Retrieval of Cloud Properties for the Copernicus Atmospheric Missions Sentinel-4 (S4) and TROPOMI / Sentinel-5 Precursor (S5P) using deep neural networks

Fabian Romahn[1], Diego G. Loyola[1] , Adrian Doicu[1], Victor Molina Garcia[1], Ronny Lutz[1], Athina Argyrouli[2]

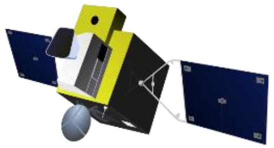
(1) German Aerospace Center (**DLR**), (2) Technical University of Munich (**TUM**)

Copernicus Satellite Sentinel-5 Precursor (S5P) and Sentinel-4 (S4)

Sentinel-4 (S4) and Sentinel-5 Precursor (S5P) are passive earth observation satellites for trace gas retrieval of the Copernicus programme:

Sentinel-4

- launch date in 2025
- geostationary orbit facing europe
- spectral range: UV/VIS/NIR
- spatial resolution: 8 km x 8 km

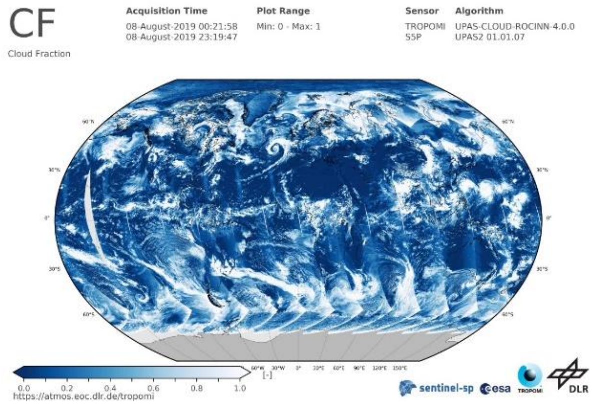


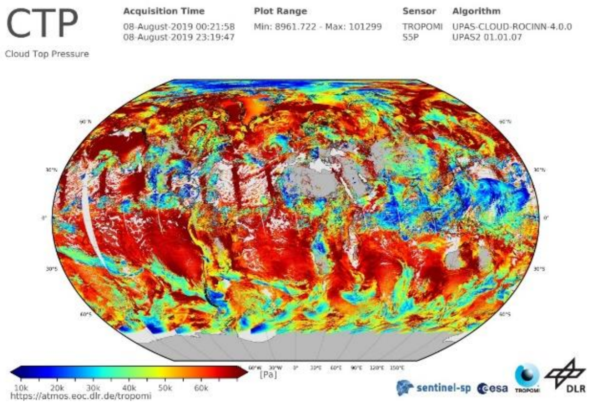
Sentinel-5 Precursor

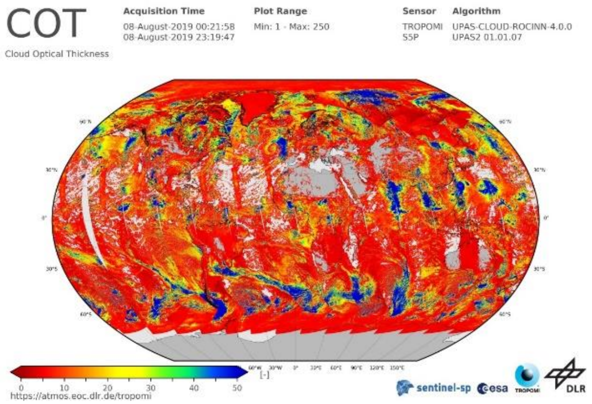
- launched in october 2017
- sun synchronous orbit at ~ 820 km
- spectral range: UV/VIS/NIR/SWIR
- spatial resolution: 5.5 km x 3.5 km

A requirement for trace gas retrieval is accurate cloud information

→ DLR is responsible for the operational **CLOUD product**







Challenges:

- Large amounts of data
- Near real time requirements (NRT)

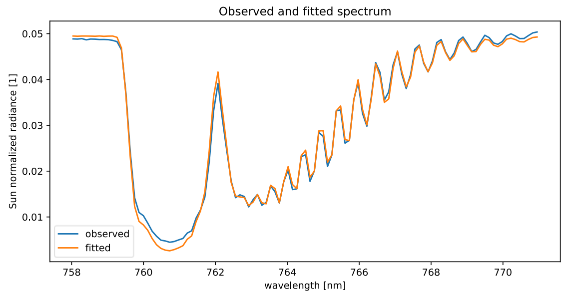
→ Application of machine learning techniques to improve performance compared to classical algorithms

Application of neural networks

Problem:

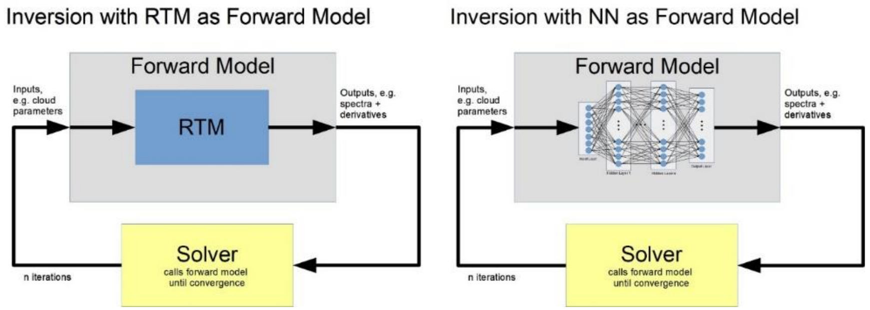
Find parameters x that minimize residual $\|F(x) - y\|_2$ between a known vector y and the mapping of the parameters $F(x)$ – where F is a predefined function

in remote sensing:
x: state of atmosphere, y: measured spectrum, F: radiative transfer model (RTM)

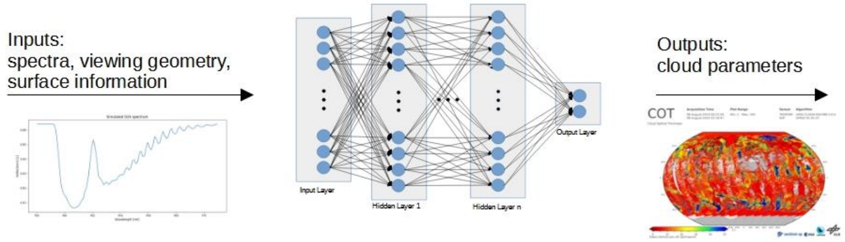


Two approaches:

1. NN as **forward model** of a spectral fitting algorithm:
   - $F: X \rightarrow Y$ *state of atmosphere → spectrum*
   - substitutes and approximates the RTM
   - gradients (w.r.t to retrieval pamareters) usually need to be provided for solver
   - called in each iteration



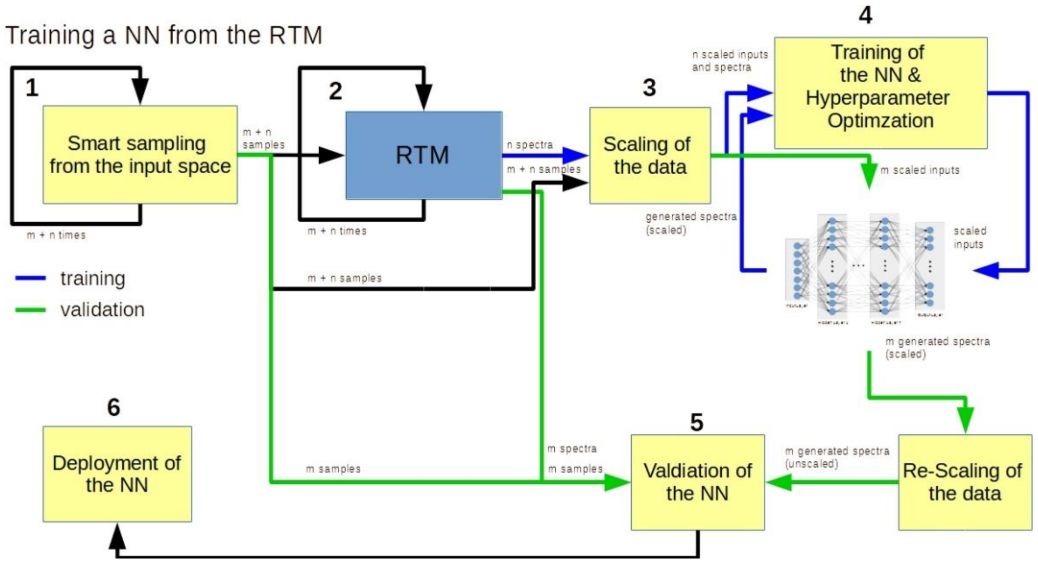
2. NN for **direct inversion**:
   - $F^{-1}: Y \rightarrow X$, *spectrum → state of atmosphere*
   - F^{-1} is generally unknown, can only be inferred through samples
   - No gradients needed after learnnig
   - called only once



NN as forward model

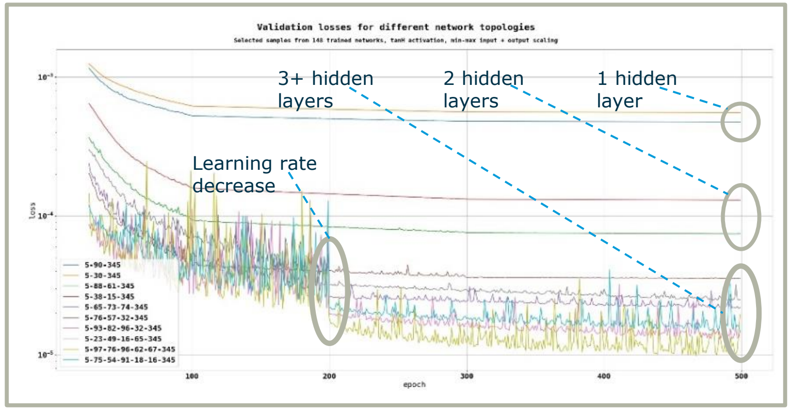
1. How to get from RTM to NN?

→ **NN Lifecycle chain**:
General procedure to replace RTM of an inversion algorithm by a NN

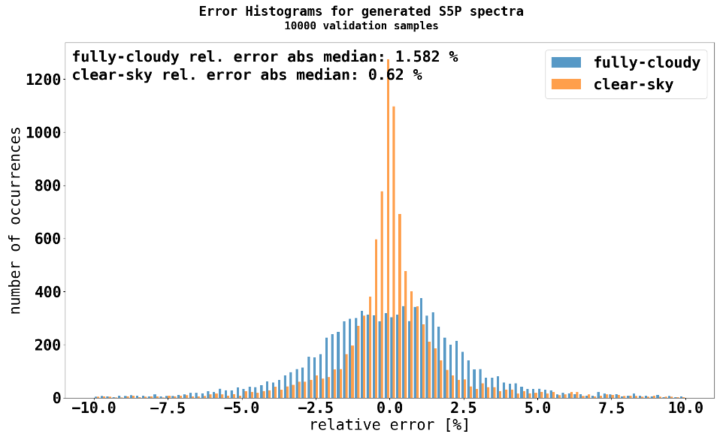


2. Finding optimal NN configuration is challenging, there are many aspects to consider:

- NN **topology**
- **activation functions**
- dataset **sampling**
- **learning algorithm**
- …



NN performances for different **topologies**



S5P NN performance - **clear-sky** and **fully-cloudy**

Spectral fitting challenges

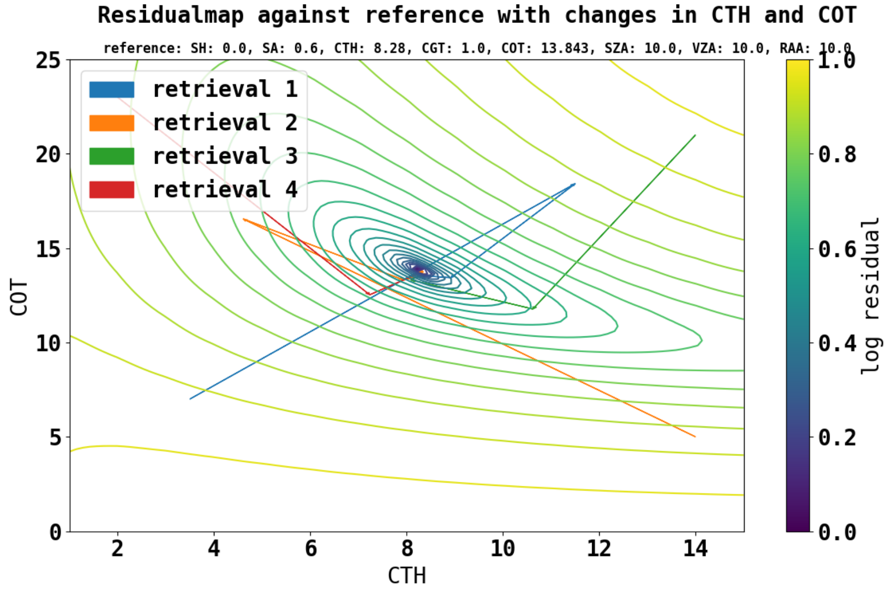
With a NN as forward model, a spectral fitting algorithm can be used for the retrieval of the atmospheric parameters

However, this is still challenging:

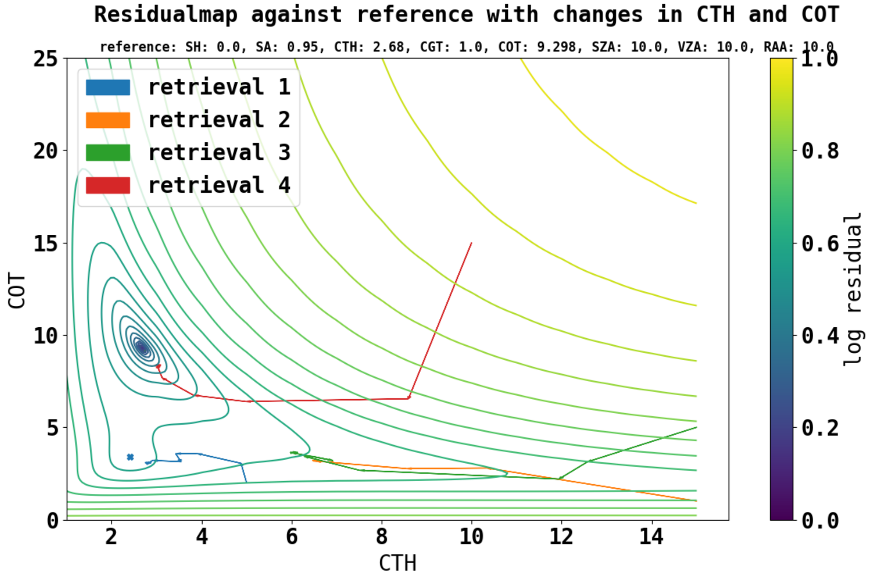
- spectral fitting problem is generally ill-posed
  → **local minima**
- real data contains noise in measurements

→ ROCINN algorithm (part of the operational CLOUD product) uses **Tikhonov Inversion**, which adds a regularization term to the problem

For difficult cases, good **a-priori** values for the retrieval parameters are still important:



Global minimum

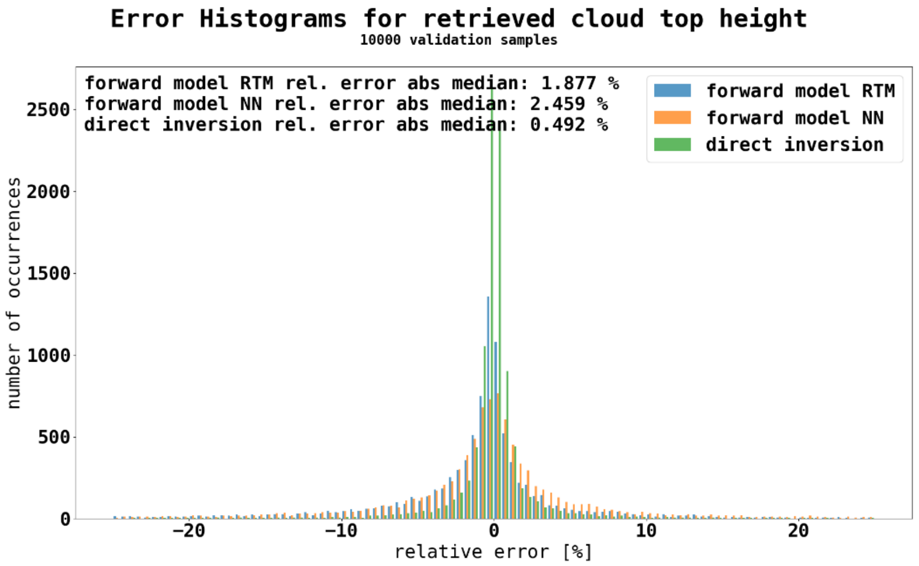


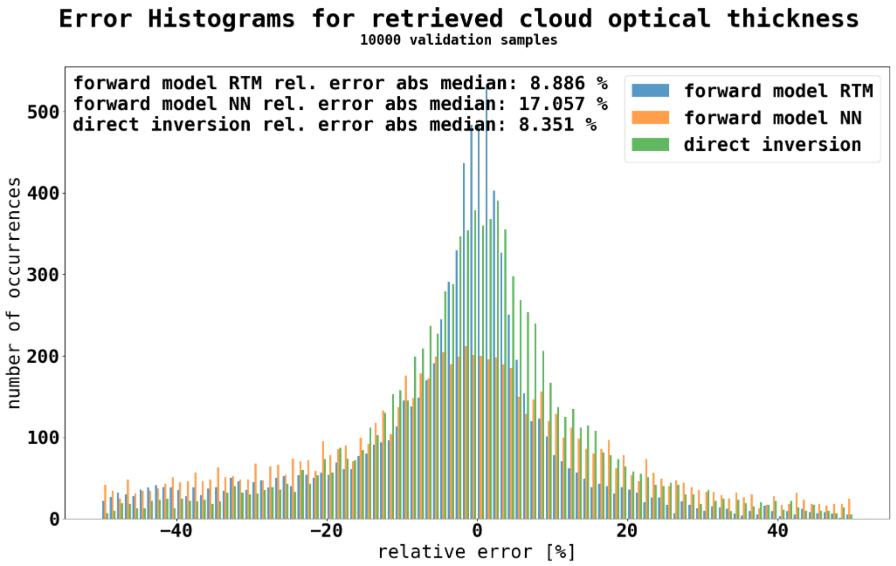
Local minima

NN for direct Inversion

- NN for direct inversion can avoid some of the issues of the spectral fitting:
  - no **fine adjustment** of the retrieval algorithm (e.g. regularization parameter, tolerances for convergence, etc.), all settings via the hyperparameters and training of the network
  - no **a-priori** necessary
  - only **one call** (iteration) per problem

- Input: spectra, viewing geometry, surface parameters, Output: cloud parameters
  (topologies: NN as FM: 7-66-77-26-89-78-94-99-107, NN for direct inversion: 112-80-80-80-80-2)
- Evaluation on validation dataset:





→ Best results for cloud top height: 0.49% vs. 2.46% (NN as FM), 1.88% (RTM as FM)
→ Best results for cloud optical thickness: 8.35 % vs 17.06% (NN as FM), 8.89% (RTM as FM)

Uncertainty Quantification

- Drawback: No indication for the quality of the results for the direct inversion NN („*blackbox*“)
- In contrast to the spectral fitting with e.g. iterations, convergence, residual, etc.

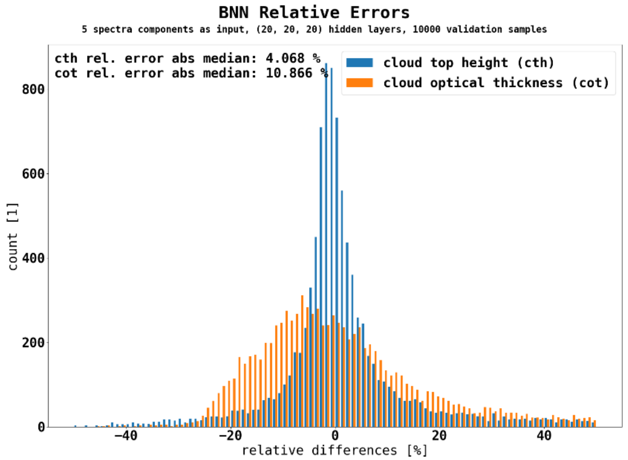
→ **Uncertainty Quantification**

Approaches:

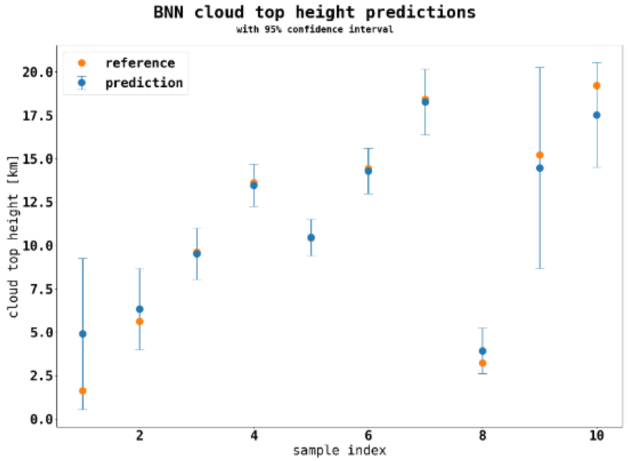
- Ensemble of NNs
    - captures model uncertainties through sampling
- Bayesian neural networks (BNN)
    - learns uncertainties in data and model parameters
    - output is a probability distribution
    - more complex and are harder to train
      → use of autoencoders to reduce complexity

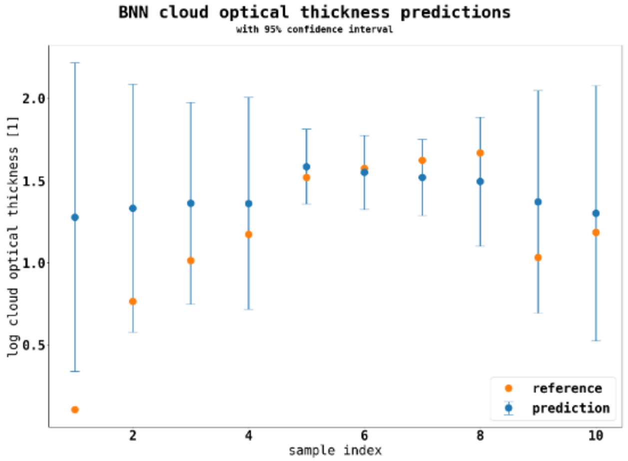
Evaluation:

1. BNN performs slightly worse than the conventional NN (taking the means as output)
    - learning is harder (much slower), current results likely not optimal
2. Standard deviation of ouptuts allows definition of a confidence interval
    - reference values are mostly inside
      → **reliable quantification of uncertainties**



BNN relative retrieval errors for CTH and COT from validation data set





Retrieved CTH (left) and COT (right) values for 10 random samples

Conclusions and Outlook

1. **NN as forward models:**
    - can improve speed of existing retrieval algorithms by orders of magnitude through substitution of the radiative transfer model → *near real time* applicable
    - NN lifecycle chain offers training and integration of specialized NNs
    - many properties from classical retrieval algorithms are inherited:
        - retrieval diagnostics
        - difficulties with ill posed problems, local minima
    - performance allows for potential in inversion algorithm improvements

2. **NN for direct inversion:**
    - easy to apply, good initial performance, no a-priori needed
    - conventional NNs are „black boxes“, no uncertainty quantification
    - Ensemble of NNs, BNNs as a possibility to overcome this:
        - provide error quantifications
        - BNNs more complex and harder to train but provide reliable error quantifications

→ NNs for direct inversion, especially BNNs with uncertainty quantification, have great potential for retrieving cloud properties for S5P as an alternative to the current forward model approach

- Further investigations in hyperparameter selection and learning have to be made
- Invertible neural networks (INN), that learn forwards and backwards and can also provide distributions are another interesting approach that should be followed

For further questions, please contact me: Fabian.Romahn@dlr.de