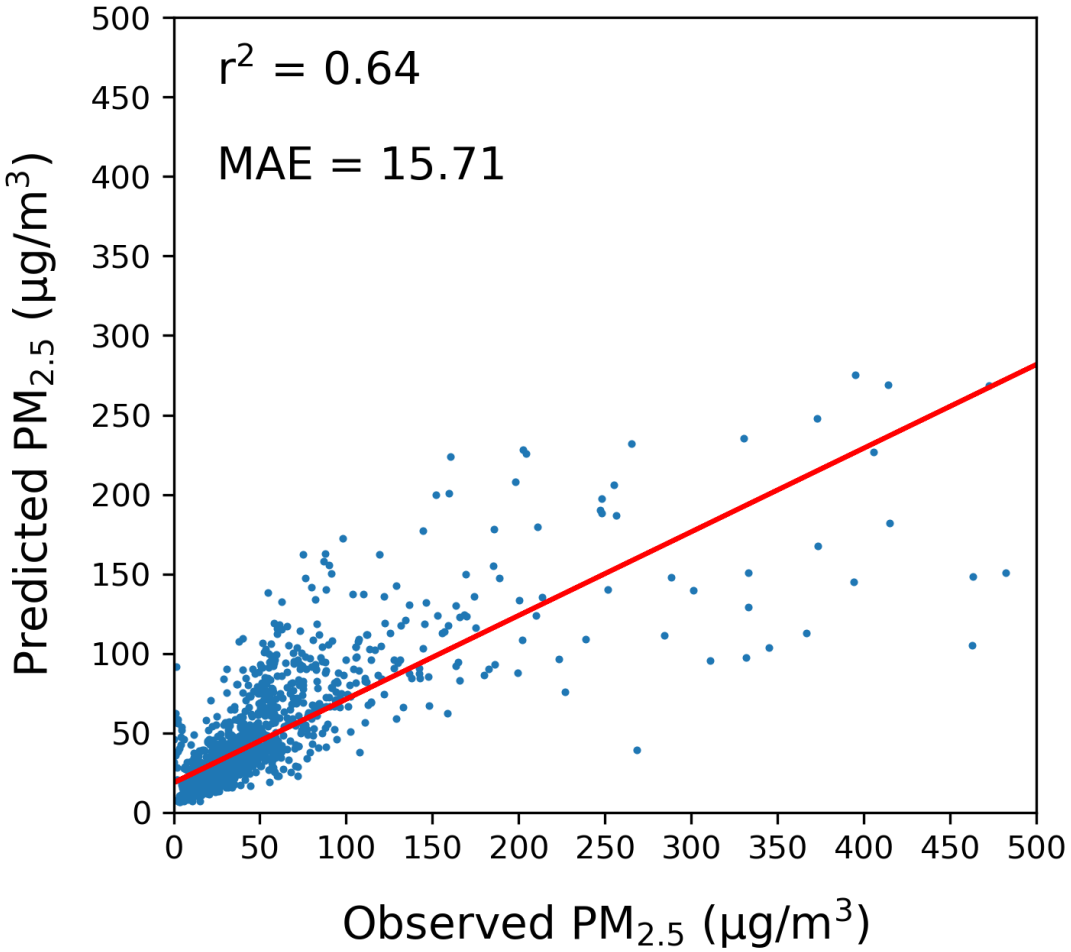


- R^2 ~ 0.7
- MAE ranges from 2-40 µg m^{-3}

Westervelt et al., 2024, in prep

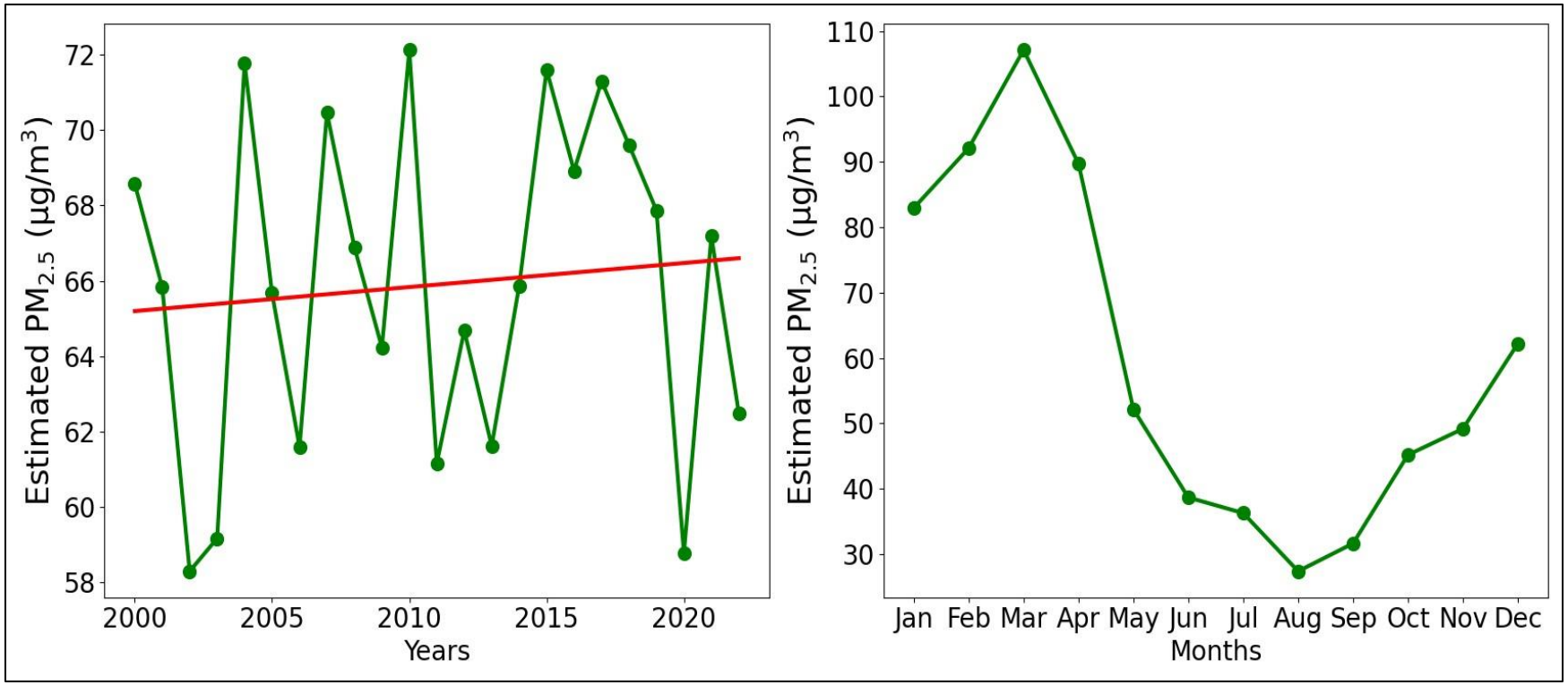
Predicted ML-satellite PM2.5 versus observations





Westervelt et al., 2024, in prep

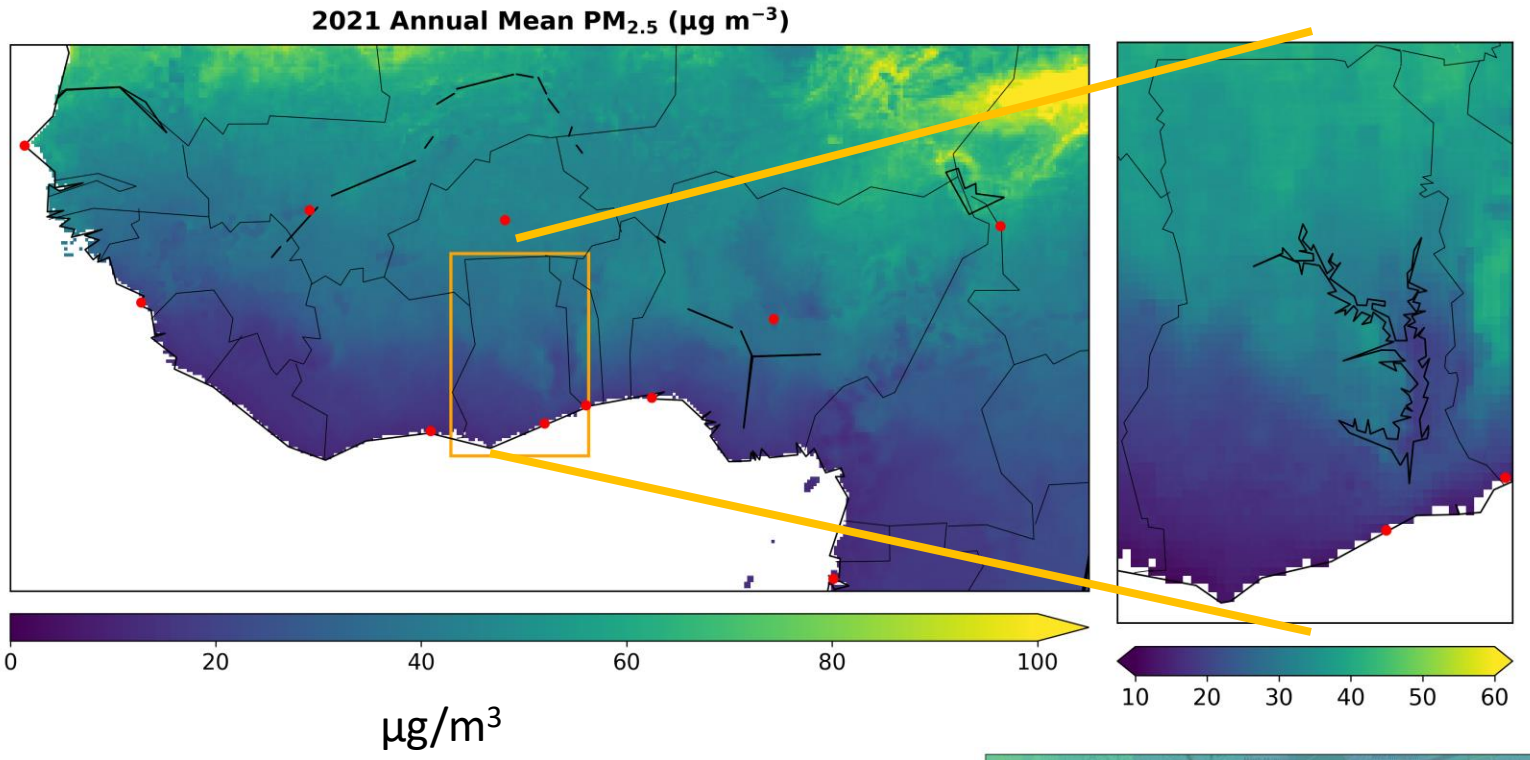
Results from Burkina Faso

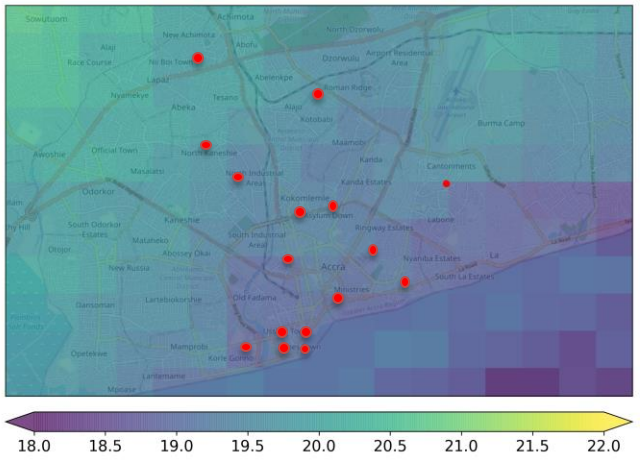


- XGBoost model has $r^2 = 0.87$ and nRMSE ~ 20%
- Evidence of slight increases in PM2.5 since 2000 (beginning of satellite record), with substantial interannual variability

Amooli et al., 2024, submitted to ES: Atmospheres

Results from Accra: 1 km^2 spatial resolution





- Can apply statistical relationship back through the satellite record to get a long term (2000-present), daily, 1km^2 dataset

Applications: Epidemiological, policy, and environmental justice studies (Ghana)

Westervelt et al., 2024, in prep

Summary

- Satellite data from agencies such as NASA provide near-complete spatial coverage, but their columnar nature is imperfect representations of surface pollution.
  - Growing hybrid observational networks of surface PM2.5 over Africa can be used to convert columnar AOD to “breathing level” PM2.5

- We demonstrate satellite-ML predictions of PM2.5, trained against air sensor data, with high r^2 (>0.7) and low error (20%)
  - Suitable for use in long-term trends, policy evaluation, epidemiology