A Physics-Informed Data-Driven Approach for Fast Atmospheric Radiative Transfer Inversion Using FORUM Simulated Measurements

Sgattoni C.[1], Chung M.[2], Sgheri L.[3]

[1] Institute of BioEconomy (IBE), National Research Council (CNR), Florence, Italy

[2] Department of Mathematics, Emory University, Atlanta, GA, USA

[3] Institute of Applied Mathematics (IAC), National Research Council (CNR), Florence, Italy

Background – Retrieval, inverse problem

- Find the atmospheric parameters x (surface temperature, temperature, water vapor, ozone, surface spectral emissivity) that best reconstruct the measured spectrum y.

! VERY ILL-CONDITIONED PROBLEM

- Formulated as a Bayesian inference problem and solved using the OPTIMAL ESTIMATION METHOD[2]:

$$x = \arg\min_{x} \frac{1}{2} \left\| L_y\big(y - F(x)\big) \right\|_2^2 + \frac{1}{2} \left\| L_a(x - x_a) \right\|_2^2,$$

where $S_y^{-1} = L_y^T L_y$ and $S_a^{-1} = L_a^T L_a$ are the inverses of the variance-covariance matrices (VCM) of the measurements y and the a-priori information x_a, respectively.

- Minimization carried out using Gauss Newton + Levenberg-Marquardt technique.

- [2] Rodgers, C. D.: Inverse Methods for Atmospheric Sounding, World Scientific, https://doi.org/10.1142/3171, 2000.

Objectives: the RETRIEVAL problem

The computational cost of a full-physics method is too large to get Near Real Time (NRT) data analysis

↓

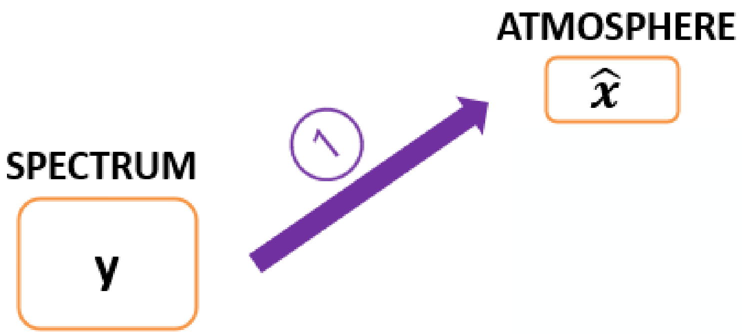
use of data-driven techniques to speed up the inversion.

Development of innovative and fast mathematical techniques to:

- exploit the huge amount of data that will be available;
- provide a flexible method, easy to apply given a database of measurements and some a-priori information.

New method: scheme

1) Approximation of the RT inverse operator with a linear operator trained using a database of FORUM simulated measurements (**completely data-driven phase**);



New method: scheme

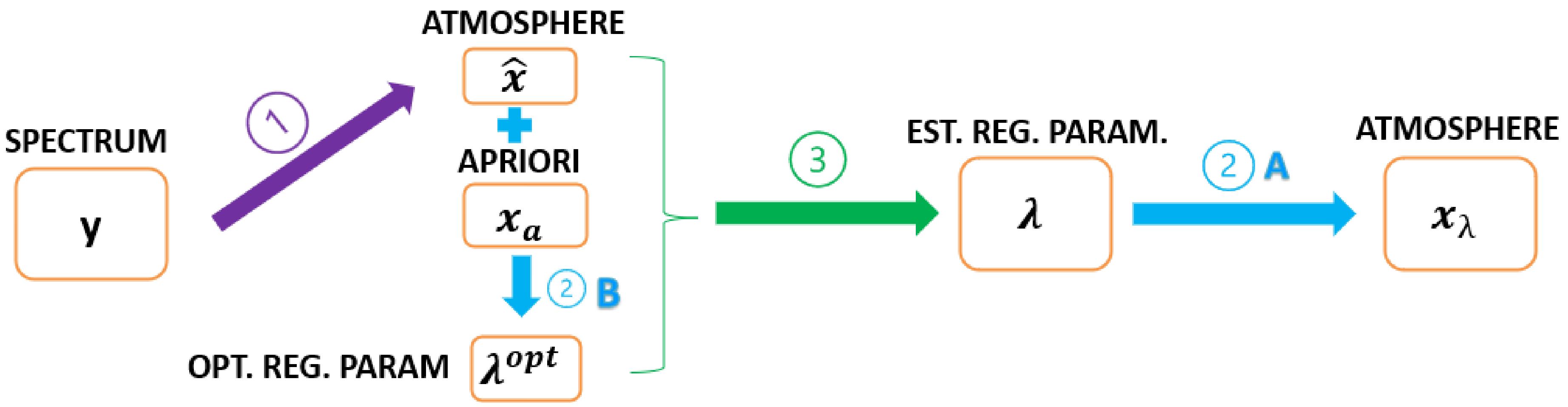
1) Approximation of the RT inverse operator with a linear operator trained using a database of FORUM simulated measurements (**completely data-driven phase**);
2) Incorporation of a-priori information into data-driven solution (**regularization technique**);

New method: scheme

1) Approximation of the RT inverse operator with a linear operator trained using a database of FORUM simulated measurements (**completely data-driven phase**);
2) Incorporation of a-priori information into data-driven solution (**regularization technique**);
3) Estimation of the optimal regularization parameters using a neural network trained with a database of pre-computed optimal parameters (**second training phase**).

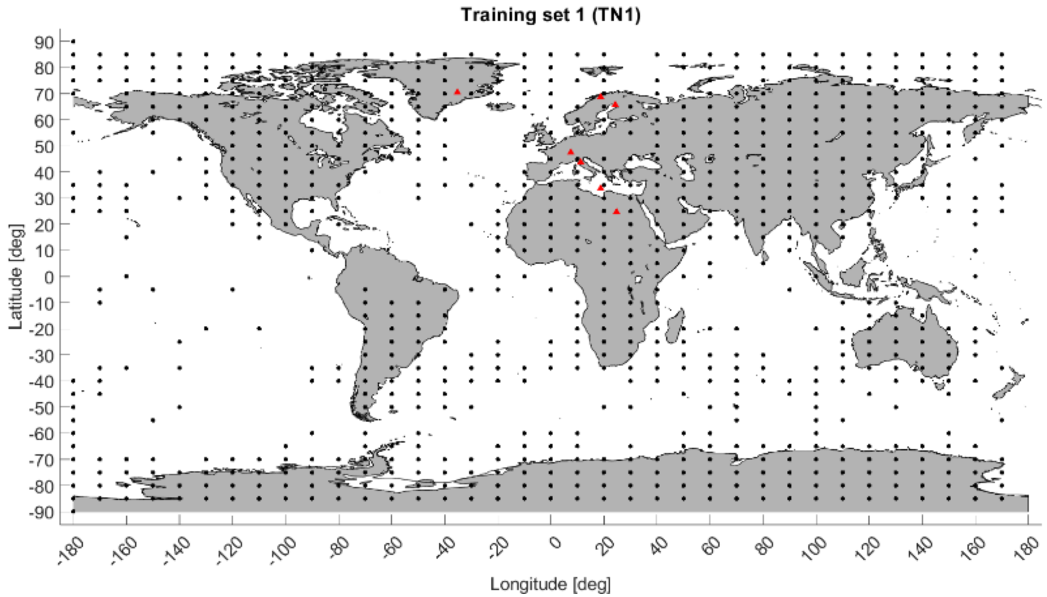
New method: scheme

1) Approximation of the RT inverse operator with a linear operator trained using a database of FORUM simulated measurements (**completely data-driven phase**);
2) Incorporation of a-priori information into data-driven solution (**regularization technique**);
3) Estimation of the optimal regularization parameters using a neural network trained with a database of pre-computed optimal parameters (**second training phase**).



1. Data-driven model

- Approximation of the RT inversion with a linear operator Z trained with simulated FORUM measurements
- Training set 1 (January and July 2021, 12:00, **clear sky**)[3]
  - X = $[x_1, x_2, \ldots, x_N]$ → N atmospheric scenarios
  - Y = $[y_1, y_2, \ldots, y_N]$ → N simulated FORUM spectra.



[3] H. Hersbach et al. "The ERA5 global reanalysis". In: Quarterly Journal of the Royal Meteorological Society 146.730 (2020), pages 1999–2049. doi: https://doi.org/10.1002/qj.3803

- Method

$$\min_{Z} f(Z) = \min_{Z} \|X - ZY\|_F^2$$

$$\frac{\delta f}{\delta Z} = -2L_y^T L_y X Y^T + 2L_y^T L_y Z Y Y^T$$

A minimizer $\hat{Z}$ of f solves $\hat{Z} Y Y^T = X Y^T$.

We can express:

$$\hat{Z} = \mathrm{XY}^{+}.$$

Then,

$$\hat{x} = \hat{Z} y.$$

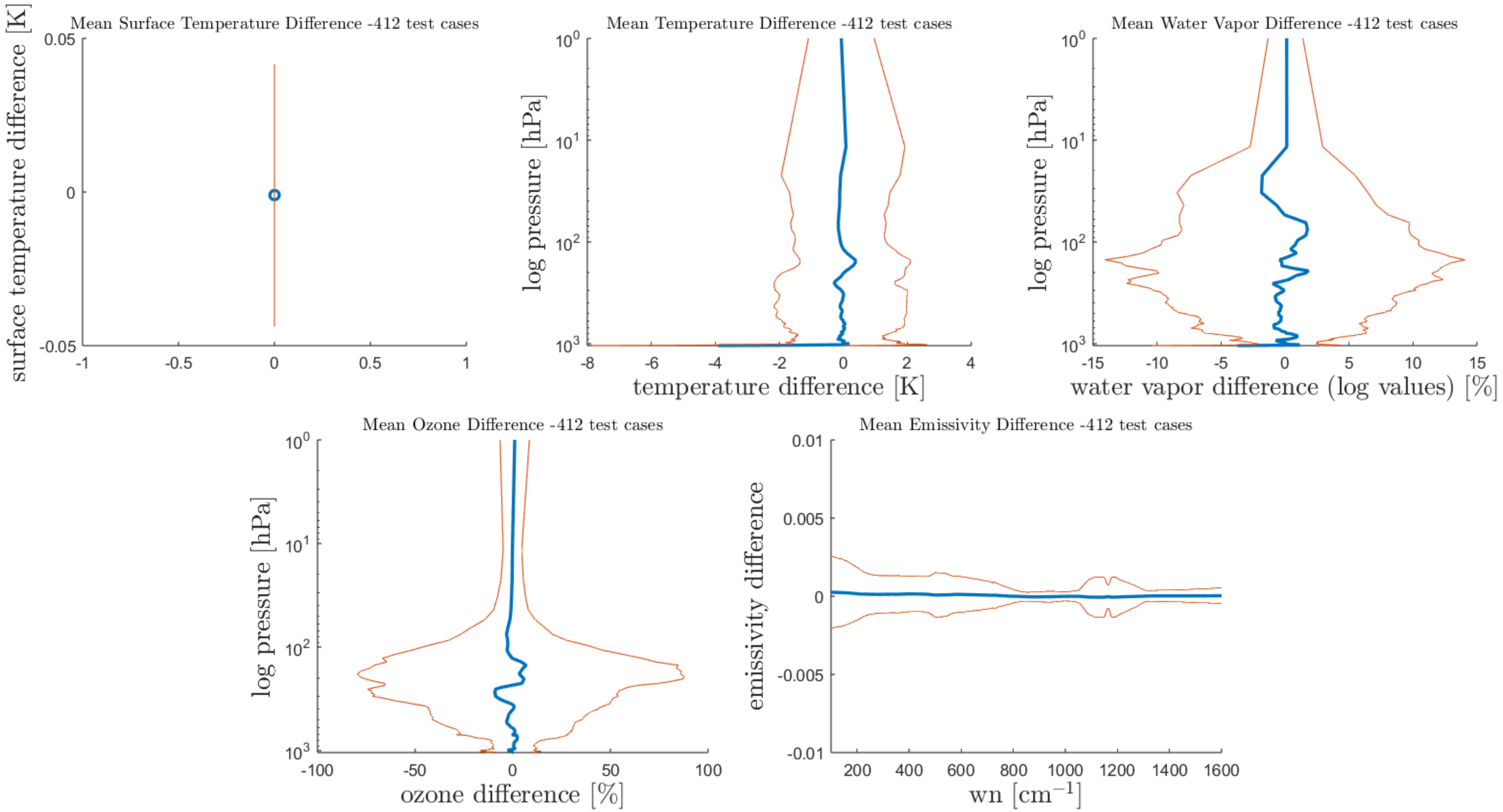
*** Moore-Penrose pseudoinverse**

Let M be a matrix of rank k with singular value decomposition $M = U \Sigma V^T$, the Moore-Penrose pseudoinverse of M is given by

$$M^{+} = V \, \tilde{\Sigma} \, \mathrm{U}^{\mathrm{T}},$$

$$\tilde{\Sigma} = diag\left(\frac{1}{\sigma_1}, \frac{1}{\sigma_2}, \dots, \frac{1}{\sigma_k}, 0, \dots, 0\right).$$

Mean signed (blue) and unsigned (orange) errors for the global test set 1



2. Tikhonov regularization – Bilevel Optimization problem

A. Inner problem

- Additional priori information:

$$x(\lambda) = \arg\min_{x} \frac{1}{2}\left\|L_x\left(x - \hat{Z}y\right)\right\|_2^2 + \frac{1}{2}\|(diag(\lambda)L_a(x - x_a)\|_2^2 \text{, with}$$

- $S_x^{-1} = L_x^T L_x$ inverse of the experimental VCM,
- x_a generated from the matrix S_a[4], with $S_a^{-1} = L_a^T L_a$.

[4]defined by the UK MetOffice for assimilation of IASI products into the operational Numerical Weather Prediction (NWP) system.

B. Outer problem

- Computation of the optimal regularization parameters for test set 1 (now training set 2):

$$\lambda^{opt} = \arg\min_{\lambda} \frac{\|x(\lambda) - x_{true}\|_2}{\|x_{true}\|_2}.$$

- Optimization carried out using interior points method.
- 5 minimizations of the inner problem changing the outer problem, one for each atmospheric component → one 5x1 parameter vector for each minimization → stored in a 5x5 matrix denoted by M.
- Strong coupling between the 5 components → aggregation of all the information in M and extraction of the most correlated parameters vector λ^{opt} with the first two left singular vectors:

$$M = USV^T, \ \lambda^{opt} = \frac{1}{4} U_1\sigma_1 + U_2\sigma_2.$$

3. Regularization parameter estimation

- Assume there exists a well-defined mapping $\Phi(\hat{x}, x_a) = \lambda$.
- Set a NEURAL NETWORK parametrized by θ to approximate Φ.

unique network for all 5 components

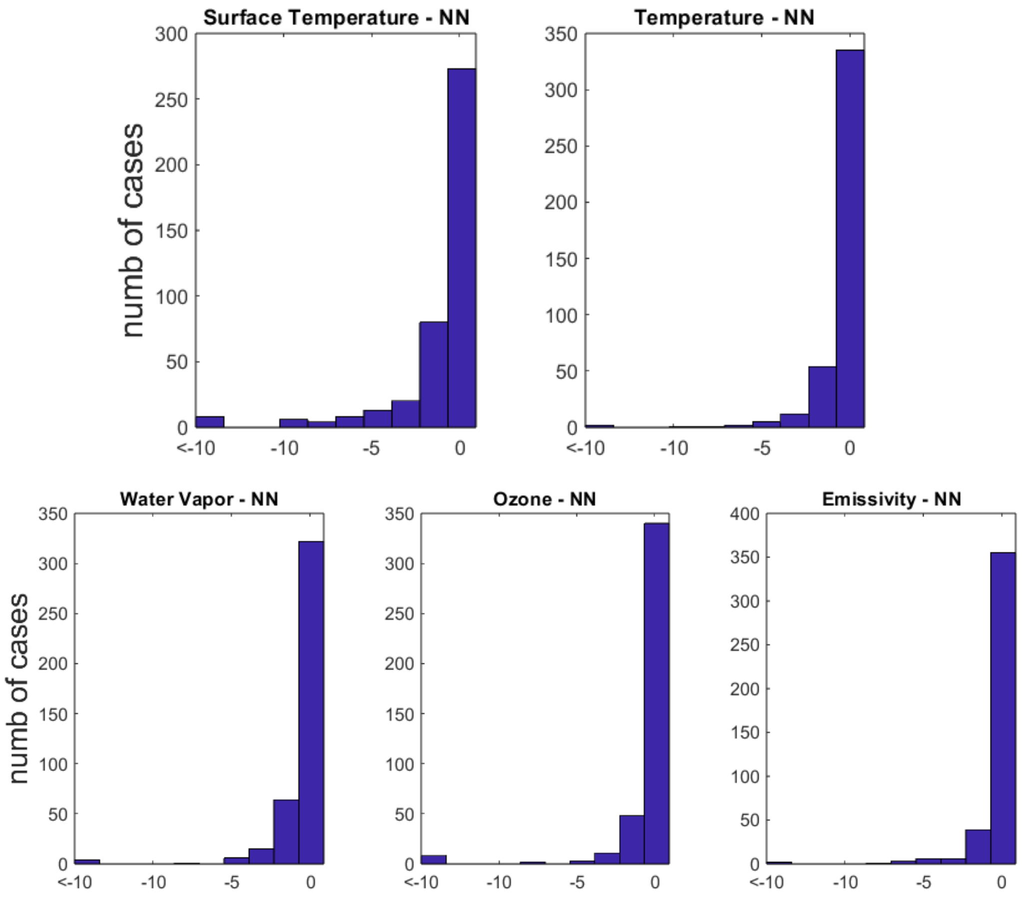
- Given training data $\left(\hat{x}_j, (x_a)_j, \lambda_j^{opt}\right)_{j=1}^{J}$ the following equation is solved:

$$\tilde{\theta} = \arg\min_{\theta} \frac{1}{J} \sum_{j=1}^{J} \left\| \Phi\left(\widehat{x_j}, (x_a)_j, \lambda_j^{opt}, \theta\right) \right\| .$$

!

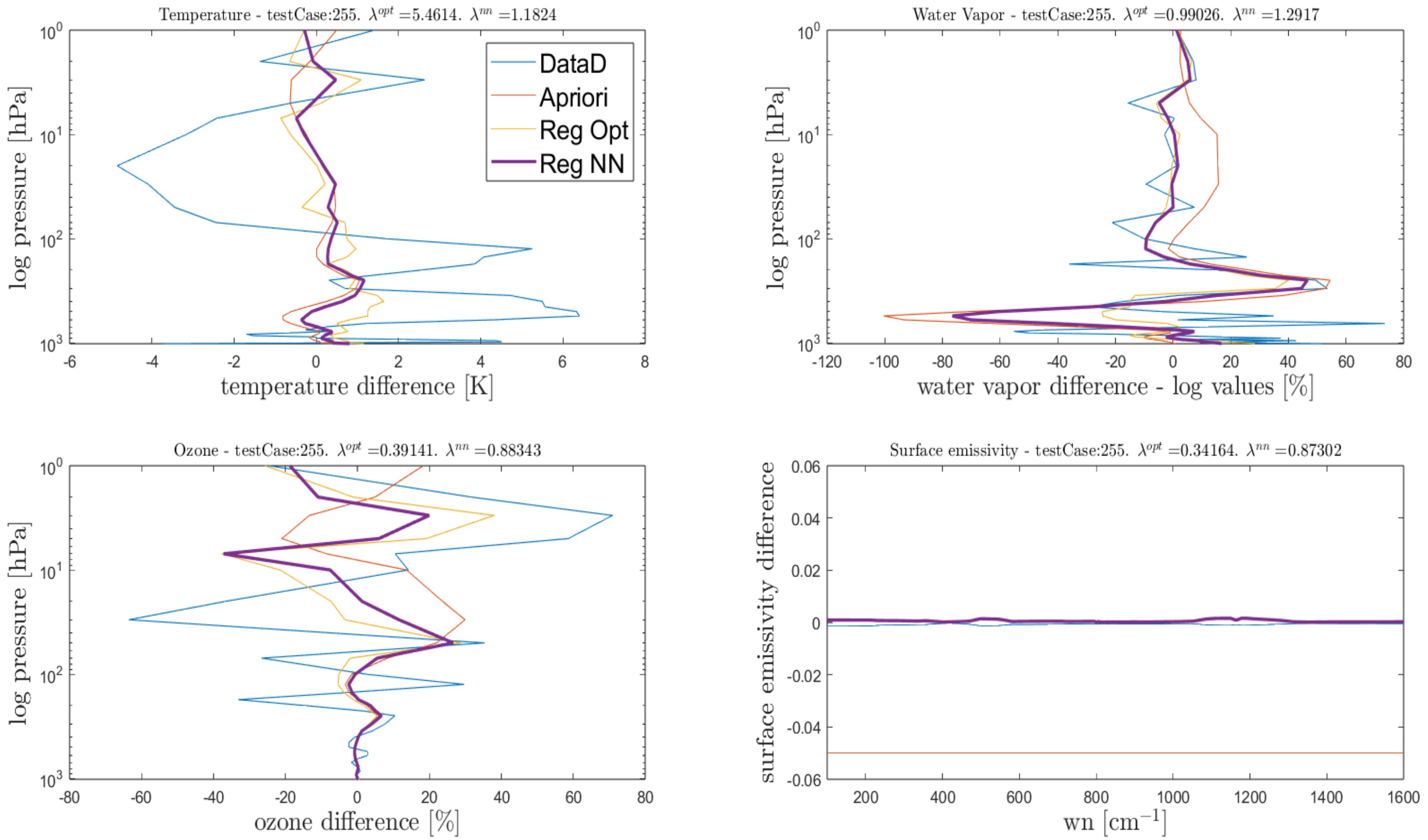
Neural Network INPUT:	$\hat{x} - x_a$
Neural Network OUTPUT (prediction):	$\log(\lambda^{nn})$
Neural Network ARCHITECTURE:	3 layers (dim: 15,10,5).

Neural Network performance for training set 2



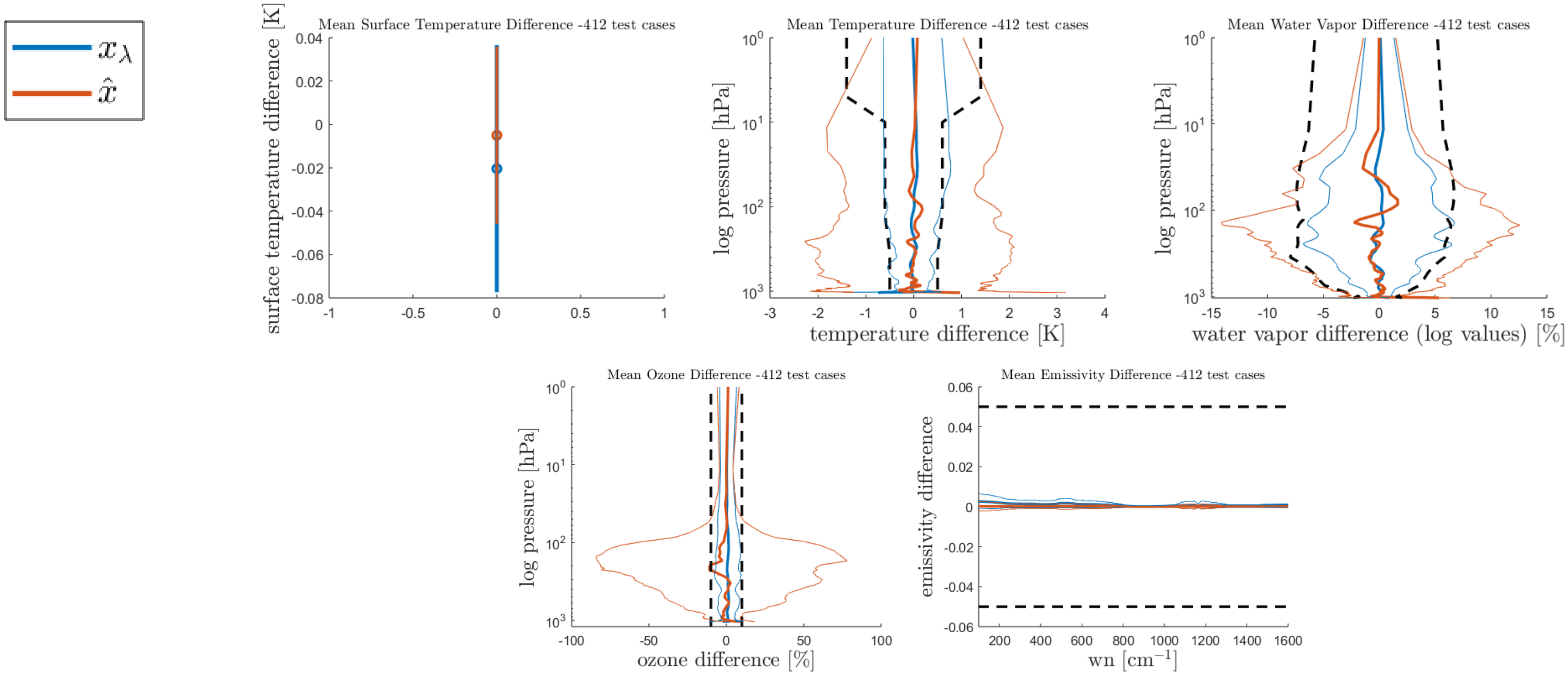
$(\lambda^{OPT} - \lambda^{NN}) / \lambda^{NN}$

Comparison for a single case in the training set 2

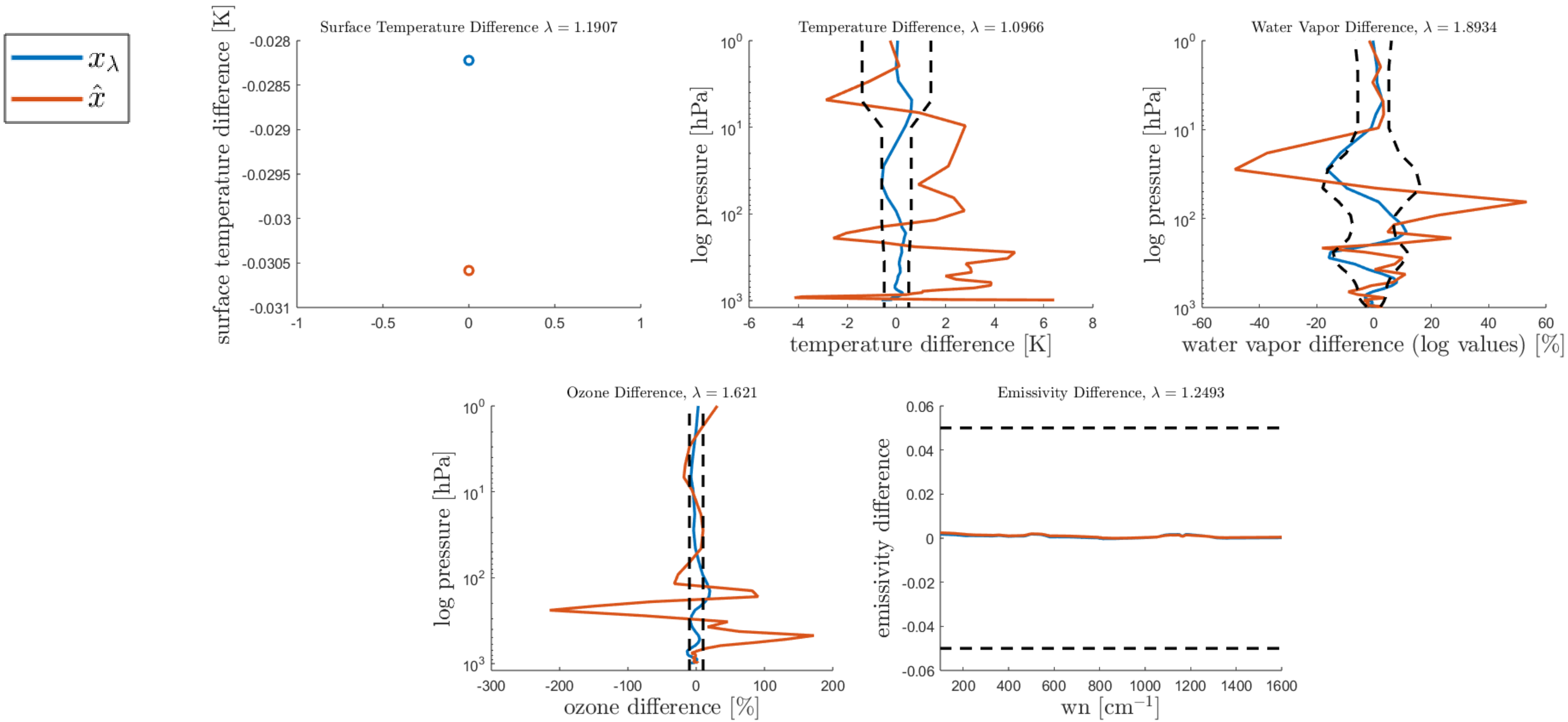


Results – Aggregated cases

Mean signed (normal) and unsigned (bold) errors for a global test set 2 vs Apriori errors (dotted)



Results - CASE 5 in test set 2



Future directions

- Comparison with full-physics methods.
- Extension to all-sky conditions.
- Adaptation for use with different instruments.
- Application of a similar data-driven approach with additional a-priori information to the direct problem.
- Incorporating this work into the analysis of fast radiative transfer models for data assimilation techniques into climate and meteorological models as part of the PNRR-EMM project.

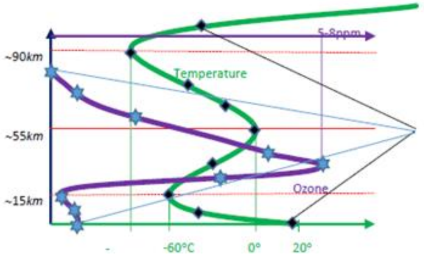
Thanks to INdAM-GNCS, Math Research Group, for financial support

Introduction – FORUM mission

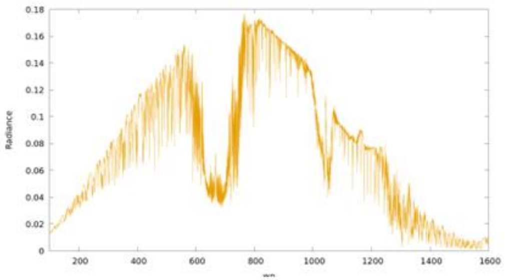
- FORUM[1] (Far-infrared Outgoing Radiation Understanding and Monitoring) is a Fourier Transform Spectrometer (FTS) selected as the ninth Earth Explorer mission by the European Space Agency in 2019.
- It will provide interferometric measurements in the Far-InfraRed (FIR) spectrum (100-1600 cm^{-1} region), constituting 50% of Earth's outgoing longwave flux.
- Accurate Top Of the Atmosphere (TOA) measurements in the FIR are crucial for improving climate models.

[1] L. Sgheri et al. "The FORUM end-to-end simulator project: architecture and results". In: Atmospheric Measurement Techniques 15.3 (2022), pages 573–604. doi: 10.5194/amt-15-573-2022. url: https://amt.copernicus.org/articles/15/573/2022/

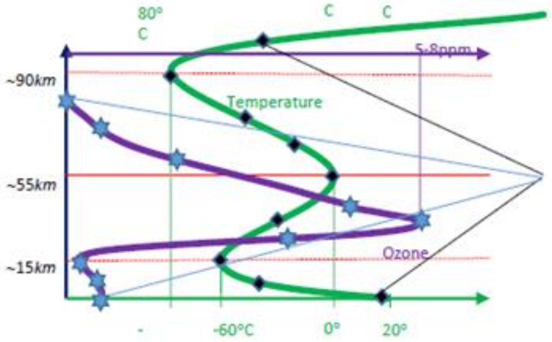
Background



$\mathbf{x} \xrightarrow{\mathbf{F}} \mathbf{y} = \mathbf{F}(\mathbf{x})$



$\min \|y - F(x)\|$



Background – Radiative Transfer (RT), forward model

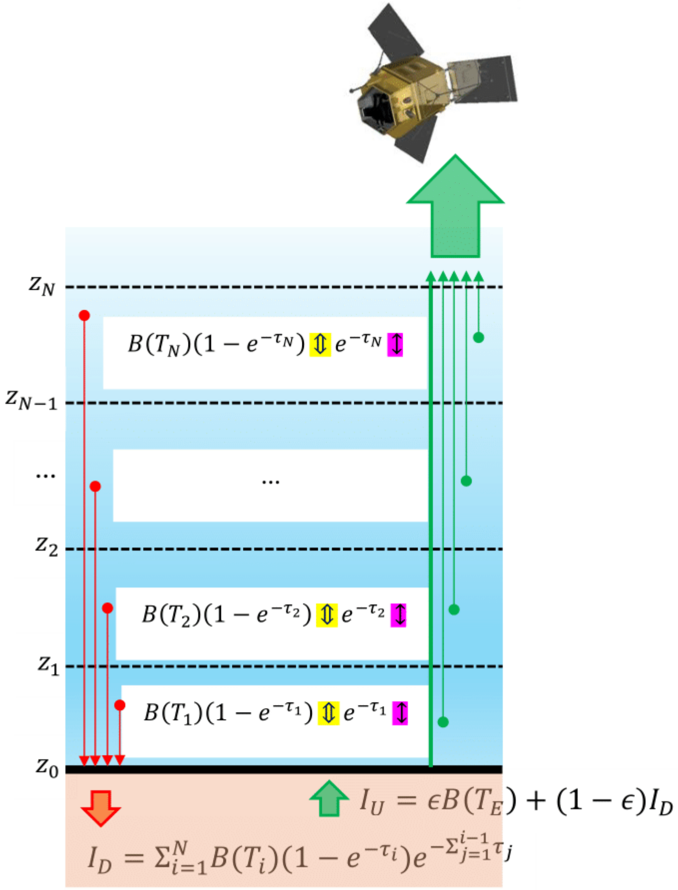
$$\begin{cases} \dfrac{\mathrm{d}I_\nu}{dz}(z) = -\alpha_\nu(p,T,c)I_\nu(z) + \alpha_\nu(p,T,c)B_\nu(T) \\ \\ I_\nu(z_0) = I_{\nu_0}, \end{cases}$$

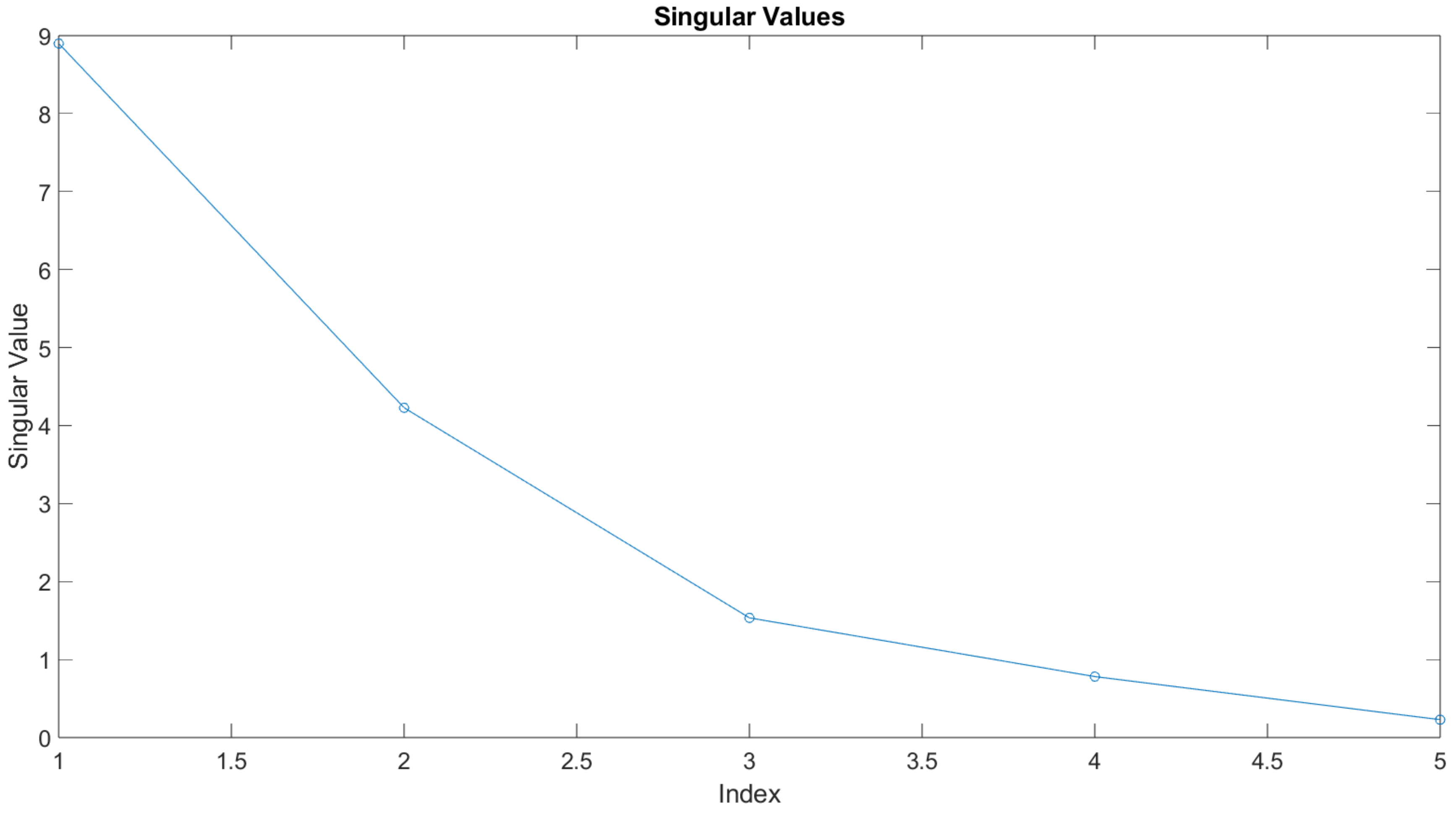
for each atmospheric layer, with, ν wavenumber, z altitude, I intensity of radiation,

B Planck function, α attenuation coefficient, p pressure, T temperature, c gases concentration.

$$I(z_N) = \left[\epsilon B(T_E) + (1-\epsilon)\left(\sum_{i=1}^{N} B(T_i)(1-e^{-\tau_i})e^{-\sum_{j=1}^{i-1}\tau_j}\right)\right]e^{-\sum_{i=1}^{N}\tau_i}$$
$$+\sum_{i=1}^{N} B(T_i)\left(1-e^{-\tau_i}\right)e^{-\sum_{j=i+1}^{N}\tau_j},$$

with τ optical depth and T_E Earth surface temperature.




Singular Values
Singular Value
Index
0
1
2
3
4
5
6
7
8
9
1
1.5
2
2.5
3
3.5
4
4.5
5