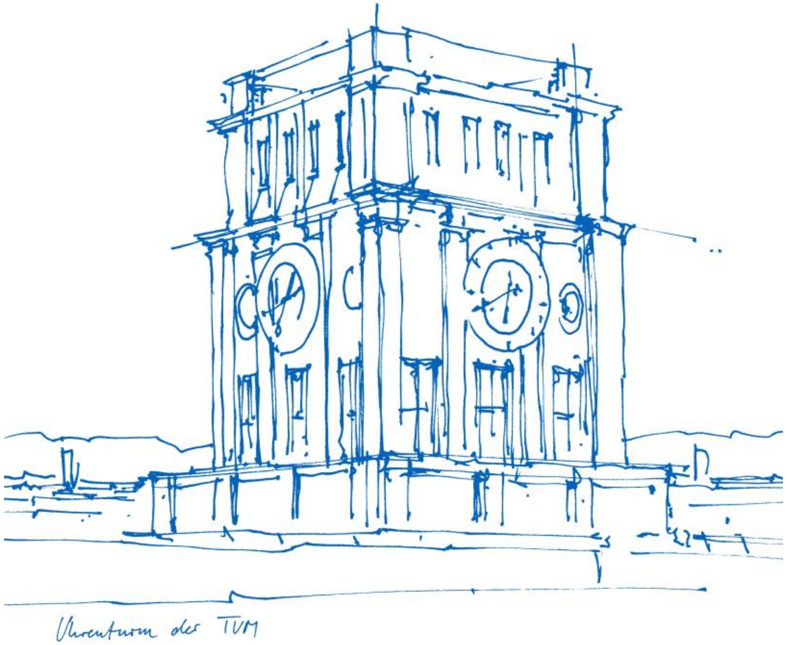
Application of deep neural networks to retrieve cloud properties for Sentinel-4 (S4) and TROPOMI / Sentinel-5 Precursor (S5P)

Doctoral Candidates‘ Day 2022

Fabian Romahn

German Aerospace Center (DLR)

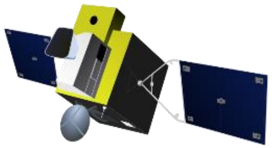


Copernicus Satellites S5P and S4

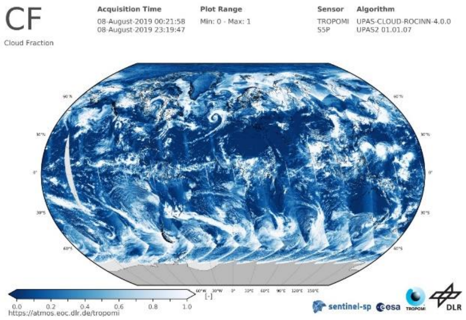
Sentinel-5 Precursor (S5P) and Sentinel-4 (S4) are passive earth observation satellites (with UV/VIS spectrometers) of the Copernicus programme:

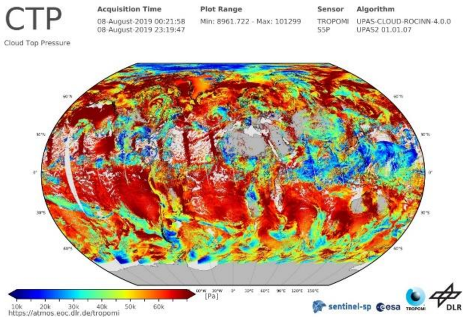
S5P:
- Launched in october 2017
- Sun-synchronous orbit at ~ 824 km

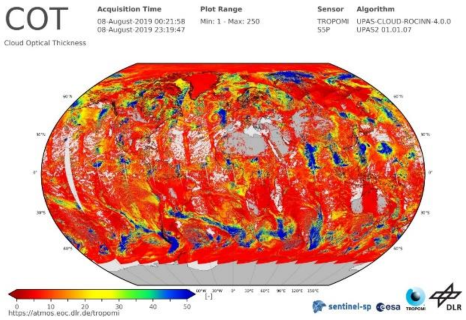
S4:
- Launch date due 2023
- Geostationary



DLR is responsible for the operational *CLOUD* product for both satellites







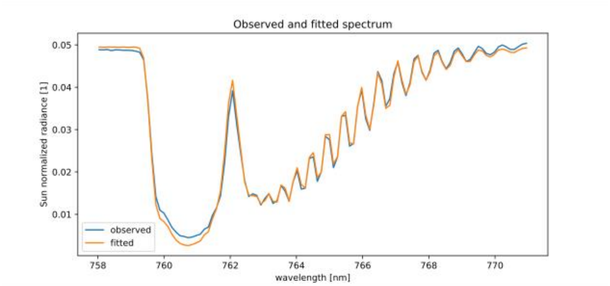
Challenges:
- Large amouts of data
- Near real time requirements (NRT)

→ 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

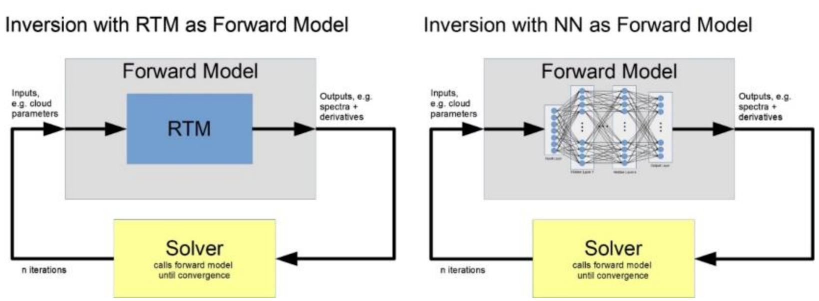


for remote sensing:

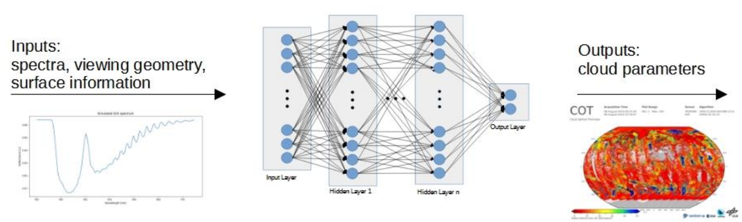
x: State of atmosphere, y: Measured spectrum, F: Radiative transfer model (RTM)

Two approaches:

1. NN as **forward model** of a spectral fitting algorithm:
   - $F: X \to Y$ *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**:
   - $F^{-1}: Y \to X$, *spectrum → state of atmosphere*
   - F^{-1} is generally unknown, can only be inferred through samples
   - No gradients needed after learnnig
   - called only once



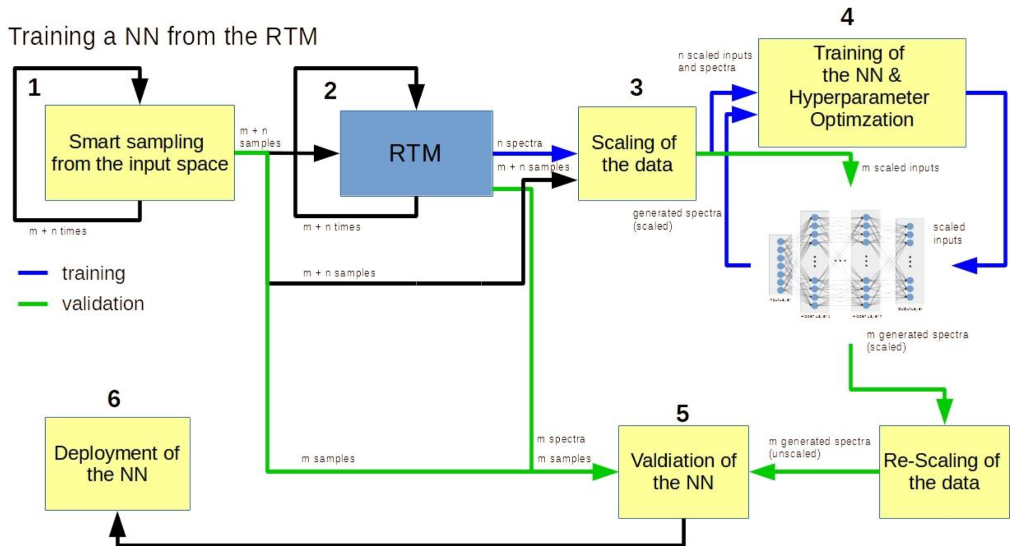
NN as forward model

How to get from RTM to NN?

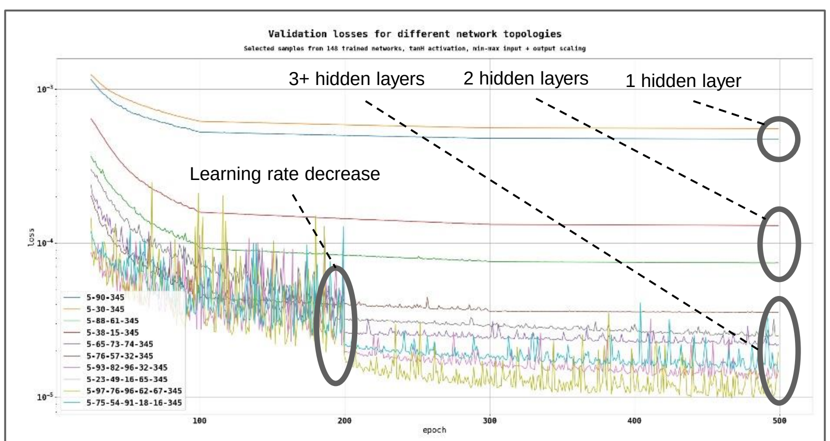
→ NN Lifecycle chain:
General procedure to replace RTM of an inversion algorithm by a NN

Finding optimal NN configuration is challenging, aspects:

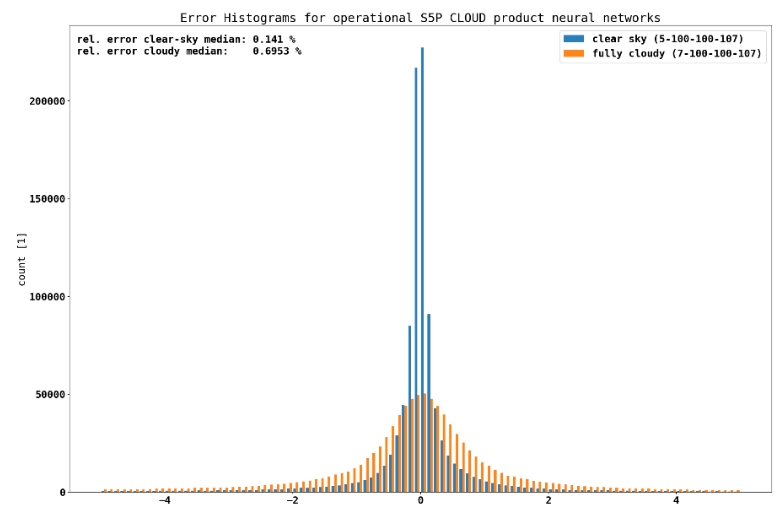
- NN **topology**
- **activation functions**
- dataset **sampling**
- **learning algorithm**
- …



<u>NN performances for different **topologies**</u>



<u>**Operational S5P** NN performance</u>

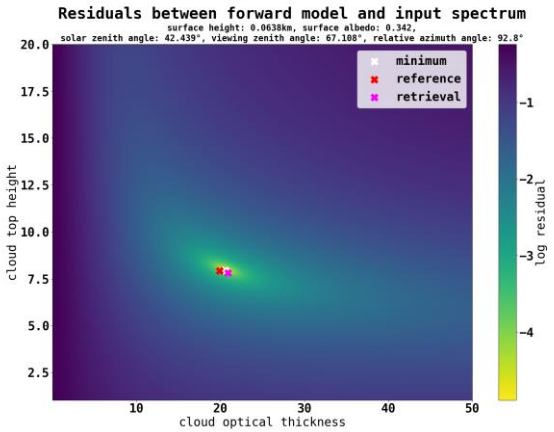


Spectral fitting challenges

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:

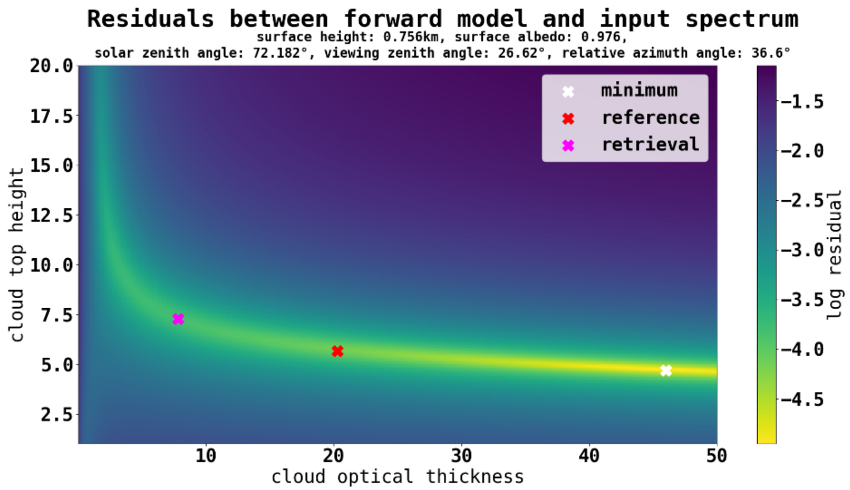
- spectral fitting problem is generally ill-posed
  → **local minima**
- real data contains noise in measurements



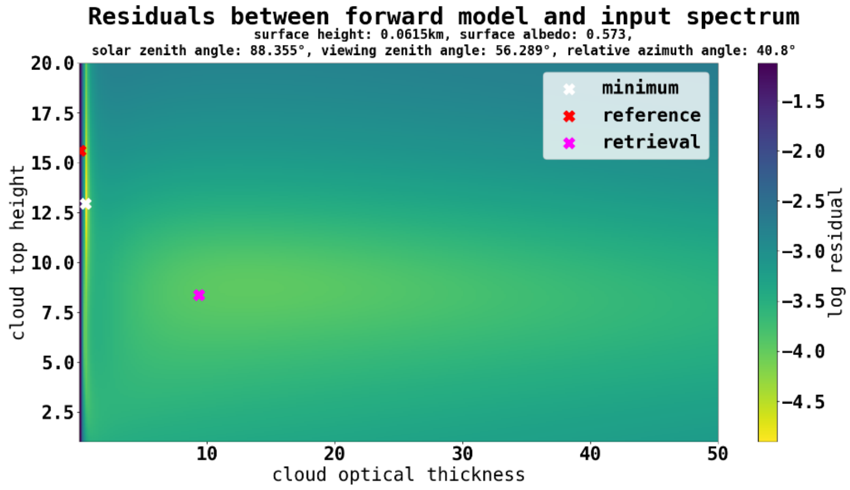
Residual map for an „easy“ problem with a well defined global minimum

→ ROCINN algorithm (part of the operational S5P CLOUD product) uses **Tikhonov Inversion**, which adds a regularization term to the optimization problem

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



‚stretched‘ minimum



local minima

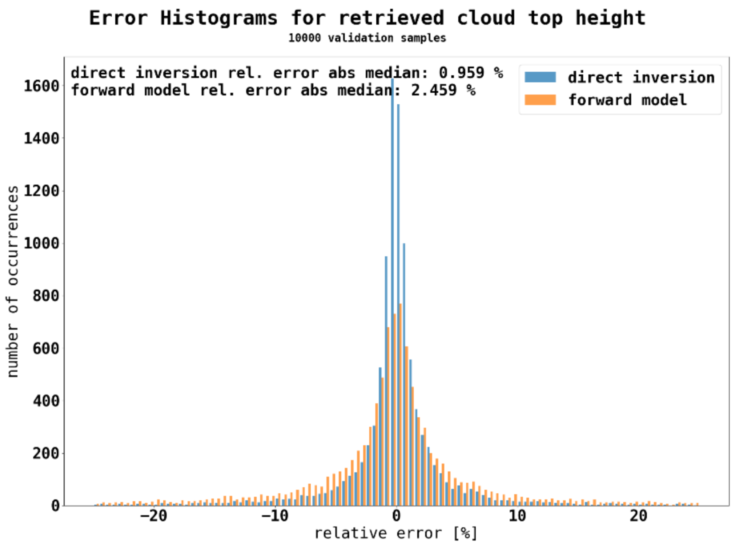
NN for direct Inversion

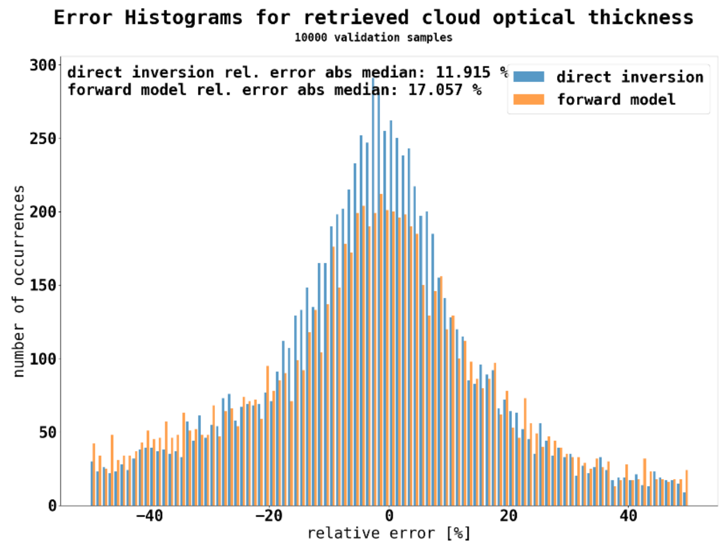
NN for direct inversion can avoid some of the issues of the spectral fitting:

- 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
- not as affected by **local minima**
- not as affected by **local minima**

Input: spectra, viewing geometry, surface parameters, Output: cloud parameters
evaluation for comparison with forward model NN in spectra fitting for validation dataset:

topologies: NN as forward model: 7-66-77-26-89-78-94-99-107
NN for direct inversion: 112-80-80-80-80-2





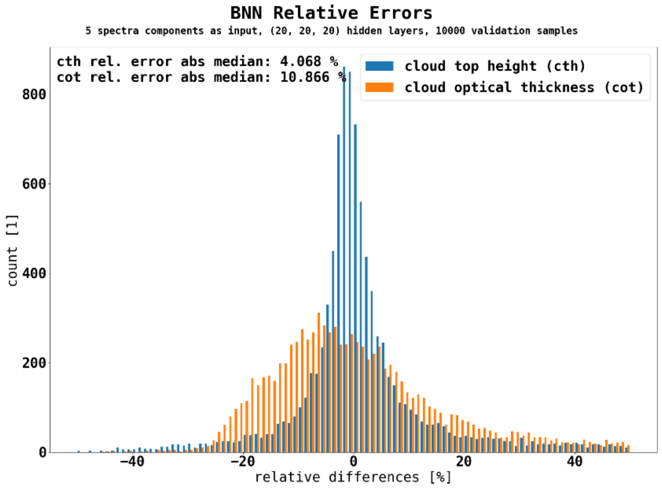
→ Better results for direct inversion NN: **CTH**: 0.96% vs 2.46%, **COT**: 11.92% vs. 17.06% (med. abs. rel. error)

Bayesian Neural Networks

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.

→**Bayesian neural networks (BNN)**:

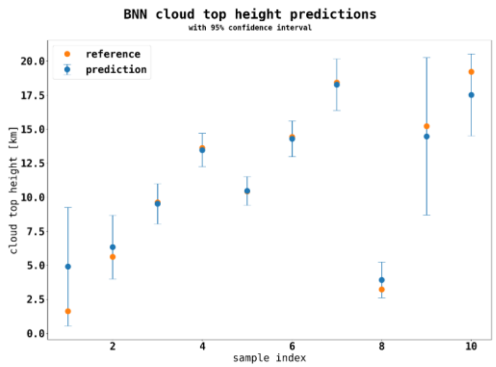
- learns uncertainties in model parameters
- output is a probability distribution
- more complex and are harder to train

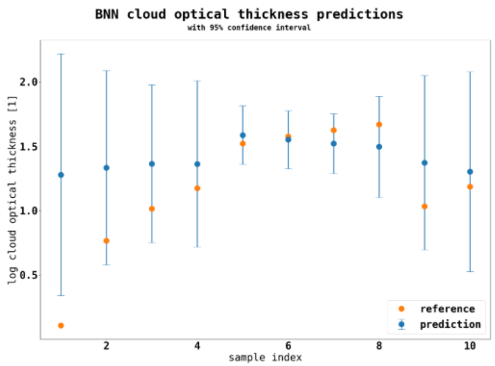


Evaluation:

1. Overall, BNN performs slightly worse than the conventional NN (taking the means as output)
   - learning is harder (much slower), current results are likely not optimal
   - for many deep topologies (> 3 hidden layers) learning is not successful

2. Standard deviation of ouptuts allows definition of a confidence interval
   - reference values are mostly inside
     → **reliable quantification of errors**





Conclusions and Outlook

1. NN as forward models:

- can improve speed of existing retrieval algorithms by orders of magnitude through substitution of existing radiative transfer model (RTM)
- many properties from classical retrieval algorithms are inherited:
  - retrieval diagnostics
  - difficulties with ill posed problems, local minima

2. NN for direct inversion:

- easy to apply, good initial performance, no a-priori needed
- conventional NNs are „black boxes“, no error quantification
- BNNs as a possibility to overcome this:
  - provide error quantifications
  - more complex and harder to train

→NNs for direct inversion, especially when using BNNs with error quantification, have great potential for retrieving cloud properties for S4 / S5P as an alternative to the current approach that uses NNs as forward models

- Further investigations in hyperparameter selction 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