Funded by the European Union

Co-ordinated by ECMWF

CORSO
CO2MVS Research on Supplementary Observations

TOWARD STATISTICAL FORWARD OBSERVATION OPERATOR TO ASSIMILATE LEVEL-1 MICROWAVE OBSERVATIONS

Sebastien Garrigues with contributions to

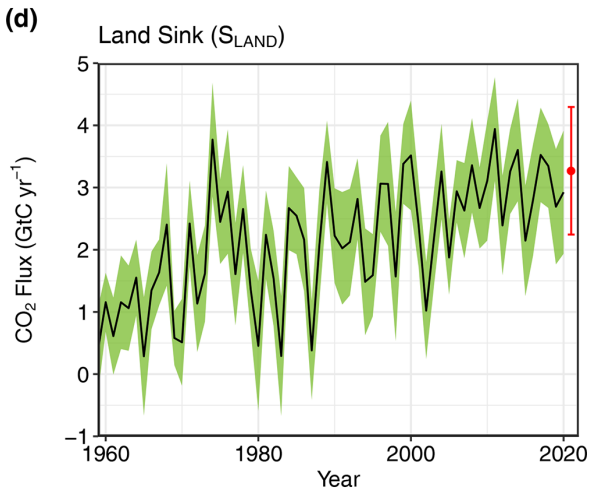
Patricia de Rosnay, Pete Weston, Ewan Pinnington, David Fairbairn

ECMWF

The CORSO project (grant agreement No 101082194) is funded by the European Union.

Introduction

Large uncertainties in biogenic $C0_2$ fluxes

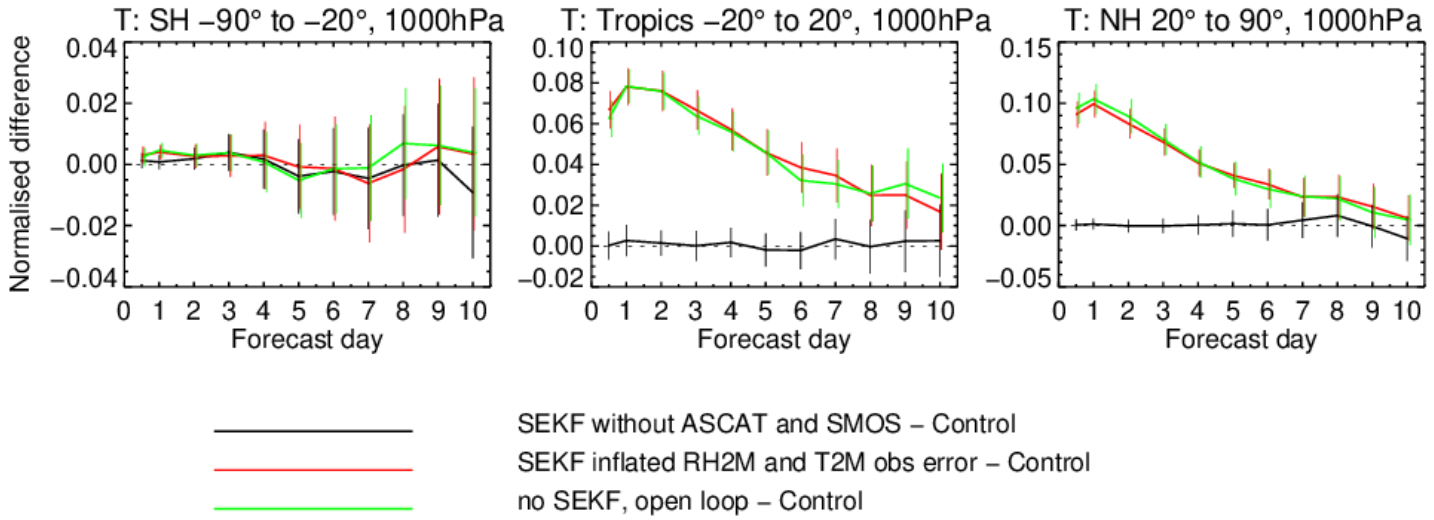


Friedlingstein et al.: Global Carbon Budget 2021

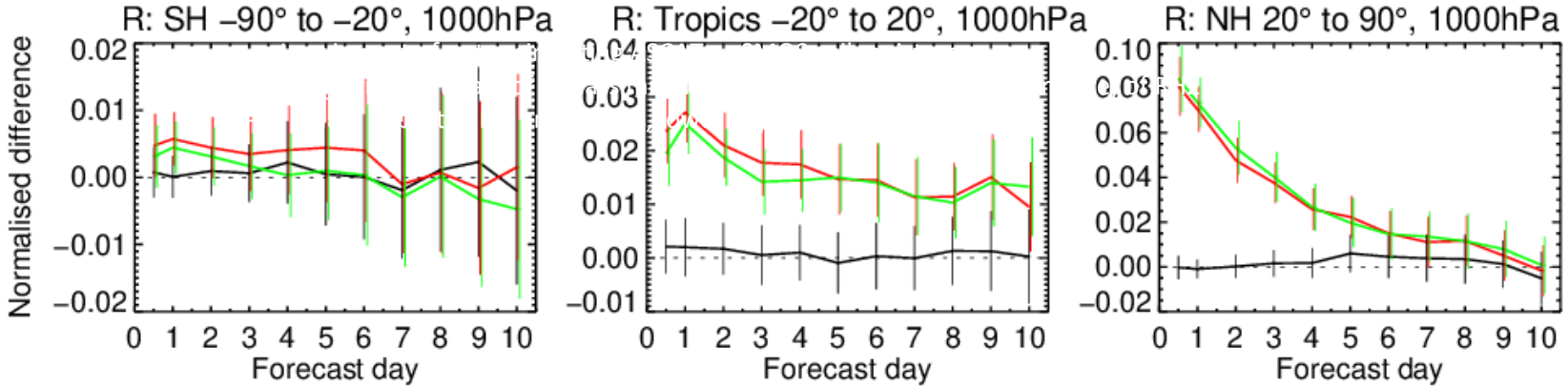
- CORSO Copernicus project
  - Preparation of the C02MVS
  - Better constraint biogenic fluxes
  - Develop statistical fwd observation operator to analyse land surface variable controlling carbon fluxes
  - Application to passive and active level-1 microwave and SIF observations
- Why assimilating level-1 satellite observations instead of soil moisture retrieval ?
  - More accurate quantification of uncertainties
  - Analyze not only soil moisture but also vegetation variables
  - Consistent with NWP radiance DA framework
- Why statistical observation operator ?
  - Large uncertainties in physical-based model for land surface: e.g. surface heterogeneity, volume scattering
  - Remove bias between model and retrieval
  - Consistent observation operator across frequencies and instrument
  - Maximize the information content of the training database

Impact of analyzing soil moisture

2m temperature



2m Relative humidity



- Low impact of not assimilating ASCAT or SMOS soil moisture
- Significant impact of not assimilating (or inflating the obs error variance) of RH2M and T2M
- Main impact in NH and tropics, lower impact in SH

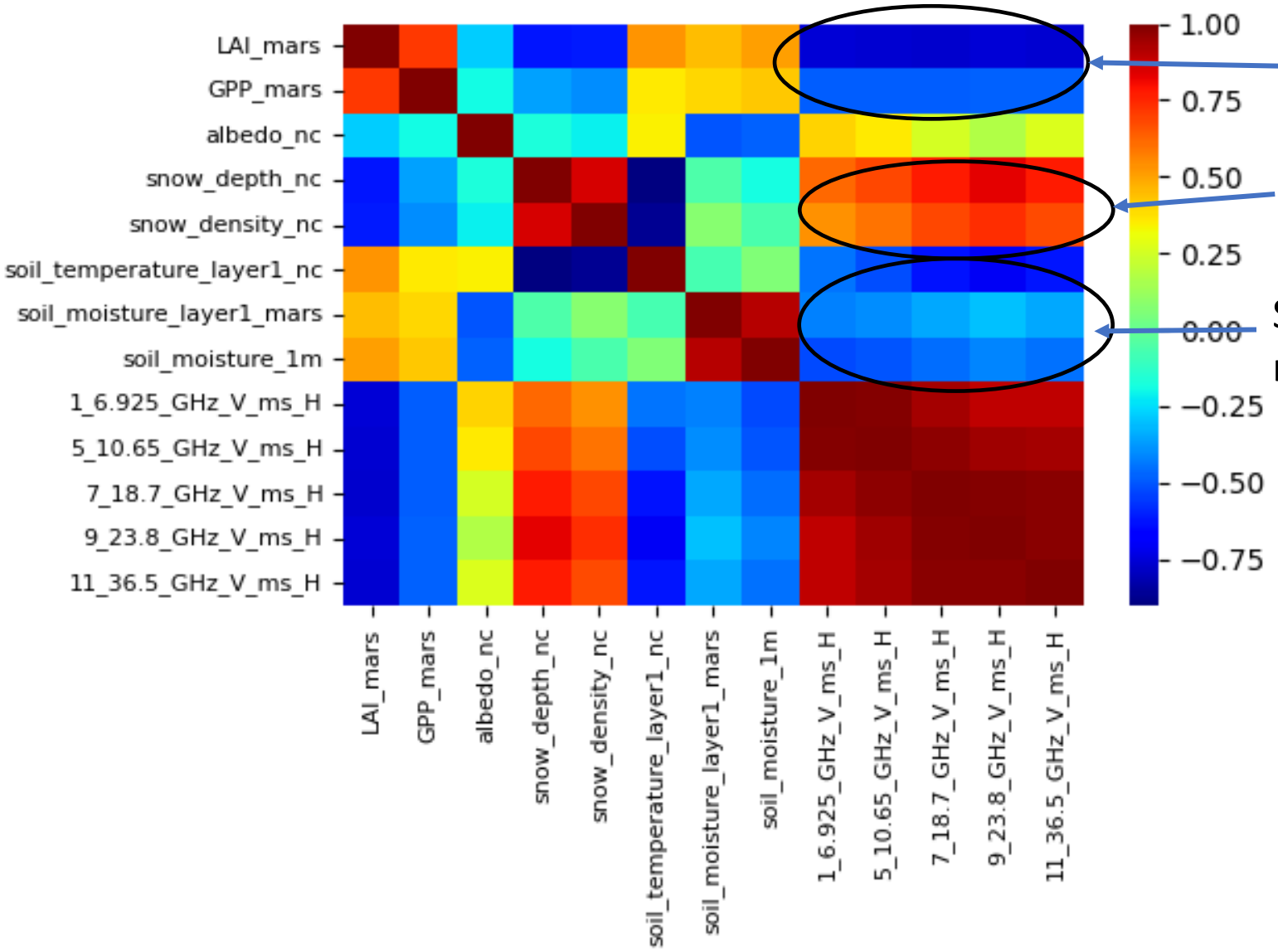
Training databases for statistical forward observation operator

- Active microwave ASCAT:
  - Credit: P. Weston see Aires, Weston et al., 2021
  - 2016-2019; *0.25°-resolution grid.*
  - *Satellite: Backscatter coefficient at 40 degree*
  - ERA-5 model variables: LAI, soil moisture, soil temperature, magnitude of 2m temperature diurnal cycle

- Passive microwave: AMSR2
  - Credit: Alan Geer
  - CERISE/CORSO joint training database
  - Model: IFS Cycle 47r1 (all-sky observations from 47r3)
  - N256 reduced Gaussian grid
  - 15-month period:2020/07/01-2021/09/30
  - Satellite: 14 channels brightness temperature, including polarization

Information content of passive microwave database (AMSR2)

Model field versus polarization index (V-H)/(V+H)



Vegetation variables: negative corr

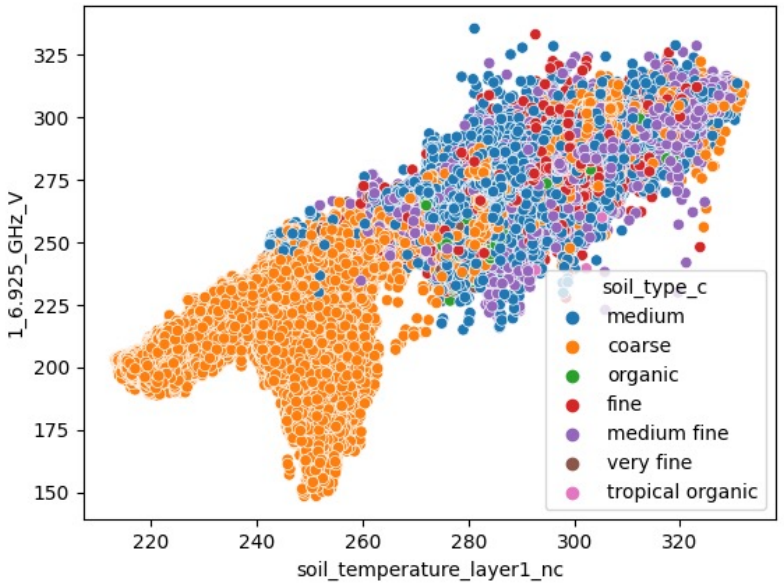
Snow variables: positive corr

Soil moisture , soil temperature: negative corr

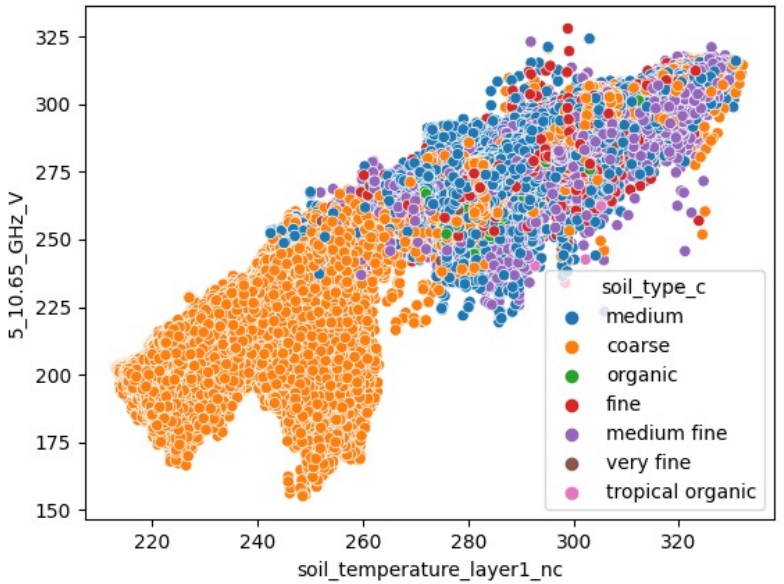
Information content of passive microwave database (AMSR2)

Soil temperature vs observation scatter-plot (soil type clustering)

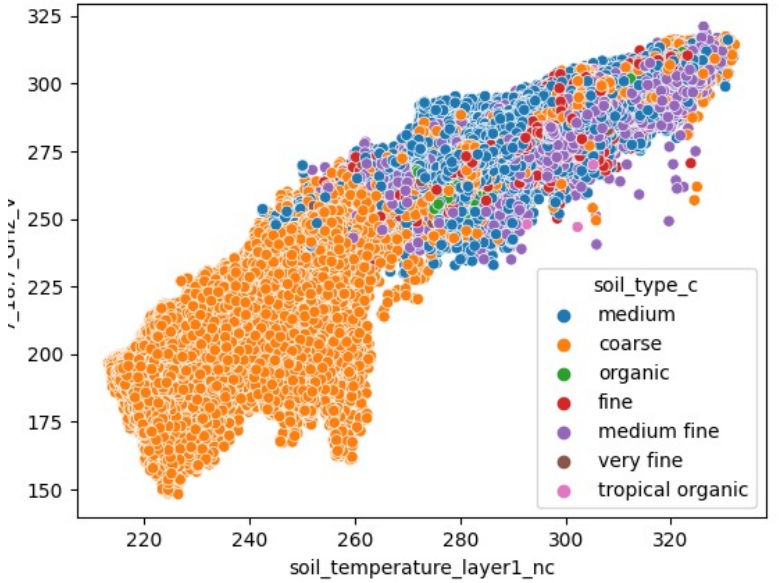
6.9 Ghz C-band (V)



10.65 Ghz X-band (V)



18.7 Ghz Ku-band (V)

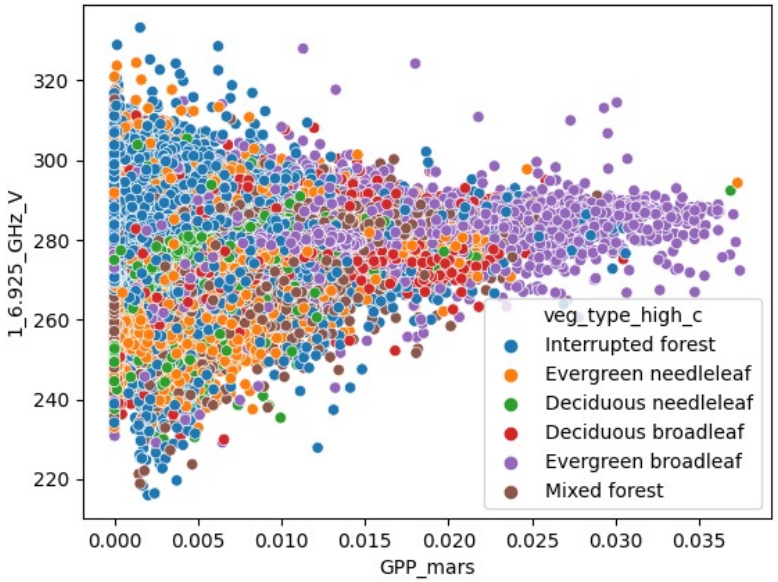


Strong influence of soil temperature at all frequencies

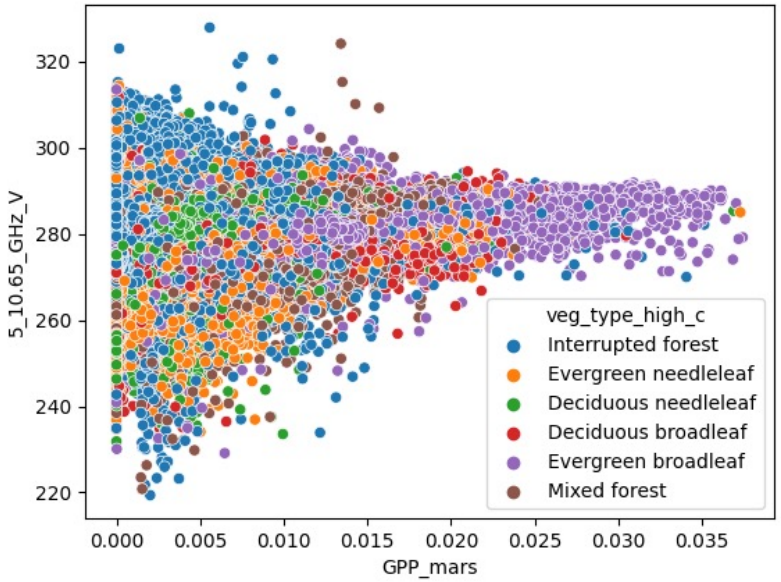
Information content of passive microwave database (AMSR2)

GPP vs observation scatter-plot (high veg type clustering)

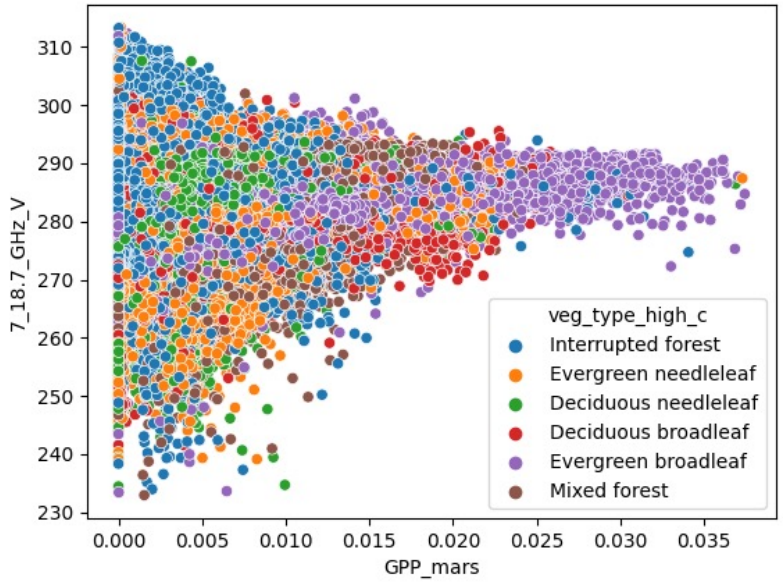
6.9 Ghz band



10.65 Ghz band (V)



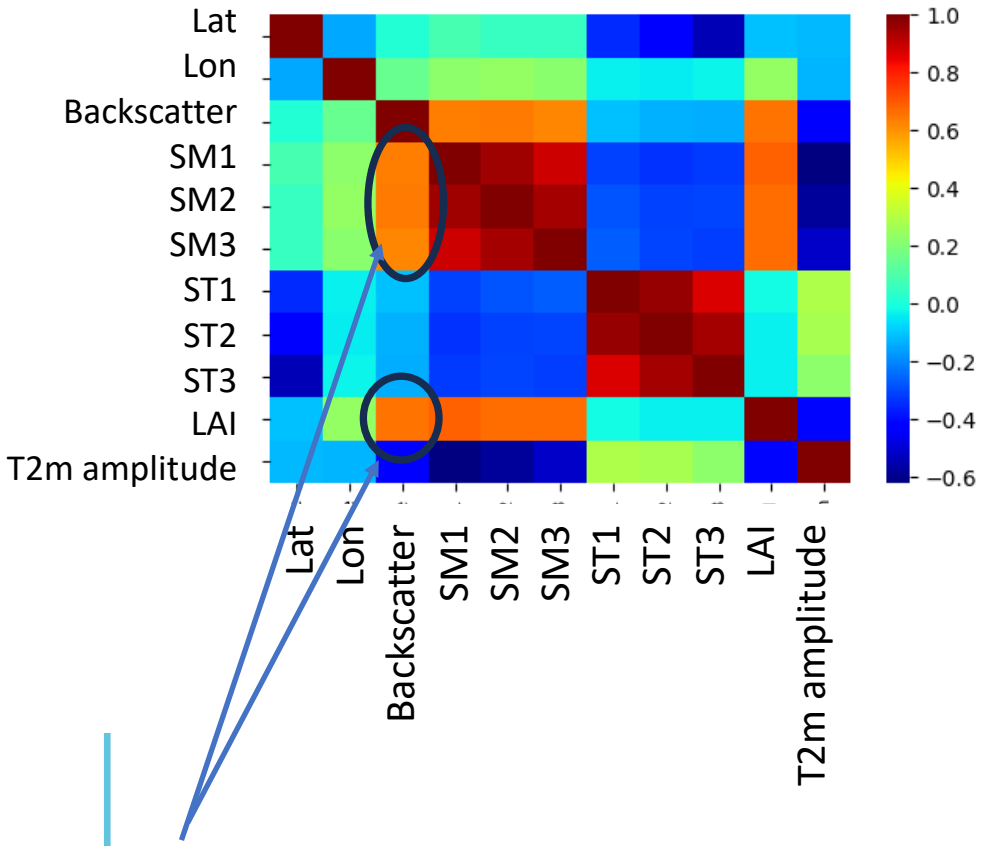
18.7 Ghz band (V)



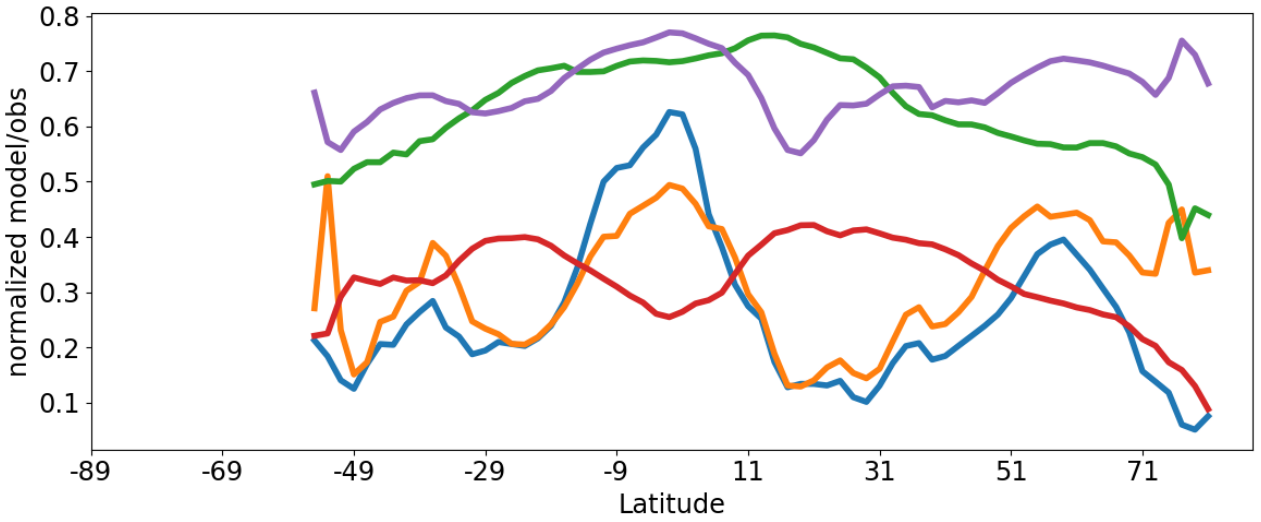
More difficult to identify microwave sensitivity to GPP across vegetation types

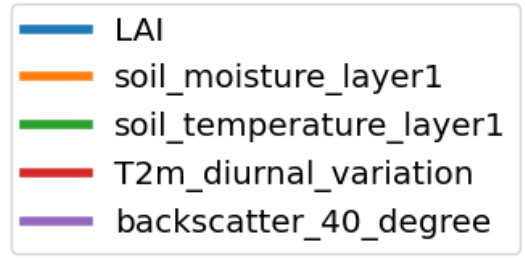
Information content of active microwave database (ASCAT)

Correlation map



Latitude transect of model variables and backscatter





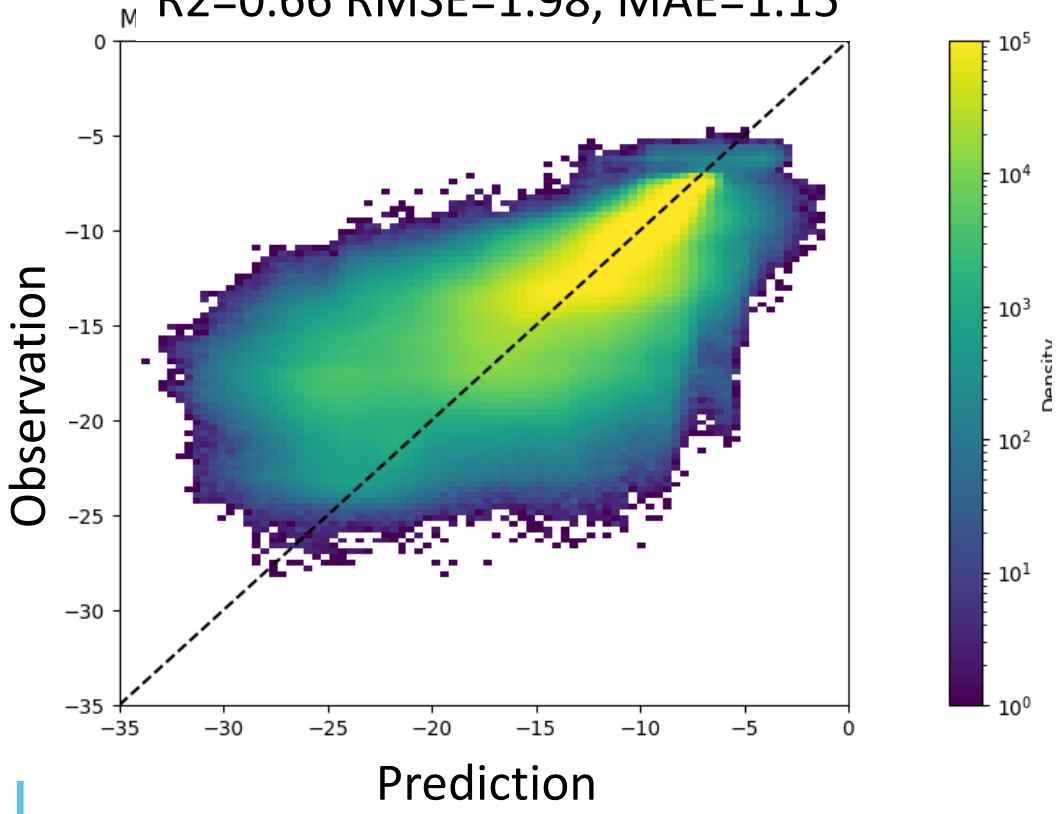
Main influent model variables on the backscatter coefficient at 40 degree

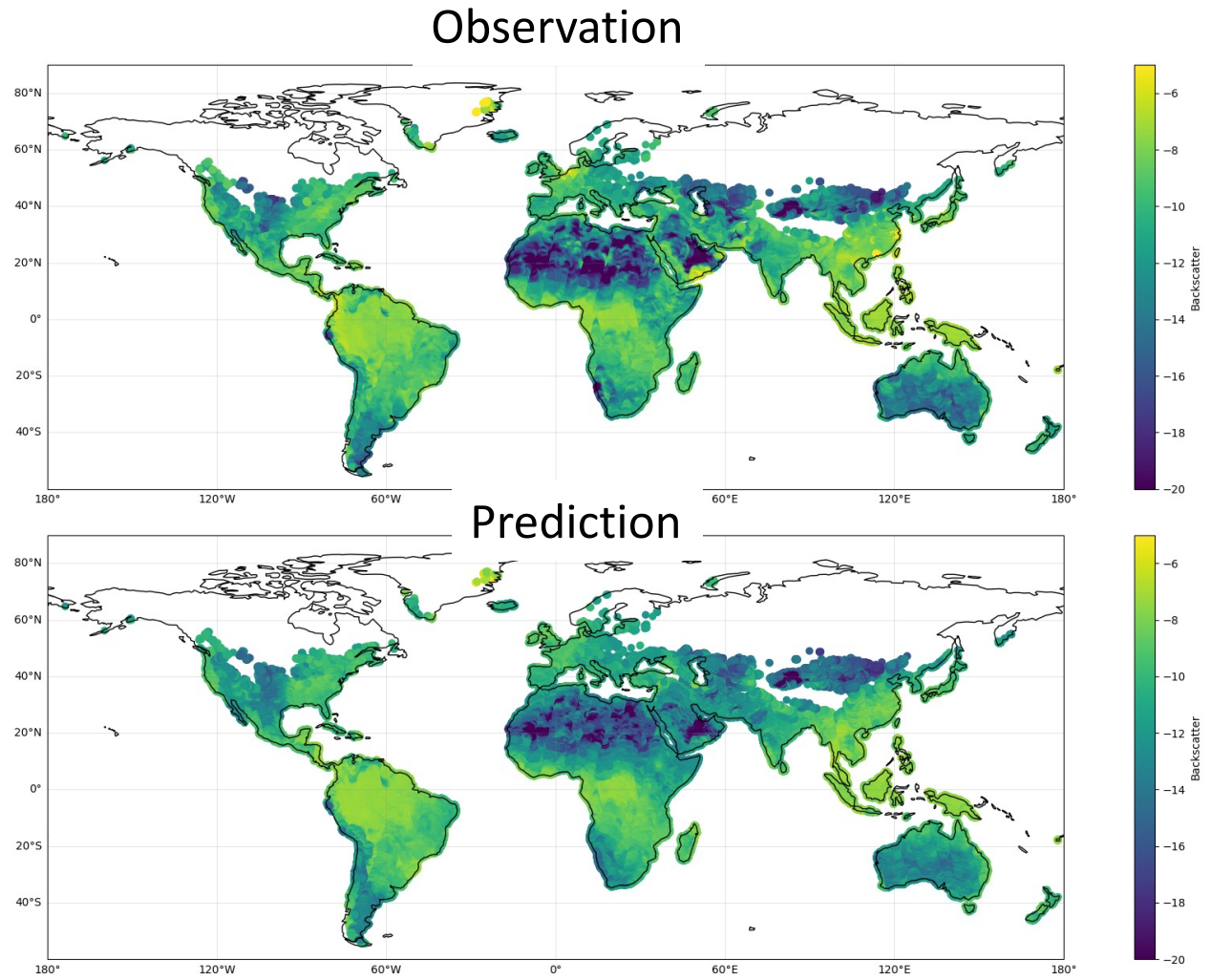
Comparison of ML approaches for forward operator for ASCAT

Gradient boosted trees (xgboost package)

Evaluation over 2017 test set

R2=0.66 RMSE=1.98, MAE=1.15



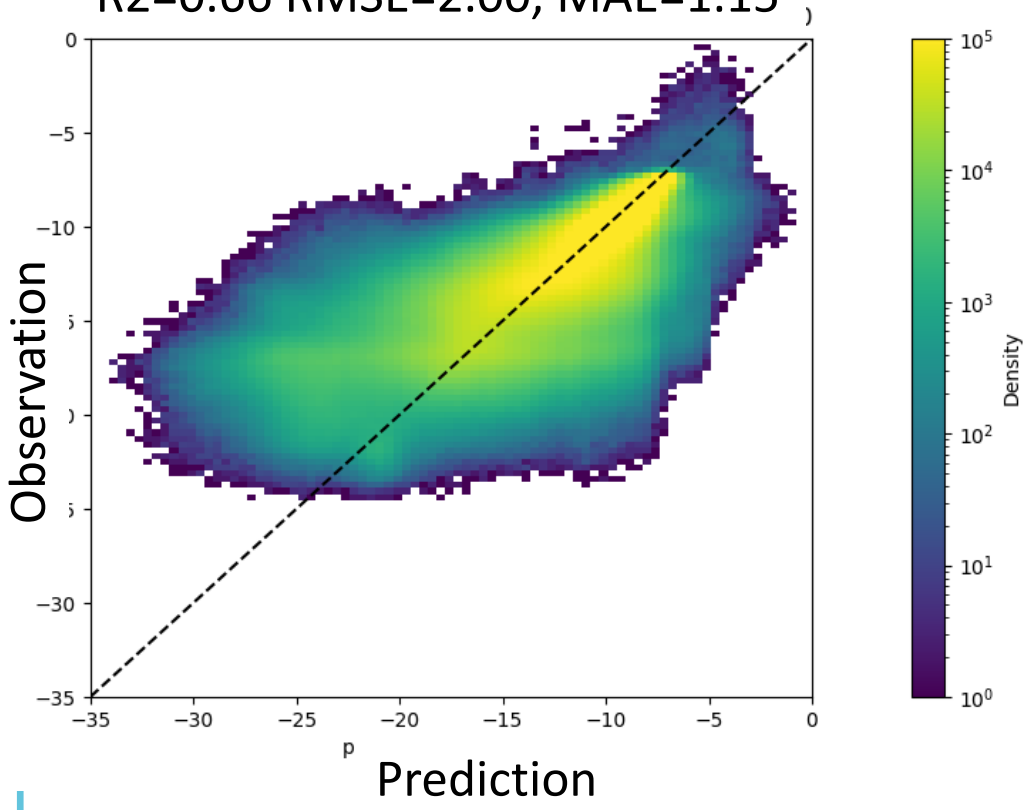


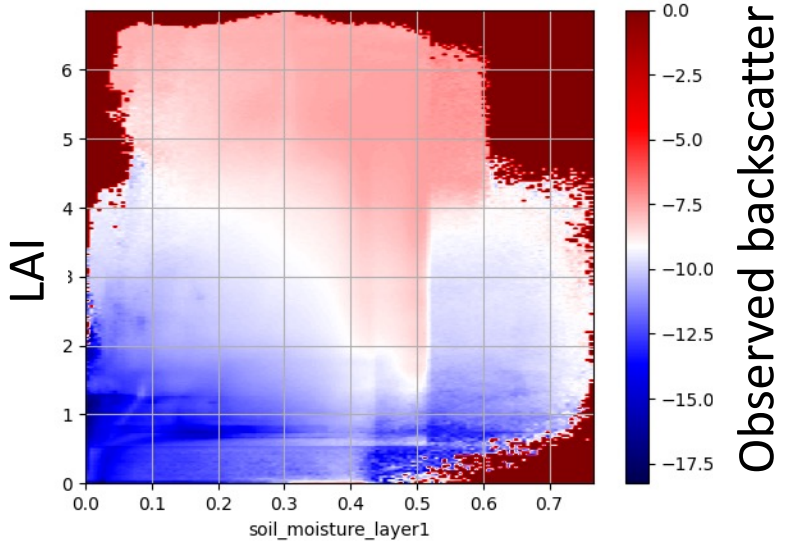
Comparison of ML approaches for forward operator for ASCAT

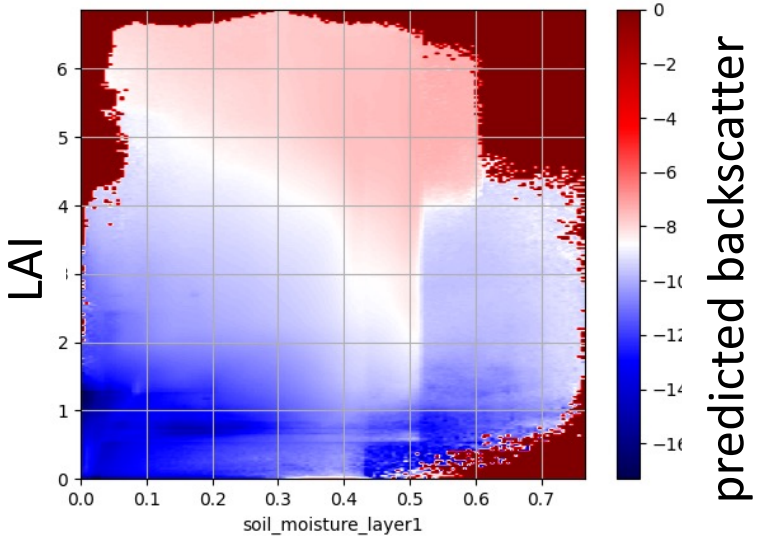
Feedforward neural network (2 hidden layers, 60 neurons, LeakyReLu)

Evaluation over 2018 test set

R2=0.66 RMSE=2.00, MAE=1.15







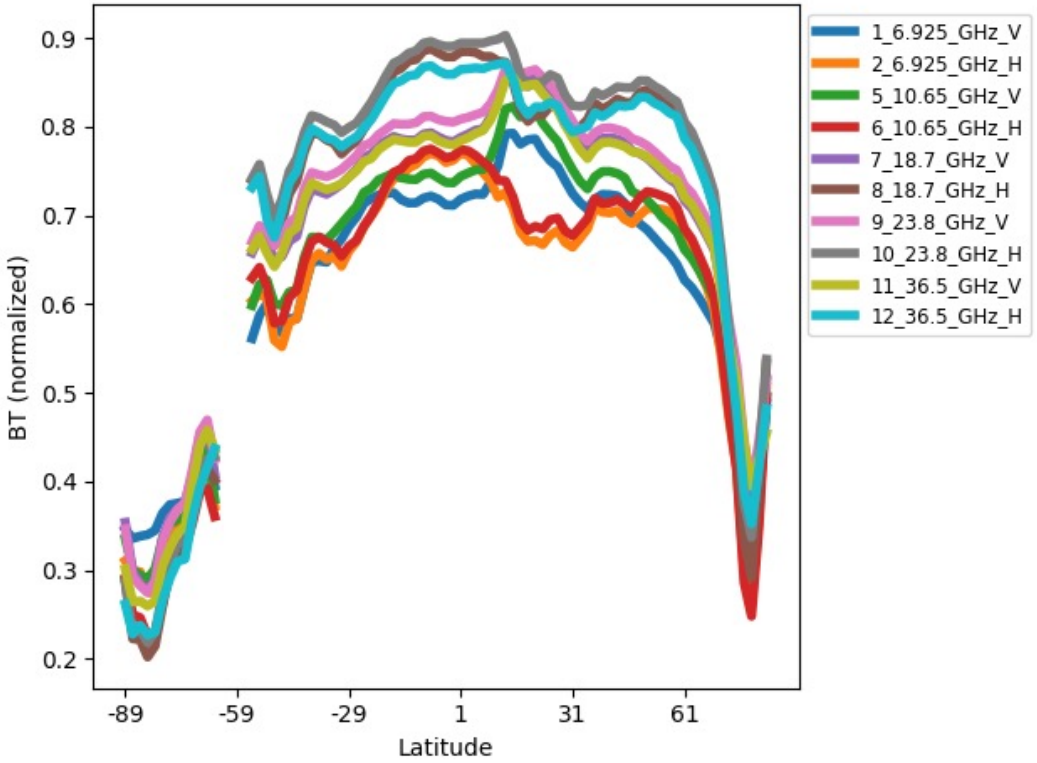
Conclusions and next steps

- Implement the observation operator in IFS to test level-1 microwave assimilation
- Application to passive microwave: AMSR-2 (C,X,Ka,Ku bands)+SMOS (L-band)
- Develop a training database and a NN for SIF
- Test other deep learning architecture: use of convolutional layer and autoencoder to deal with lack of information content and uncertainties in the model variables

Information content of passive microwave database (AMSR2)

Latitude transects (land only)

AMSR2 observations



Model fields