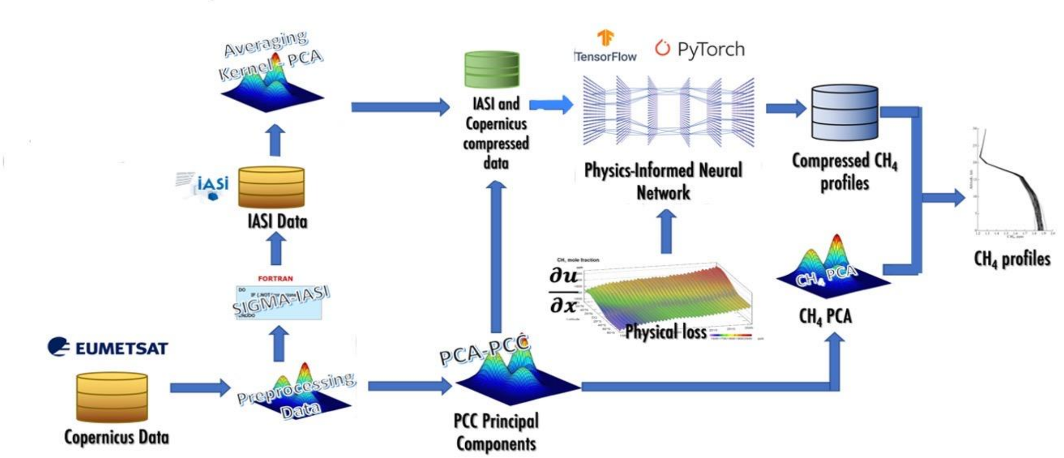
Retrieval of Vertical Profiles of Methane from IASI data using a Physically Informed Neural Network

Guido Masiello, Rocco Giosa, Giuliano Liuzzi, Carmine Serio, Marco D'Emilio, Pamela Pasquariello, Lorenzo Cassini

Finanziato dall'Unione europea
NextGenerationEU
Ministero dell'Università e della Ricerca
Italiadomani
PIANO NAZIONALE DI RIPRESA E RESILIENZA
UNIVERSITÀ DEGLI STUDI DELLA BASILICATA · 1982

Context

- Scientific gap: To date, there are no products in the literature that address the vertical profile of methane (CH_4).
- Innovative proposal: As part of the PRIN-MVP project, a Physically-Informed Neural Network (PINN) was developed ⇒ Retrieval of CH_4 vertical profiles.
- Main results submitted to GRL



Outline

- The Methodology
  - Physical Informed Neural Network
  - The Forward Model
- Applications
  - Simulation
    - Po Valley
  - IASI data
    - North Stream Sabotage
    - Saharan Dust Intrusion
- What Next

PINN Data

Atmospheric state vector

- The atmospheric state vector profiles were downloaded from the Copernicus database (CAMS-ERA5).
- Atmospheric profiles were interpolated into the 60-level pressure grid used by the **σ-IASI-F2N** radiative transfer code;

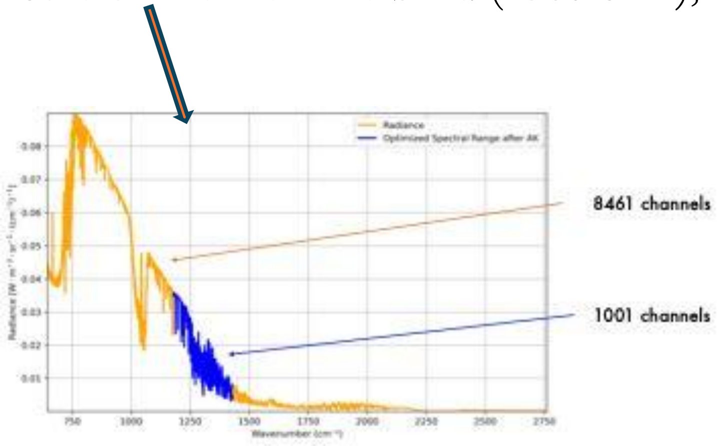
Index	Edge [hPa]	Level [hPa]
0	1013.0	
1	1004.99	1009.0
2	1000.0	1002.5
3	985.76	992.9
4	972.71	979.2
.	.	.
.	.	.
.	.	.
57	1.0	1.5
58	0.5	0.8
59	0.1	0.3
60	0.005	0.1

σ-IASI pressure grid

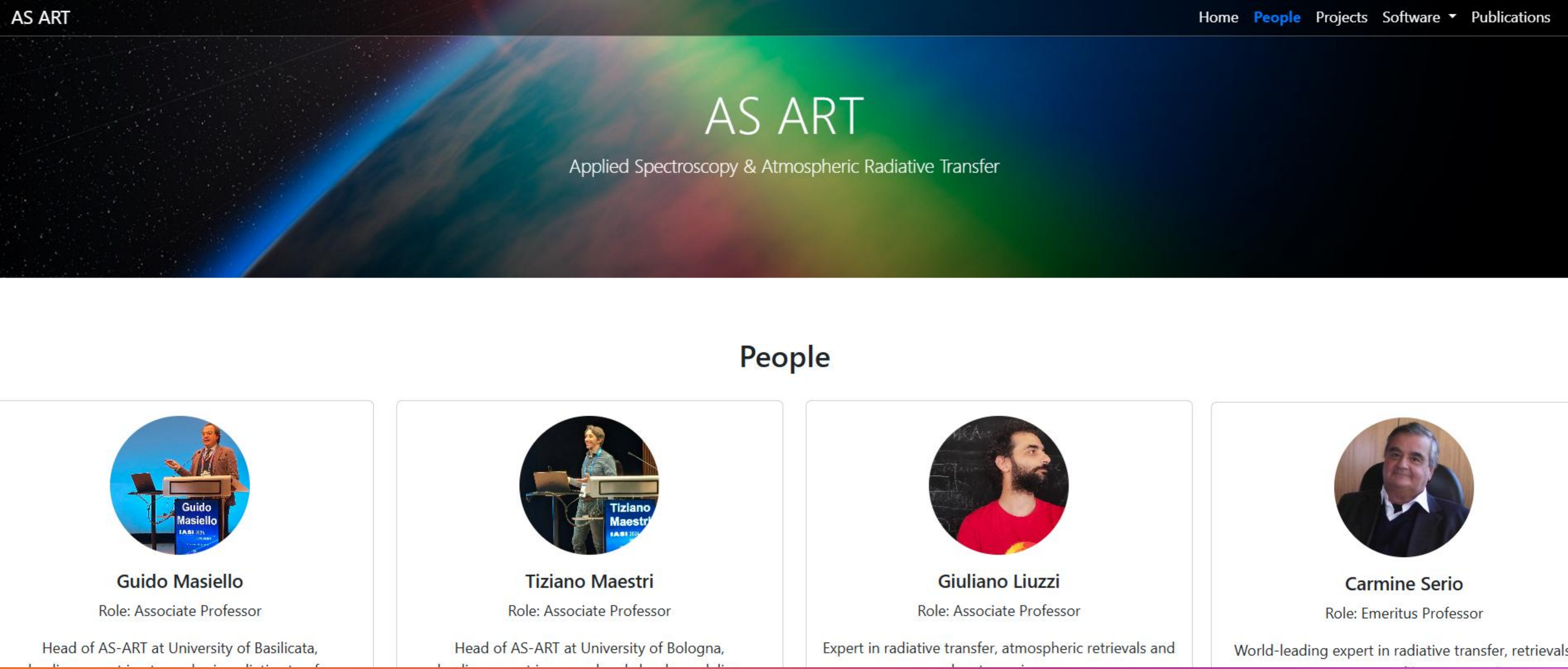
- Vertical profiles of temperature and methane were compressed using PCA, obtaining 16 PCs for the temperature profiles, and 24 PCs for the methane profiles.

Spectra

- Radiance simulation with **σ-IASI-F2N** radiative transfer code: 8461 channels;
- After performing the Averaging Kernels study, the channels are reduced to 1001 (from 1180.00 to cm^{-1} 1430.00 cm^{-1}), around the core of the methane band (1305 cm^{-1});



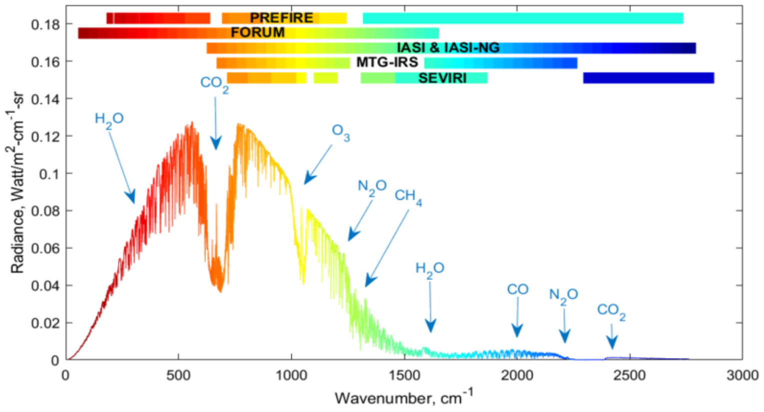
- To further reduce dimensionality and noise, the optimised spectral range was compressed through PCA, into 20 PCs.



AS ART Group

https://www.as-art.it/

σ-IASI/F2N



The RTM developed in the framework of EUMETSAT programs

- Assessment of IASI data for the Atmosphere (1996-2004) grants
  - EUM/CO/96/407/DD, EUM/CO/99/688/DD, EUM/CO/02/1053/PS
- Italian Space Agency ASI programs (2019-Now)
  - FORUM-Scienza Program of Italian Space Agency (Contract No. 2019-20-HH.0, P.I. CNR-INO)
  - FIT-FORUM (contract n. 2023-23-HH.0, CUP F33C23000240005, P.I. DIFA, University of Bologna),
  - MC-FORUM (contract n. 2023-23-HH.0, CUP F93C2300046000, P.I. IBE-CNR)
- Pseudo-monochromatic (0.01 cm^{-1})
- 5-3000 cm^{-1} spectral range
- OD databases (parametrized in T and ρ)
- Clouds and aerosols properties (parametrized in r_{eff})
- In presence of scattering layers, the code accounts for a Chou+Tang solution
- Analytical Derivatives in clear and clouds
- Masiello, G., Serio, C., Liuzzi, G., Venafra, S., Maestri, T., Martinazzo, M., Amato, U., & Grieco, G. (2023). σ-IASI (2.4). Zenodo. https://doi.org/10.5281/zenodo.8152674

PINN model construction

Step 1: AutoModel (Autokeras) → identifies the best architecture (minimizing MAE).

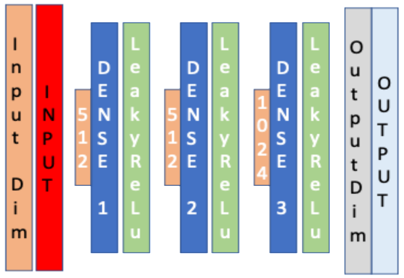
Step 2: Fine Tuning with Physical Loss → introduces a χ^2-term

- Weights differences based on uncertainty;
- Takes correlations into account.

PINN

Autokeras's architecture

Input tensor	#
Spectra PC CH_4	20
Temperature PC	16
Latitude	1
Longitude	1
Starting layer	1
Time	1
VZA	1
Total	41



Output tensor	#
Profile PC CH_4	24
Total	24

Neural network architecture obtained using the Autokeras framework

+

Physical loss

$$N = \textit{number of samples}$$

$$S_a = \text{prior } CH_4 \text{ covariance matrix}$$

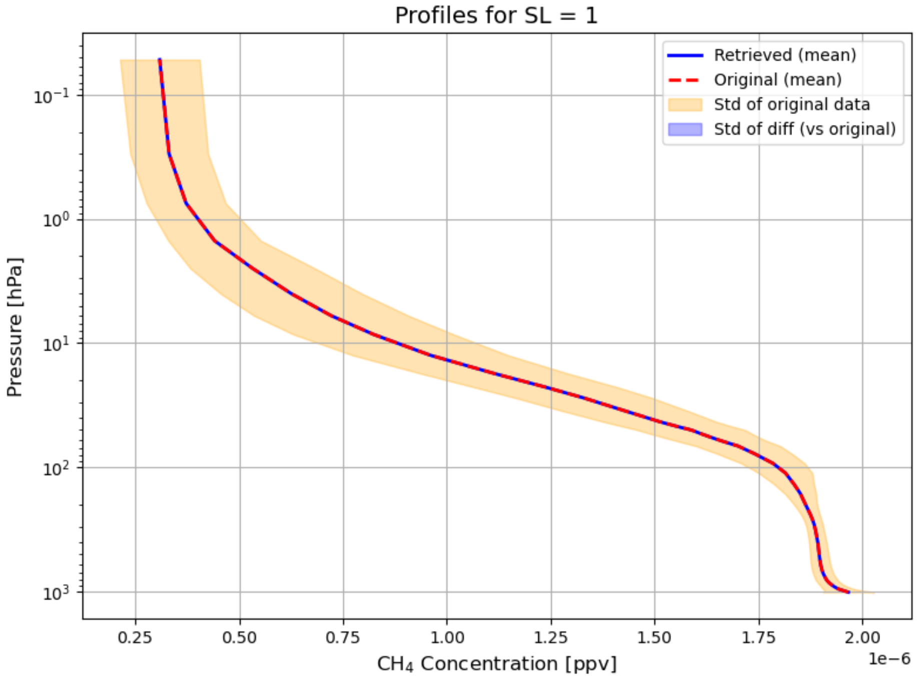
$$\textit{difference} = y_{\text{prior}}(i) - y_{\text{estimate}}(i)$$

$$\textit{loss}(i) = \boxed{\textit{difference}^{T} * S_a^{-1} * \textit{difference}} \leftarrow \chi^2$$

$$\text{Total}_{Loss} = \frac{1}{N}\sum_{i=1}^{N} \text{loss}_i$$

PINN model evaluation

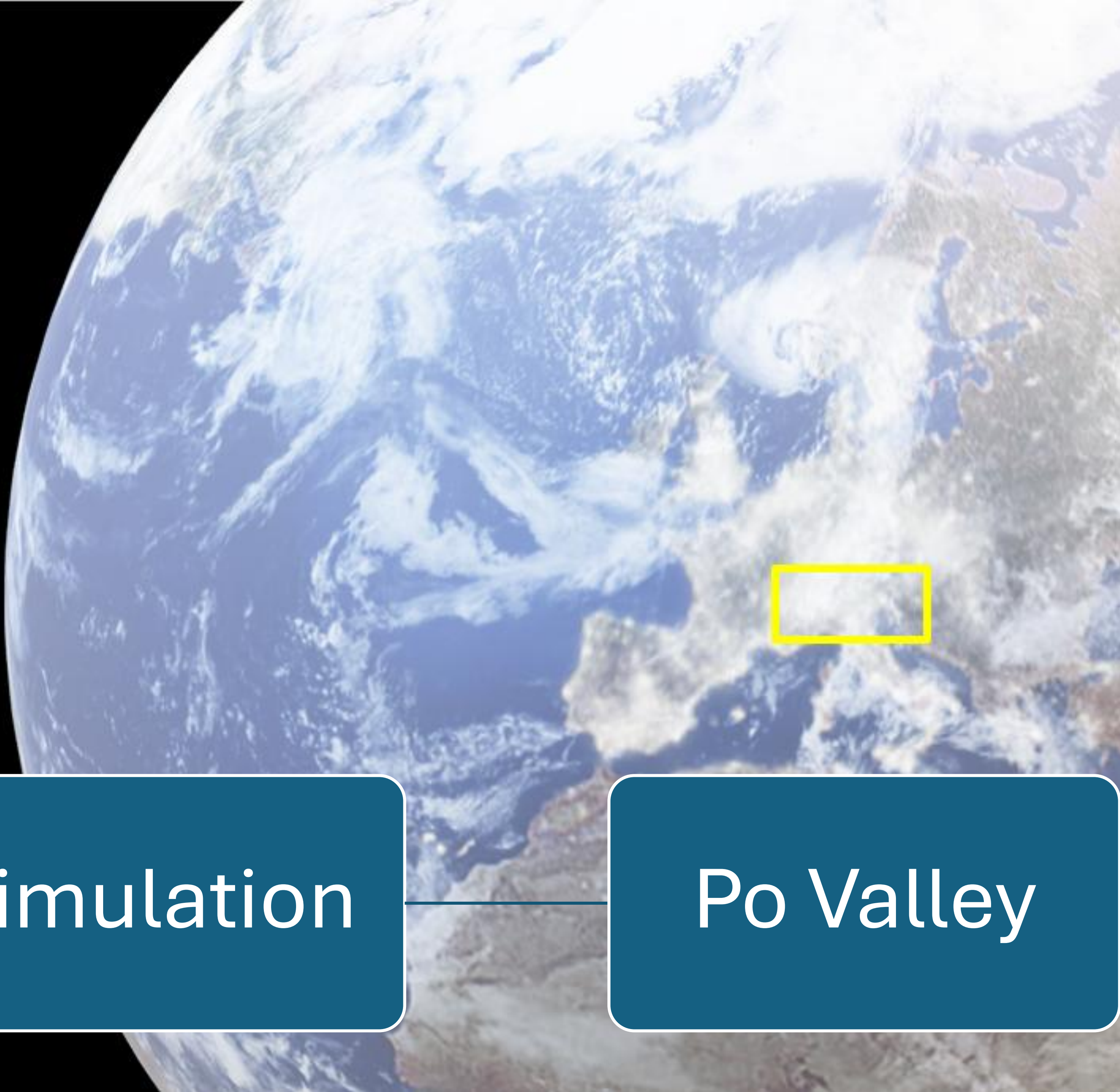
Performance testing on reconstructed profiles (from the principal component domain to vertical profiles)



Comparison between the average of the profiles retrieved and the average of the true profiles for the starting layer = 1.

- RMSE: 2.51×10^{-8} ppm → very low error, comparable to the instrument error;
- R^2 Score: 0.98 → very strong correlation.

Applications



Applications

Simulation

Po Valley

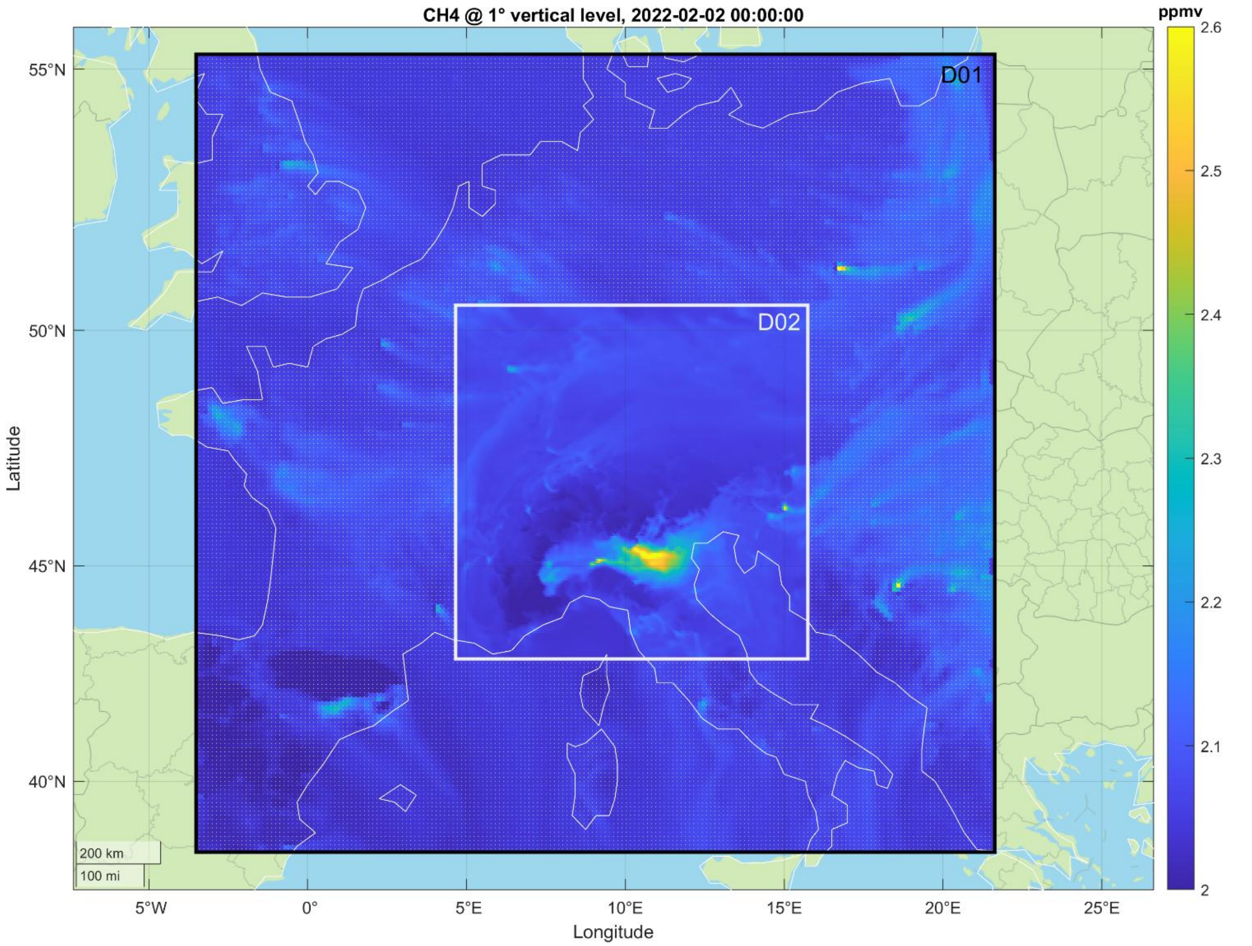
PINN Optimization using WRF-GHG simulations

We fed WRF profiles to our PINN for a second fine-tuning

Simulated dataset details:

- **1 yr** (January - December 2022)
- **2** nested domains (**we selected the innermost**)
- Input data: **GFS** forecast for climatology,

  EDGAR, CAMS, FINN, WACCM for chemistry
- **GHG** gases (CH4, CO, CO_2)
- Spatial resolution of the innermost domain is **5 km**
- Temporal resolution is **1 hr**
- Total num. of profiles: **~13M**

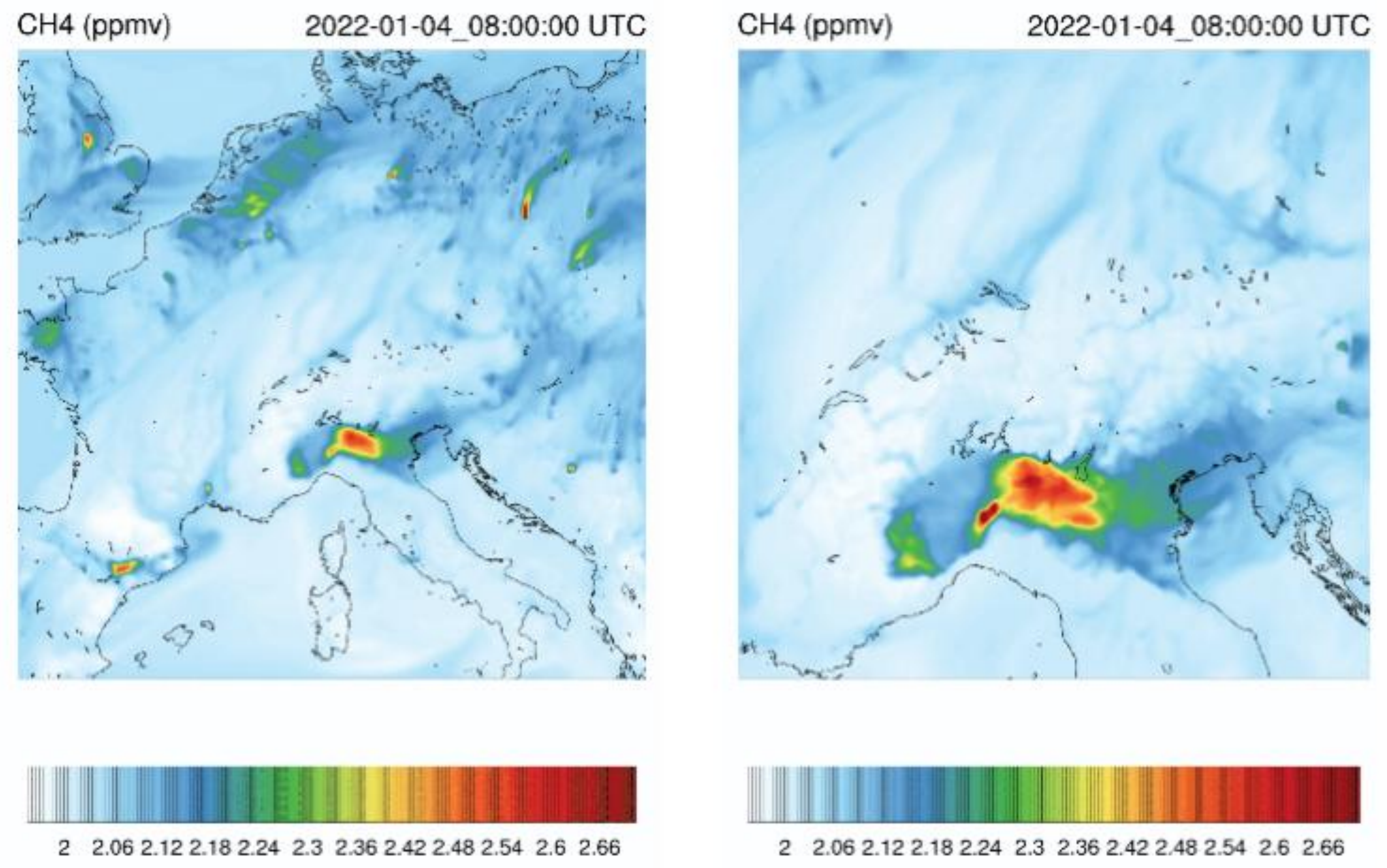


PINN Optimization using WRF-GHG simulations

Example (maps of CH_4 @ surface)

Domains details:

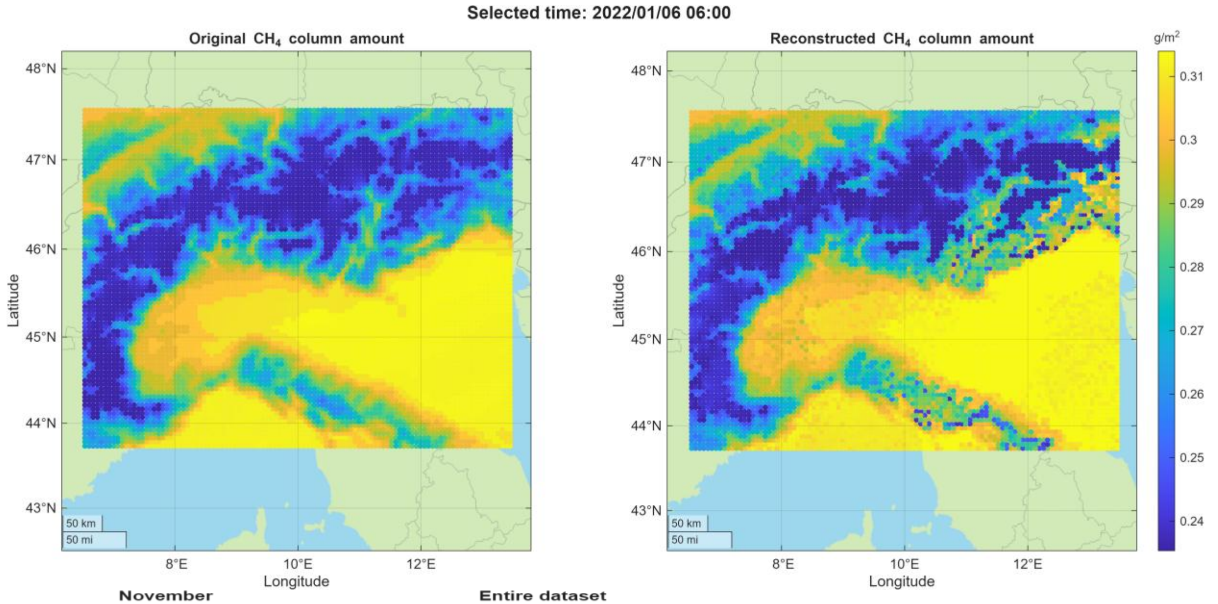
- **D01**: 10 Km
- **D02**: 5 Km

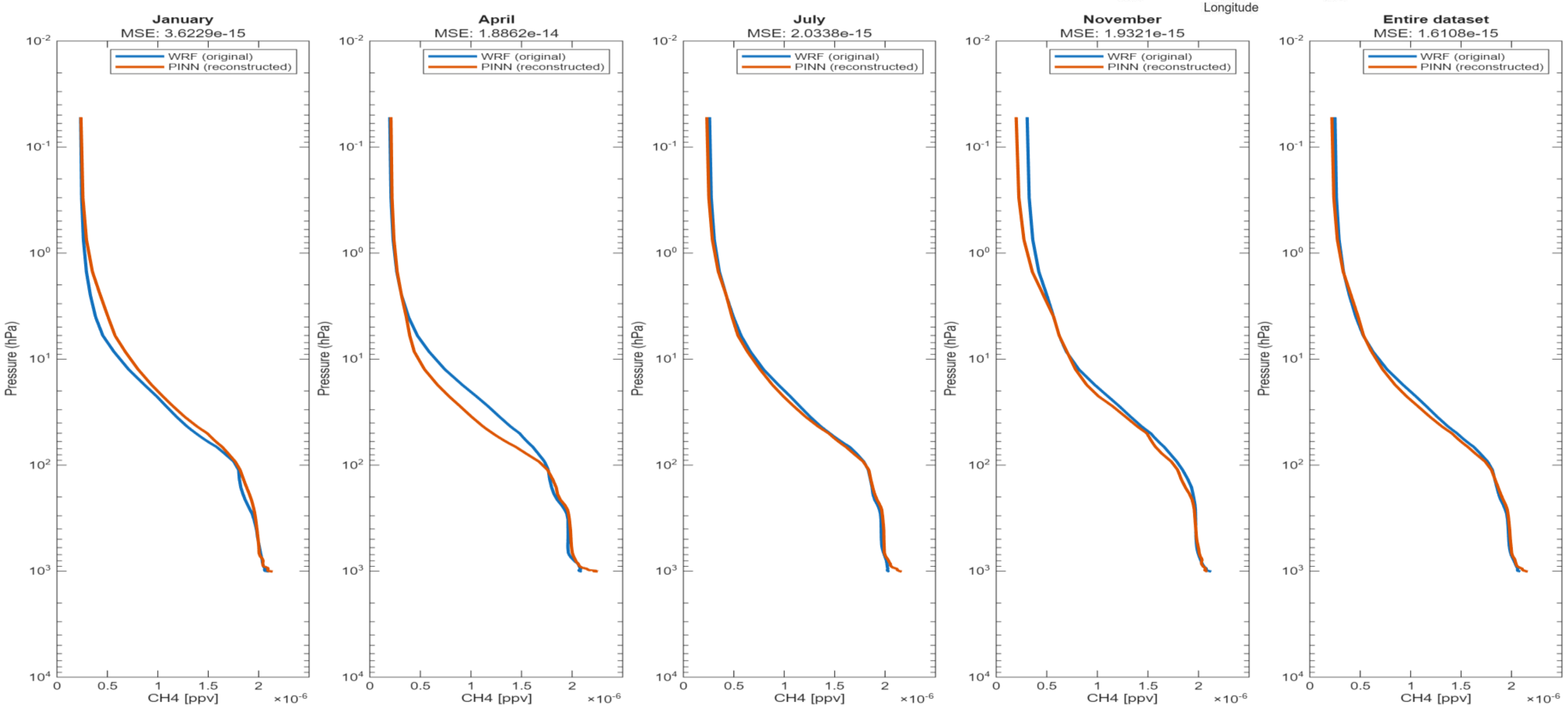


PINN Optimization using WRF-GHG simulations

Comparison between WRF data and PINN output in terms of:

- Profiles;
- Column amount.





UNIVERSITÀ DEGLI STUDI DELLA BASILICATA

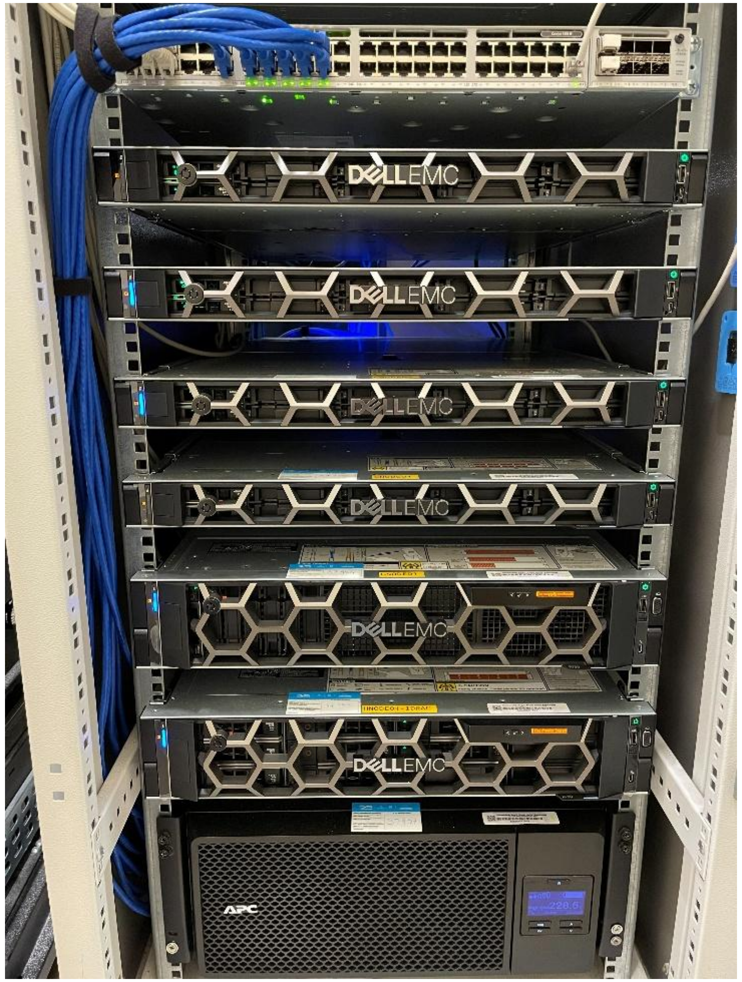
PINN Optimization using WRF-GHG simulations

We routinely run WRF simulations on our HPC Cluster.

Cluster hardware:

- 4 Compute Nodes with 56 cores and 256 GB RAM each
- 1 GPU Node with 40 cores and 256 GB RAM,

  NVIDIA Ampere A30 GPU, 24 GB HBM2

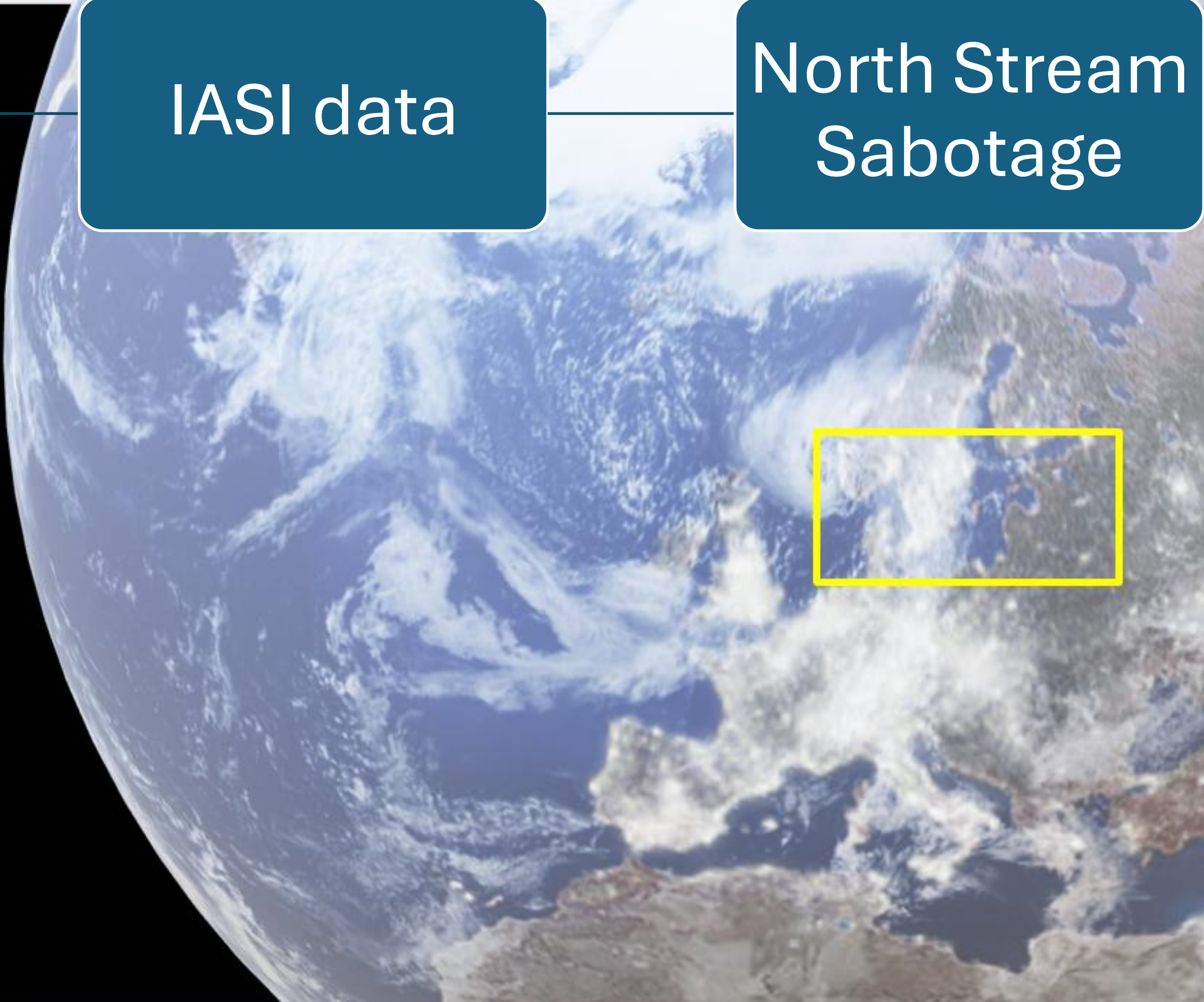


Applications

Applications

IASI data

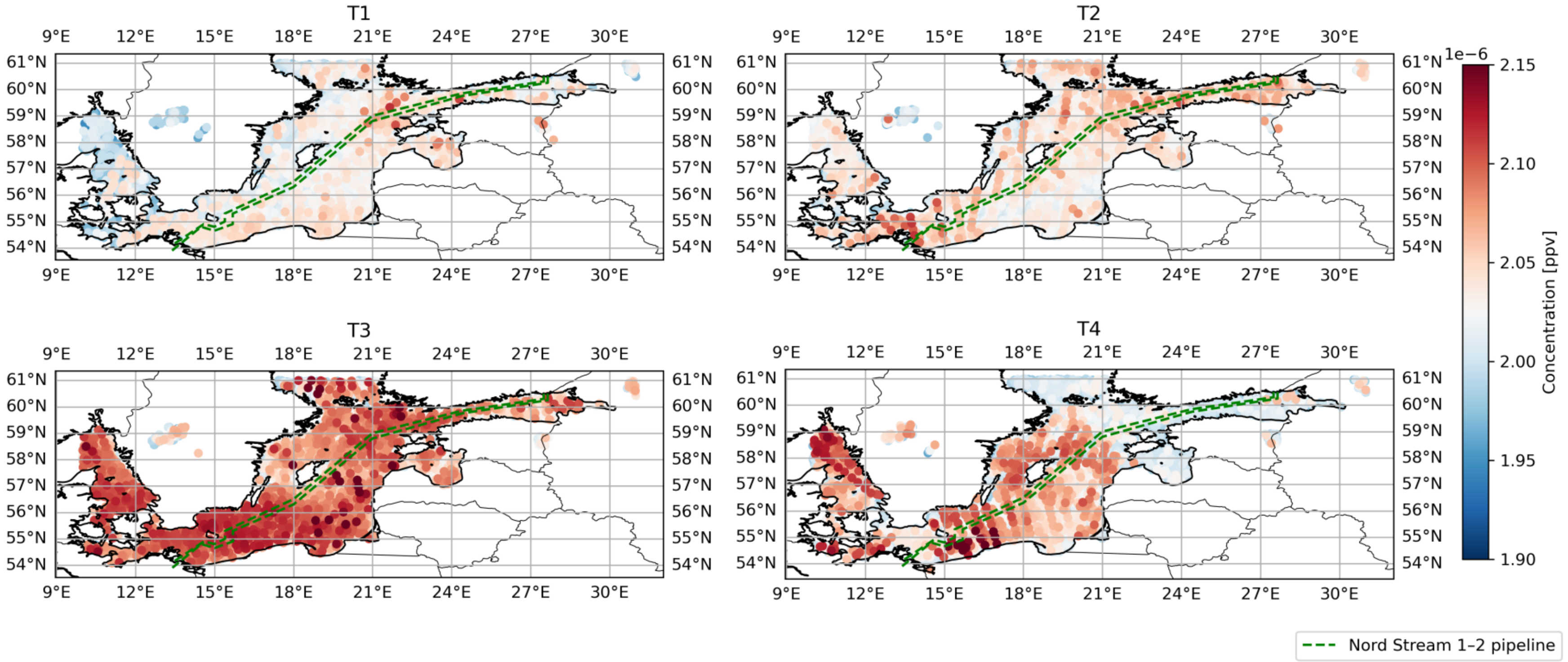
North Stream Sabotage



North Stream: CH_4 Columnar Amount

We collect IASI B and C data in 4-time intervals and apply the PINN methodology

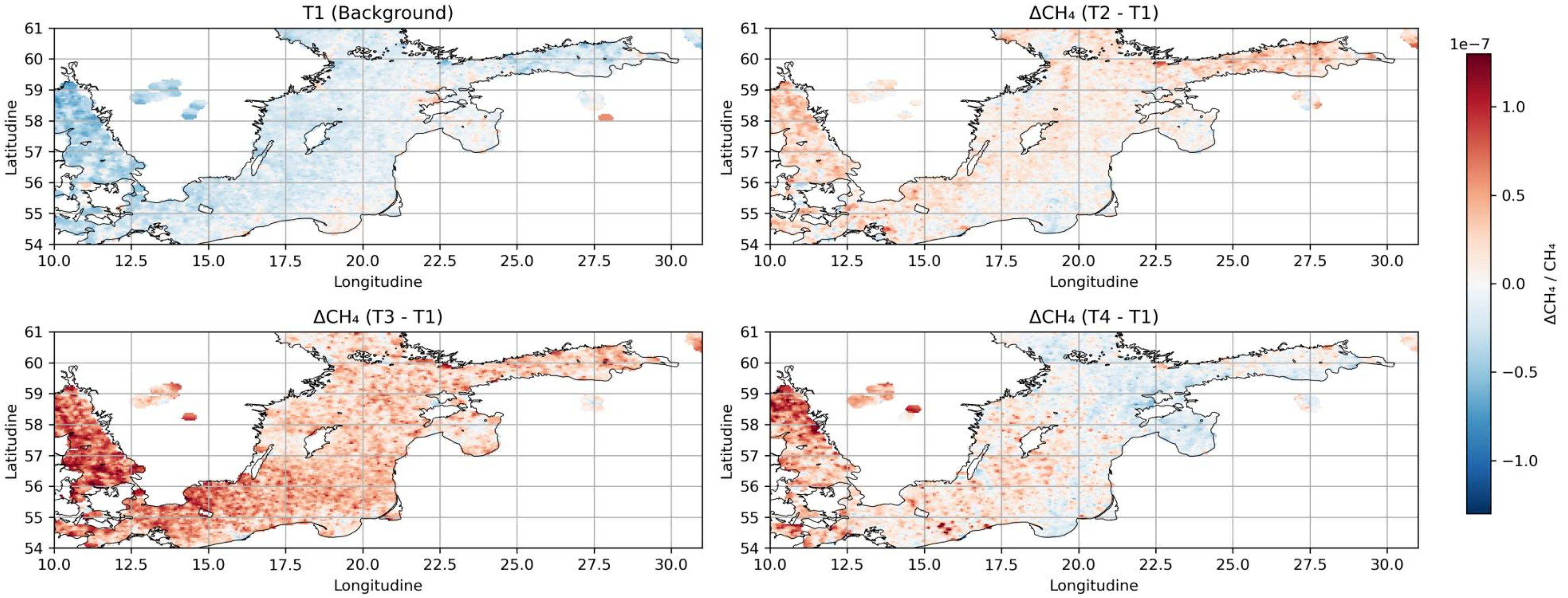
- **T1**: Sept. 19–Sept 21, pre-event phase;
- **T2**: Sept. 22– Sept. 25, phase immediately preceding the event;
- **T3**: Sept. 26– Sept. 29, period when the explosions occurred;
- **T4**: Sept. 30– Oct. 3, Post-event phase, characterized by interventions to stop gas leaks.



Maps of spatially unplaced data of methane surface concentration for the 4 time intervals chosen for analysis.

North Stream: CH_4 Columnar Amount v/s BG

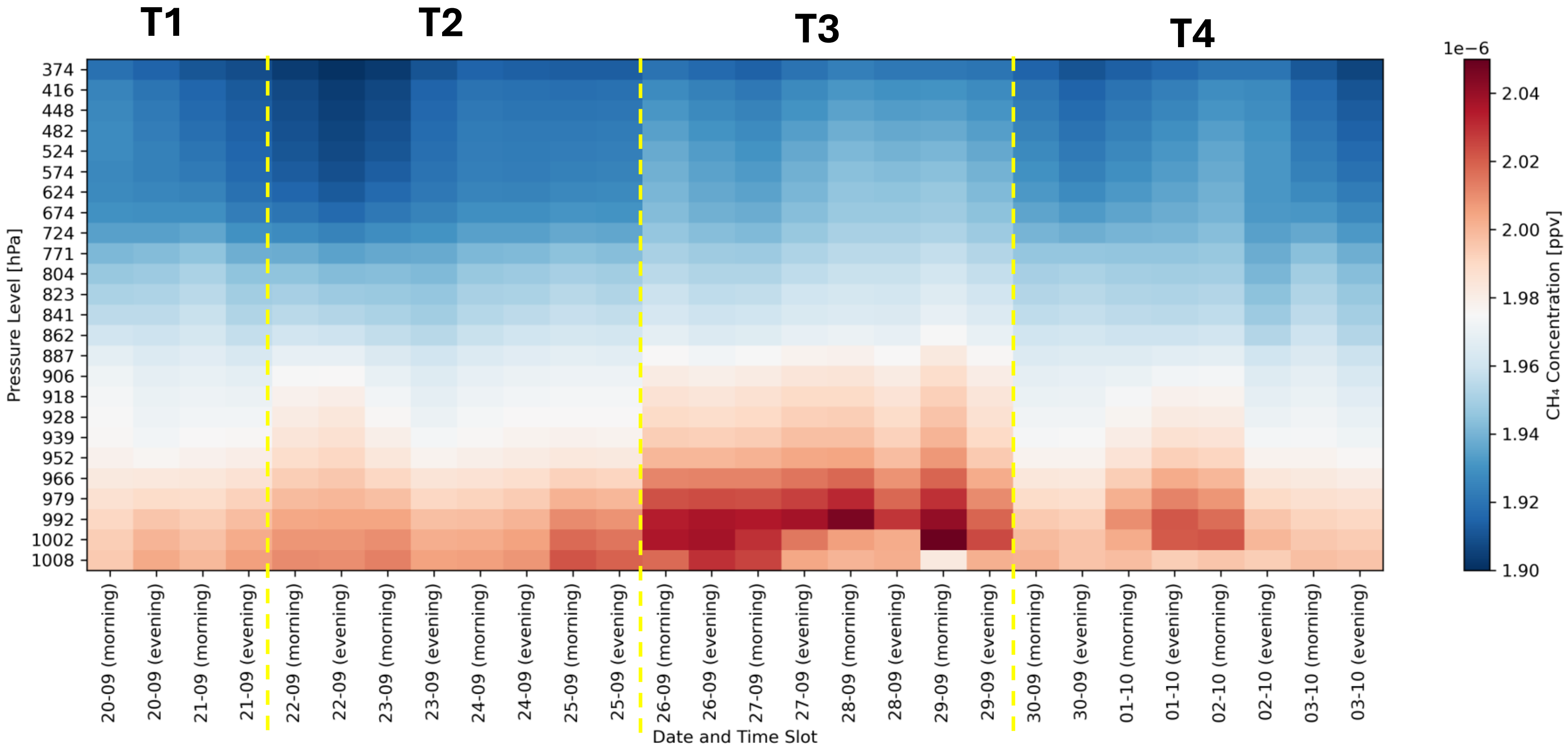
The analysis of the variation in surface methane concentration was performed by averaging each time interval, spatially locating the pixels, and using period T1 as the background.



Maps of spatially unplaced data of methane surface concentration for the 4 time intervals chosen for analysis.

North Stream: CH_4 Profiles

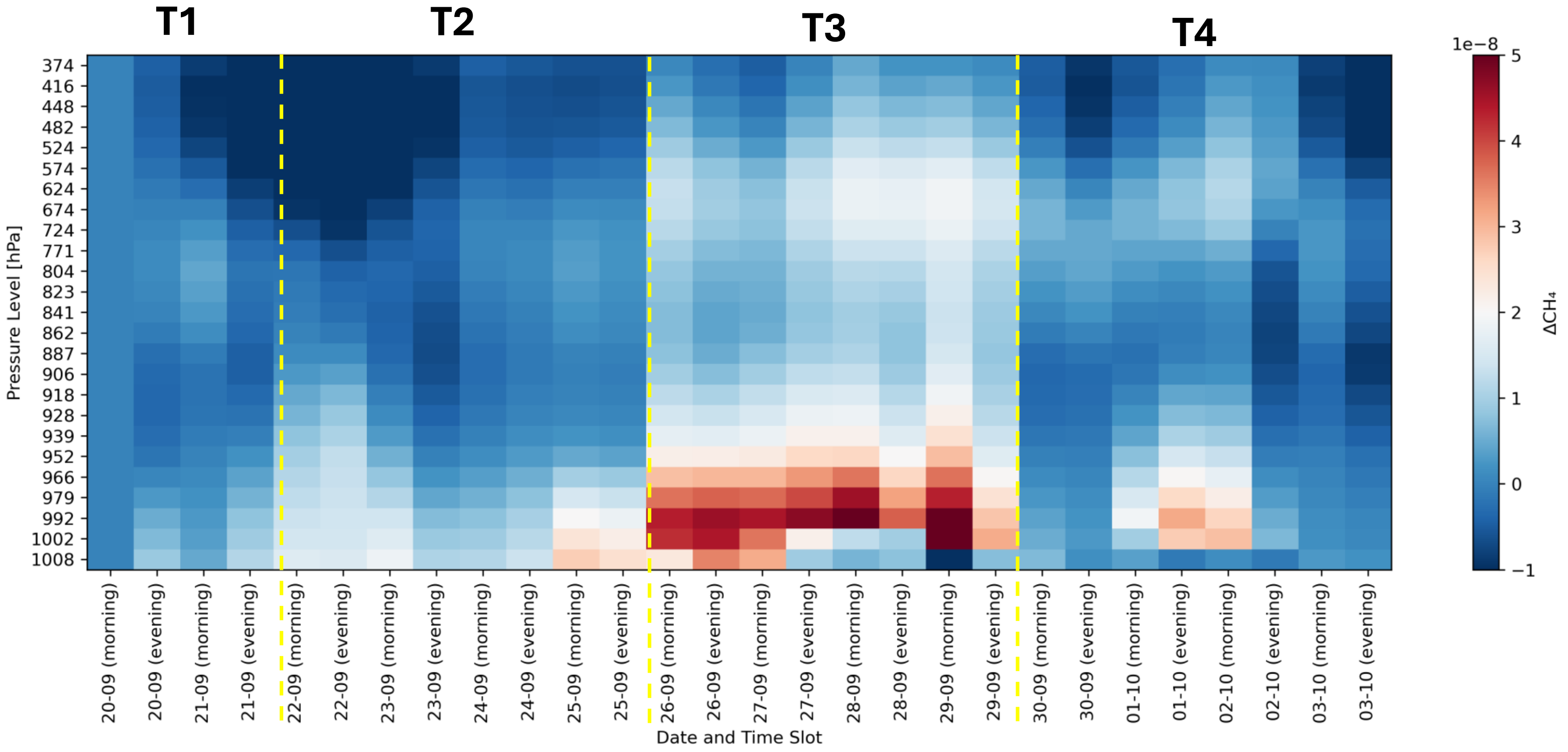
Analysis of the average profiles across the entire area analyzed shows a significant increase in methane concentration during phase T3 and T4.



Average daily profile of methane concentration in the top 25 pressure levels (1000 to 380 hPa), divided into morning and afternoon slots, over the entire study area from Sept. 20 to Oct. 3.

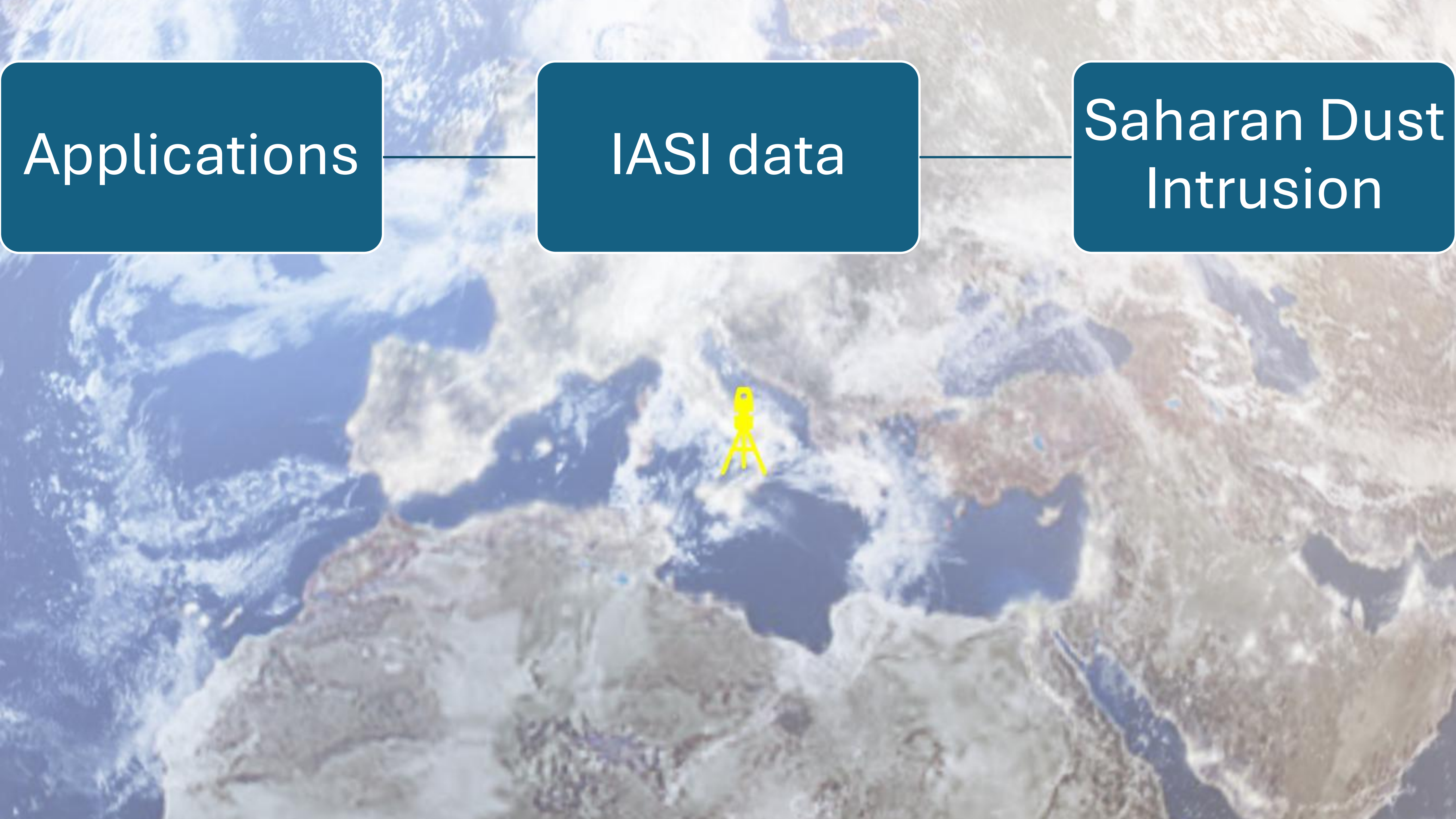
North Stream: CH4 Profiles v/s BG

The analysis of concentration variation was also performed for average daily profiles, using the profile of the morning of September 20 as a background.



Average daily profile of methane concentration in the top 25 pressure levels (1000 to 380 hPa), divided into morning and afternoon slots, over the entire study area from Sept. 20 to Oct. 3.

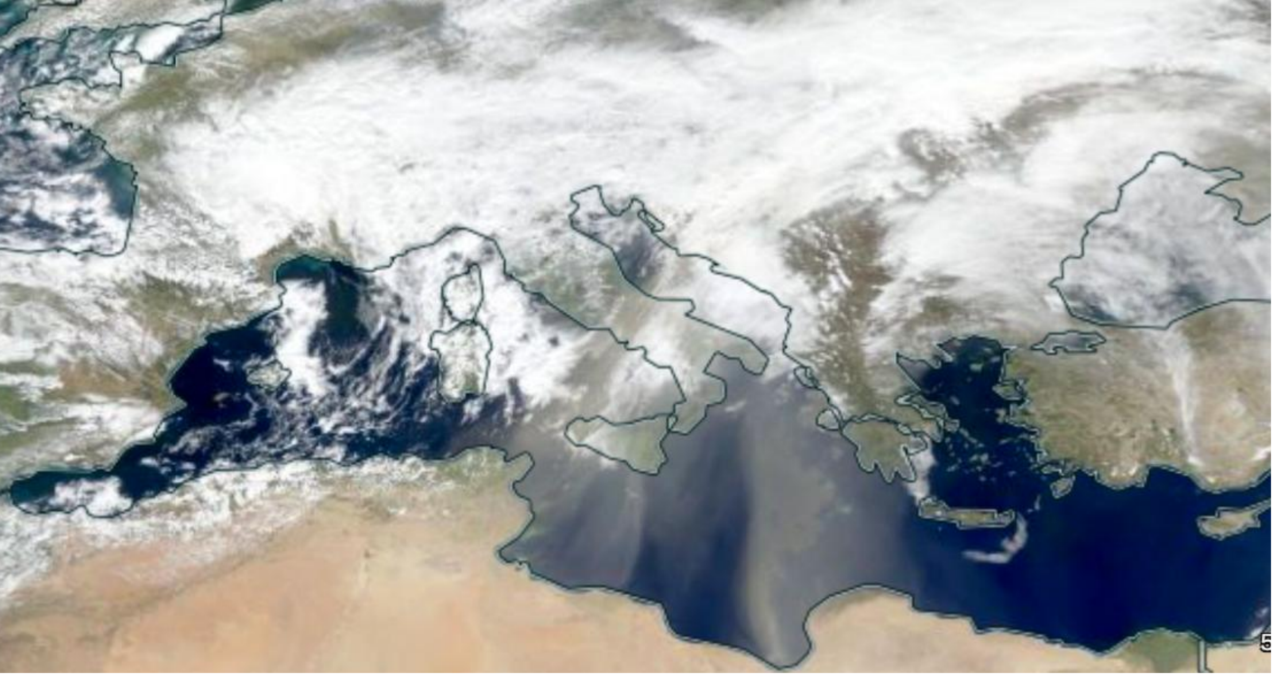
Applications

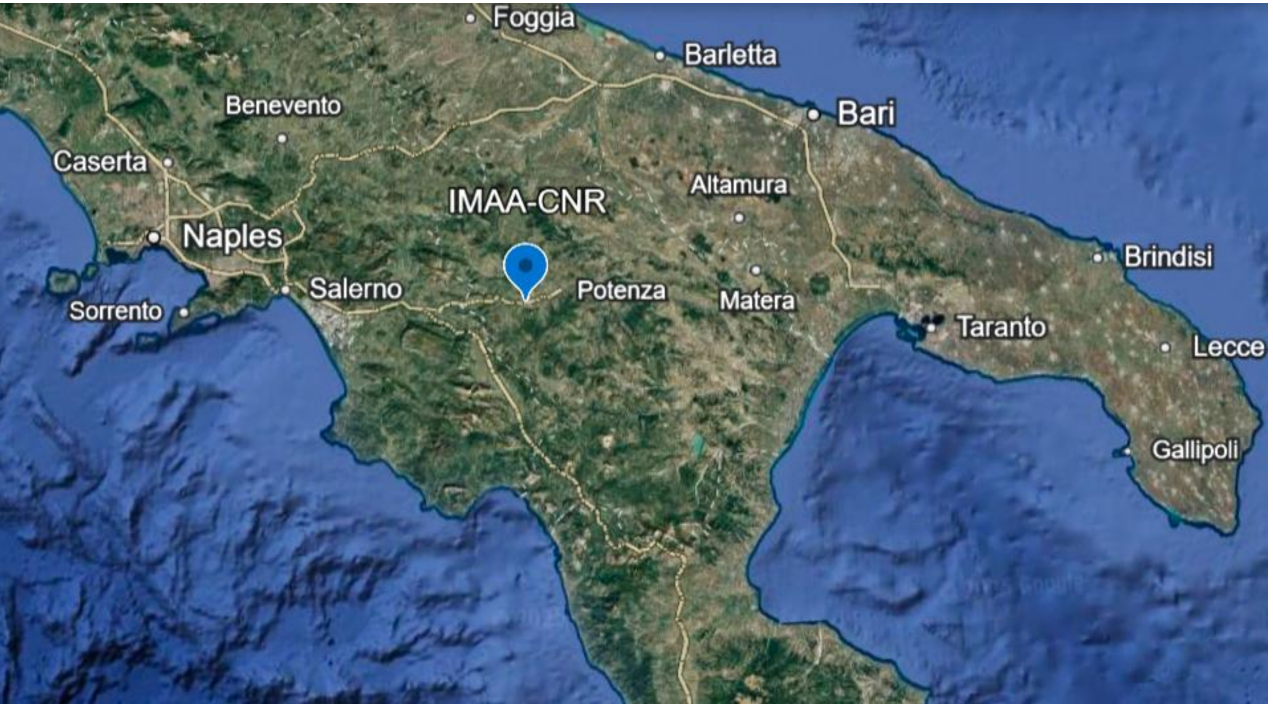

Applications
IASI data
Saharan Dust Intrusion

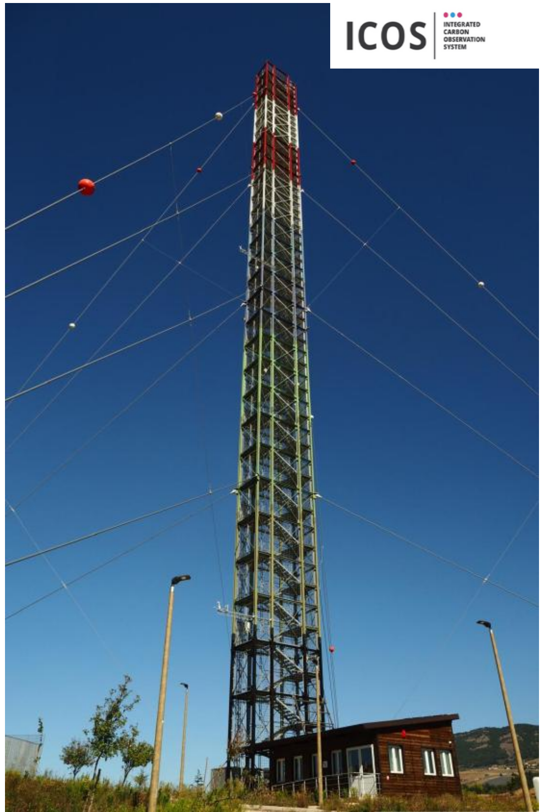
IMAA-CNR

The CNR at Integrated Carbon Observation System (ICOS) POT station in Tito Scalo (Italy), revealed isotopic shifts and a change in CH_4 concentrations during Saharan dust intrusion.

A decrease in CH_4 concentrations was observed, indicative of accelerated oxidation in the presence of Saharan dust aerosols and sea spray.







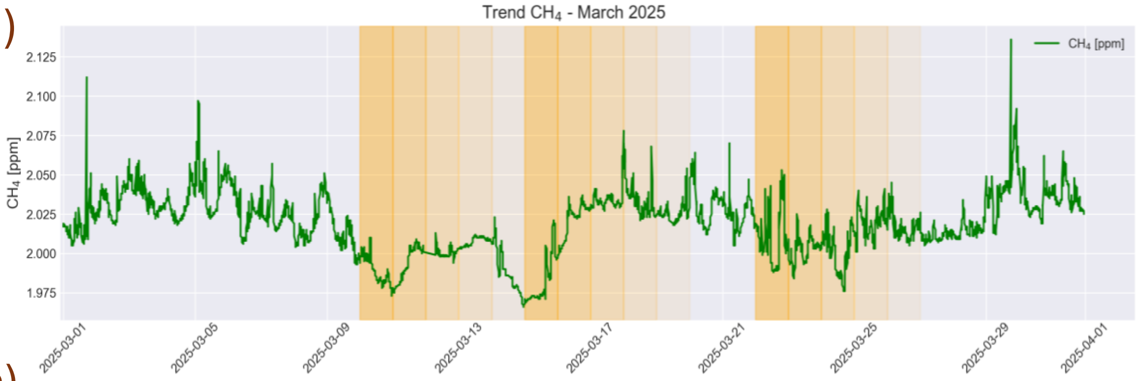
$^{13}C, ^{12}C$ (CO_2, CH_4)

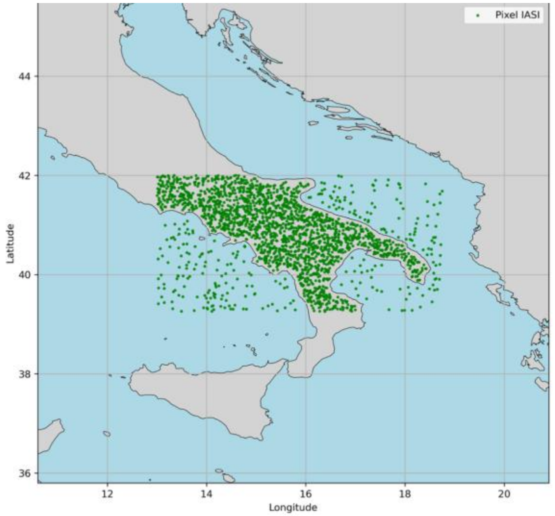
Saharan dust intrusion

Comparison PINN-POT data

a) Ground-based Data – POT station

a)





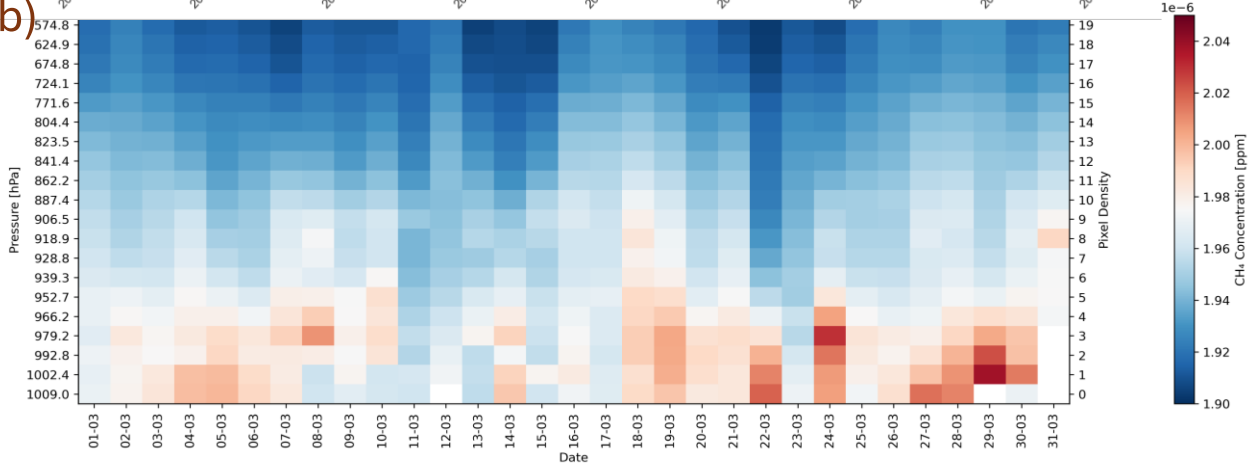
b) CH4 profiles – PINN retrieved

We collect IASI B/C data during March 2025

From these data we retrieved CH_4 profiles using PINN model.

From CH4 profiles we extracted the first layer concentration and compare with GB mesasurements

b)



a) CH4 gas surface trend derived from data from the ICOS-POT ground station in March 2025 in Tito. The yellow bands identify time intervals characterized by intrusions of Saharan dust.

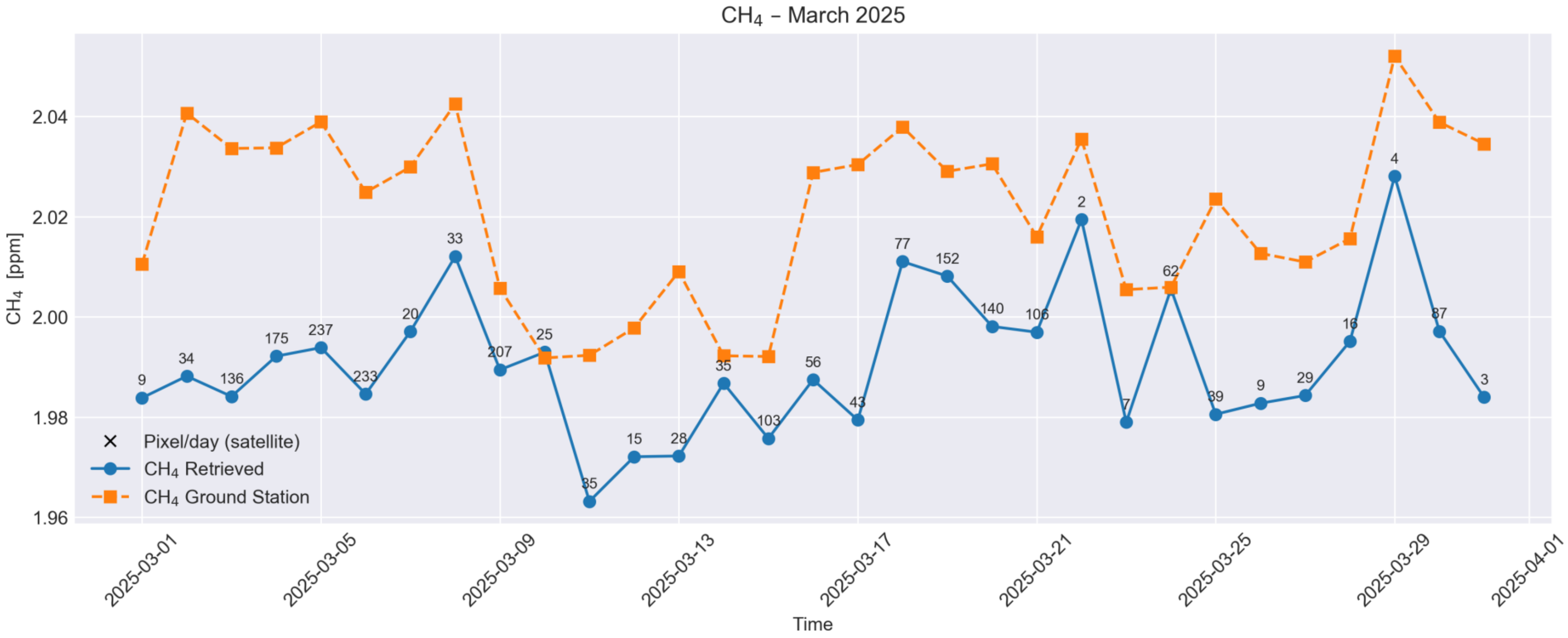
b) Average vertical CH_4 concentration profiles estimated using the PINN model in March 2025, corresponding to episodes of Saharan dust intrusion. The profiles cover a pressure range between approximately 1000 and 550 hPa.

Saharan dust intrusion

Comparison PINN-POT data

The comparative analysis was carried out by comparing the daily averages of surface concentration, obtaining the following results:

- Pearson correlative coefficient R = 0.605;
- p-value = 3.16×10^{-4} (< 0.05)



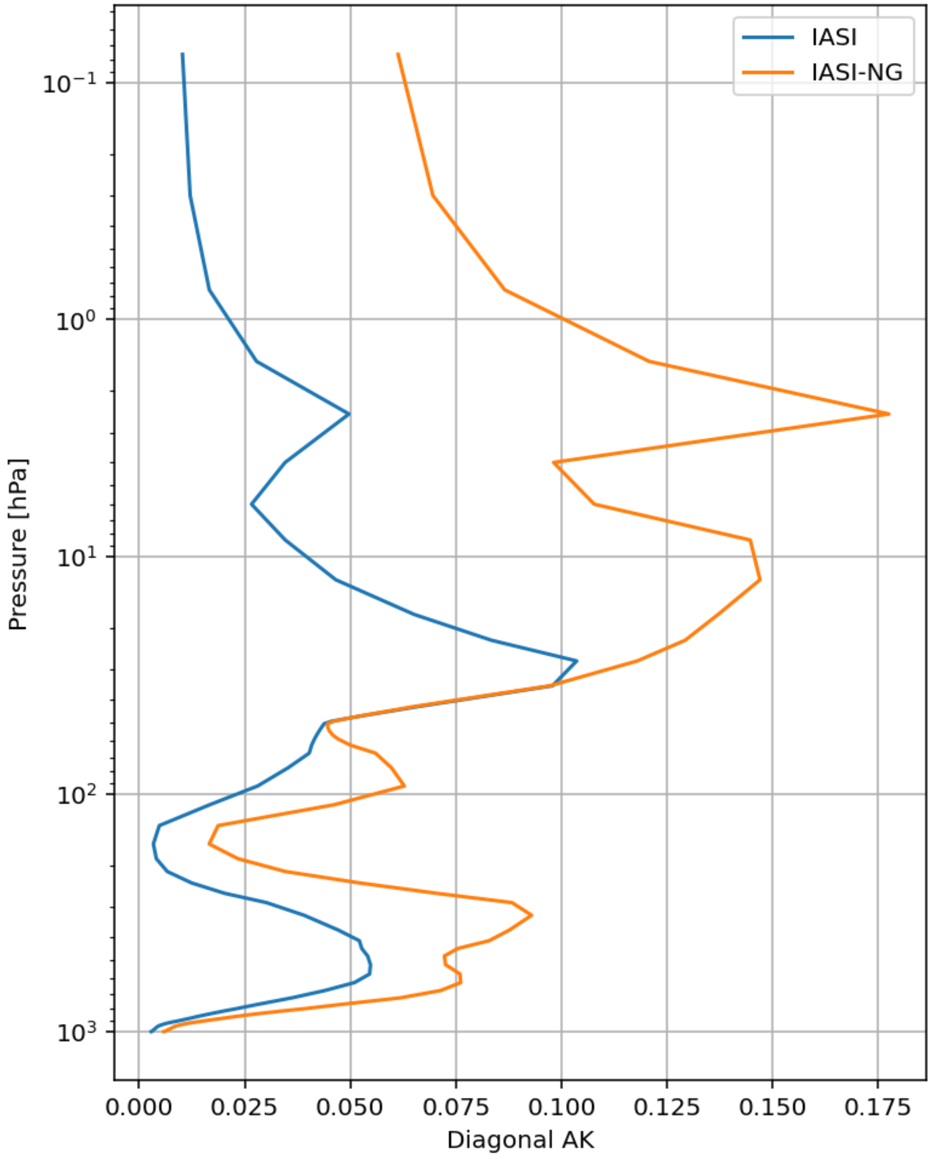
Comparison of CH_4 trends based on ground-based measurements and Physics-Informed Neural Network (PINN) estimates. For each day, the number of satellite observations (pixels) used to calculate the daily average is reported.

From IASI to IASI-NG

The analysis of the Averaging Kernels regarding methane, was performed on both IASI and IASI-NG data:

- dof IASI: 1.93;
- dof IASI-NG: 3.56;

What results would we obtain if the model were trained with approximately twice the spectral information?



Considerations

The PINN model reconstructs vertical methane profiles with a high degree of accuracy.

Future developments

- Further optimizations to reduce the underestimation of absolute concentrations (e.g., training set balancing, additional physical corrections).
- Extension of the study to CO_2.
- IASI-NG potentially will double the DOF for the Methane