Koninklijk Nederlands Meteorologisch Instituut
Ministerie van Infrastructuur en Waterstaat

esa

EUROPEAN WEATHER CLOUD
CLOUD COMPUTING-BASED INFRASTRUCTURE, FOCUSED ON THE NEEDS OF THE METEOROLOGICAL COMMUNITY

An algorithm of scattering cloud retrieval based on Neural Network for TROPOMI using Oxygen absorption band

Xiaoyun Zhang, Ping Wang, Tao Xie, Piet Stammes,
Gijsbert Tilstra, Olaf Tuinder, Maarten Sneep, Feng Lu

ping.wang@knmi.nl

xiaoyun.zhang@knmi.nl

July 4, 2024

Outline

- Motivation
  - FRESCO+ - - FRESCO-S - - FRESCO-N
- Introduction
  - DISAMAR & FRESCO-N model
- Algorithm - - FRESCO-N
  - Forward model - - NN Sequential model
    - DISAMAR & Training datasets
  - Retrieval model - - Optimal Estimation
  - TROPOMI datasets & case scenario data
- Results
  - Predicted High resolution spectra
  - Retrieval
  - Compare with Fresco, DISAMAR, Suomi NPP Cloud Products
- Summary

Motivation

➢ FRESCO+ - - FRESCO-S - - FRESCO-N

Fast Retrieval Scheme for Clouds from the Oxygen A-band

https://doi.org/10.5194/acp-8-6565-2008

14 Nov 2008

FRESCO+: an improved O_2 A-band cloud retrieval algorithm for tropospheric trace gas retrievals

P. Wang, P. Stammes, R. van der A, G. Pinardi, and M. van Roozendael

Abstract. The FRESCO (Fast Retrieval Scheme for Clouds from the Oxygen A-band) algorithm has been used to retrieve cloud information from measurements of the O_2 A-band around 760 nm by GOME, SCIAMACHY and GOME-2. The cloud parameters retrieved by FRESCO are the effective cloud fraction and cloud pressure, which are used for cloud correction in the retrieval of trace gases like O_3 and NO_2. To improve the cloud pressure retrieval for partly cloudy scenes, single Rayleigh scattering has been included in an improved version of the algorithm, called FRESCO+. We compared FRESCO+ and FRESCO effective cloud fractions and cloud pressures using simulated spectra and one month of GOME measured spectra. As expected, FRESCO+ gives more reliable cloud pressures over partly cloudy pixels. Simulations and comparisons with ground-based radar/lidar measurements of clouds show that the FRESCO+ cloud pressure is about the optical midlevel of the cloud. Globally averaged, the FRESCO+ cloud pressure is about 50 hPa higher than the FRESCO cloud pressure, while the FRESCO+ effective cloud fraction is about 0.01 larger.

The effect of FRESCO+ cloud parameters on O_3 and NO_2 vertical column density (VCD) retrievals is studied using SCIAMACHY data and ground-based DOAS measurements. We find that the FRESCO+ algorithm has a significant effect on tropospheric NO_2 retrievals but a minor effect on total O_3 retrievals. The retrieved SCIAMACHY tropospheric NO_2 VCDs using FRESCO+ cloud parameters (v1.1) are lower than the tropospheric NO_2VCDs which used FRESCO cloud parameters (v1.04), in particular over heavily polluted areas with low clouds. The difference between SCIAMACHY tropospheric NO_2 VCDs v1.1 and ground-based MAXDOAS measurements performed in Cabauw, The Netherlands, during the DANDELIONS campaign is about -2.12×10^{14}molec cm^{-2}.

Wang, P., Stammes, P., van der A, R., Pinardi, G., and van Roozendael, M.:
FRESCO+: an improved O_2 A-band cloud retrieval algorithm for tropospheric trace gas retrievals,
Atmos. Chem. Phys., 8, 6565–6576, https://doi.org/10.5194/acp-8-6565-2008, 2008.

Fresco+

- Cloud Effective Fraction
- Cloud Pressure
- Lambertian Clouds(albedo 0.8) & Single Rayleigh Scattering
- Fit limit wavelengths around 760nm 758-759, 760-761, 765-766nm

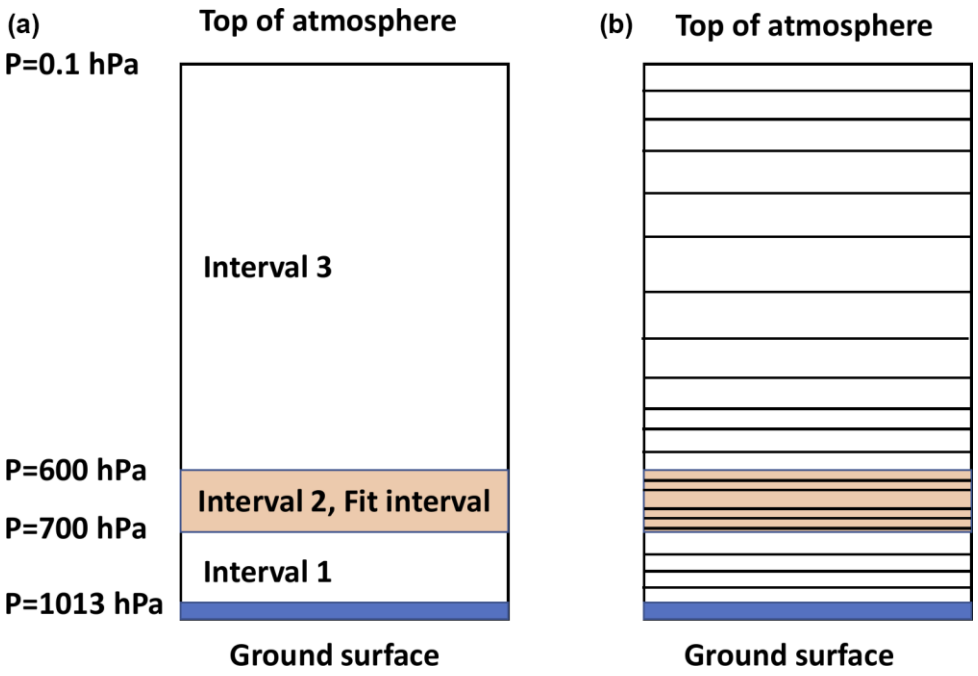
$$R_{\text{sim}} = (1-c)T_s A_s + (1-c)R_s + cT_c A_c + cR_c.$$

Fresco-N

- Cloud Fraction
- Cloud Pressure
- Cloud Optical Thickness
- Scattering Clouds(HG-Phase Function)
- High resolution spectra from TROPOMI O_2-A: 757-770nm & O_2-B: 685-691nm
- A NN-model & OE

Introduction

- DISAMAR - Scattering Cloud - HG phase function



In DISAMAR the atmosphere is vertically divided into pressure intervals.
Each interval is divided into a number of homogeneous layers

DISAMAR Forward Model

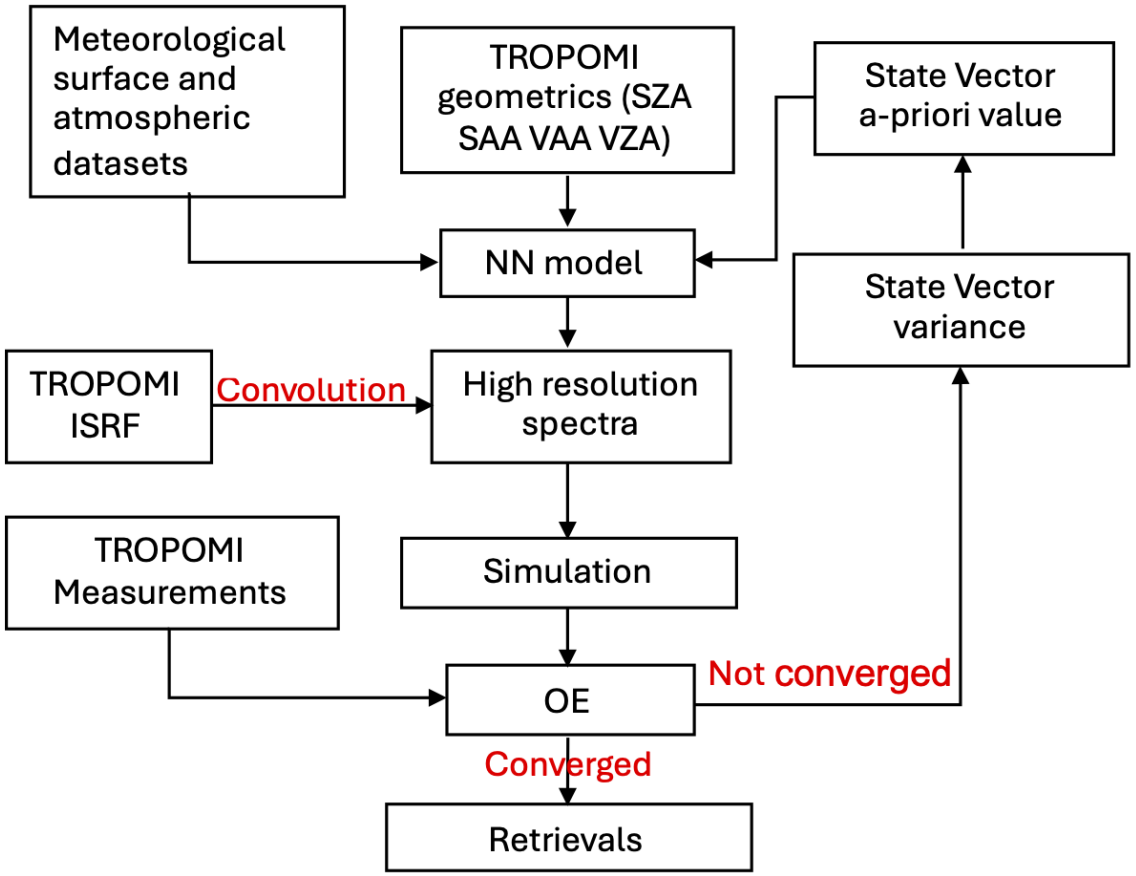
- Layers
  - *an opaque Lambertian reflector*
  - a layer of scattering particles
    - *Mie phase function*
    - *Henyey-Greenstein phase function*
- *Double Adding Method* ⇩ g = 0.85, number of streams = 20 (10 Gaussian Points)
  - simulated radiance & irradiance
  - derivatives of the reflectance with respect to the state vector elements.

DISAMAR Retrieval Model

- *Optimal Estimation (OE)*
- *Differential Optical Absorption Spectroscopy (DOAS)*
- *Differential and Smooth Absorption Separated (DISMAS) methods*

Introduction

- FRESCO-N



FRESCO-N flow chart

Meteorological Datasets
- Surface pressure
- Temperature profile

  - - From ERA5 hourly datasets

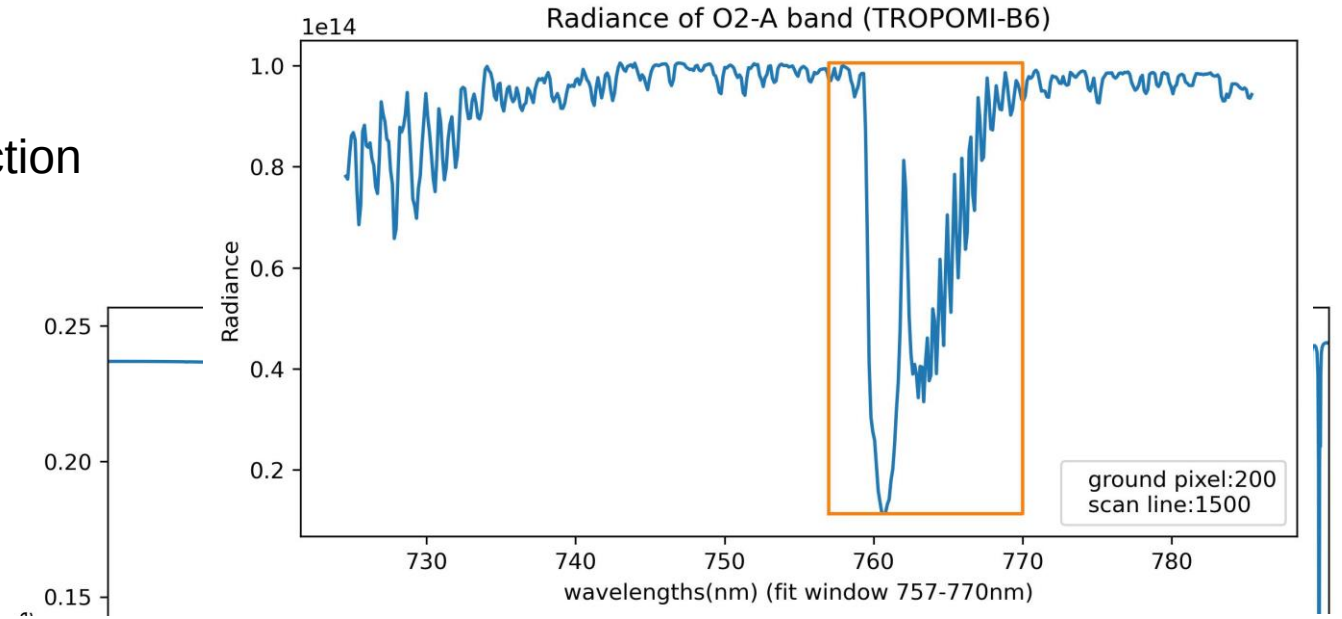
TROPOMI - ground pixel & scanline
- O_2-A band Radiance
- Irradiance
- ISRF
- SZA SAA VAA VZA

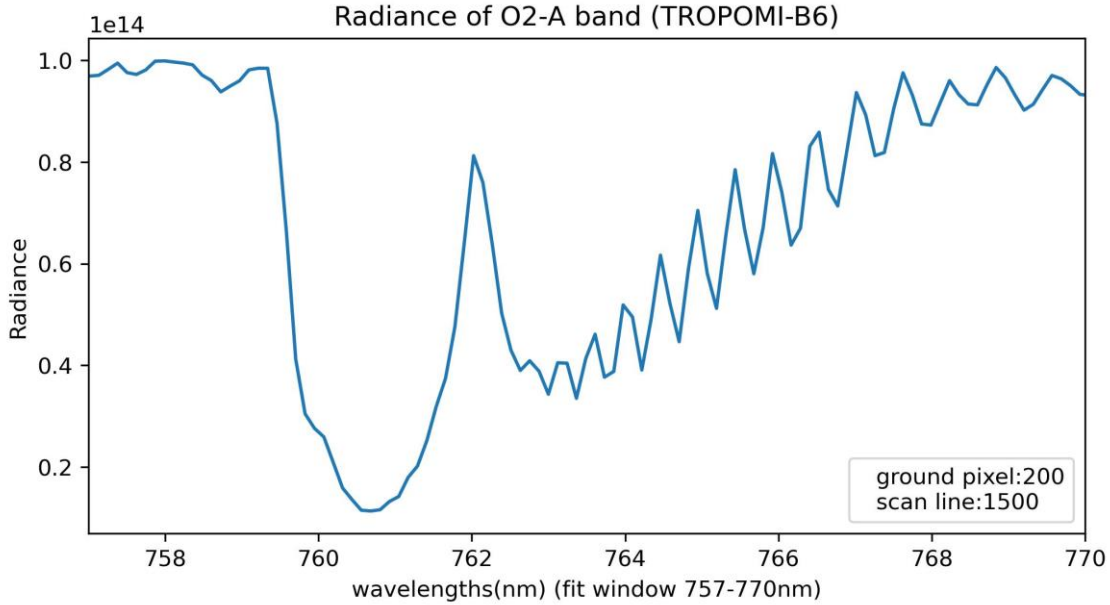
State vector (fit parameters)
- Cloud optical thickness (COT)
- Cloud pressure (CLP)
- Cloud fraction (CLF)

Spectra predicted by NN models (Cloud Free & Fully cloudy)
- Reflectance
- Derivatives of Reflectance with respect to state vectors

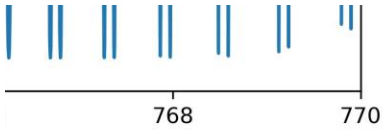
FRESCO-N

- ➢ Forward Model - training dataset collection
- Fit Window
  - O_2-A : from 757.0nm to 770.0nm
- Cross Section Database
  - HITRAN2020
- Henyey-Greenstein phase function
  - Asymmetric factor (g): 0.85
  - Gaussian points: 10





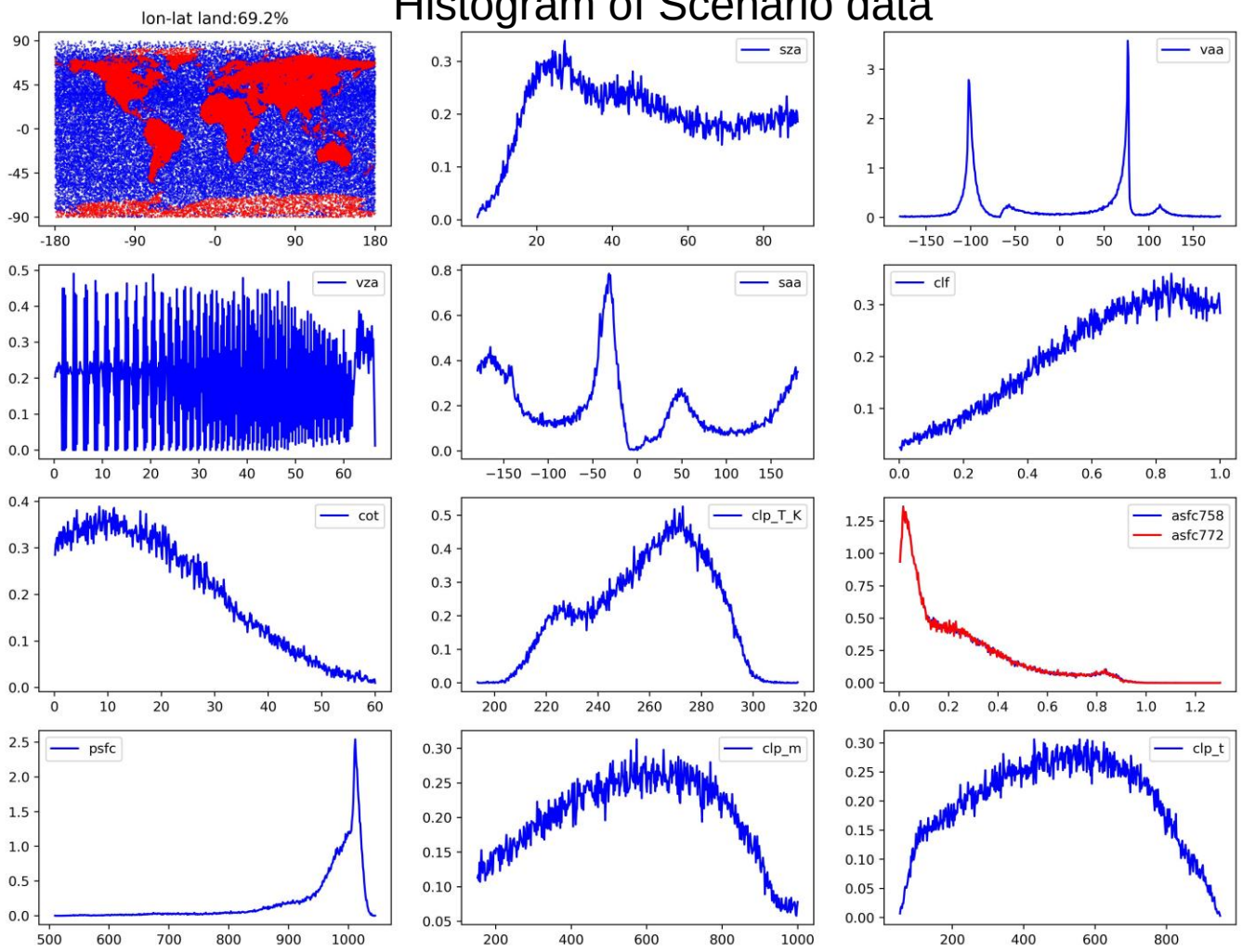
Radiance from
TROPOMI Band 5 & Band 6
orbit 24701
ground pixel number 200
scan line number 1500



FRESCO-N

- Training Datasets preparation - - Scenario data for spectra simulation

Histogram of Scenario data



400, 000 cases in total
Land surface cases: about 70%
Water surface cases: about 30%
Meteorological data read from ERA5

Geometrics of TROPOMI:
Selected randomly from one TROPOMI Orbit

State Vector (Fit parameters)
Cloud Fraction (CLF)
Cloud Optical Thickness (COT)
Cloud Pressure (middle) (CLP)
Cloud Pressure Thickness (intervalDP)

O_2-A 757-770nm
HIT2020

FRESCO-N

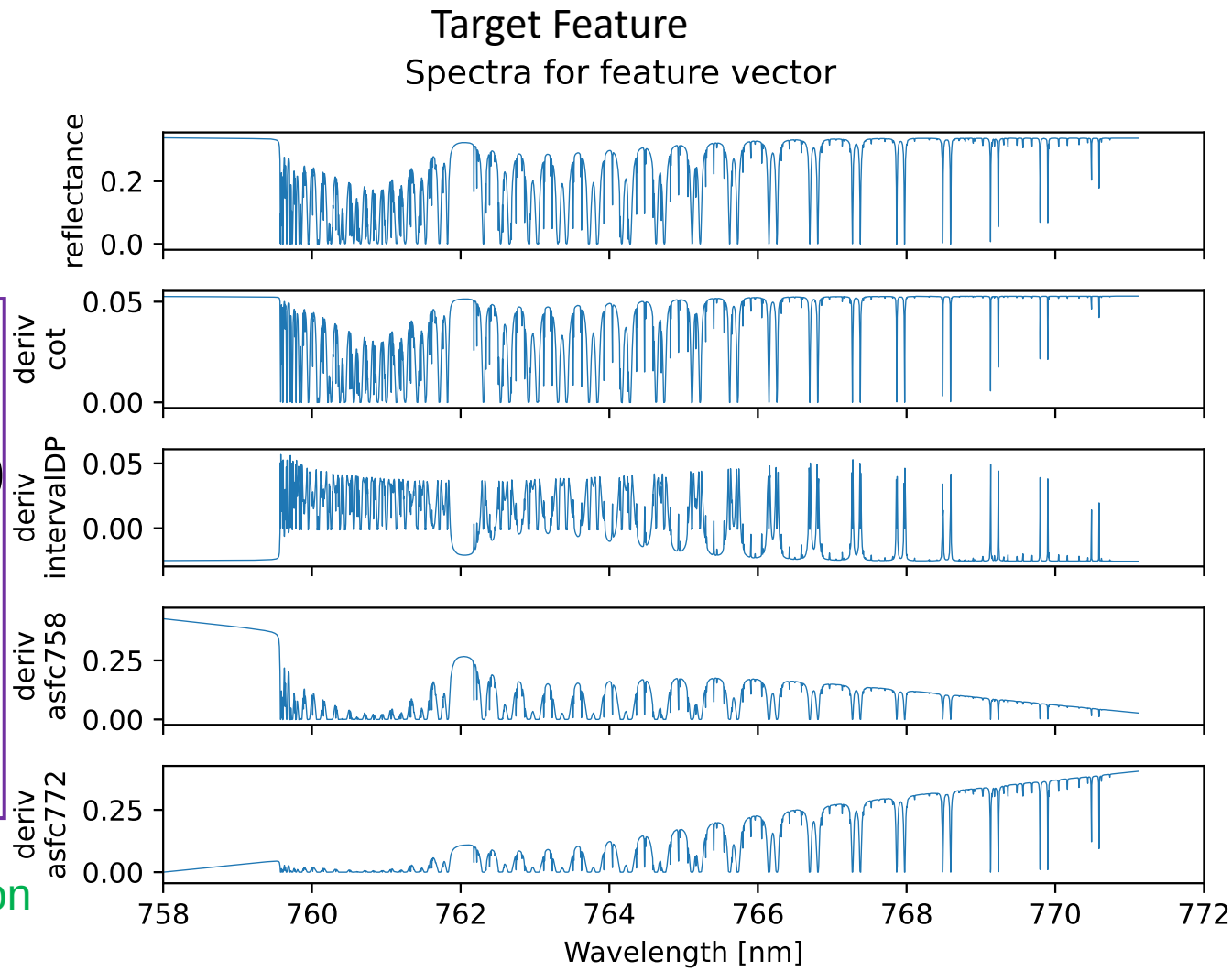
- Forward model - - NN Sequential model

Input Feature Vector:

- surface pressure psfc (hPa) REA5
- cloud pressure CLP (cloud middle) (hPa)
- cloud temperature CLP_T_K(P) (K)
  - - interpolate from temperature profile
- mu: $\cos SZA$ TROPOMI
- mu0: $\cos VZA$
- scattering angle θ (degree)
  $\cos^{-1}(\sin SZA * \sin VZA * \cos(180 - (SAA - VAA)) - \cos SZA * \cos VZA)$
- cloud optical thickness
- HG phase function
  $(1 - g^2) \times (1 + g^2 - 2g\cos\theta)^{-1.5} \times 0.25/\pi$
- transmission
  $e^{(-1*COT*\frac{1}{\cos VZA}+\frac{1}{\cos SZA})}$
- surface albedo at 758nm & 772nm

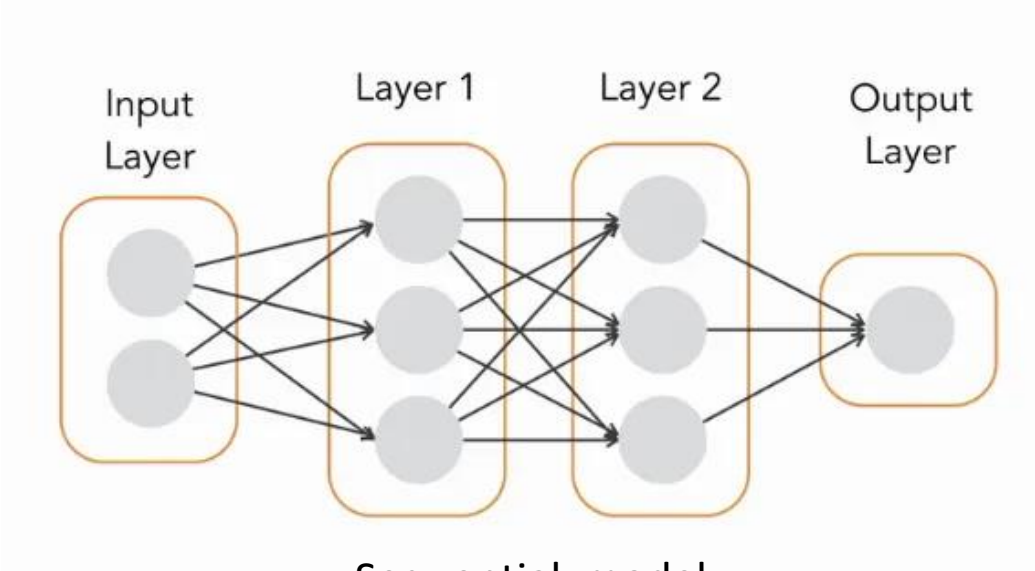
Created follow standard normalize distribution

- reflectance
- derivatives of reflectance with respect to state vectors



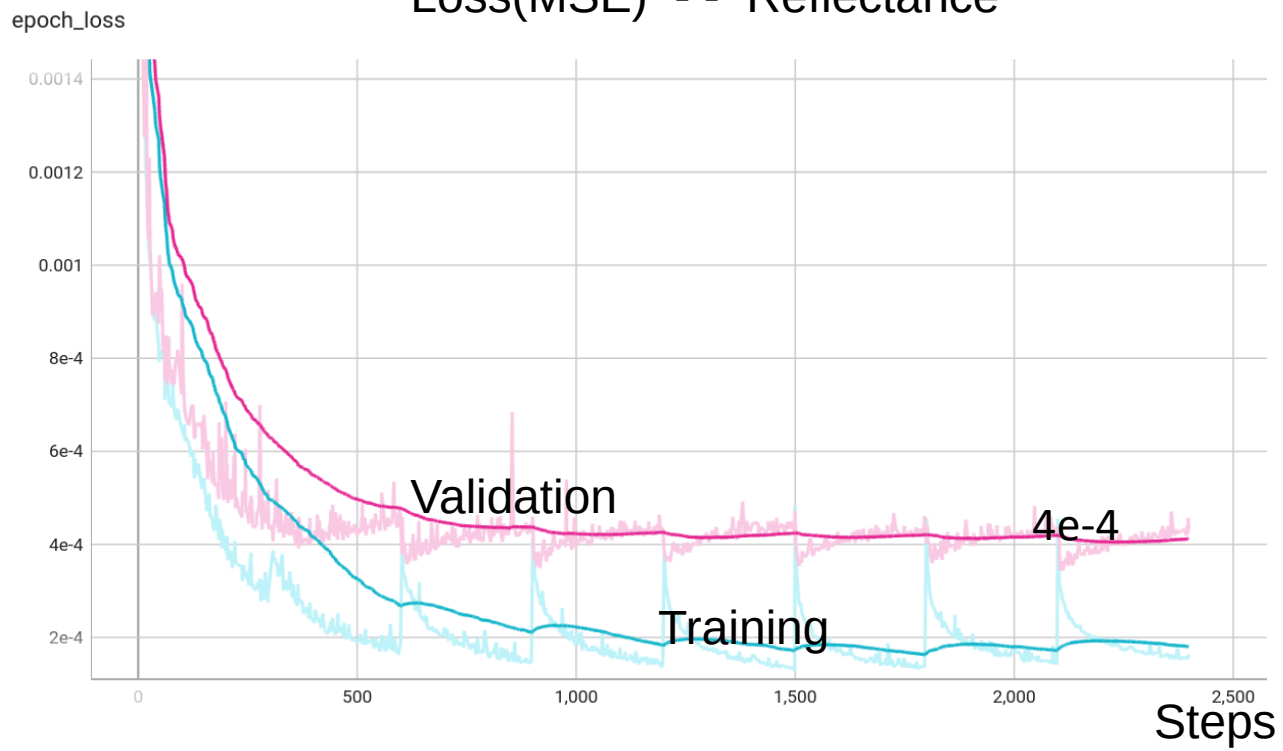
FRESCO-N

- Forward Model - - NN Sequential model



Sequential model

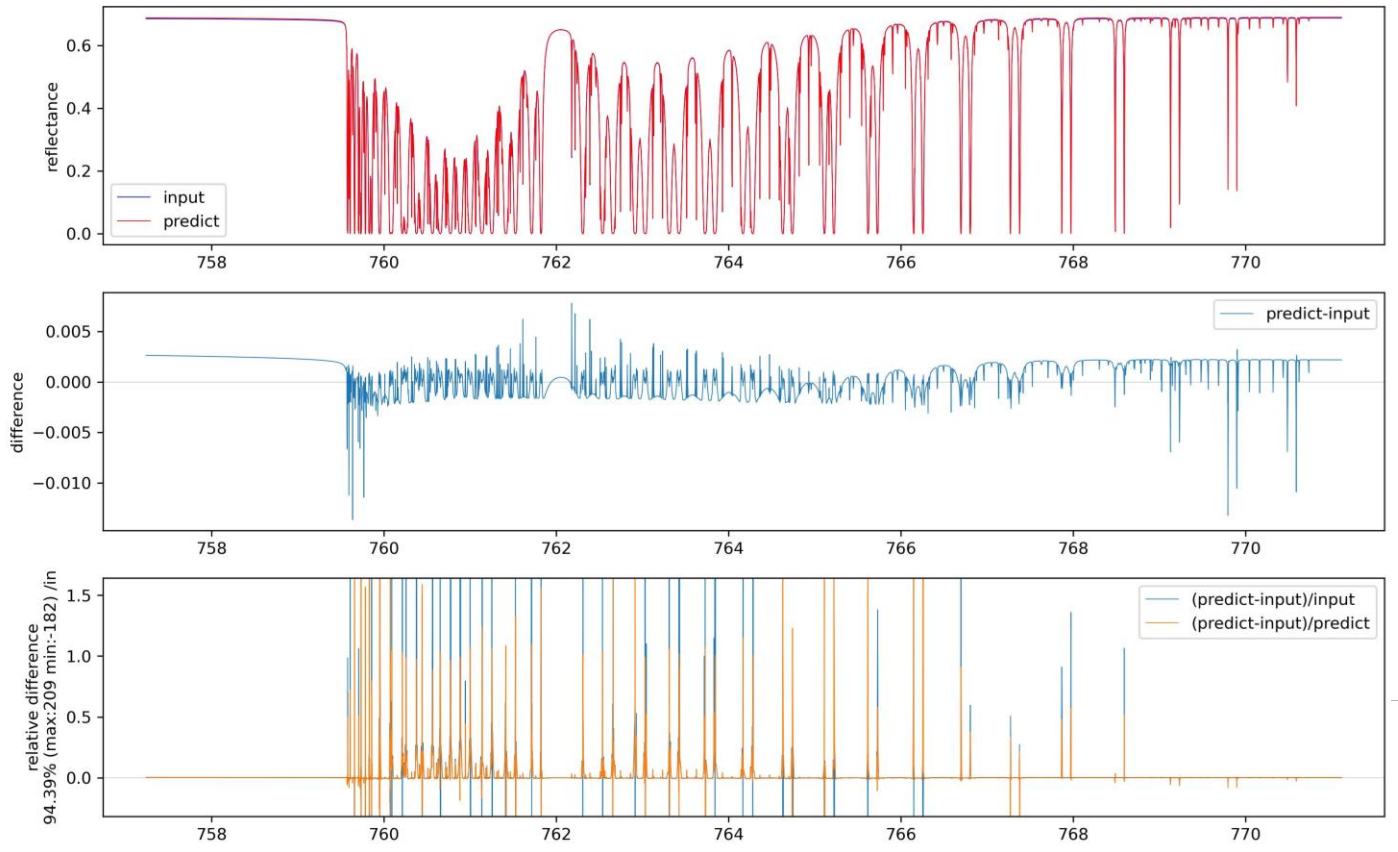
Loss(MSE) - - Reflectance

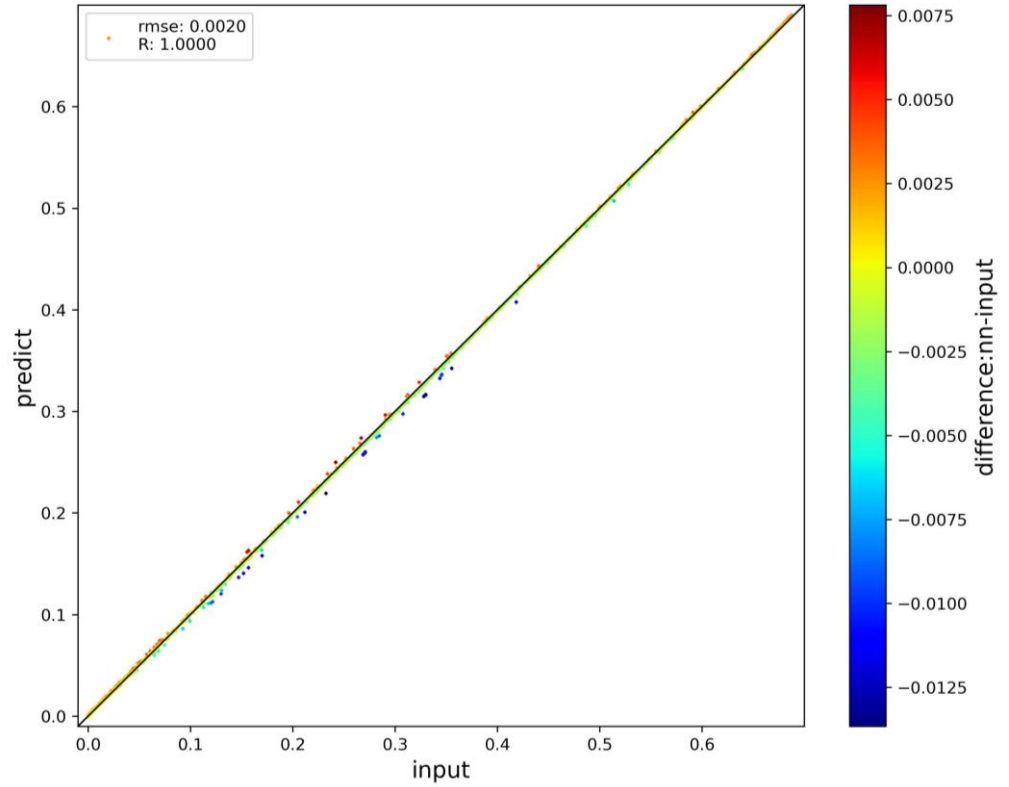


Layers: [256,256,256] epoch:300, batch size:600

FRESCO-N

- Forward Model - - Predicted High resolution spectra





scat_angle: 131.4
mu: 0.41
mu0: 0.63

asfc758: 0.086
psfc: 1014.1hPa

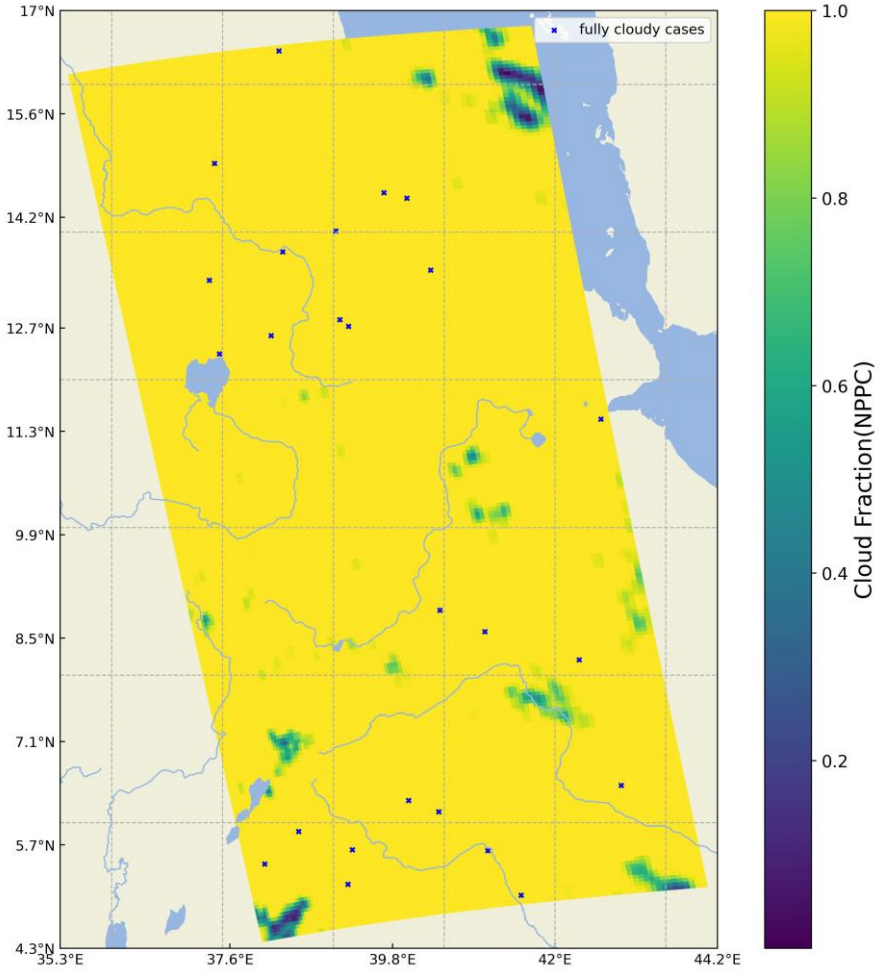
clp_T_K: 285.7K
clp_m: 741.84hPa
cot: 22.45
hg_phase_func: 4.86e-3
trans_cld: 5.4e-40

Results

- Test FRESCO-N model -- Fully cloud

Cases distribution:

- Skip cloud edge
- SZA<75
- Fully cloudy
- Thick clouds
- 40 cases



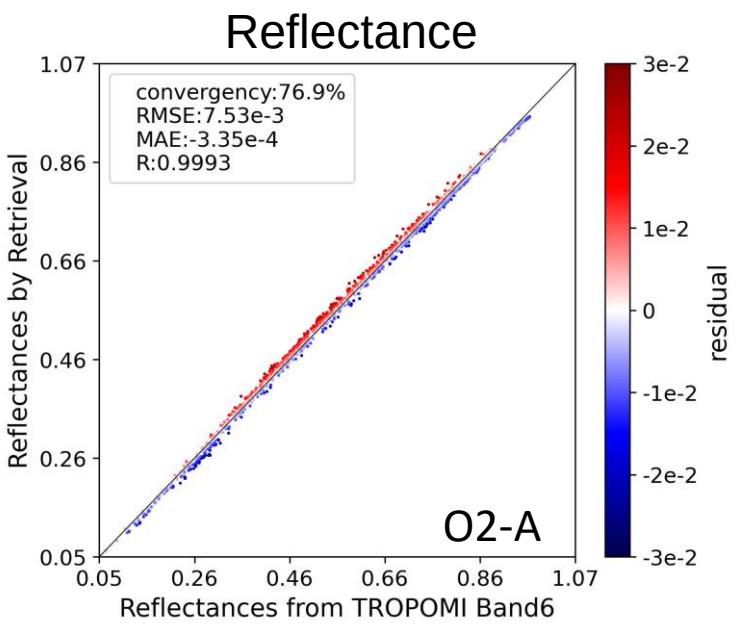
Input case data for DISAMAR:

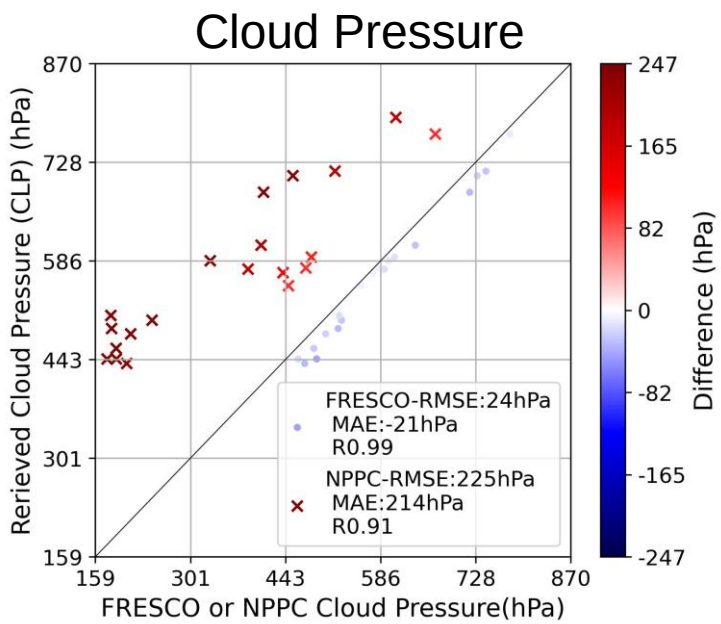
- Surface albedo (DLER)
- Surface Pressure (ERA5)
- Temperature profile (ERA5)
- Geometrics from TROPOMI:
  - SZA
  - SAA
  - VZA
  - VAA
- TROPOMI Radiance & Irradiance
- Slit Function (calculated per ground pixel) (ISRF)

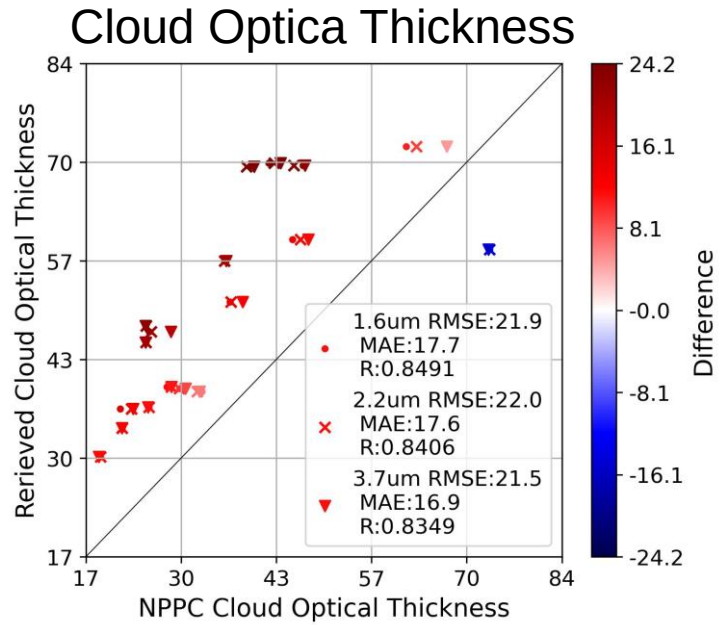
- Interval top & base pressure (pressure thickness of fit interval)
  - 100hPa
- A-priori value and variance of COT
  - 30/5

Results

- DISAMAR Retrievals compared to Fresco & NPPC products



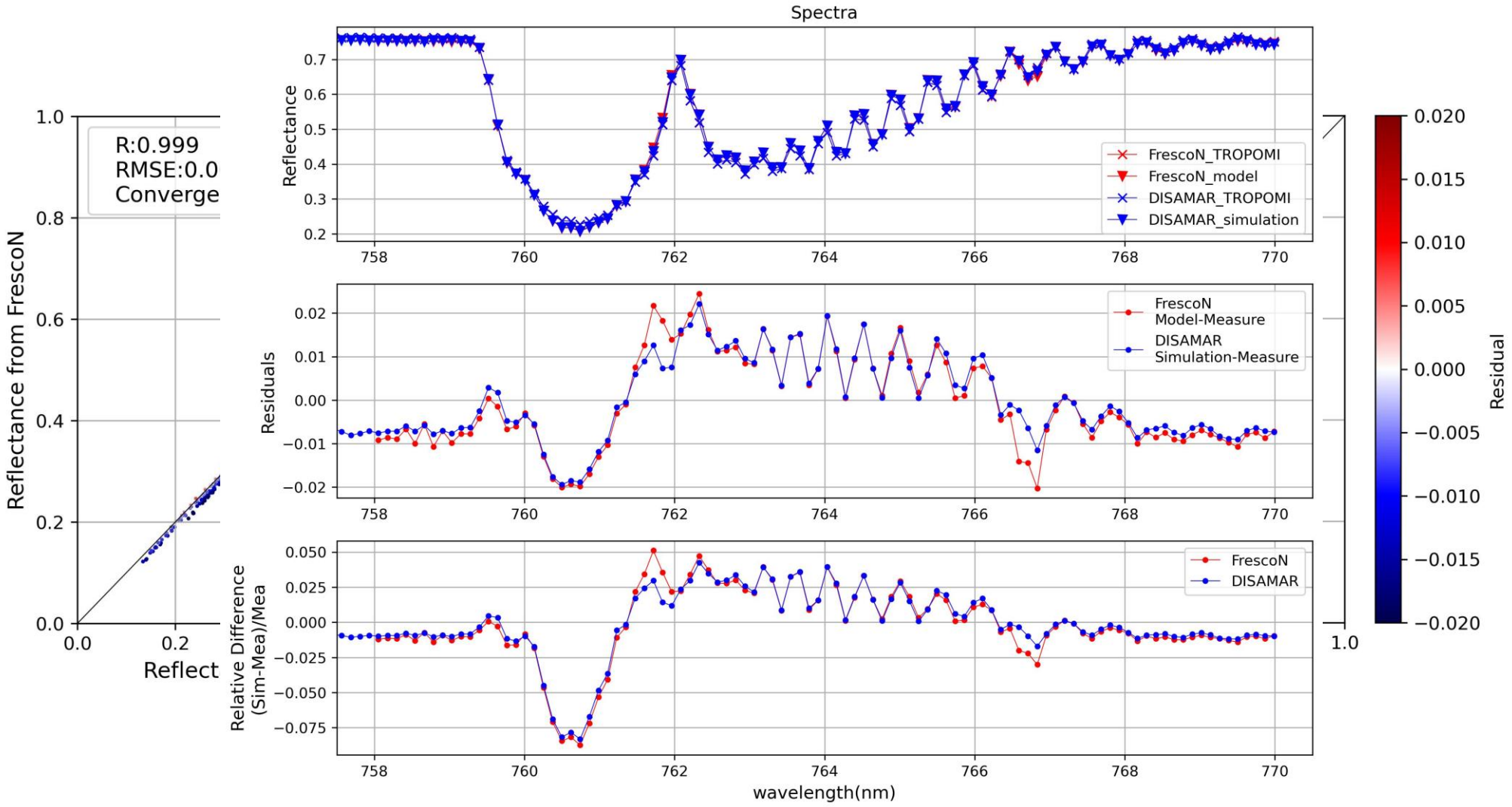




g=0.85 & 20 Gaussian points

Results

- FRESCP-N Retrievals compared to DISAMAR & Fresco & NPPC products

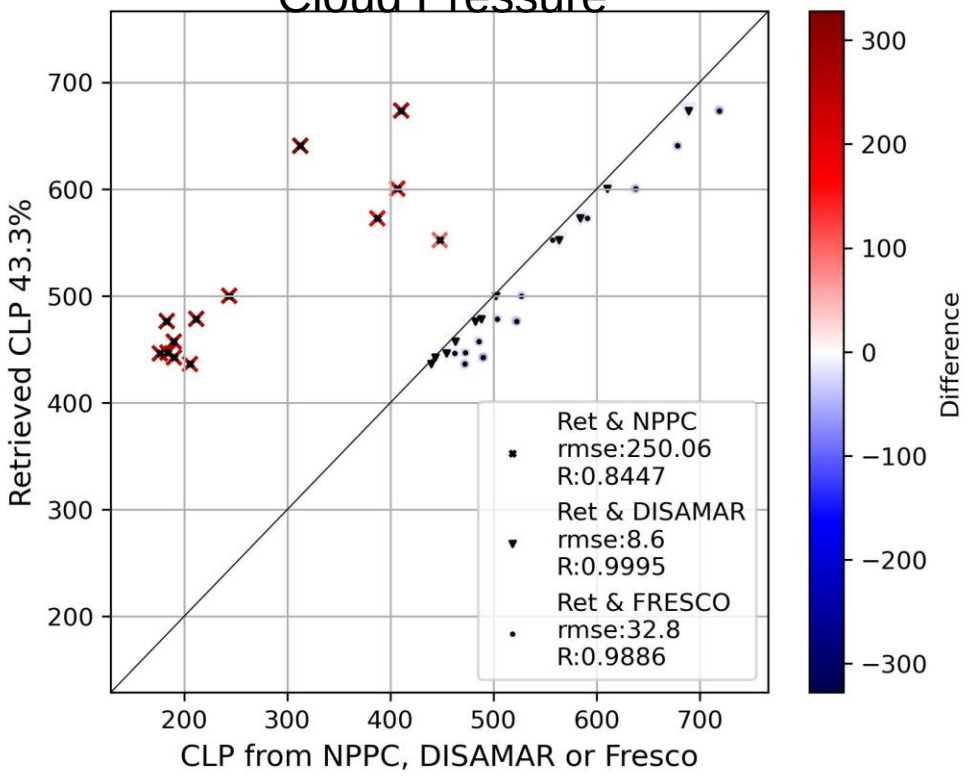


Results

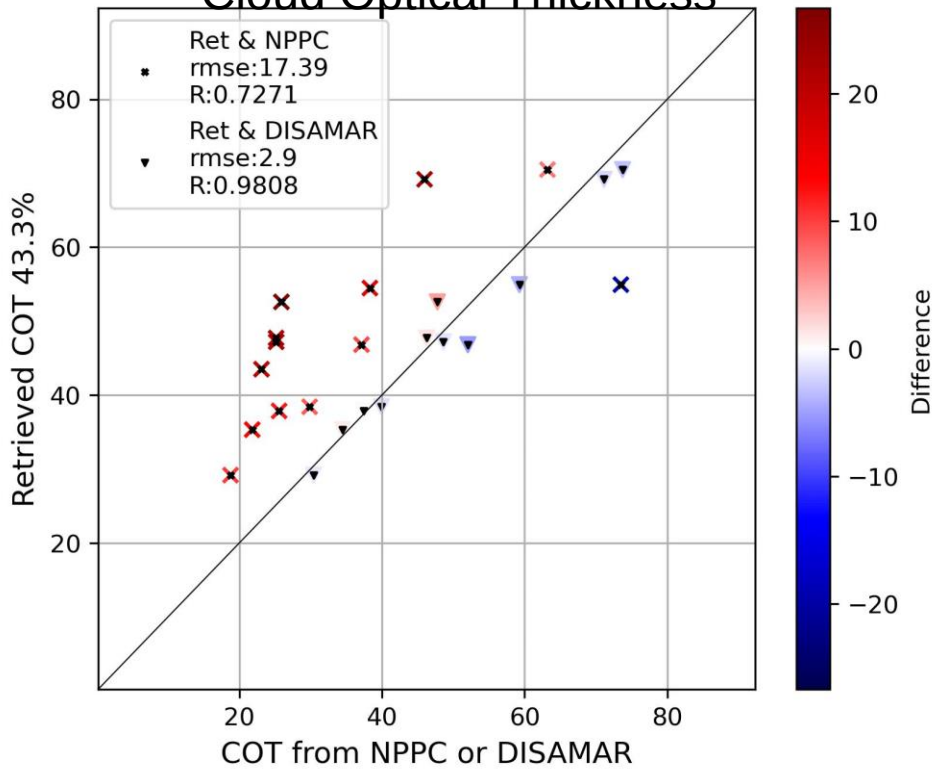
- FRESCP-N Retrievals compared to DISAMAR & Fresco & NPPC products

State Vector (fit parameters)

Cloud Pressure



Cloud Optical Thickness



Summary

- Generally, replacing the line-by-line radiative transfer model (DISAMAR) with a neural network (NN) model can significantly reduce the time consumption. The runtime for one case decreases from over 3 hours to just several seconds.

- Although the spectra predicted by the NN model show a high correlation with the DISAMAR simulation, the loss fluctuates during the training process. There is still a lot of work to be done to train a more reliable NN model.

- The target features (spectra) of the training datasets were calculated with an asymmetric factor of 0.85 and 10 Gaussian points. Although this approach is much faster than using 20 Gaussian points, the convergence in FRESCO-N becomes worse. This is similar to what we did in the DISAMAR retrieval.

- The retrievals of COT show a correlation coefficient of 0.98 with the NPP-COT data, while CLP shows a correlation coefficient of 0.98 with the values retrieved by FRESCO and DISAMAR.

Thanks for your attention!