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

Methane source identification from hyperspectral infrared satellite observations: a physically informed neural network-based inversion approach

living planet symposium VIENNA 23–27 June 2025

R. Giosa[1], P. Pasquariello[1], M. D'Emilio[1], G. Masiello[1], C. Serio[1], M. Ragosta[6], F. Carbone[2], L. Cassini[1,3], F. Della Rocca[4], C. N. Gencarelli[5], I. De Feis[4], G. Liuzzi[1]

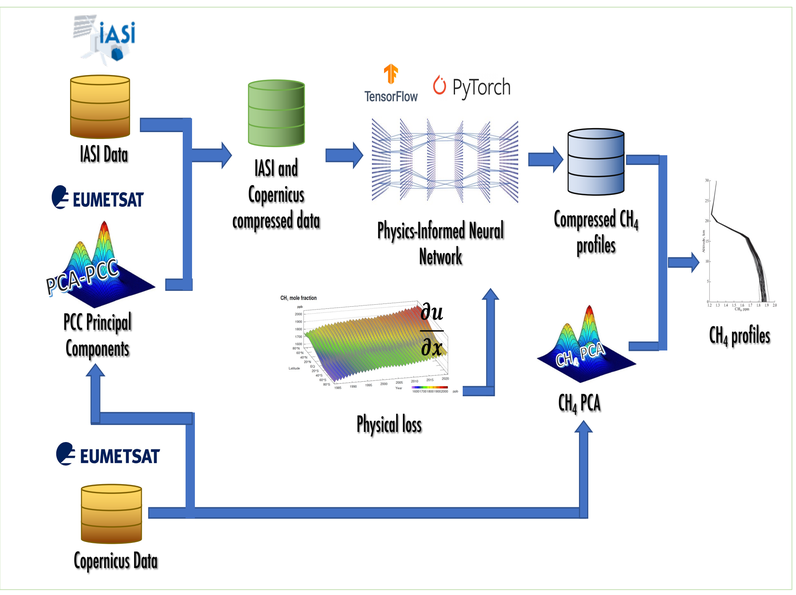
esa

(1) University of Basilicata, Department of Engineering, (2) IAA, National Council of Research, (3) University "La Sapienza" Department Of Civil, Building And Environmental Engineering, (4) IAC, National Council of Research, (5) IGAG, National Council of Research, (6) University of Basilicata, Department of Health Sciences

Context and motivation

Monitoring greenhouse gas emissions, particularly methane (CH_4) and carbon dioxide (CO_2), is essential to counter climate change. Methane, more potent than CO_2 in terms of greenhouse effect and with important anthropogenic sources, is a priority target for rapid mitigation actions.
In the PRIN-MVP project, an innovative methodology based on Physics-Informed Neural Networks (PINNs) [1] was developed and validated to obtain vertical CH_4 profiles from satellite data, allowing accurate identification of methane sources at the surface. We have successfully applied to IASI data; this technique leverages approximately 1,700,000 simulated clear-sky spectra generated by the σ-IASI/F2N radiative model [2], along with ancillary information. Averaging Kernel are used to optimize spectral intervals [3], while PCA reduces dimensionality and noise to achieve consistent projections at each atmospheric level.
The results demonstrate the effectiveness of the PINN model in reconstructing complete and consistent CH_4 profiles validated against reference measurements and source identification against a test case.

General Outline



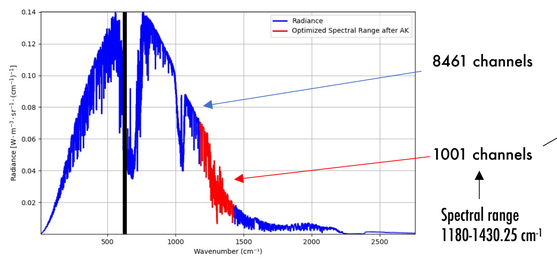
Data

Radiances were simulated using the radiative transfer code σ-IASI/F2N, which uses a fixed pressure grid. Input profiles and parameters are downloaded from the Copernicus databases. The dataset includes clear-sky measurements for the European area for the year 2020, as well as targeted measurements for major emission events such as: Deepwater Horizon Platform Disaster (2010), San Bruno Gas Pipeline Explosion (2010), Elgin Platform Accident (2012), Permian Basin (2015, 2016) and Emissions in the European interest area (2014-2018). All measurements refer to the 4 synoptic times (00-06-12-18 UTC).

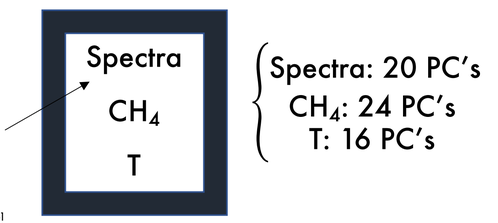
Compression Data

The AK matrix, was calculated starting from the core of the CH_4 band in the infrared (wavelength number 1305 cm^{-1}) by broadening the spectral range under analysis step by step. The study showed that the maximum amount of information, relating to CH_4 profiles, is obtained between the wavenumbers 1180 cm^{-1} and 1430.25 cm^{-1} (1001 channels). PCA was applied to both spectra and vertical profiles of methane and temperature.

Averaging Kernel Analysis



PCA



Inflated Physics Informed Neural Network (Inflated PINN)

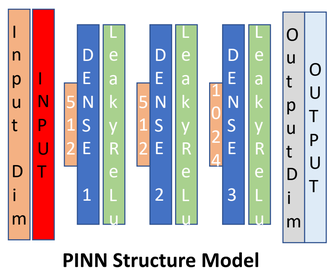
The basic structure of the PINN was derived through the Autokeras framework. Starting from this structure, a PINN was trained with the following loss function for each input tensor, where we will call sl the starting layer of generic example, yr the predicted output tensor and y_{true} the reference output tensor. The covariance matrices (one was constructed for each starting layer) were inflated by a factor of 10 for the first 10 values of the first 10 rows. This is intended to provide more freedom to the network to find the minimum error in the first pressure layers (about 1000 hPa), in order to help the network identify surface methane emissions. For this reason, the name of the model is INFLATED PINN. The loss function for a generic example is:

$$S_{a,sl}(p_i,p_j) = \begin{cases} inv\left(10 \cdot cov_matrix_{sl}(p_i,p_j)\right), & if\ 1 \le i,j \le 10 \\ inv\left(cov_matrix_{sl}(p_i,p_j)\right), & otherwise \end{cases}$$

$$diff = y_{true} - yr$$

$$loss = diff * S_{a,sl} * diff^{\mathsf{T}}$$

Input Tensor	#
Spectra PC	20
Temperature PC	16
Latitude	1
Longitude	1
Starting Layer	1
Time	1
Vza	1
Input Dimension	41



PINN Structure Model

Output Tensor	#
CH_4 PC	24
Output Dimension	24

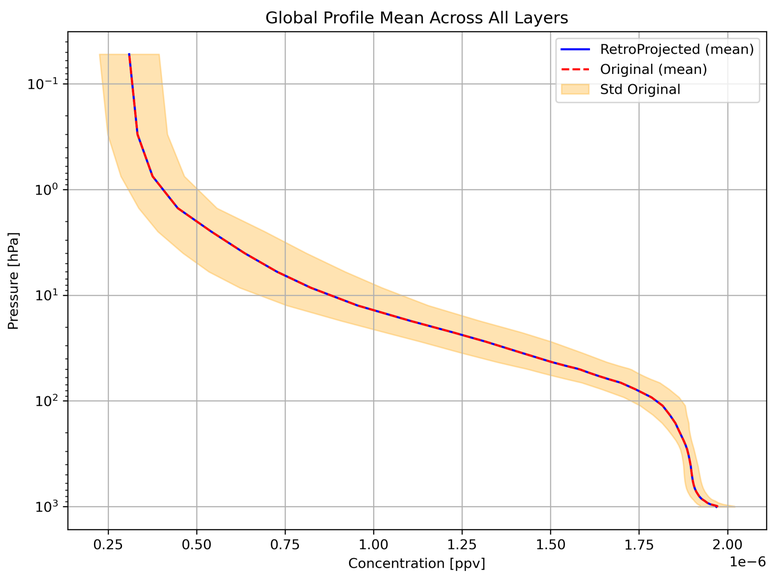
Each output tensor ($yr_{j,sl}$) is subjected to the backprojection process in order to obtain the profiles in the pressure domain. Given the matrix $U_{CH4,sl}$ containing the principal components associated with the sl-th starting layer, and the vector of predicted scores $yr_{j,sl}$, the normalized reconstructed profile $qr_{j,sl,norm}(p)$ is obtained by projecting the scores back into the original feature space through matrix multiplication. Specifically, the reconstructed profile $qr_{j,sl}(p)$ is obtained as follows:

$$qr_{j,sl,norm}(p) = U_{CH4,sl} \cdot yr_{j,sl}$$

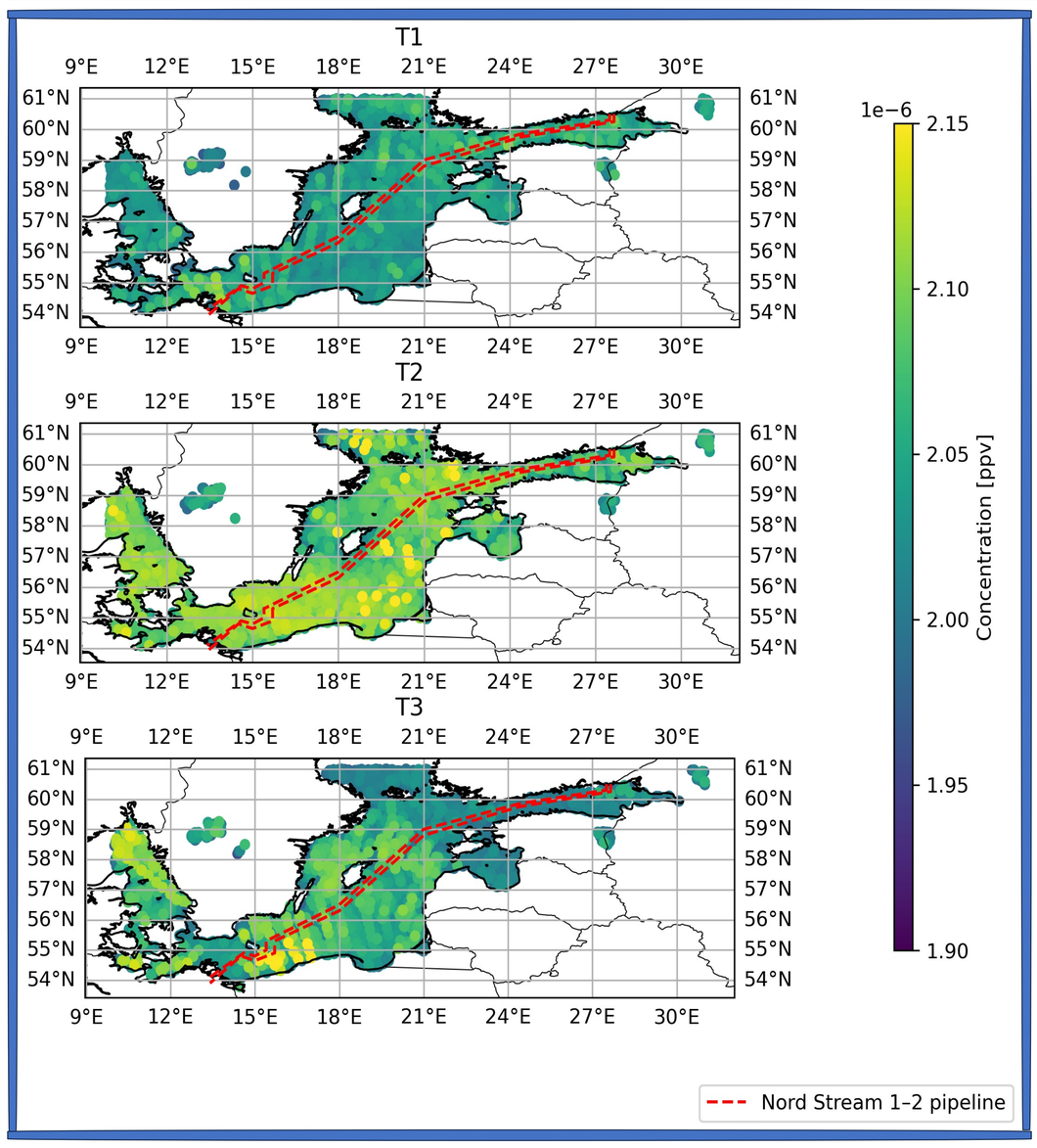
$$qr_{j,sl}(p) = qr_{j,sl,norm}(p) \cdot q_{std,sl} + q_{mean,sl}$$

PINN Inflated evaluation

The performance of the PINN Inflated model can be summarized with the values of the following metrics: RMSE: 4.5147 e-08 ppv; Relative Error: 2,79%; R^2 score: 99,18%.
It can be seen how, in the lower left figure, the average of the actual profiles coincides with the average of the reconstructed profiles related to the test-set, also indicating the variability of the tested data indicated through the standard deviation of the same.
This indicates the correct finding of the vertical profiles by the PINN inflated model.



Test on source identification for real data: the case of "North Stream pipelines sabotage"



A further analysis (blue box) was conducted using real IASI observations and associated profiles (temperature, surface pressure, focusing on the Baltic Sea region in the vicinity of Bornholm Island for the period when submarine explosions involving North Stream 1 and 2 pipelines were recorded.
In detail, 3 time intervals were analyzed:
T1: Sept. 22– Sept. 25, phase immediately preceding the event;
T2: Sept. 26– Sept. 29, period when the explosions occurred;
T3: Sept. 30– Oct. 3, Post-event phase, characterized by interventions to stop gas leaks.
The pixels analyzed refer only to clear sky pixels over the sea.

The «North stream test» assesses the behavior of the inflated PINN under stress conditions to identify CH_4 surface sources, showing encouraging results in that the highest emissions occur at the time interval T2 (period of the explosion), and to a lesser extent at T3 (pipeline safety work), while at the interval T1 the emissions are minimal and attributable to the usual daily emissions.

[1] Cuomo et al. (2022), GSC, doi: 10.1007/s10915-022-01939-z
[2] Masiello et al. (2024), JQSRT, doi:10.1016/j.jqsrt.2023.108814
[3] Serio et al. (2019), JQSRT, doi: 10.1016/j.jqsrt.2018.10.020

Consiglio Nazionale delle Ricerche
Methane Vertical Profiling
UNIVERSITÀ DEGLI STUDI DELLA BASILICATA
diing DIPARTIMENTO DI INGEGNERIA