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

Living Planet Symposium 2022

Fabian Romahn[1], Diego G. Loyola[1], Adrian Doicu[1],
Victor Molina Garcia[1], Ana del Águila[1], Ronny Lutz[1]
(1) German Aerospace Center (**DLR**)

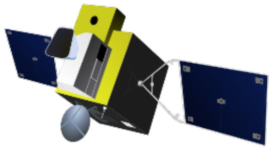
DLR

Copernicus Satellites S5P and S4

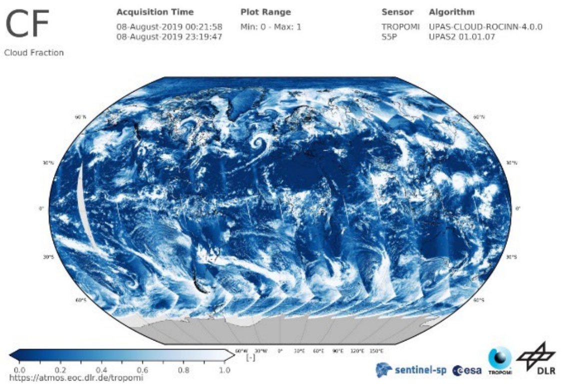
Sentinel-5 Precursor (S5P) and Sentinel-4 (S4) are passive earth observation satellites (with UV/VIS spectrometers) of the Copernicus programme:

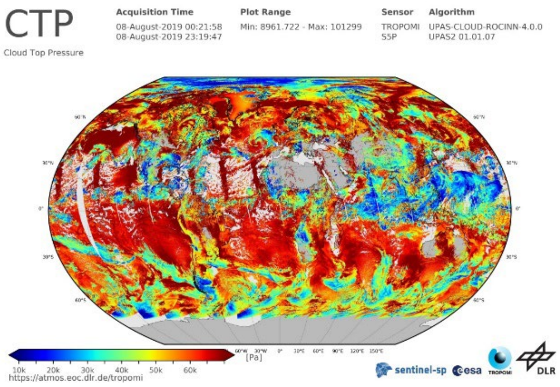
- **S5P**
  - Launched in october 2017
  - Sun-synchronous orbit at ~ 824 km

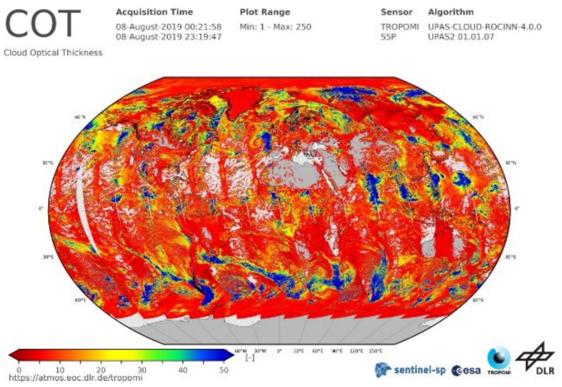
- **S4**:
  - Launch date due 2023
  - Geostationary



DLR is responsible for the operational CLOUD product for both satellites







Machine Learning in Remote Sensing

Why Machine Learning?

- Dramatically increased amount of data with latest generations of earth observation satellites
- Near real time requirements (NRT) for many products

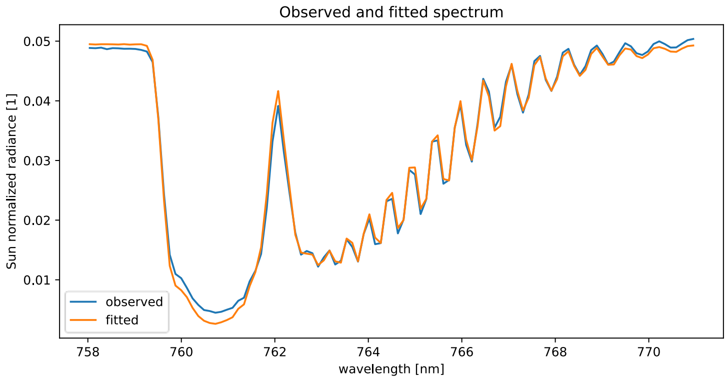
→ Retrieval algorithms have not only to be accurate but also to be very fast
→ Application of machine learning techniques to improve performance compared to classical algorithms

Machine Learning for Inversion Problems:

- Atmospheric retrieval problems can be formulated as inversion problems:

  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

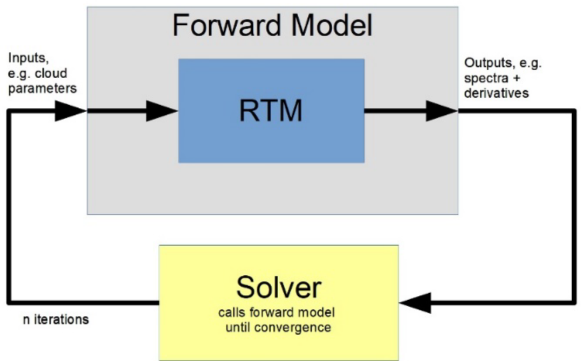
- In context of atmospheric retrieval:
  - x: State of atmosphere
  - y: Measured spectrum
  - F: Radiative transfer model (RTM)



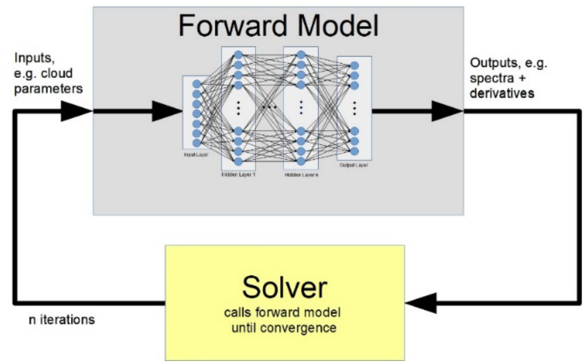
Two approaches for neural networks use

1. NN as **forward model** of a spectral fitting algorithm:
   - implements $F: X \to Y$, *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

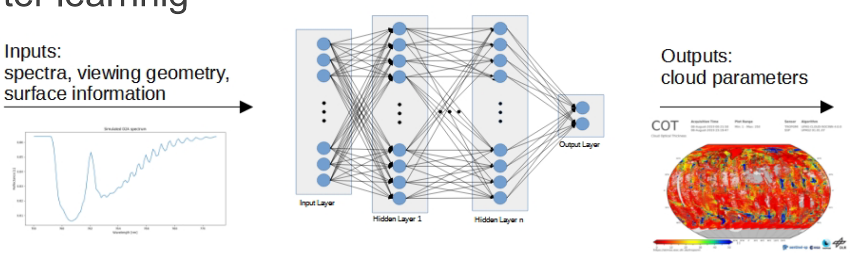
Inversion with RTM as Forward Model



Inversion with NN as Forward Model



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

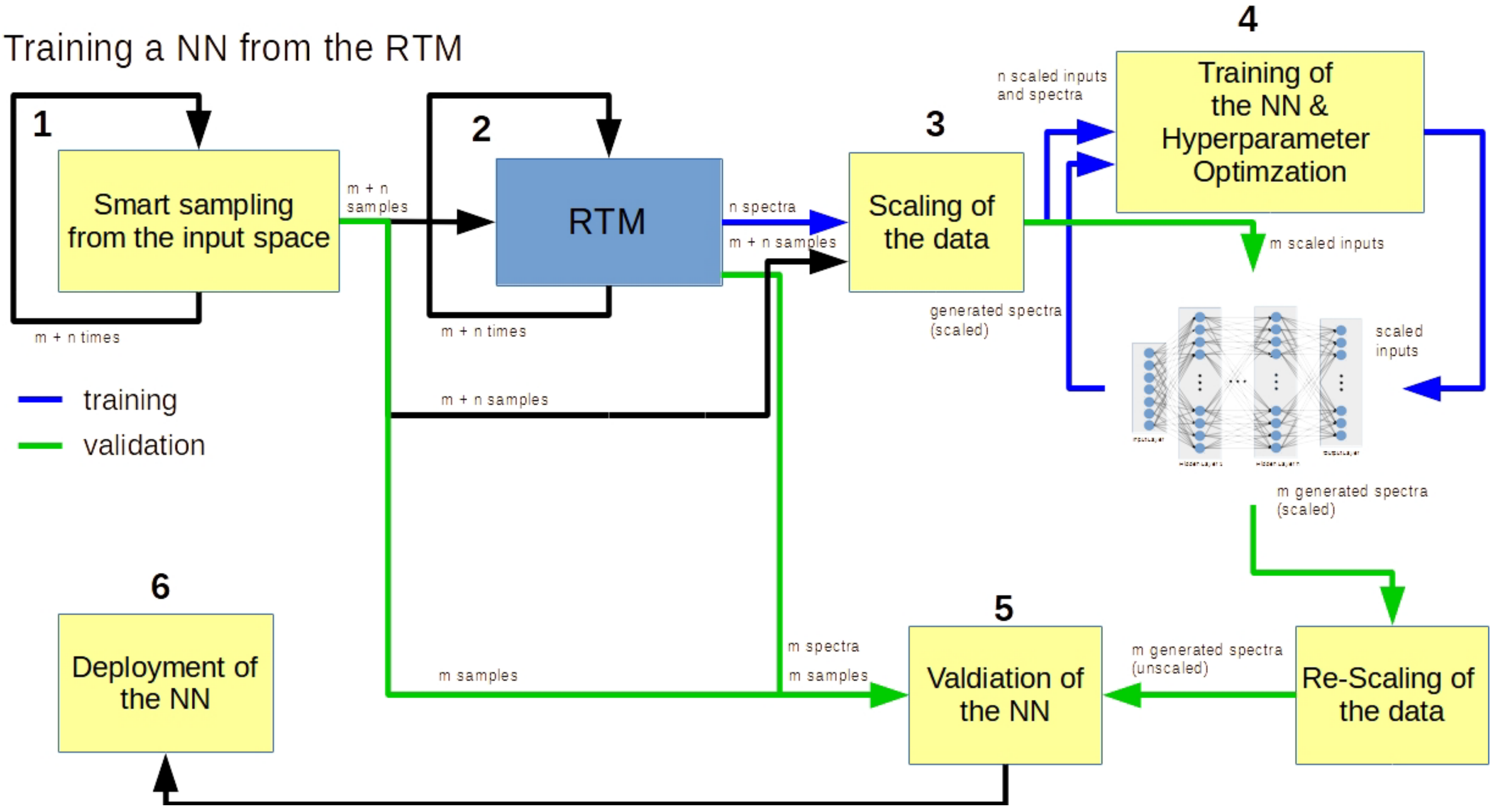


1. NN as forward model - lifecycle chain

How to get from RTM to NN for algorithms in S4, S5P, …?

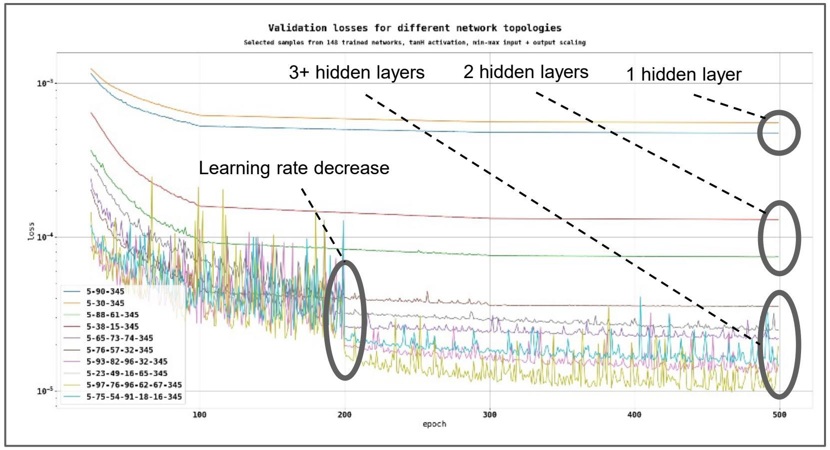
→ General procedure to replace RTM of an inversion algorithm by a NN:

Training a NN from the RTM

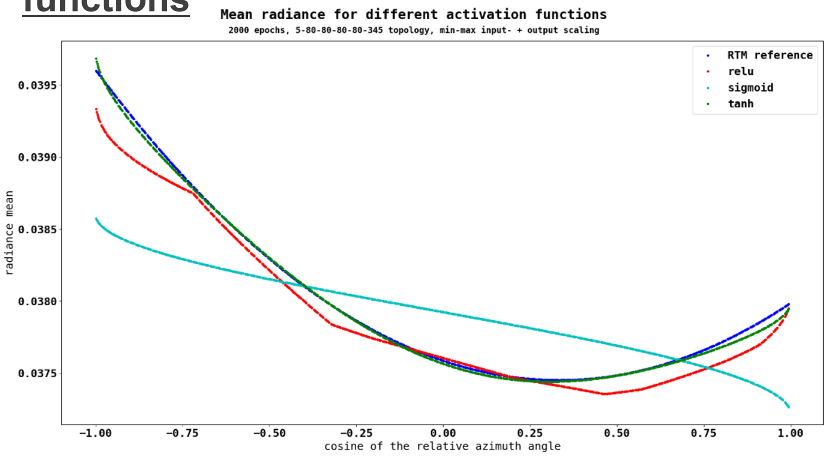


Evaluation of NNs

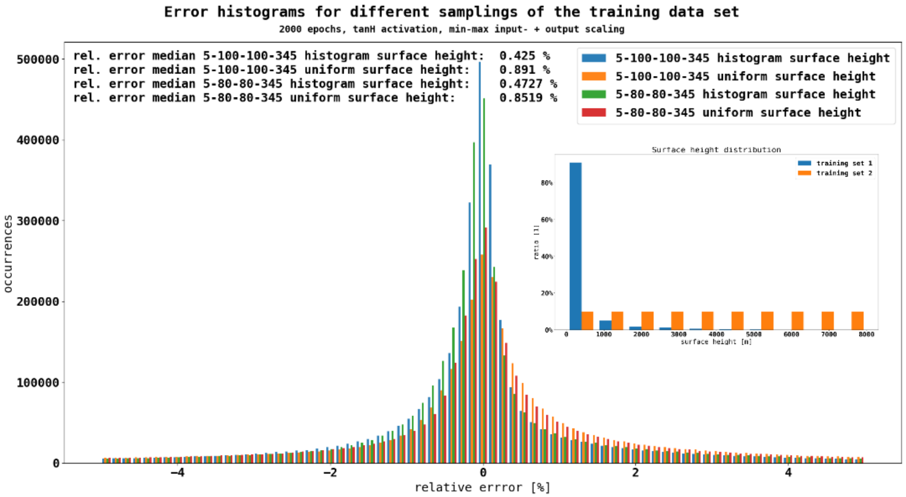
1. NN performances for different **topologies**



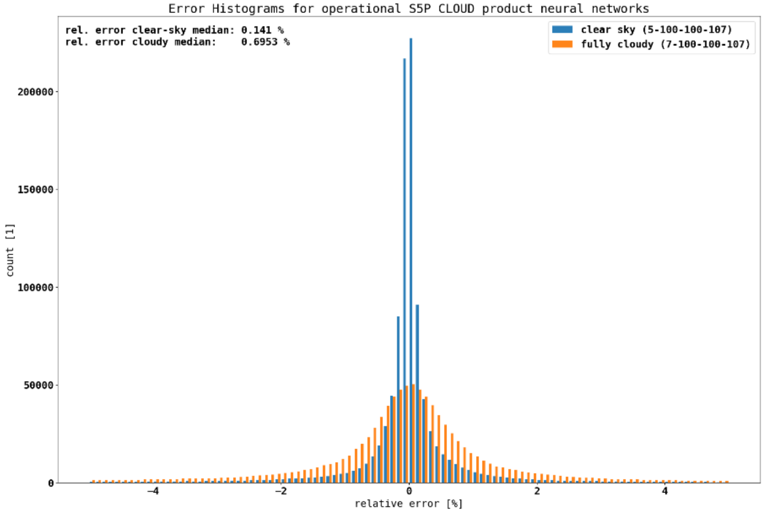
2. NN performances for different **activation functions**



3. NN performances for different dataset **samplings**



4. **Operational S5P** NN performance



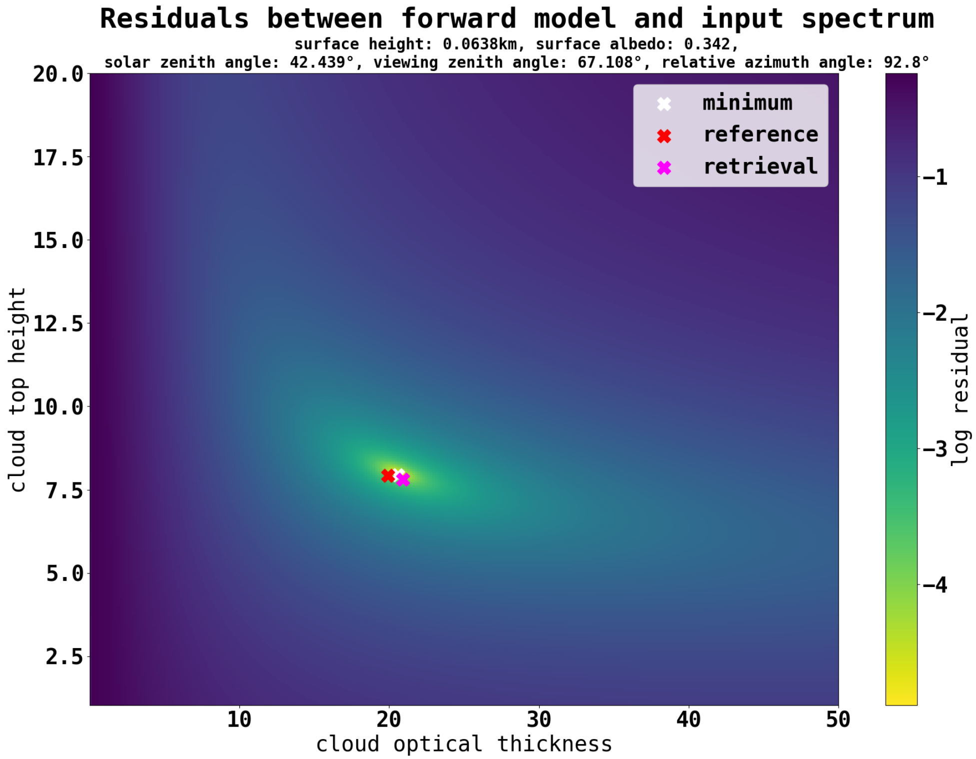
Spectral fitting challenges

- With a neural network 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 S5P CLOUD product) uses **Tikhonov Inversion**, which adds a regularization term to the optimization problem

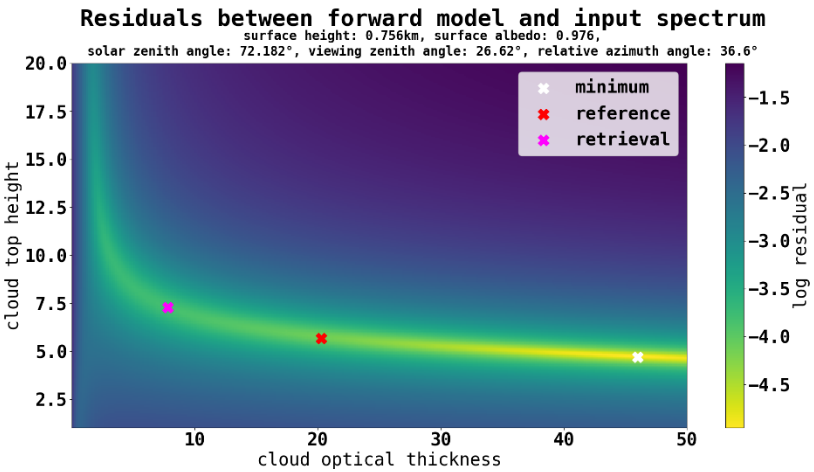
- For difficult cases, good **a-priori** values for the retrieval parameters are still important



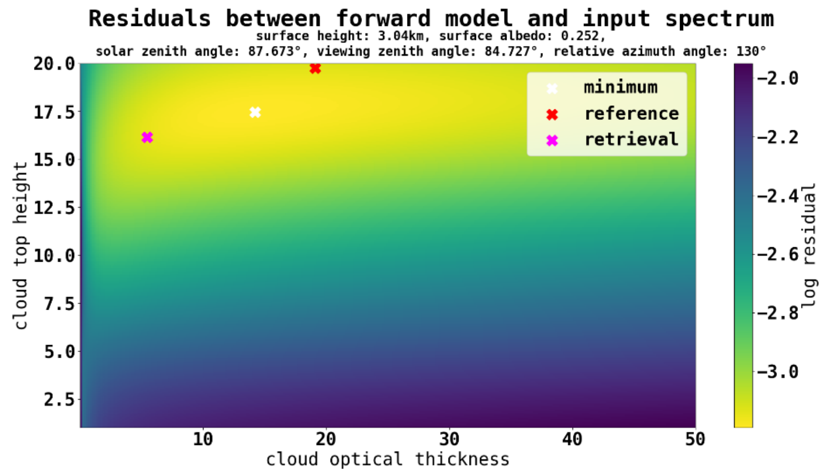
Residual map for an „easy“ problem with a well defined global minimum

Spectral fitting challenges - examples

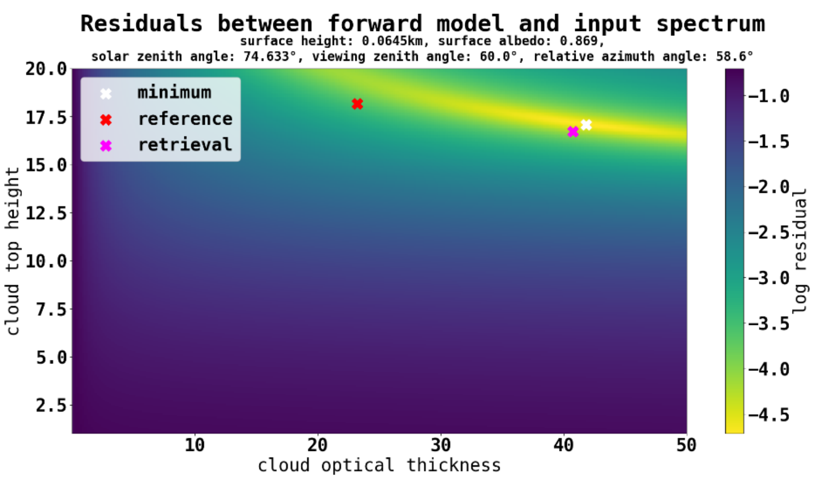
‘stretched‘ minimum:



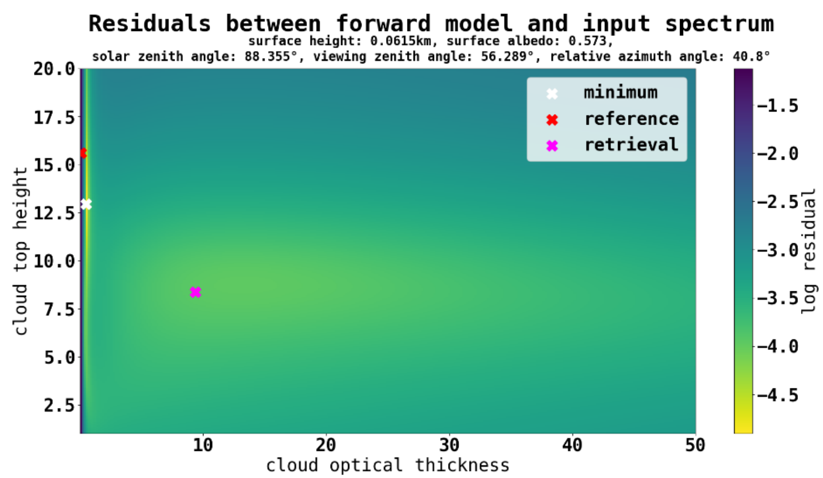
diffuse minimum:



forward model error:



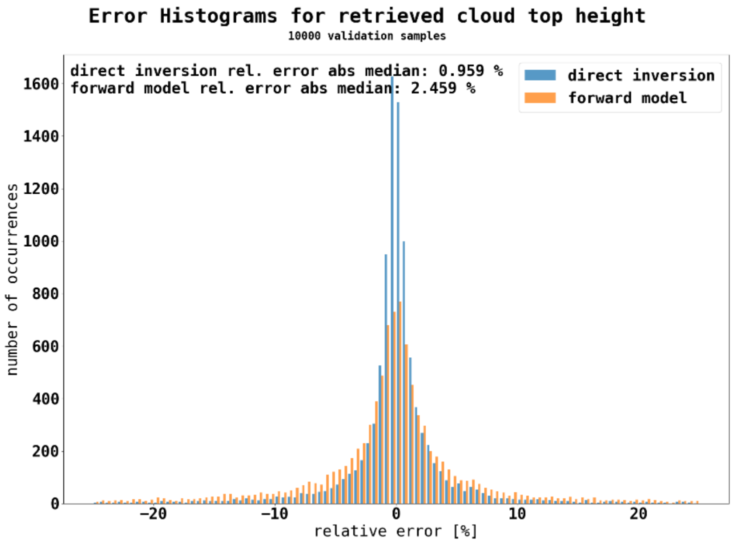
local minima

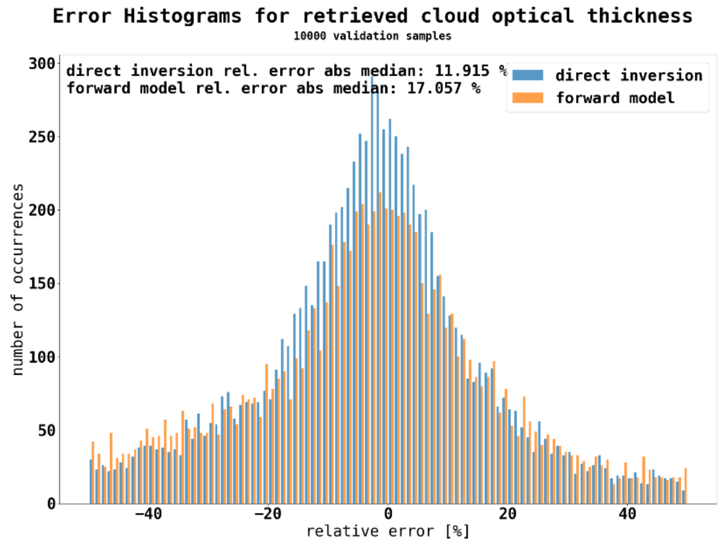


2. NN for Direct Inversion

- 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**
  - only **one call** (iteration) per problem

- 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

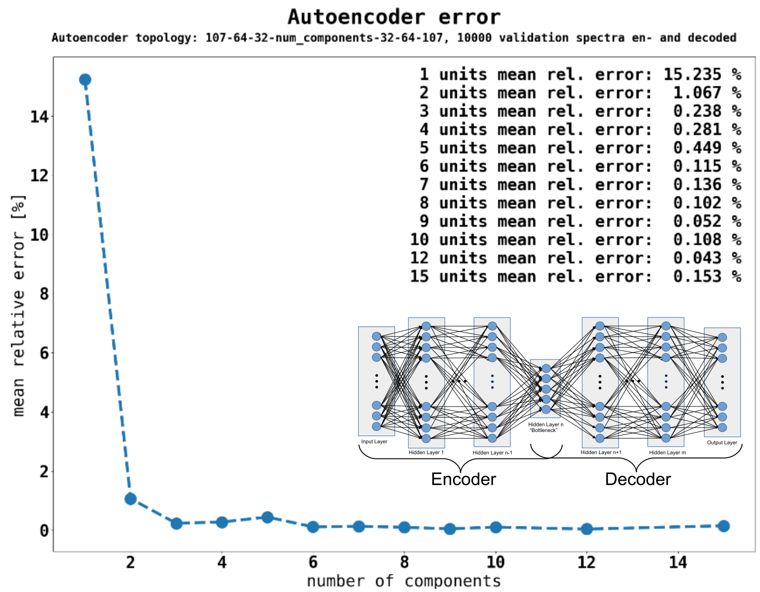
- Drawback: No indication for the quality of the results for the direct inversion NN („*blackbox*“)
- In contrast to the spectral fitting with e.g. iterations, convergence, residual, etc.

→ **Bayesian neural networks (BNN)**:

  - learns uncertainties in model parameters
  - output is a probability distribution
  - more complex and are harder to train:
    - example: network with (112, 20, 20, 20, 2):
      - NN: 3,142 parameters
      - BNN: 2,735,179 parameters

    → use **autoencoders** to reduce input complexity

- Autoencoder can reduce dimensionality of input spectrum to a few components
  → reduces the complexity of first layer in BNN:
  2,735,179 → 202,335 parameters
  5 instead of orginally 107 components for input spectrum

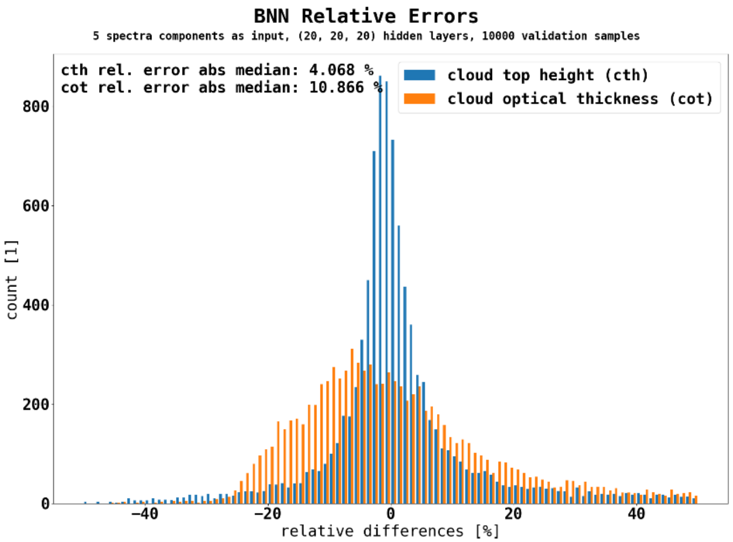


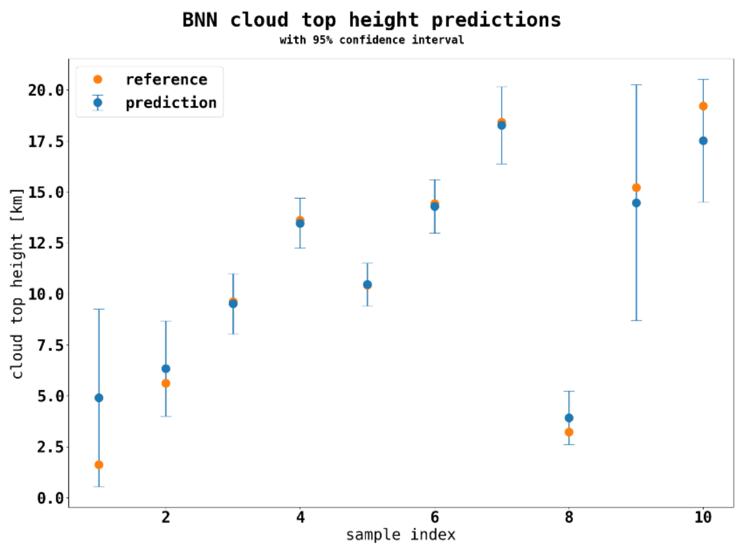
Autoencoder errors for different numbers of components

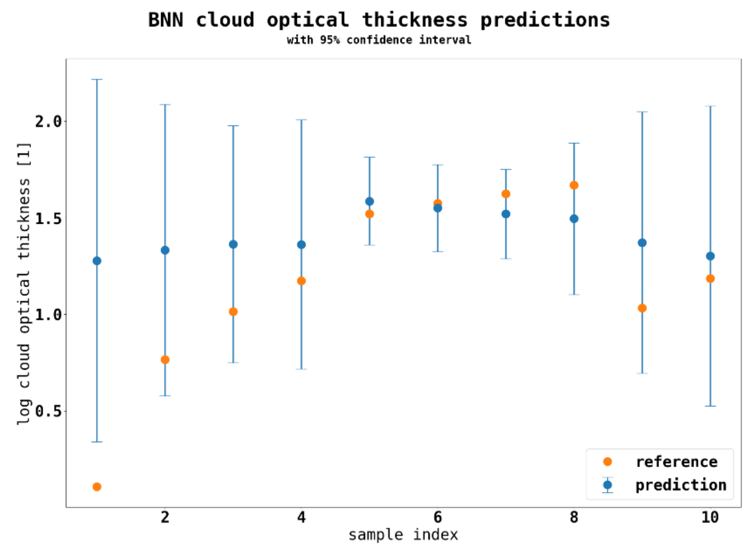
Bayesian Neural Networks - Results

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