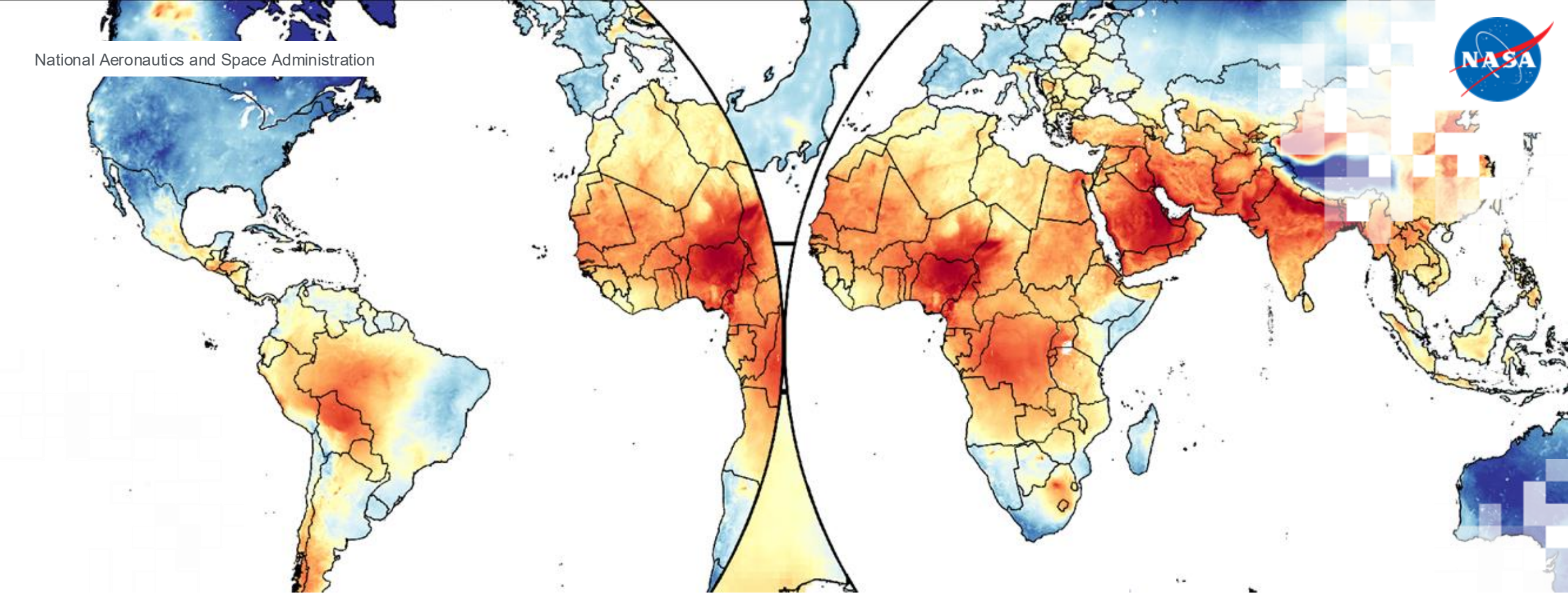
National Aeronautics and Space Administration



Estimating Surface $PM_{2.5}$ using Satellite Data and Other Information Sources

Part 3: Case Study – Comparing Available Satellite-Derived $PM_{2.5}$ Products to In-Situ Measurements

Carl Malings (Morgan State University, MSU), Pawan Gupta (NASA GSFC), Aaron van Donkelaar (Washington University in St. Louis), Junhyeon Seo (MSU), Sebastián Diez (Universidad del Desarrollo), & Ana Prados (University of Maryland Baltimore County)

July 22, 2026

Part 3 – Trainers

Carl Malings

Assistant Research Scientist

Morgan State University

Sebastián Diez

Research Scientist

Centro de Investigación en Tecnologías para la Sociedad

Universidad del Desarrollo, Chile

Part 3 Objectives

By the end of Part 3, participants will be able to:

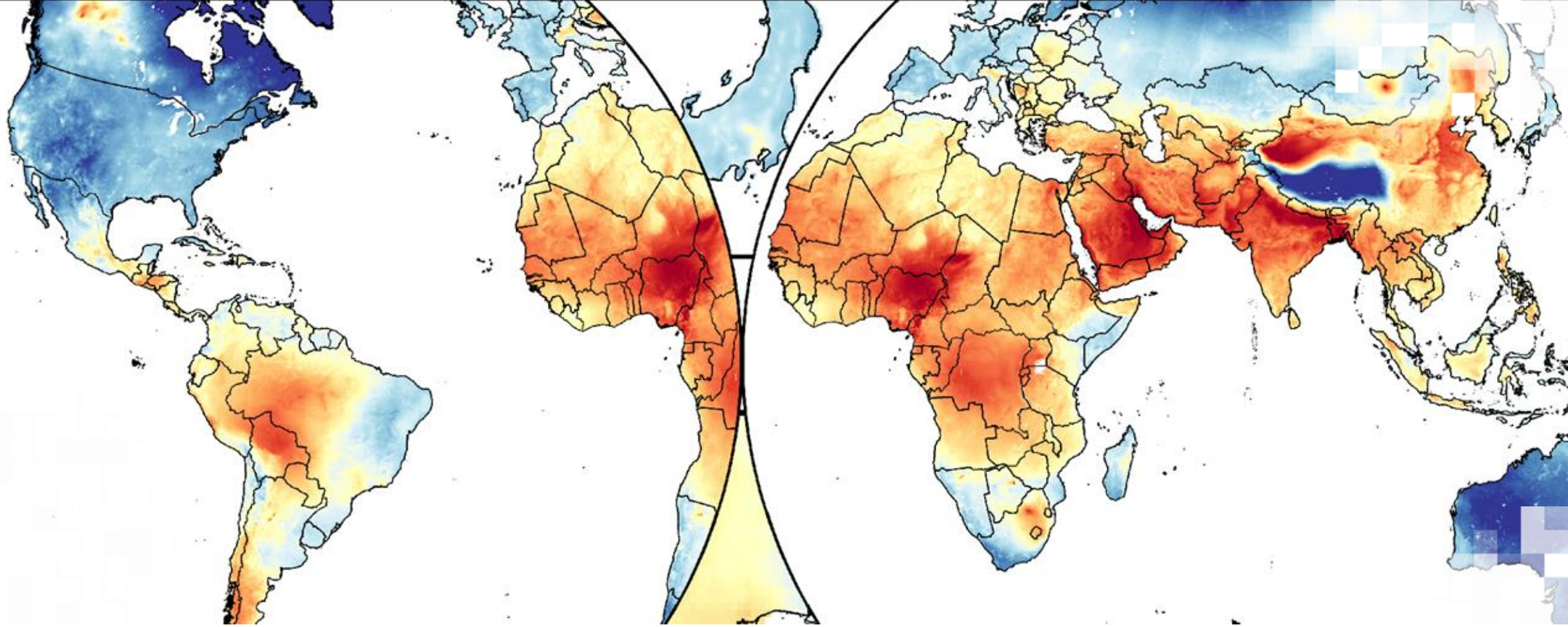
- Use NASA Earthdata, the SatPM website, OpenAQ, and Python code to access surface $PM_{2.5}$ data products for the location and time period of this case study.
- Compare these satellite-derived $PM_{2.5}$ estimates to in-situ measurements using appropriate co-location strategies and evaluation metrics.

Review of Prior Knowledge

- Washington University in St. Louis's Sat$PM_{2.5}$ uses a hybrid geophysical method to combine satellite AOD and global CTM information to estimate monthly-average $PM_{2.5}$ at ~1km spatial resolution.
- NASA's Bias-Corrected MERRA-2 $PM_{2.5}$ uses machine learning to correct for biases in the MERRA-2 reanalysis, producing $PM_{2.5}$ estimates at the same spatial (~50km) and temporal (hourly) resolution.
- Ground-based monitors provide the best data on $PM_{2.5}$ concentrations; evaluating the quality of AOD-derived $PM_{2.5}$ should be done with comparisons to ground-based monitor data.

How to Ask Questions

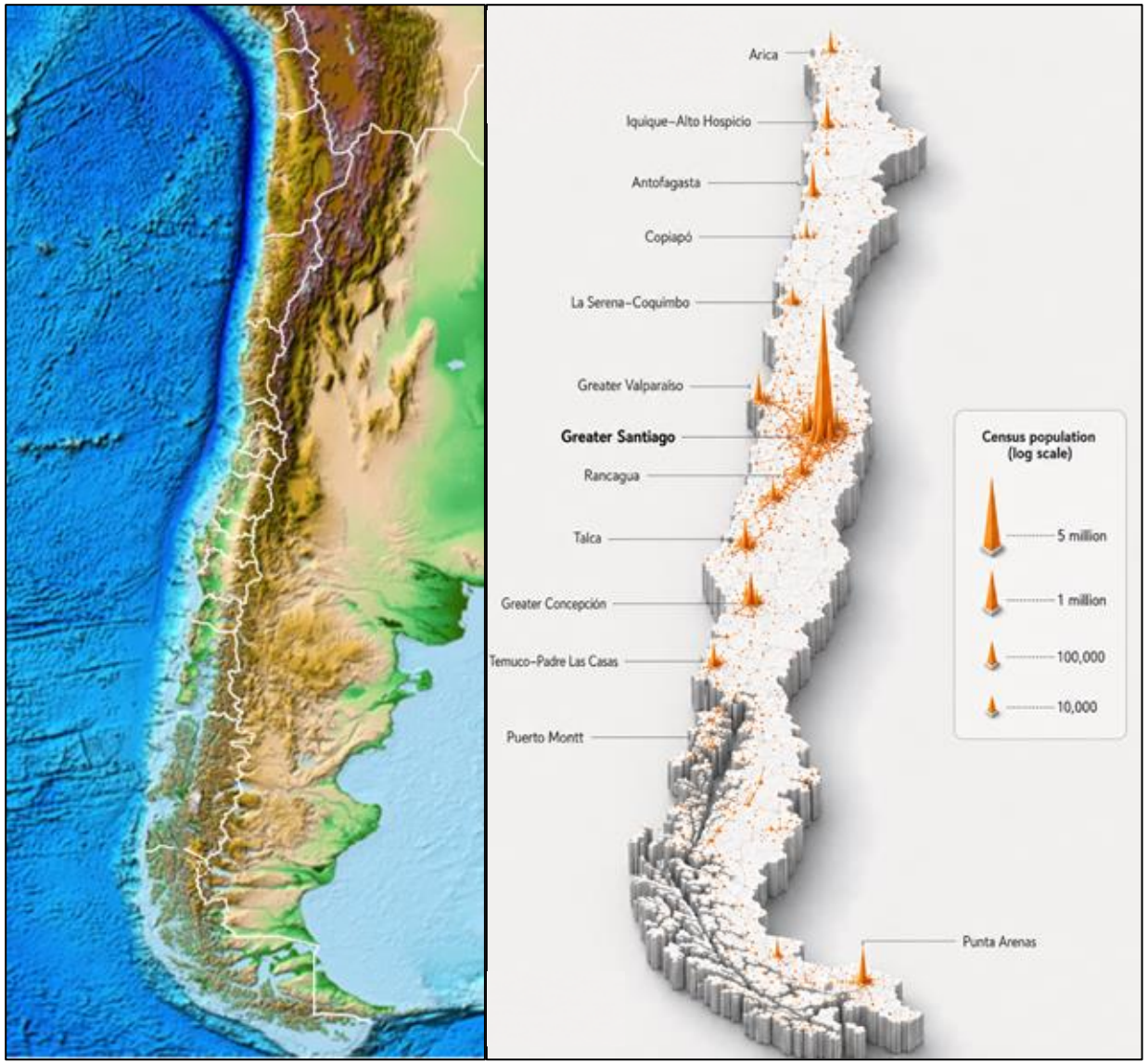
- To ensure we see your question, please locate the three dots ... to the bottom right of your training window.
- You will see a Slido/Q&A option. You may also see a standalone Slido tab or app.
- Please place all questions there and we will answer them during the Q&A session.
- We will try to get to all of the questions during the Q&A session after the webinar. The remainder of the questions will be answered in the Q&A document, which will be posted to the training website about a week after the training.



Part 3:

Introduction to the Case Study:
Effect of the 2023 wildfires on air quality in Chile

Chile is a natural laboratory for $PM_{2.5}$.



Extreme Gradients

- A 4,300 km ribbon between the Andes and the Pacific: desert North, Mediterranean Center, rainy South.

Trapped Basins

- Central valleys hemmed in by two mountain ranges concentrate both people and pollution.

High Spatial Variability

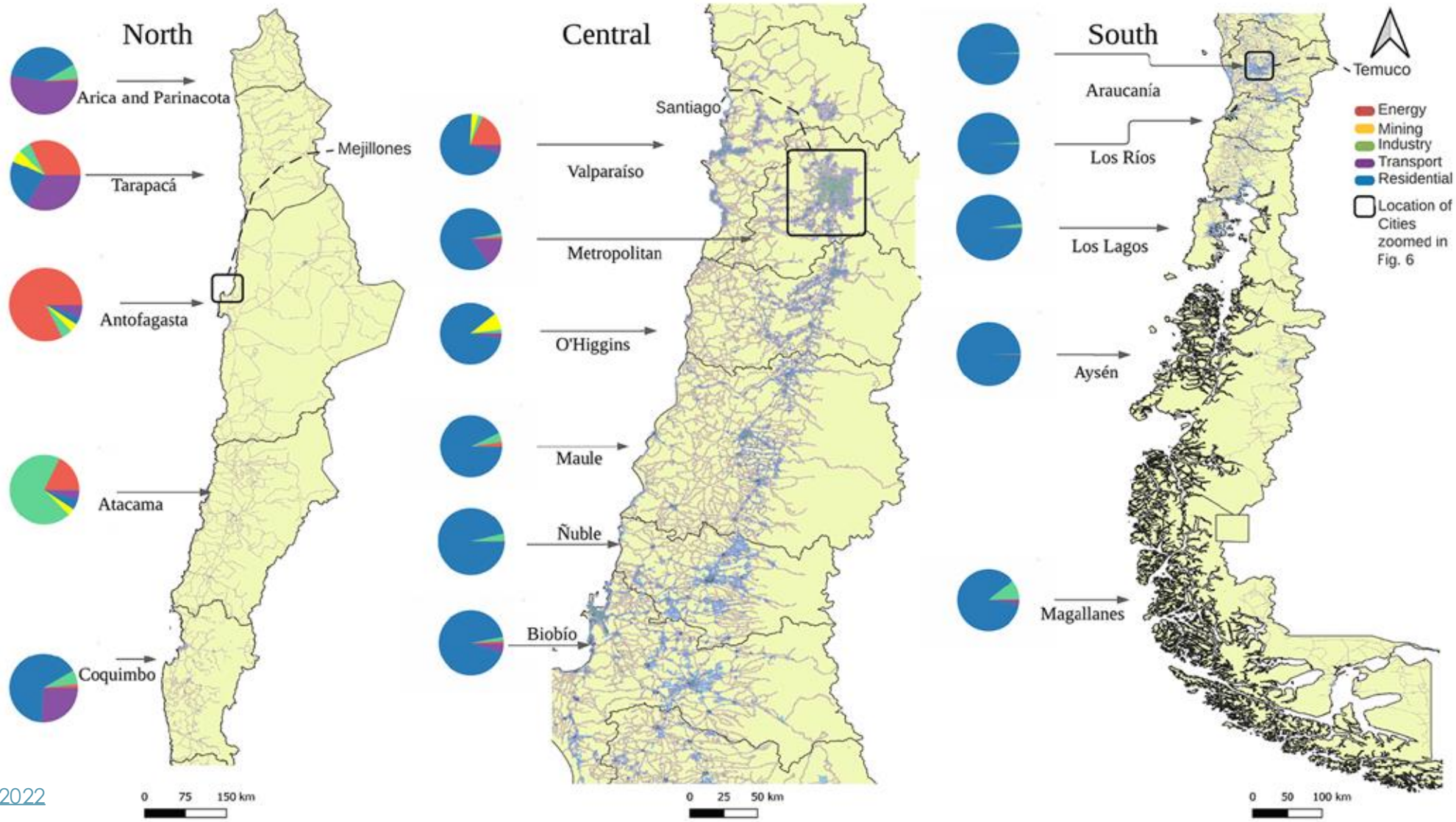
- $PM_{2.5}$ changes sharply with latitude and terrain, exactly what makes satellites useful, and hard to validate.

Sources: Wikipedia (Relief Map of Chile), INE Chile, 2024 Census.

Wood smoke dominates PM emissions in Chile.

- In many Central and Southern cities, **residential** (wood burning) is >80% of fine PM
- Mainly a winter problem

Sources:
Energy
Mining
Industry
Transport
Residential

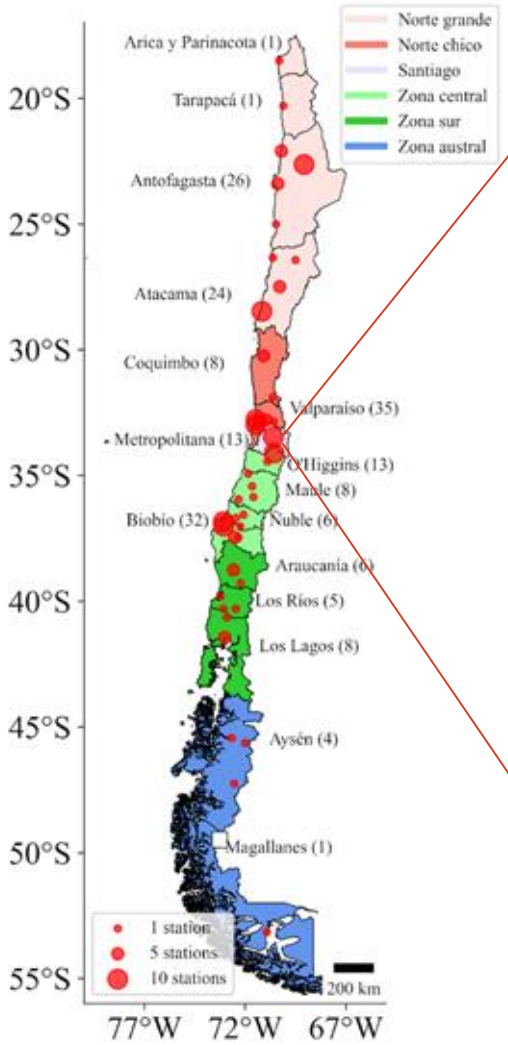


Source: Álamos et al., 2022

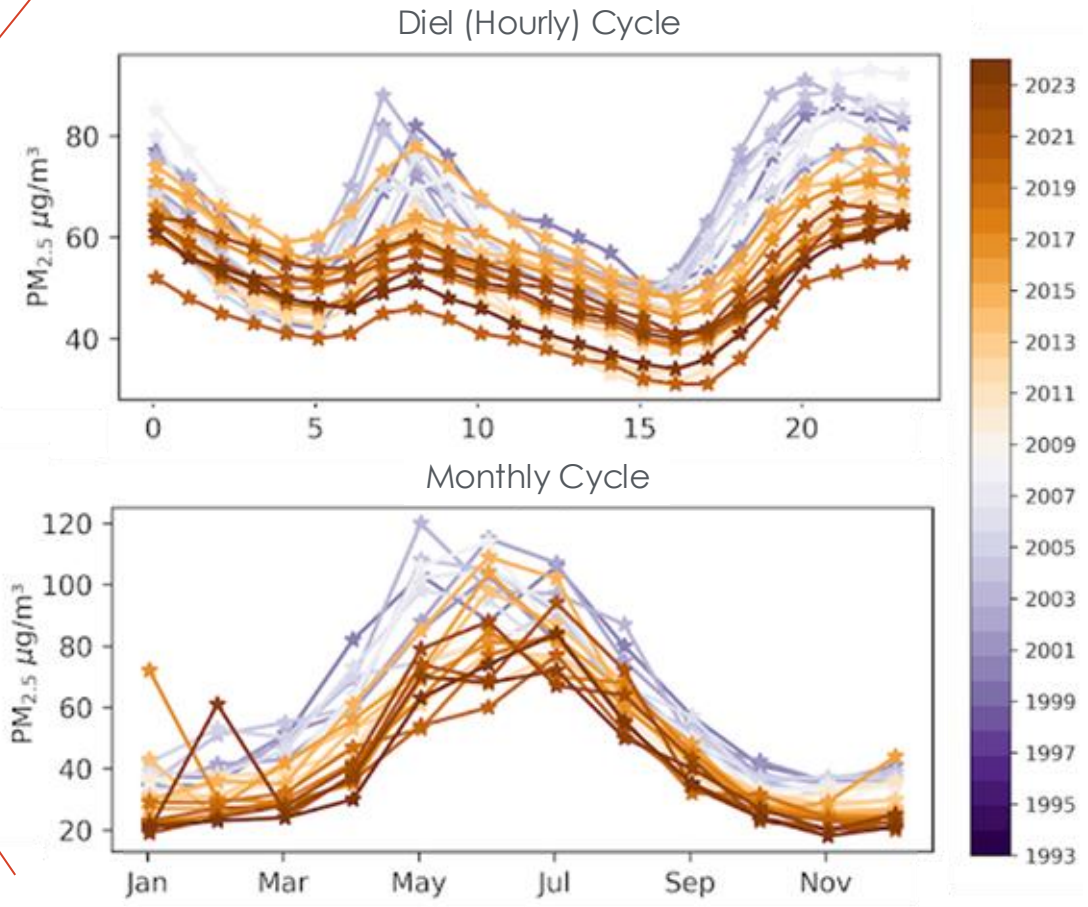
SINCA is Chile's official air quality monitoring network.

- SINCA = Sistema de Información Nacional de Calidad del Aire
- ~200 monitoring stations
- <20% of municipalities are monitored
- ~11M people live in areas without a station

SINCA Monitoring Stations



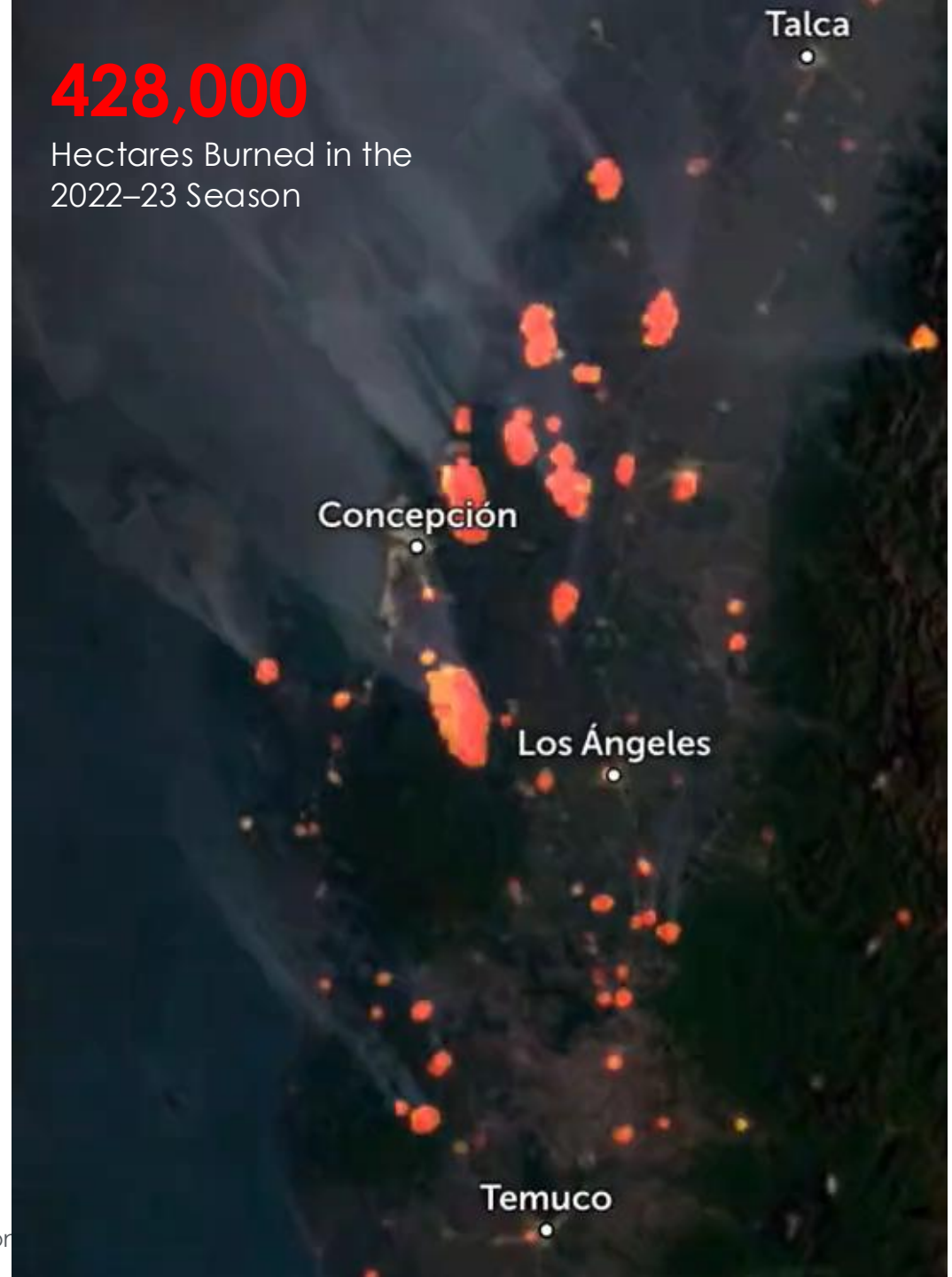
$PM_{2.5}$ Diel & Monthly Profiles in Santiago from SINCA



Sources: SINCA, UDD, Basoa et al., 2025

Case Study Event: February 2023 Fires.

- One of the worst fire seasons on record: ~428,000 ha burned, +244% vs. the previous season.
- Heavy human toll: 26 people killed and ~2,450 homes destroyed.
- A megadrought backdrop: over a decade of drought and hotter, drier summers left the landscape primed to burn.
- Concentrated in one week: early February 2023, across Maule, Ñuble, Biobío and La Araucanía regions.



Sources: Zoom Earth, NOAA/NESDIS/STAR/GOES-East, NASA FIRMS

Factors Leading to the February 2023 Fires.

Heat

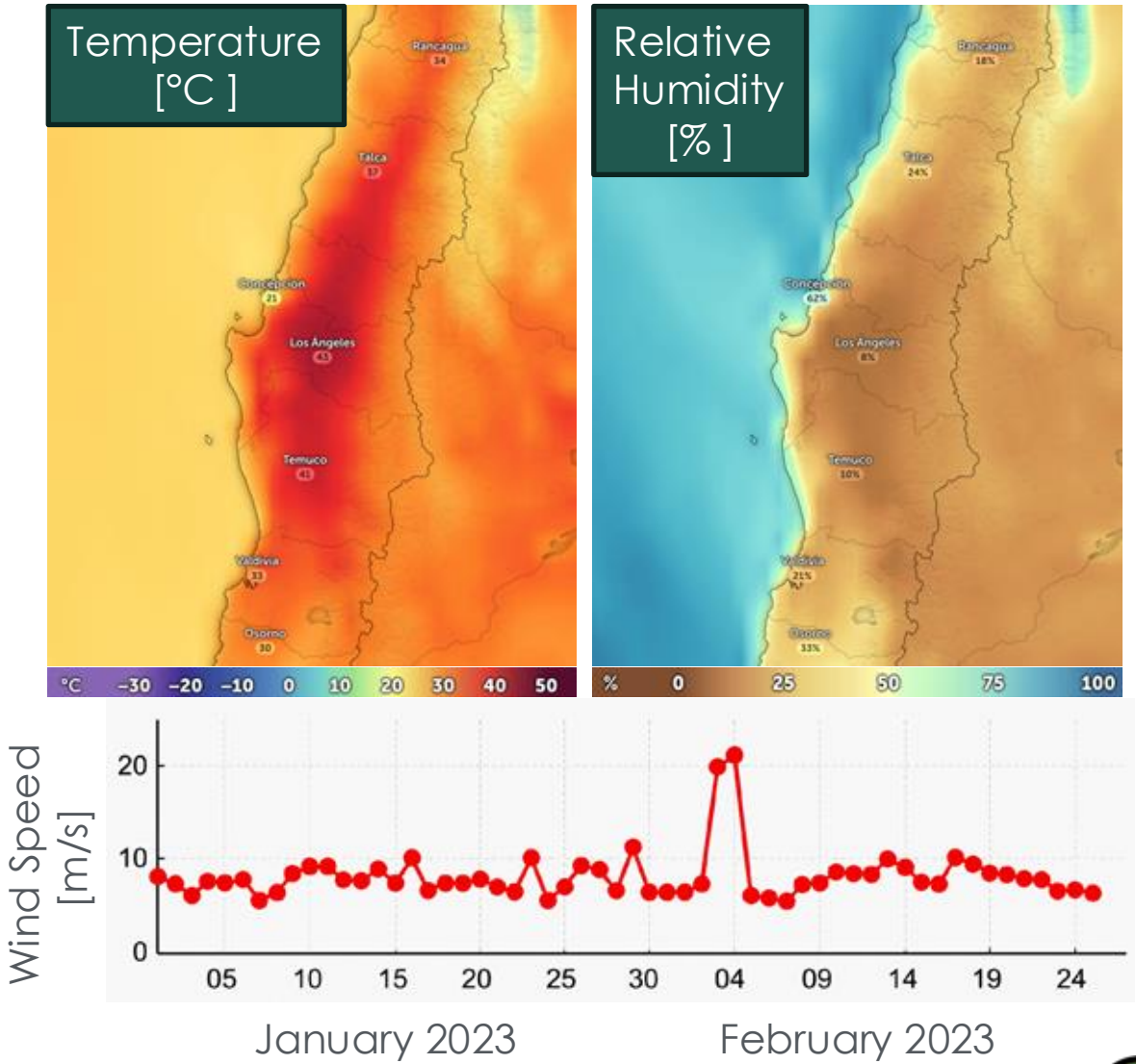
- Chillán hit a record 41.6°C & RH fell below 10% on 3 February 2023.

Wind

- The hot, dry "Puelche" fanned the flames and pushed smoke hundreds of km.

Fuel

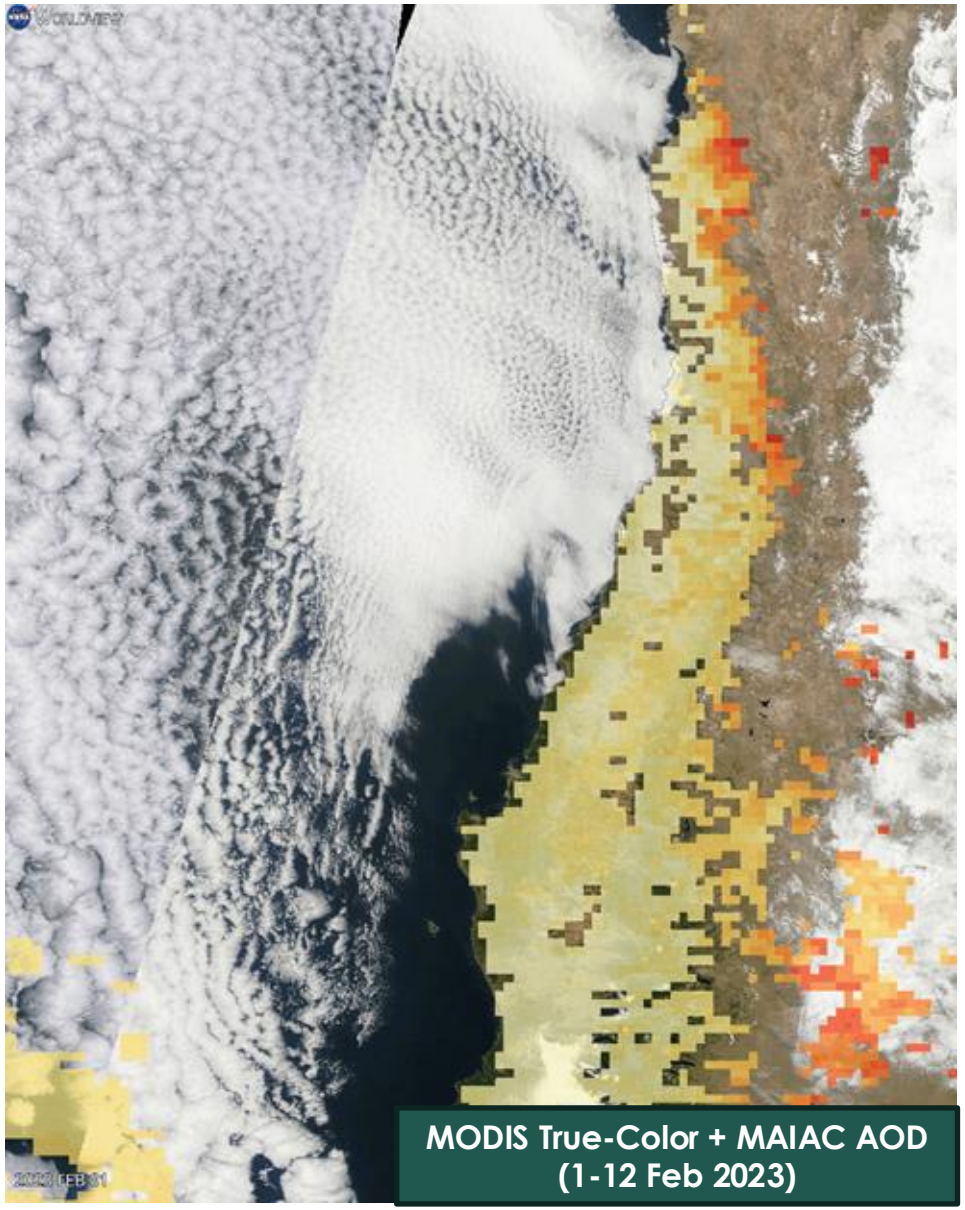
- Dense pine and eucalyptus plantations spread fire unchecked.



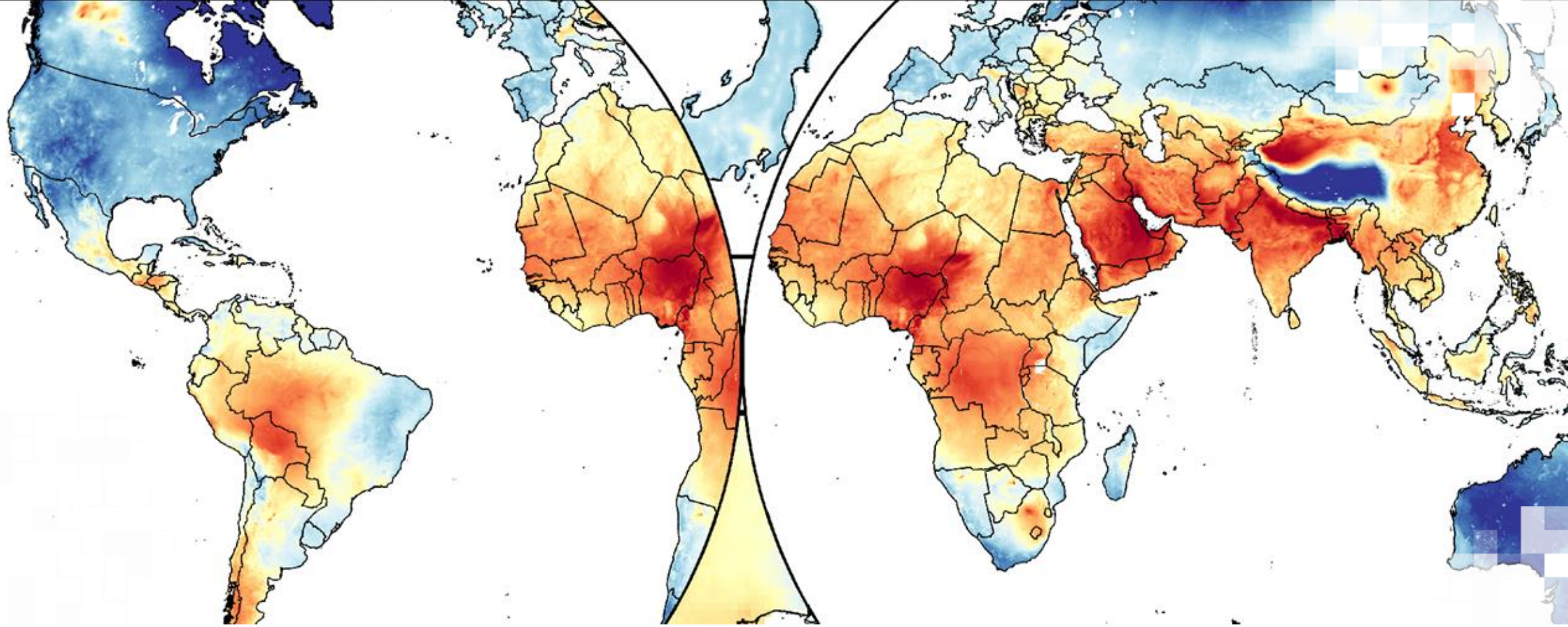
Sources: Zoom Earth, CR2

February 2023 Fire Smoke as Seen from Space with Satellite AOD.

- **Satellites Capture the Full Plume:** Its regional reach over the Pacific, right where no ground station sits.
- AOD shows smoke, not how much $PM_{2.5}$ people breathe at the surface.
- $PM_{2.5}$ satellite derived products are only useful if we trust them.
- **To verify, we should compare against reference ground data.**



Source: NASA Worldview



Part 3:

Case Study: Comparing SatPM, MERRA-2 CNN, and SINCA Data

Case Study Outline

- Setup
  - Access Python Notebook from ARSET Github
  - Download a copy to your Google Drive
  - Set up folders for downloaded data in your Google Drive
  - Orientation to Google Colab controls
  - Import required packages
  - Enable Google Drive access
  - Provide NASA Earthdata and OpenAQ credentials
  - Define useful functions

Case Study Outline (Continued)

- Download Data
  - Download SatPM data, subset to region of interest, apply basic quality control, and plot
  - Download MERRA-2 CNN data, subset to region of interest, apply basic quality control, and plot
  - Download ground monitor data from OpenAQ and SINCA network, apply basic quality control, plot, and compare
- Analyze Data
  - Compare SatPM, MERRA-2 CNN, and ground monitor data
  - Assess using common metrics: bias, accuracy, correlation
  - Assess using different time bases: monthly, daily, and hourly

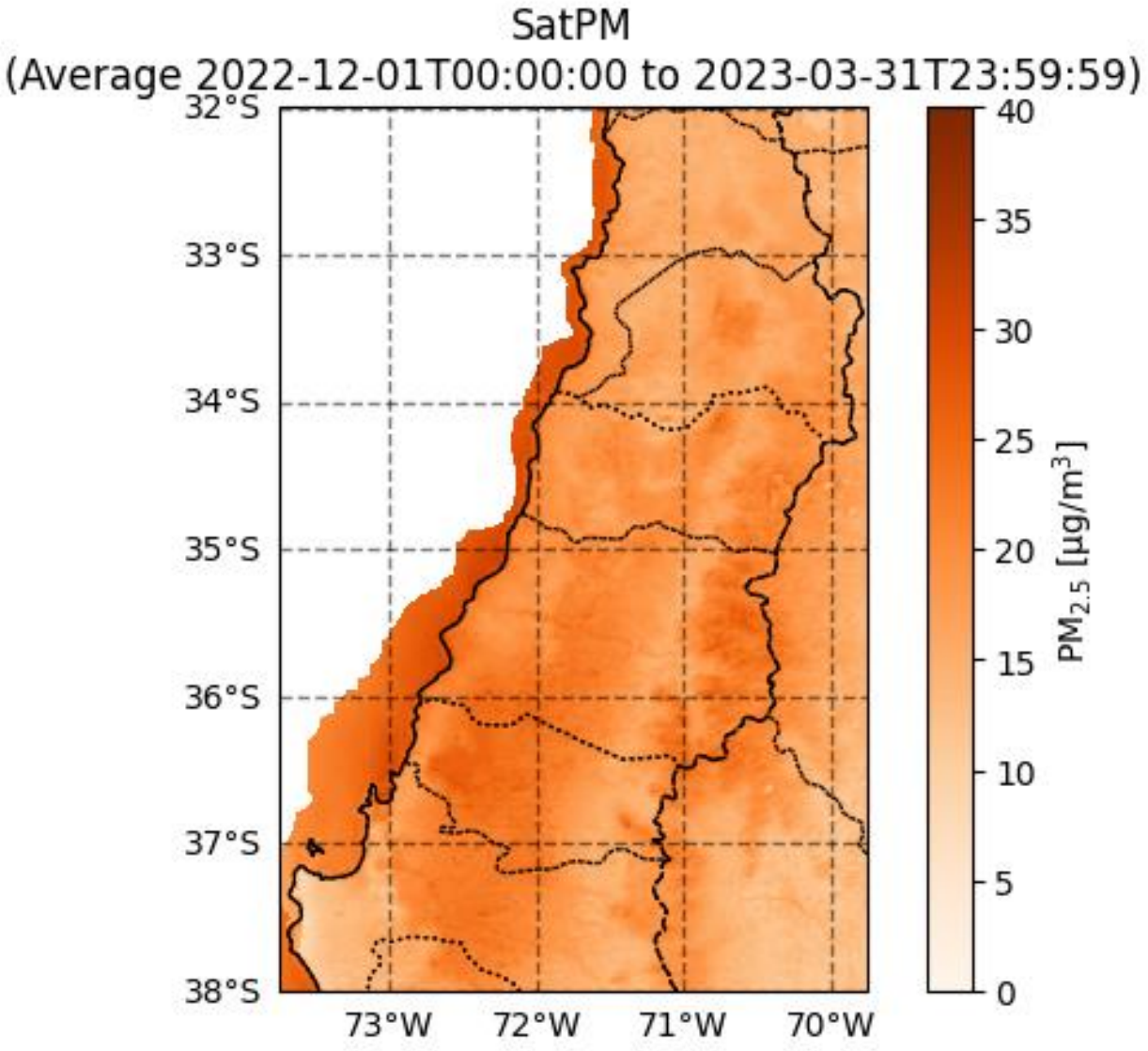
Common Metrics

- **Bias:** Measures the average difference between the estimates and the true values
  - Range: -∞ to +∞
  - Best Value: 0
  - Same units as measured quantities
- **Mean Normalized Bias (MNB):** Bias divided by the average of the true values
  - Range: -∞ to +∞
  - Best Value: 0
  - Unitless
- **Root Mean Square Error (RMSE):** Measures the accuracy of the estimate compared to the true value, as the square-root of the average of the squared errors between the estimates and true values
  - Range: 0 to +∞
  - Best Value: 0
  - Same units as measured quantities
- **Mean Normalized RMSE (NRMSE):** RMSE divided by the average of the true values
  - Range: 0 to +∞
  - Best Value: 0
  - Unitless

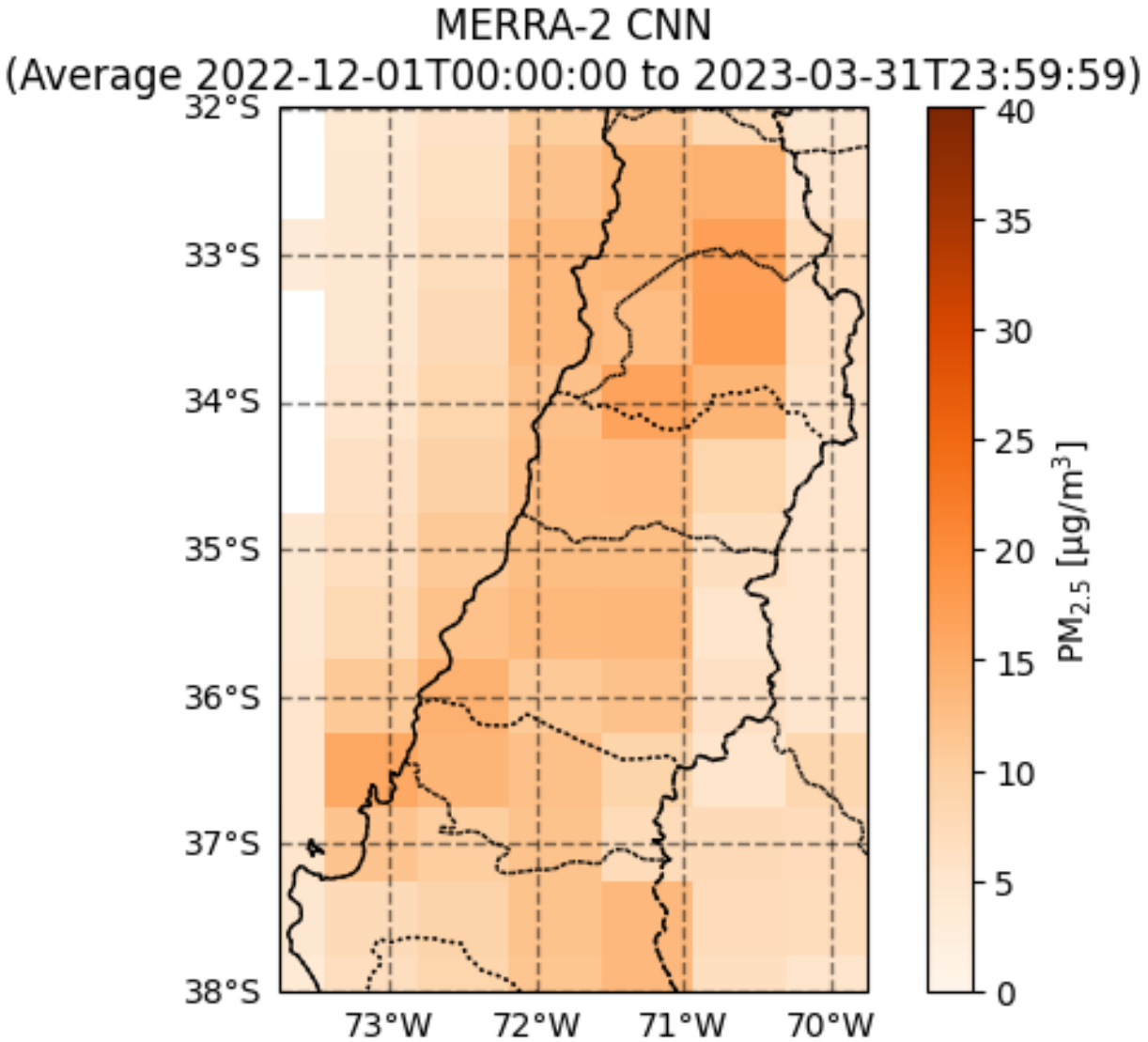
Common Metrics (Continued)

- **Coefficient of Determination (R^2):** Measures the fraction of variance in the true values explained by variance in the estimates
  - Range: -∞ to 1
  - Best Value: 1
  - Unitless
- **Spatial Correlation (r_{space}):** Measures the linearity in the relationship between the spatial pattern of the estimates and the spatial pattern of the true values
  - Range: -1 to 1
  - Best Value: 1
  - Unitless
- **Temporal Correlation (r_{time}):** Measures the linearity in the relationship between the temporal pattern of the estimates and the temporal pattern of the true values
  - Range: -1 to 1
  - Best Value: 1
  - Unitless

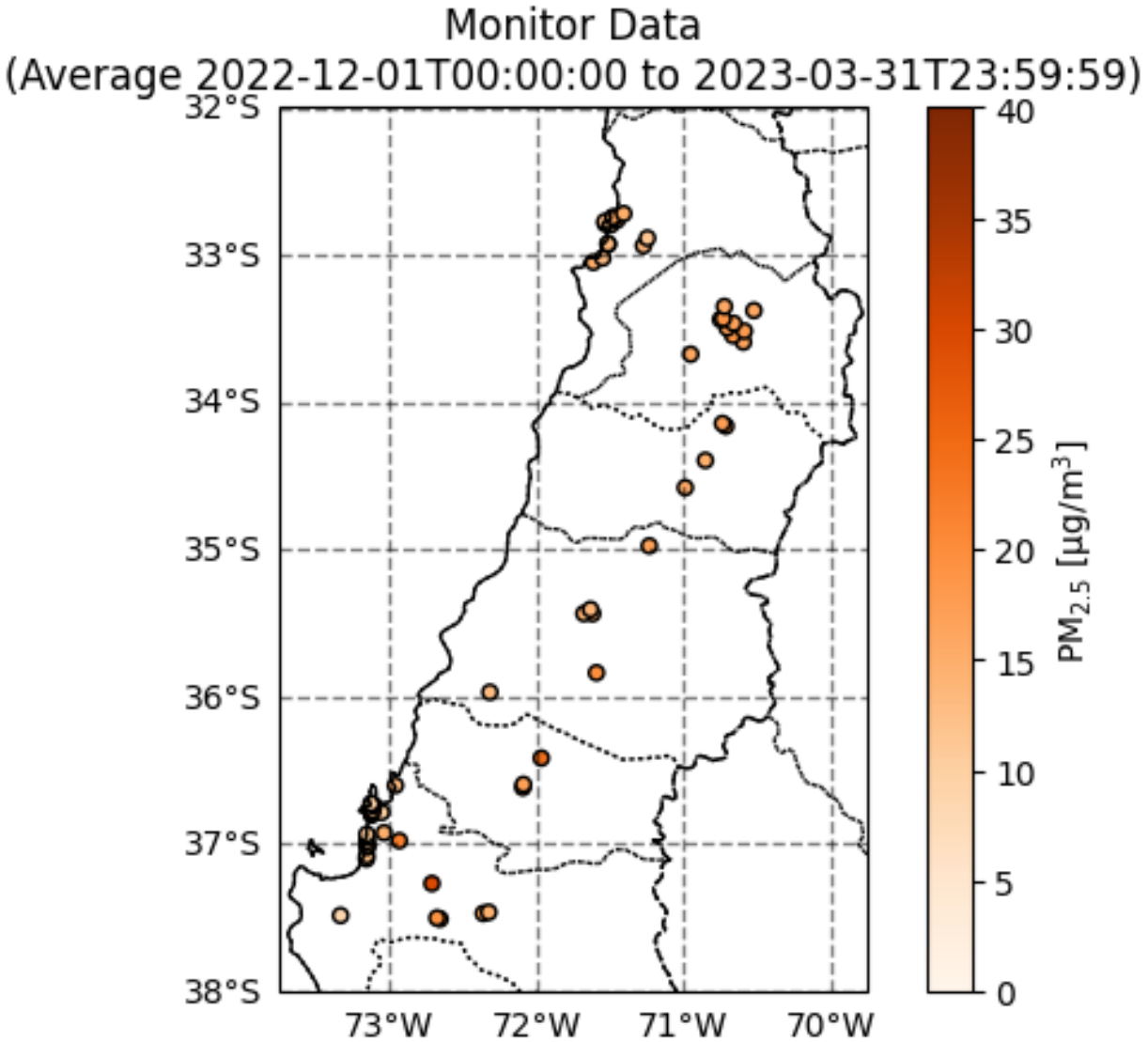
Download and Plot SatPM$_{2.5}$ Data for the Region of Interest



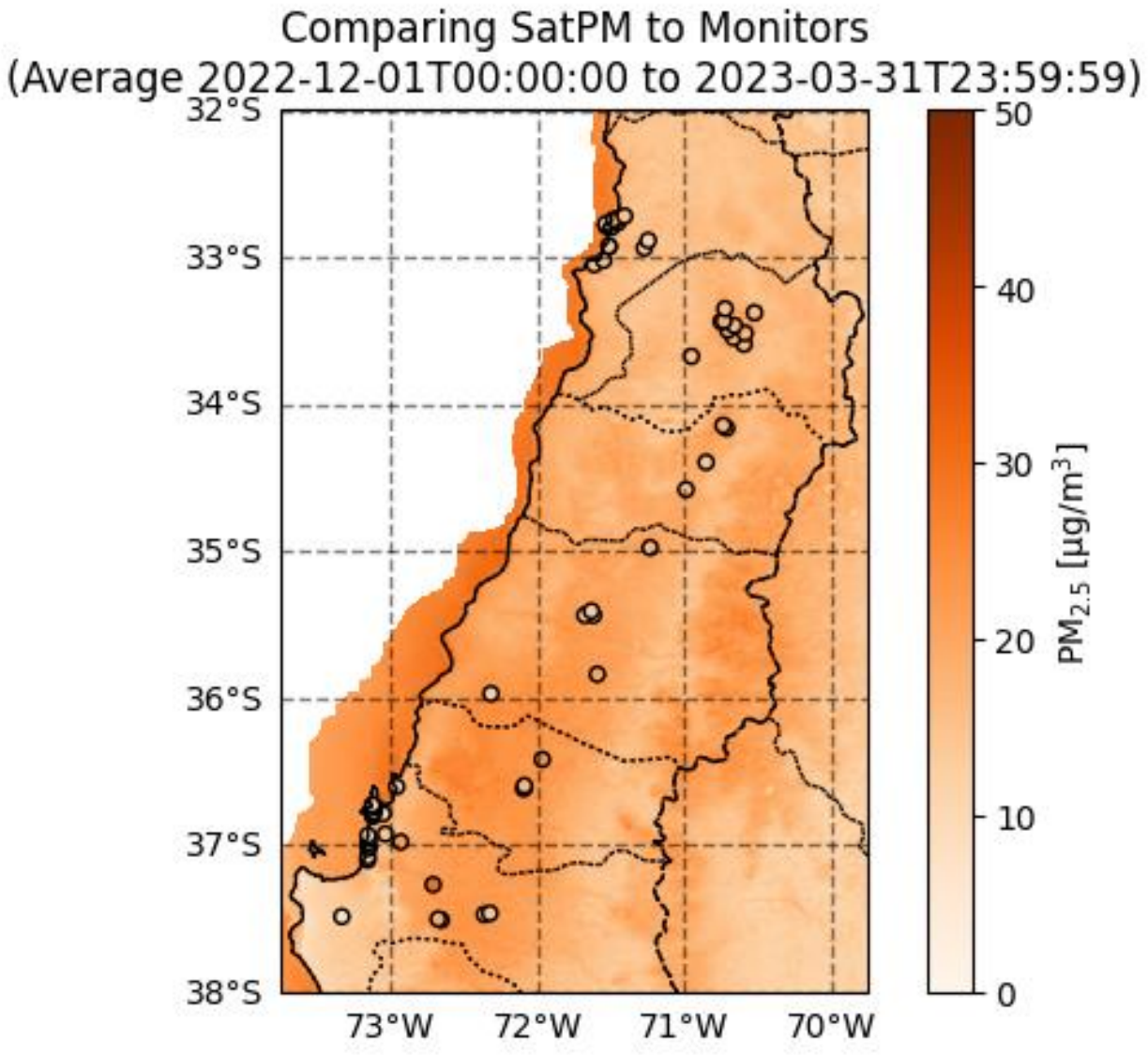
Download and Plot MERRA-2 CNN Data for the Region of Interest

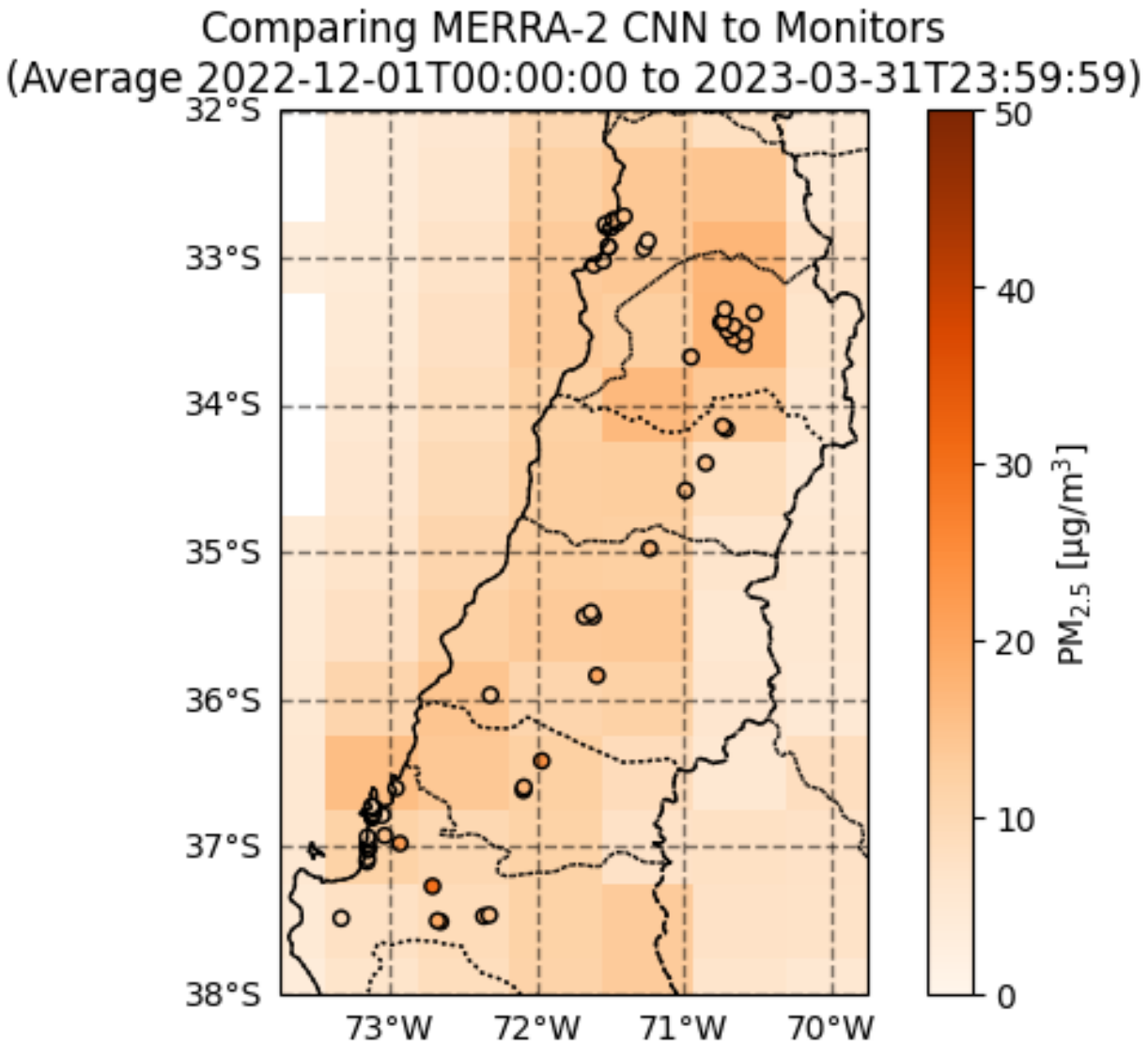


Download and Plot SINCA Monitor Data for the Region of Interest

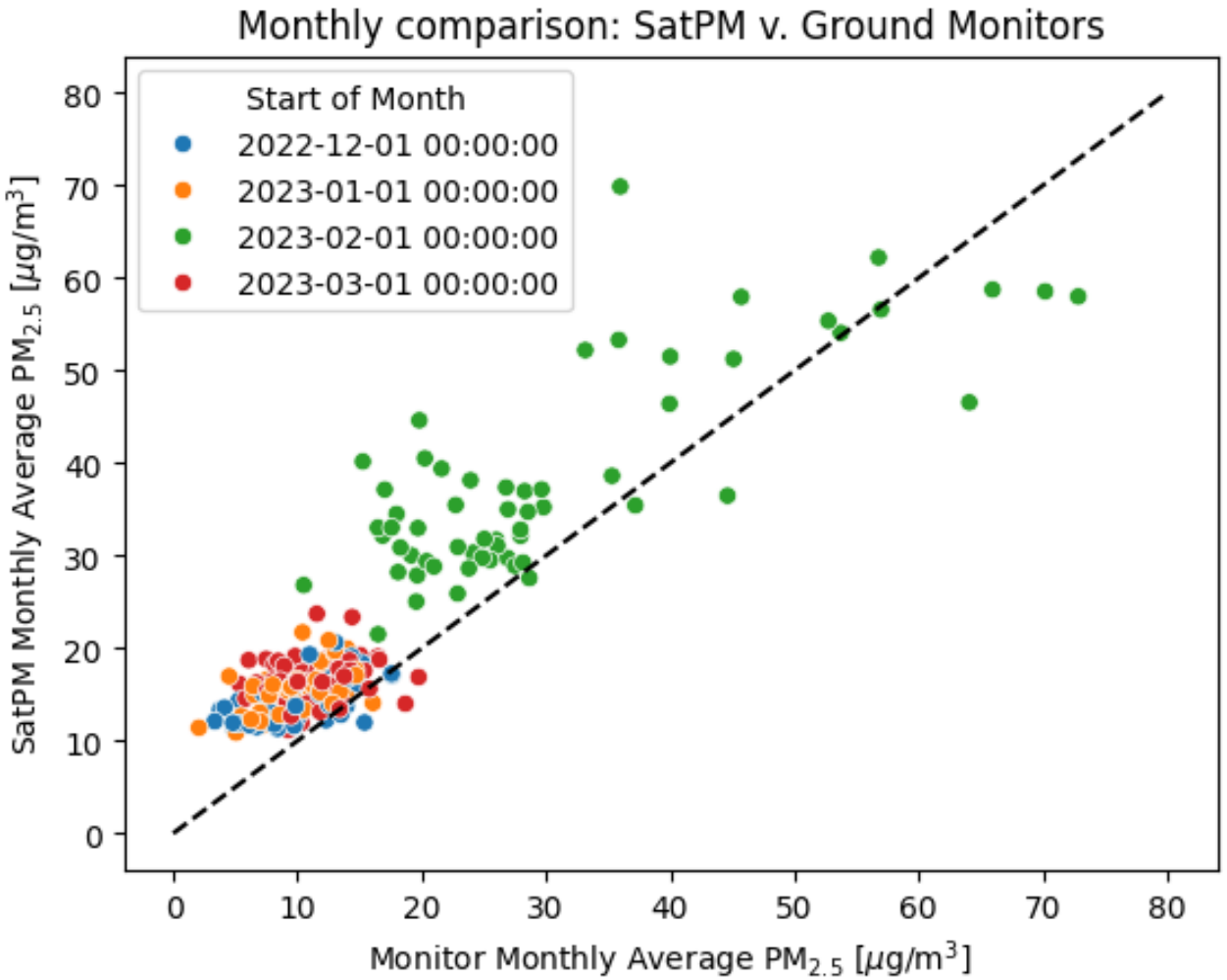


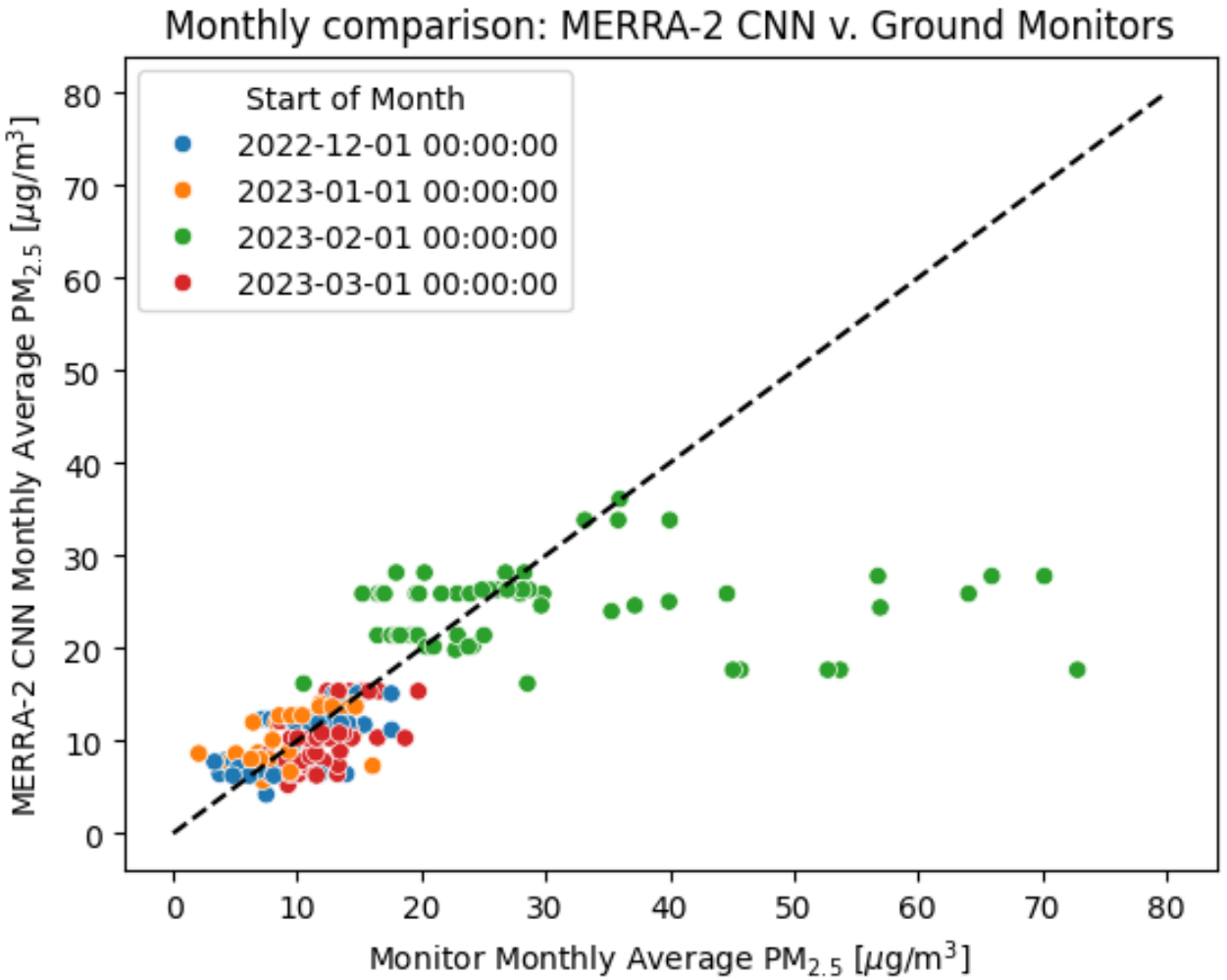
Compare Datasets – Plot Maps



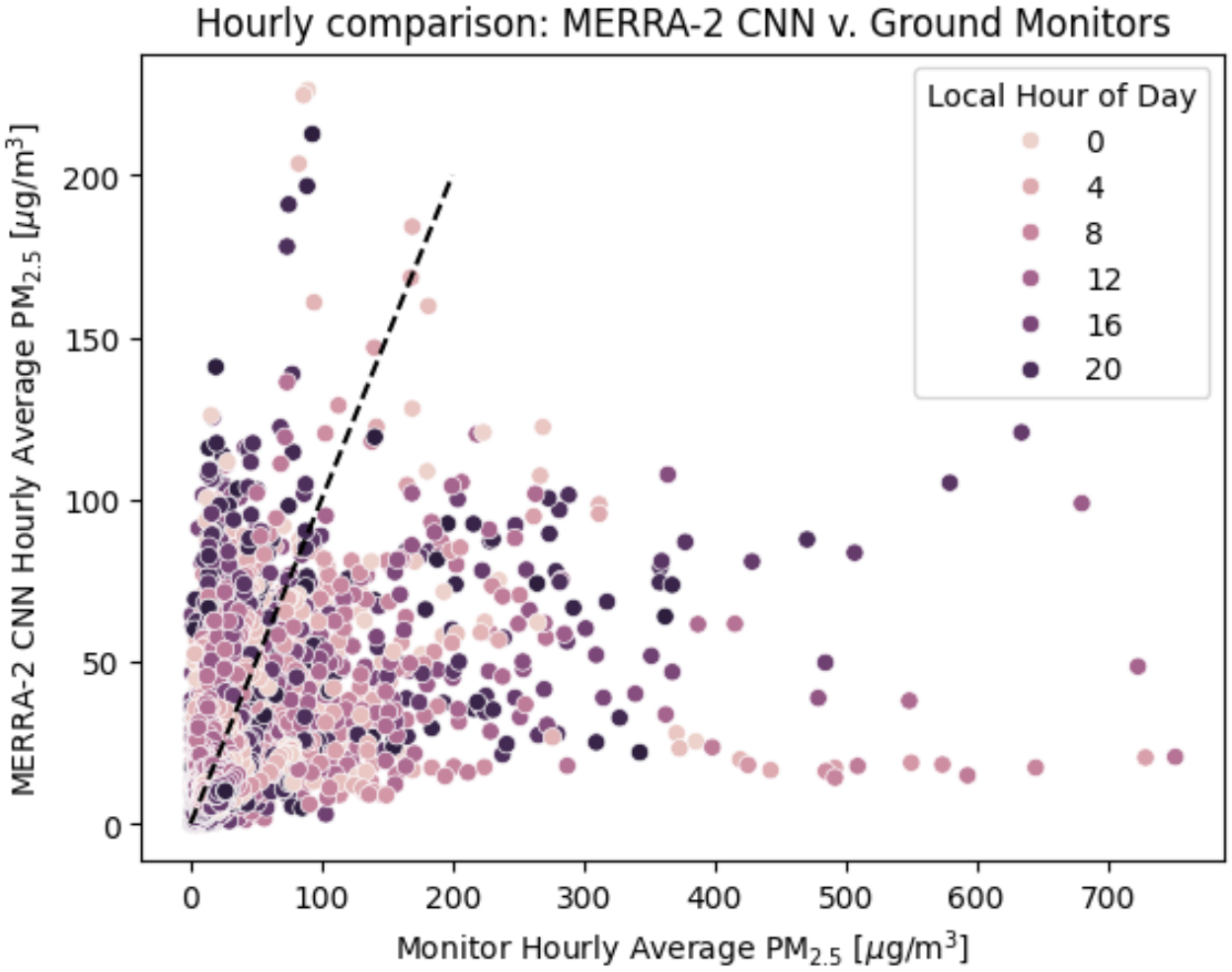


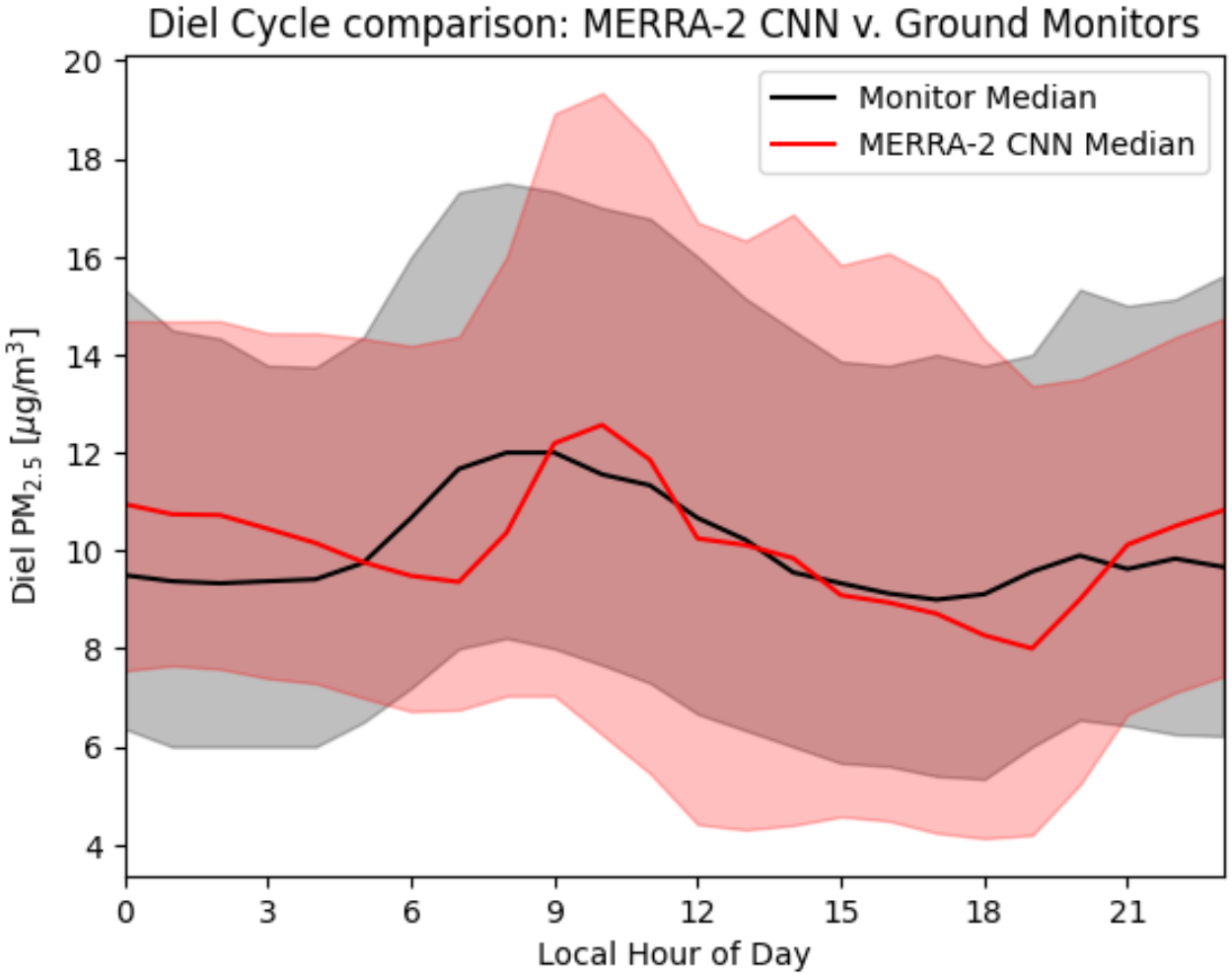
Compare Datasets – Monthly Scatterplots



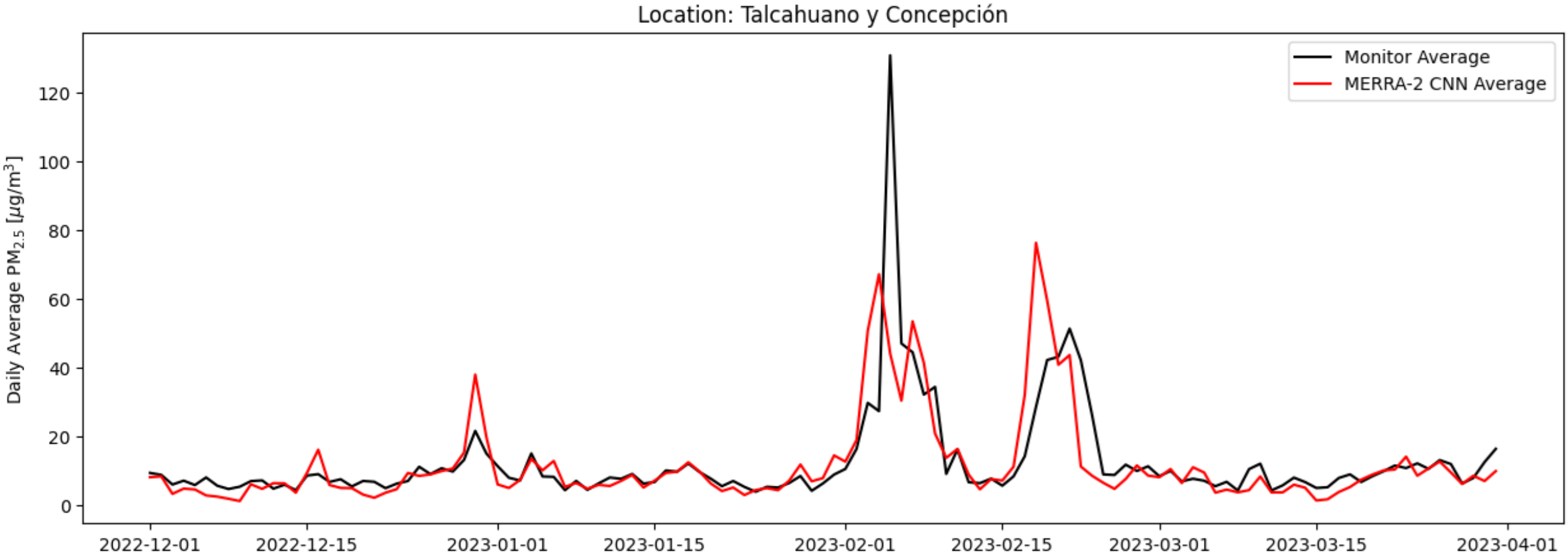


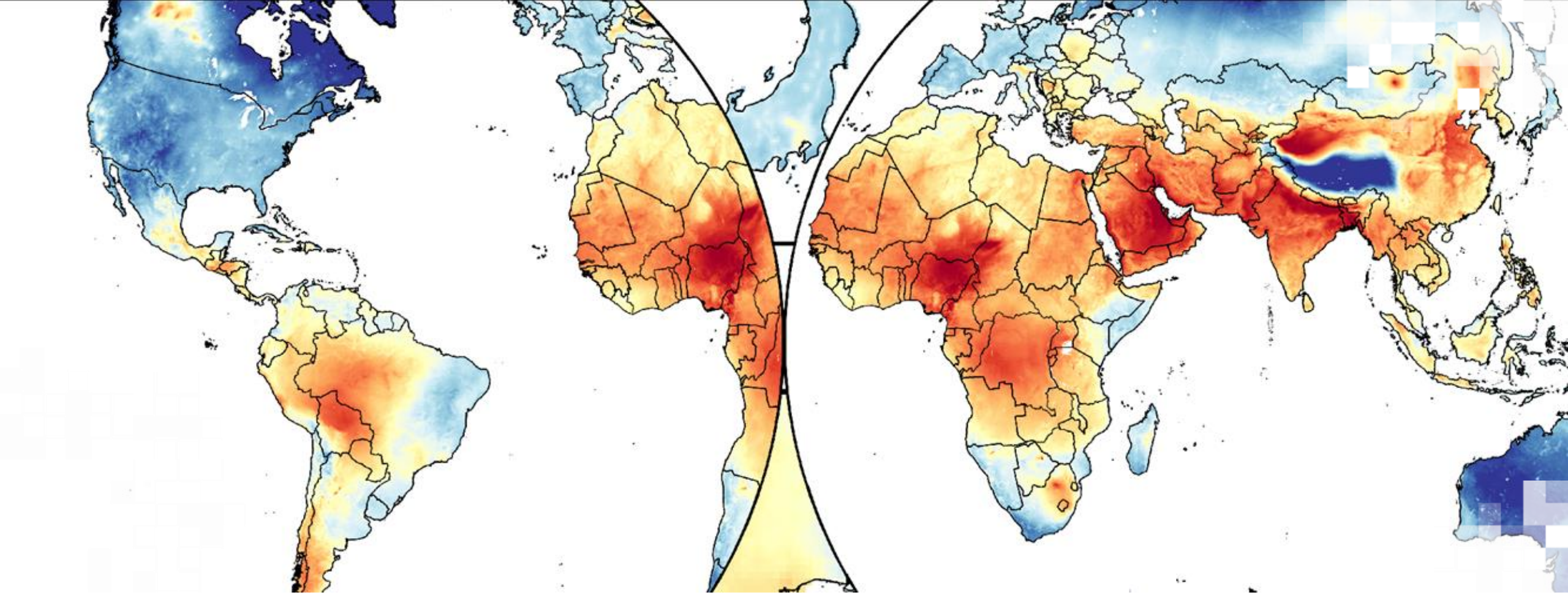
MERRA-2 CNN – Hourly and Time-of-Day Comparisons





MERRA-2 CNN – Daily Time Series Plot





Estimating Surface $PM_{2.5}$ using Satellite Data and Other Information Sources

Summary

Training Summary

- Satellite AOD data can be combined with surface $PM_{2.5}$ measurements and CTM outputs to estimate surface $PM_{2.5}$ in locations without ground-based monitors.
- Various approaches (linear regression, geographically weighted regression, machine learning techniques, and hybrid methods) can be applied to estimate $PM_{2.5}$ from these inputs.
- The Sat$PM_{2.5}$ dataset and Bias-Corrected MERRA-2 $PM_{2.5}$ dataset are two publicly-available $PM_{2.5}$ datasets with global coverage.
  - **Sat$PM_{2.5}$** provides higher spatial (**~1 km**) but lower temporal (**monthly**) resolution
  - **MERRA-2 CNN** provides lower spatial (**~50 km**) but higher temporal (**hourly**) resolution.
- Datasets should be compared to ground-based $PM_{2.5}$ data where possible to evaluate their suitability for regional applications.
  - Bias
  - Accuracy (root mean square error)
  - Correlation (overall, spatial, temporal)

Homework and Certificates

- **Homework:**
  - One homework assignment
  - Opens **Today**
  - Access from the training webpage
  - Answers must be submitted via Google Forms
  - **Due by August 5, 2026**

- **Certificate of Completion:**
  - Attend all three live webinars (attendance is recorded automatically)
  - Complete the homework assignment by the deadline
  - You will receive a certificate via email approximately two months after completion of the course.

Contact Information

Trainers:

- Carl Malings
  - carl.a.malings@nasa.gov
- Sebastián Diez
  - sebastian.diez@udd.cl

- ARSET Website
- ARSET YouTube

For questions, comments, or to share how you have applied our trainings to your work or studies, email nasa.arset@gmail.com.

Join our quarterly newsletter to stay up-to-date on our latest trainings:

1. Send an email with no subject line to arset-join@lists.nasa.gov.
2. Follow the instructions sent in response.

Resources

- Github page for the case study exercise presented in this part of the training

References

- Álamos et al., 2022, "High-resolution inventory of atmospheric emissions from transport, industrial, energy, mining and residential activities in Chile", *Earth System Science Data*. doi: 10.5194/essd-14-361-2022
- Basoa et al., 2025, "Current Status, Trends, and Future Directions in Chilean Air Quality: A Data-Driven Perspective", *Atmosphere*. doi: 10.3390/atmos16060733


APPLIED REMOTE SENSING TRAINING PROGRAM
TRAIN - EMPOWER - ADVANCE

Thank You!