

An aerial photograph of a dense evergreen forest, likely a spruce or fir forest, with a high density of trees. The canopy is thick and green, with some areas showing more yellowish-green, possibly due to sunlight filtering through. The text is overlaid on the center of the image.

GPP prediction of needleleaf evergreen forests

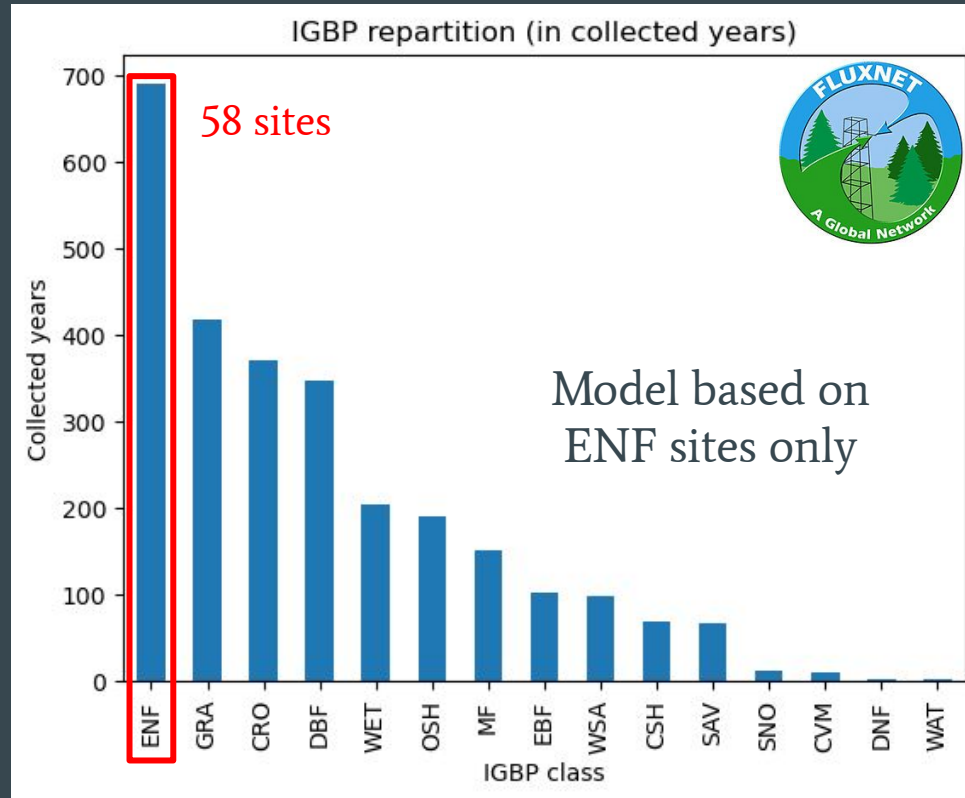
...

Mayeul Marcadella

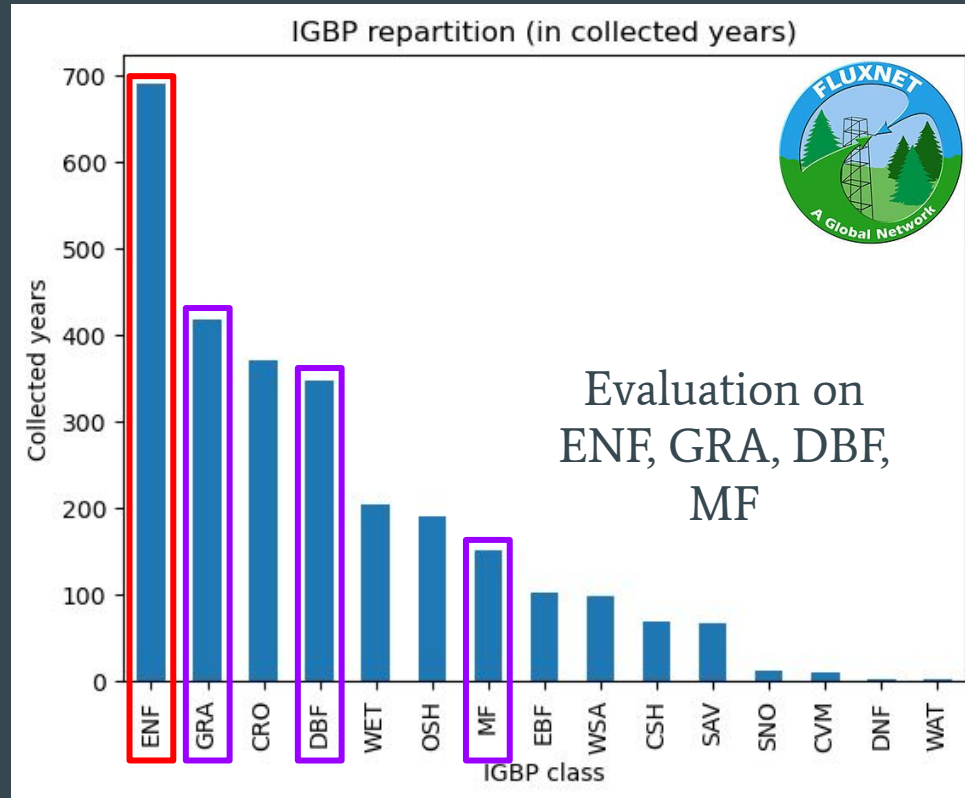
Objectives

- Model GPP for one plant functional type using local and/or remote sensed data.
- Do models generalize through time and across functional types?

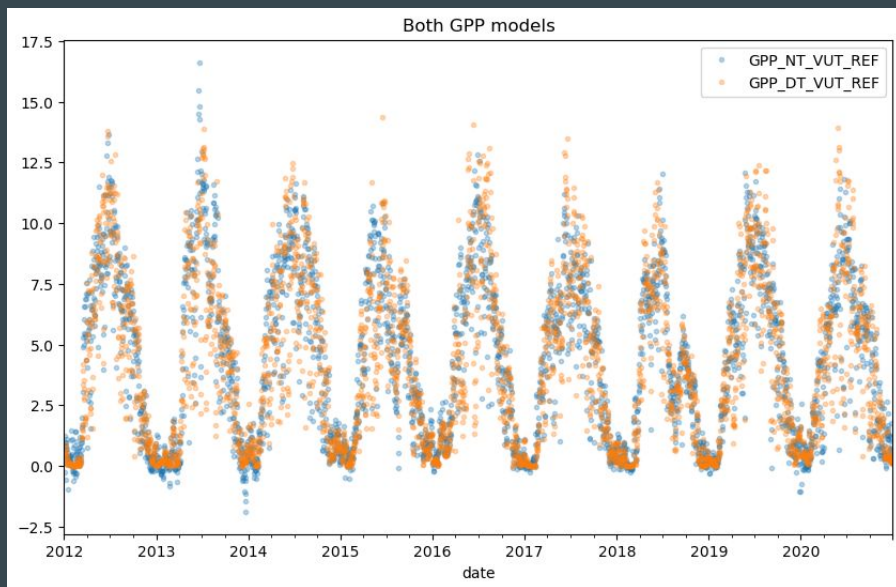
Methods - Data selection



Methods - Data selection

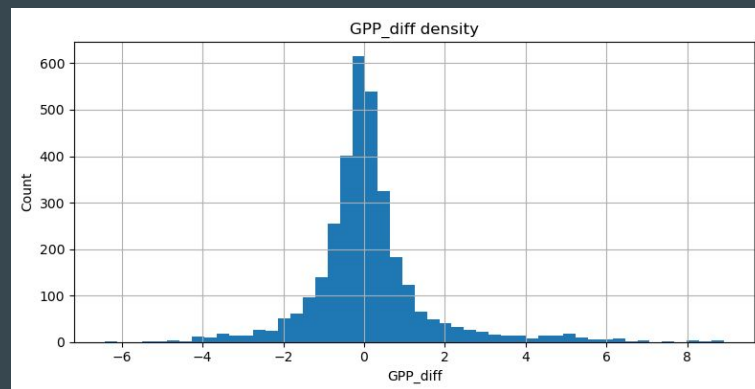


Methods - CO₂ flux partitioning

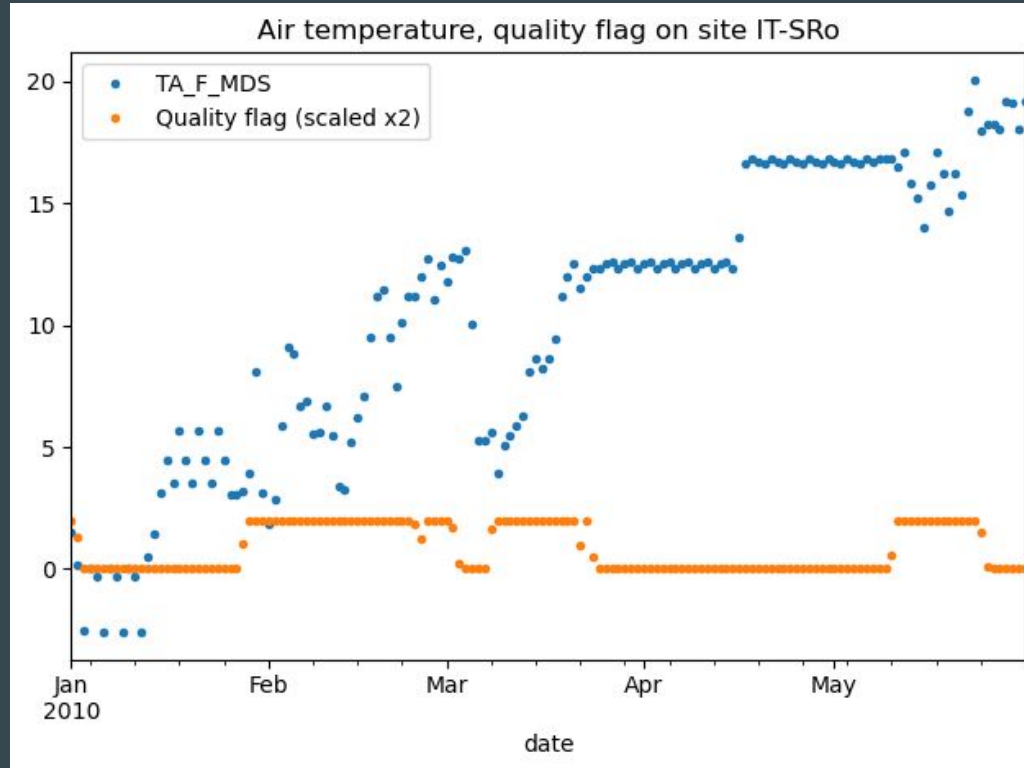


$$(GPP, Reco) = P(NEE, env)$$

$$GPP = \text{mean}(P_{NT}, P_{DT})$$

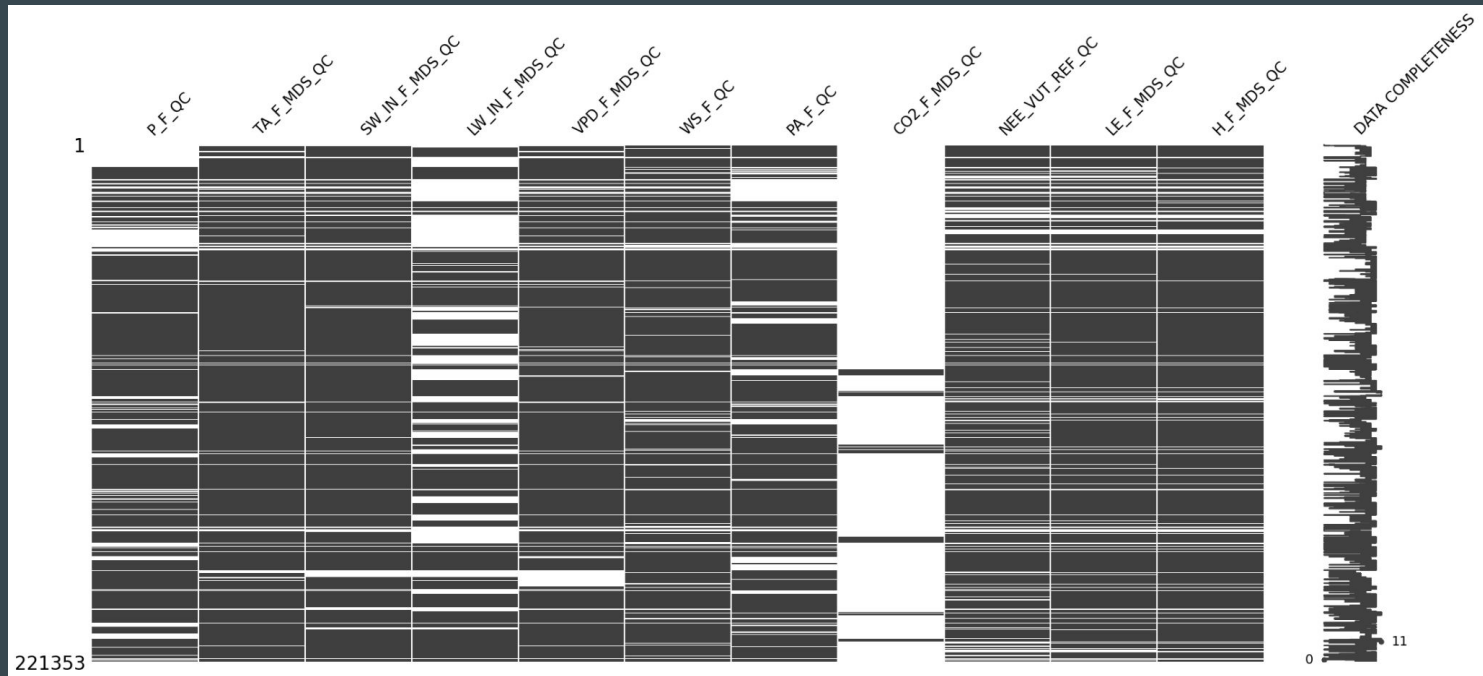


Methods - Data cleaning



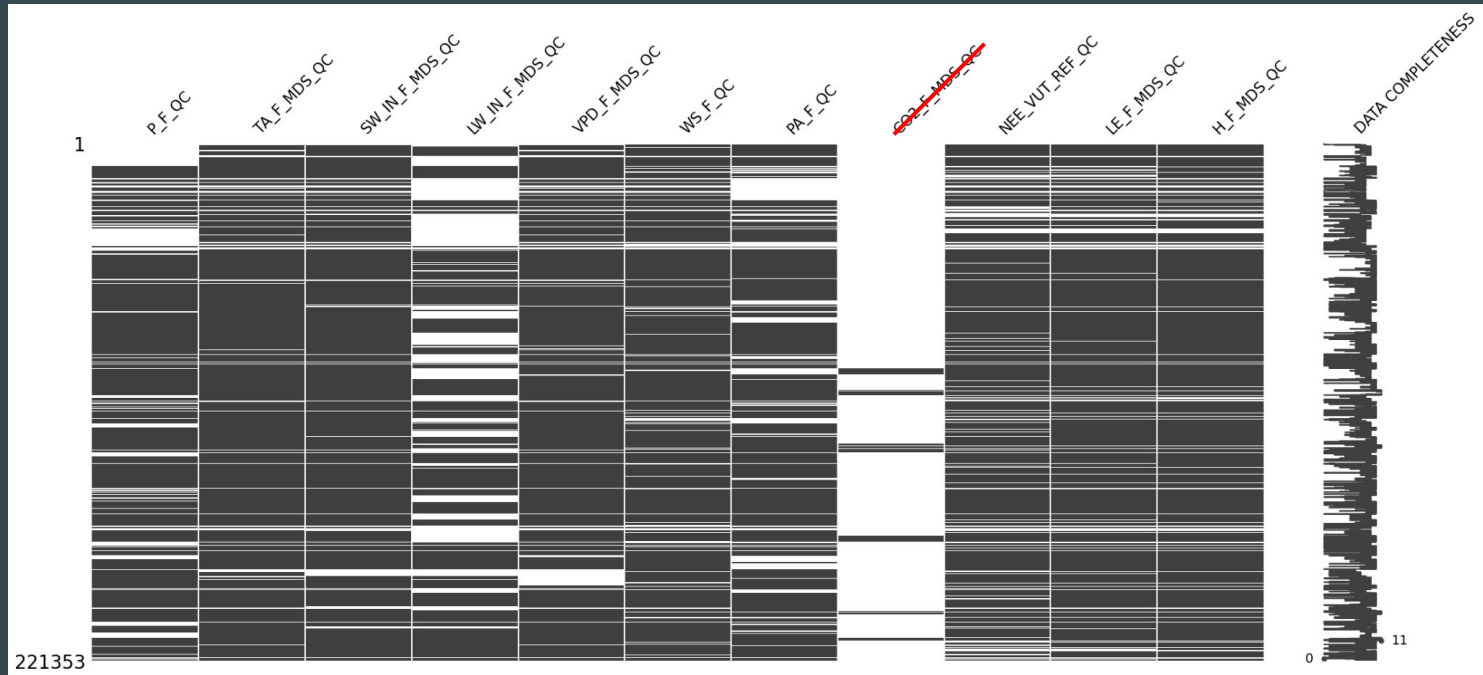
Methods - Data cleaning

In white: $QC < 0.5$



Methods - Data cleaning

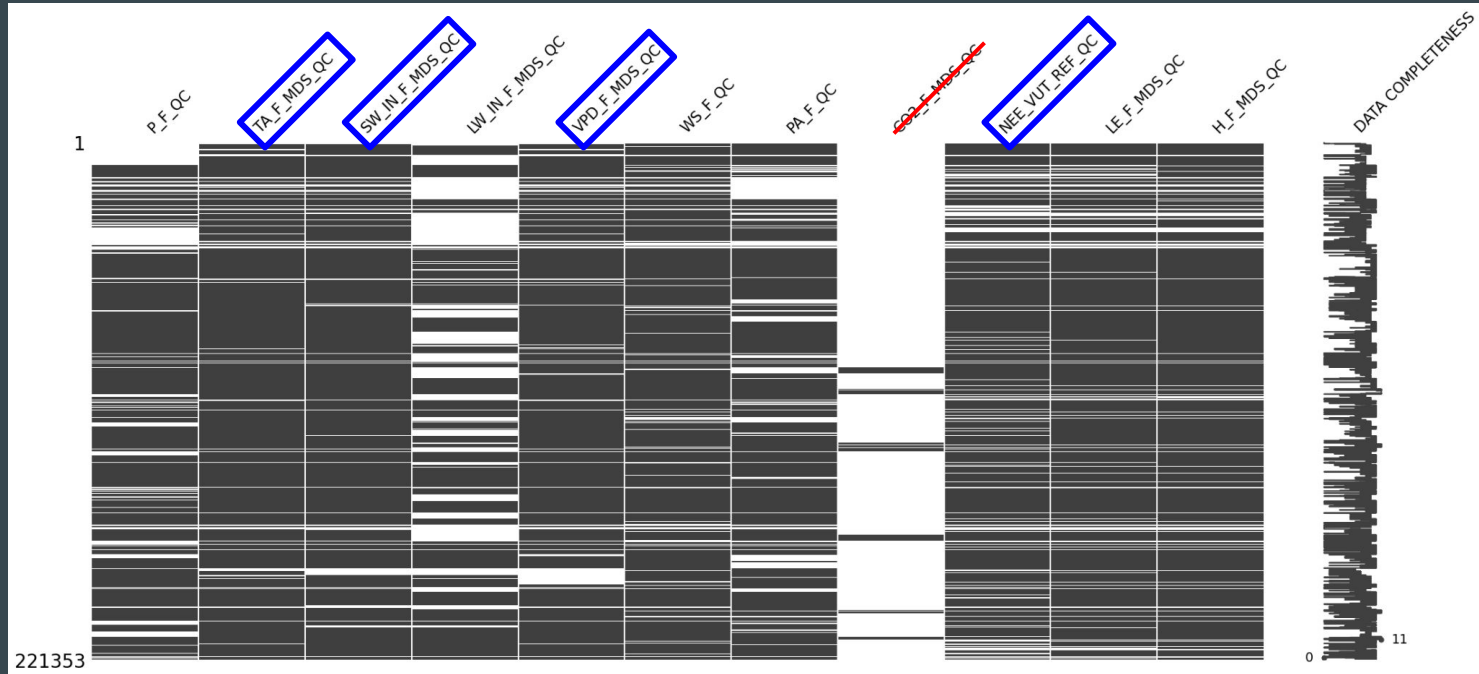
In white: $QC < 0.5$



Methods - Data cleaning

In white: $QC < 0.5$

→ 172k samples



Methods - Feature selection

Measured

- Temperature*
- VPD*
- Precipitation*
- Atmospheric pressure
- Wind speed
- SW incoming radiation
- LW incoming radiation
- Latent heat flux
- Sensible heat flux
- LAI
- FPAR

Computed

- $\text{apar} = \text{FPAR} * \text{SW_in}$
- Temperature squared
- Daylength
- Solar altitude

* With rolling windows for atemporal models (see next slide).

Note: I discarded Reco as modelling the flux partitioning methods was not the goal. But why not use NEE??

Methods - Atemporal vs temporal

Atemporal models

Added context to the data with 1 and 4 weeks averaging rolling windows for:

- temperature
- precipitation
- vapor deficit

221k → 172k

Temporal models

- Split data in blocs of 64 contiguous days.
- Discard any bloc which contains any sample not meeting the QC criteria.

221k → 172k → 92k

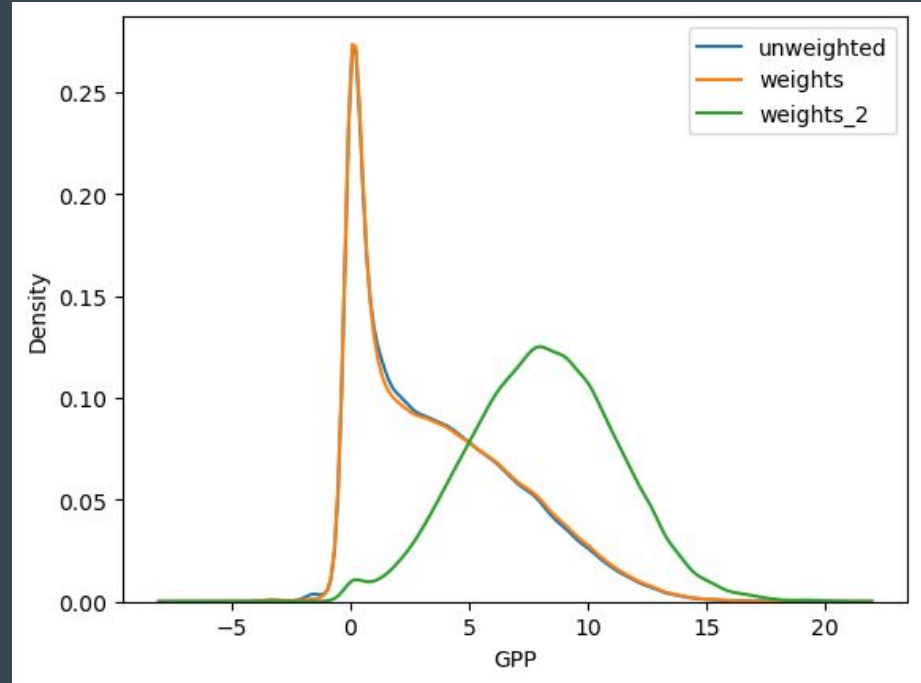
Methods - Weight based alternative

- $\text{weight} = \sum w_i Q_i$
- $\text{weight_2} = (\text{GPP}+1)^2 \cdot \sum w_i Q_i$

→ Same results as with atemporal models.

→ Might be worth to try with temporal models (saves a lot of data!)

→ Not used in the remaining of this presentation.



ML Models

Baseline (mean)

2-parameter models

- Ridge 2p (FPAR & LAI)
- kNN 2p (PCA)

Atemporal models

- Ridge
- Random forest
- SVR
- DNN (1 and 2 hidden layers)

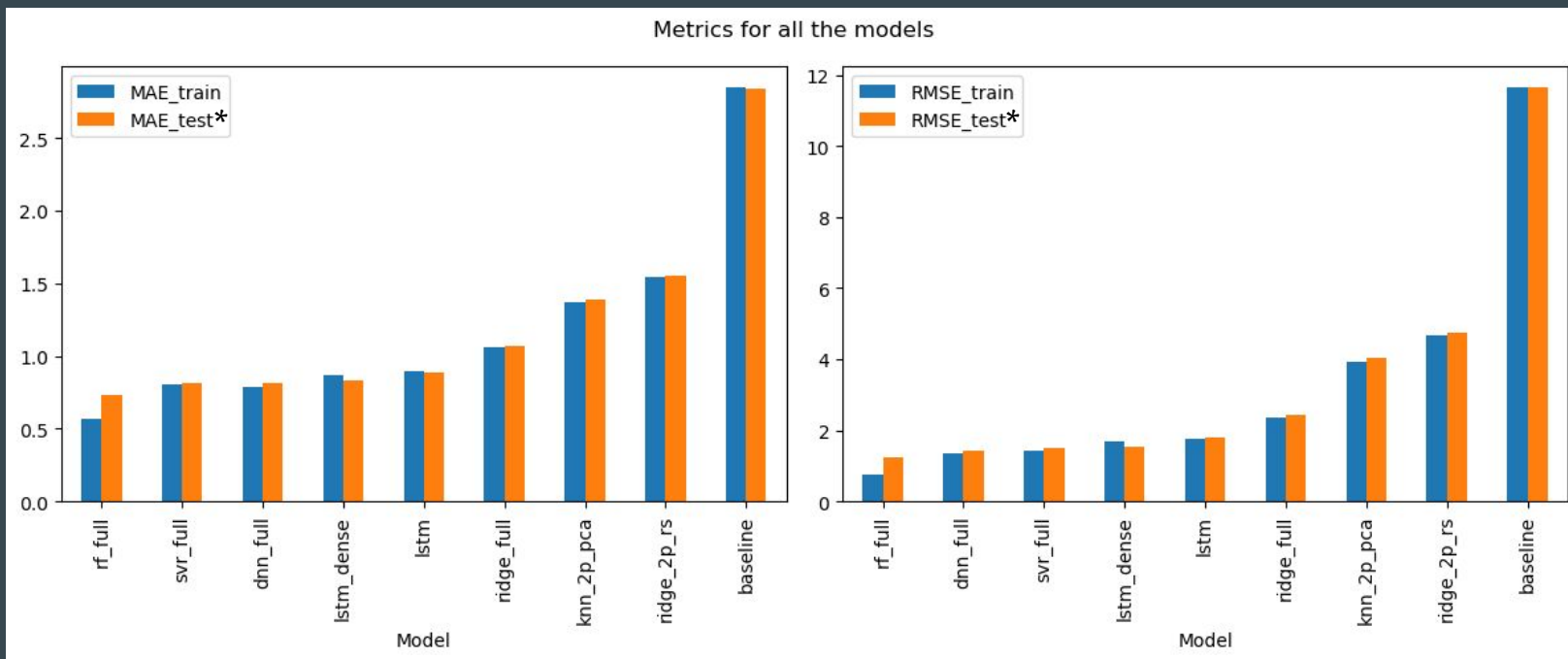
Temporal models

- LSTM
- LSTM + Dense

ML Models - Strategy

- Train-validate / Test splits: 0.7/0.3 with shuffling
- No outlier detection
- Scaling: Quantile transform due to heavy bias
- Grid search with CV
 - 3 splits (shuffle split): 0.7/0.3
 - metric: -MAE
- Tuning hyperparameters two by two, focusing on not overfitting.

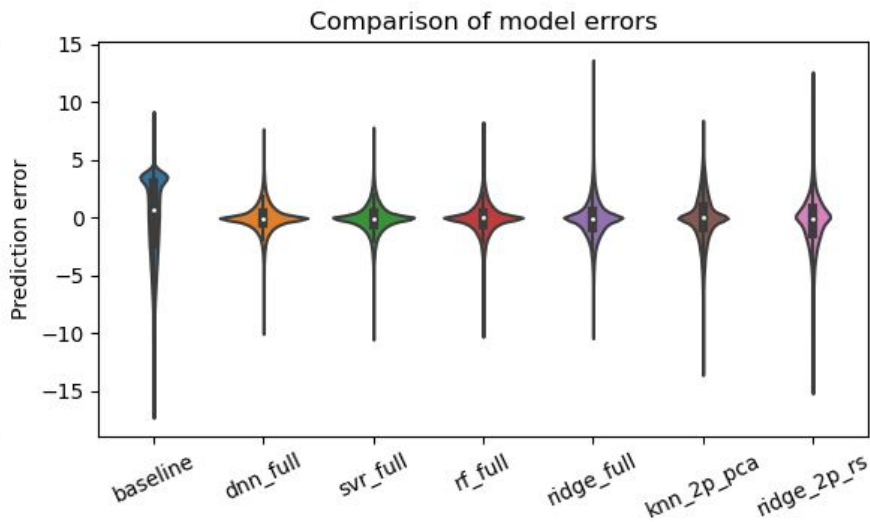
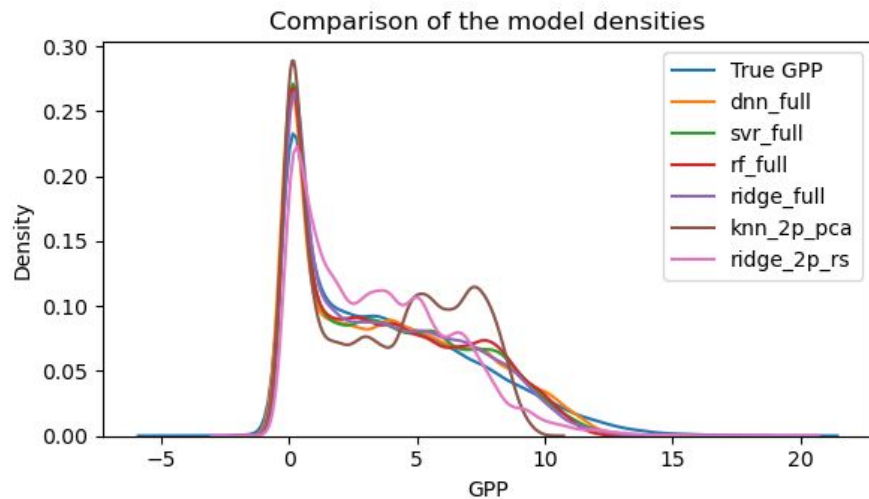
Results - All models



* Out of sample

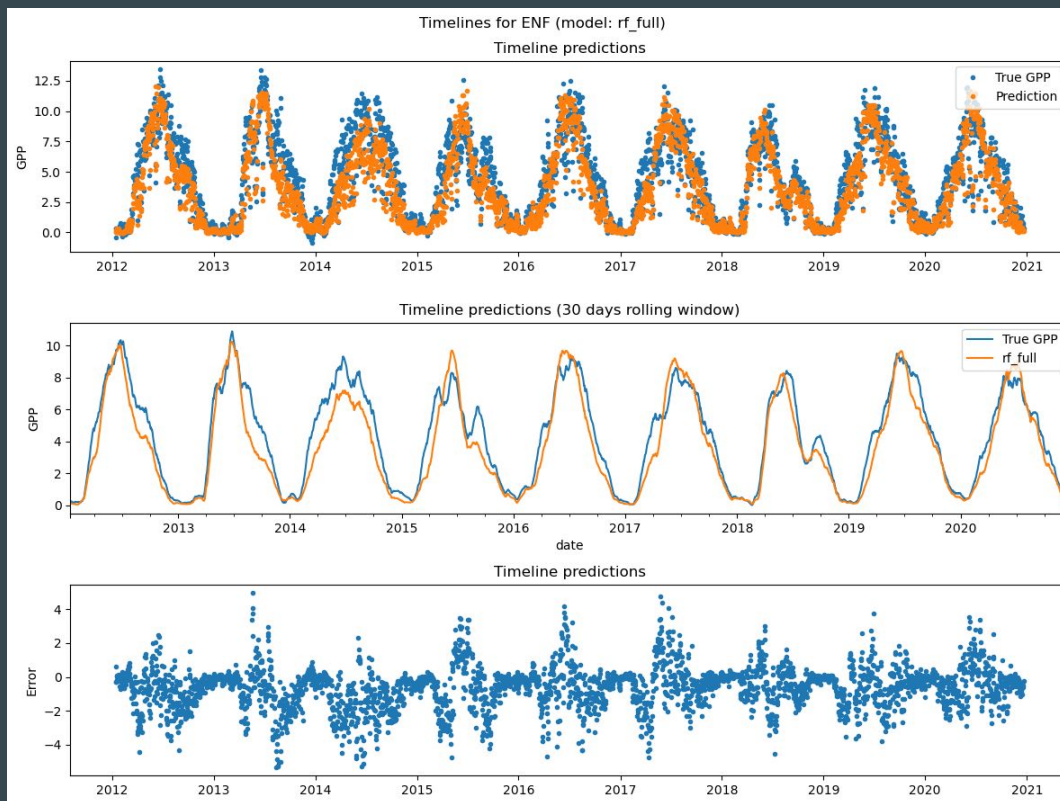
Results - All models

Predicted densities and errors for ENF



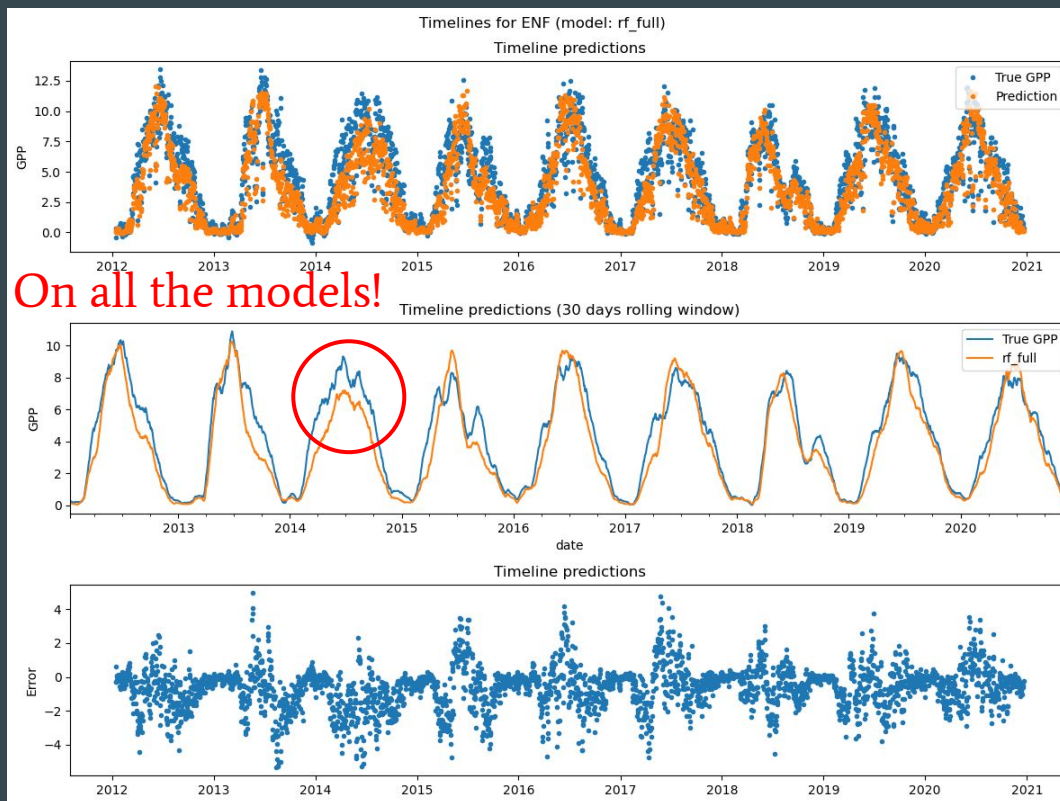
(Out of sample)

Results - Random forest



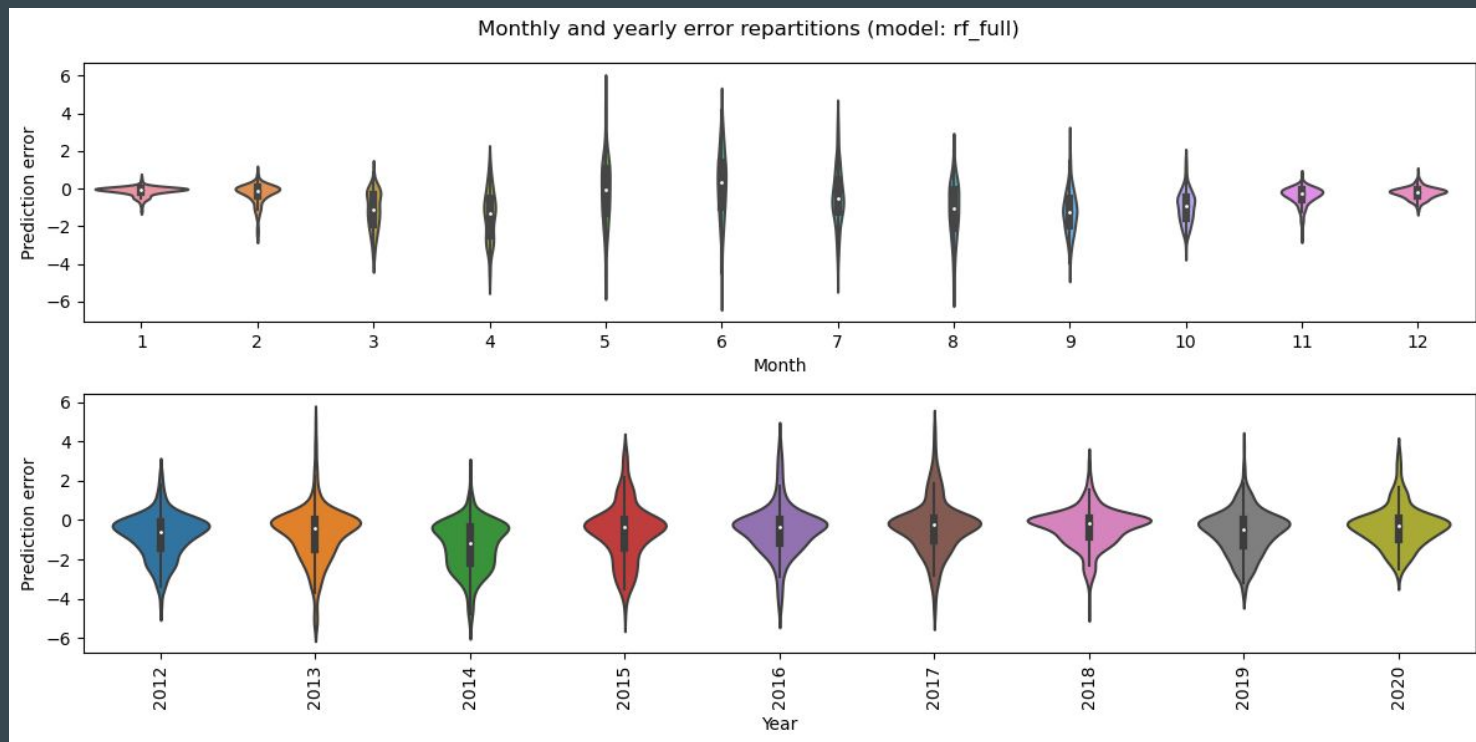
(Out of sample)

Results - Random forest

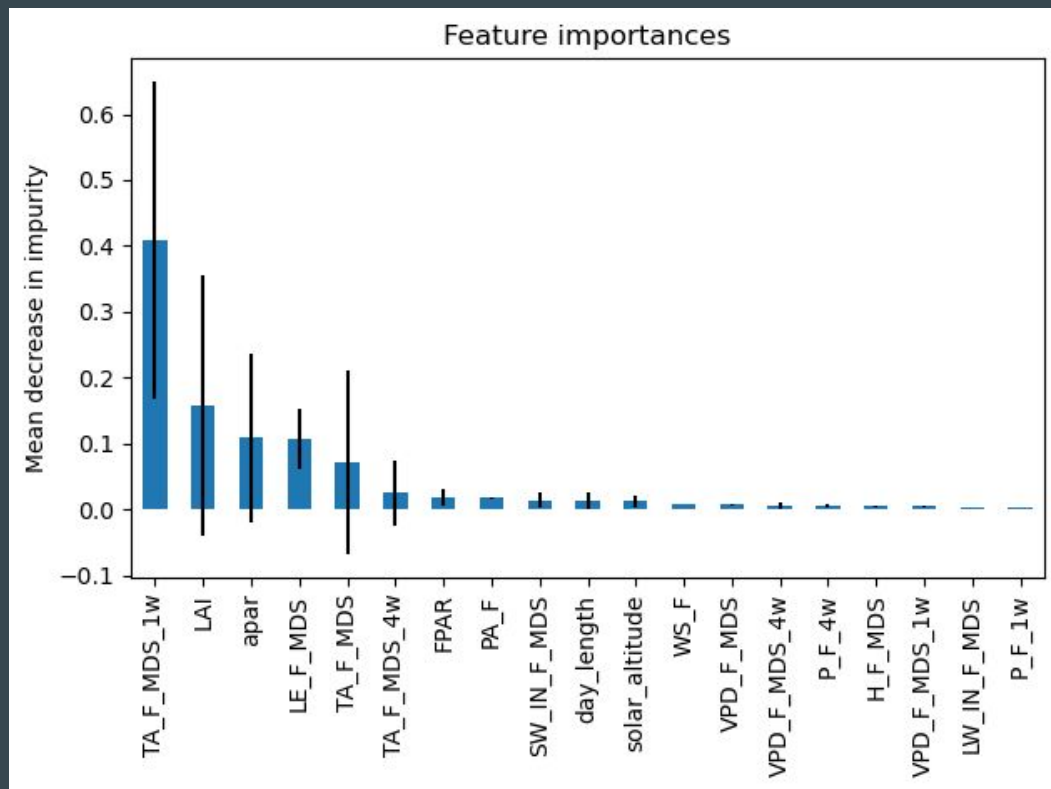


(Out of sample)

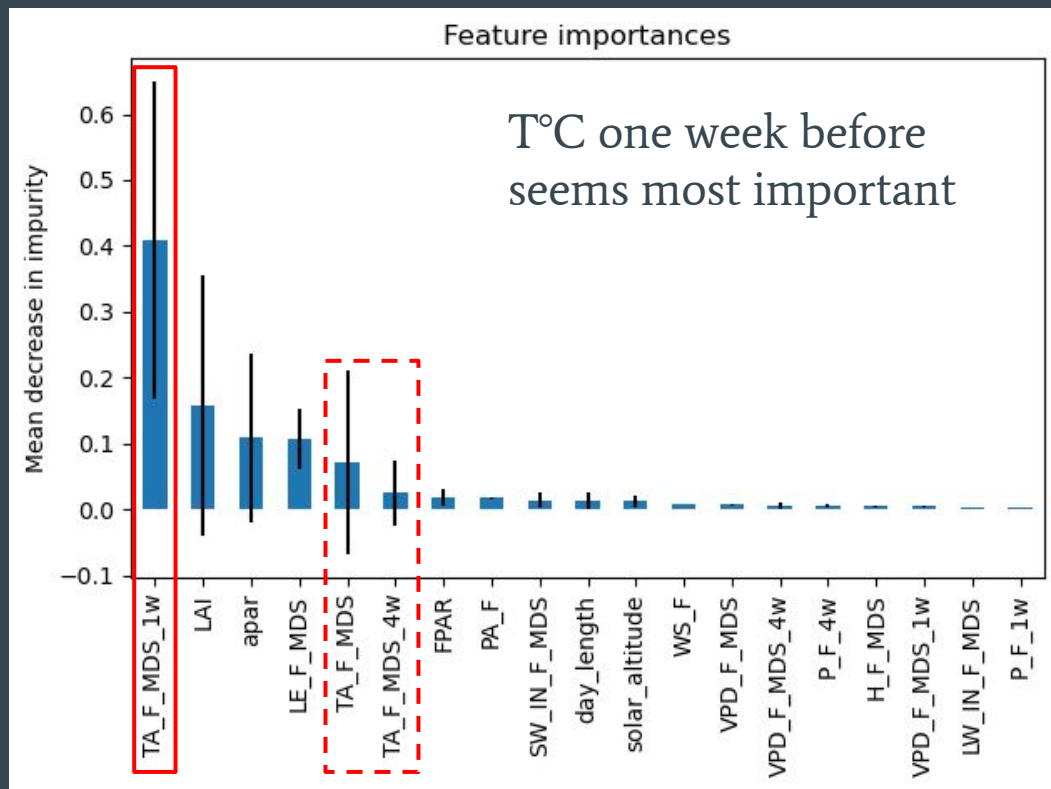
Results - Random forest



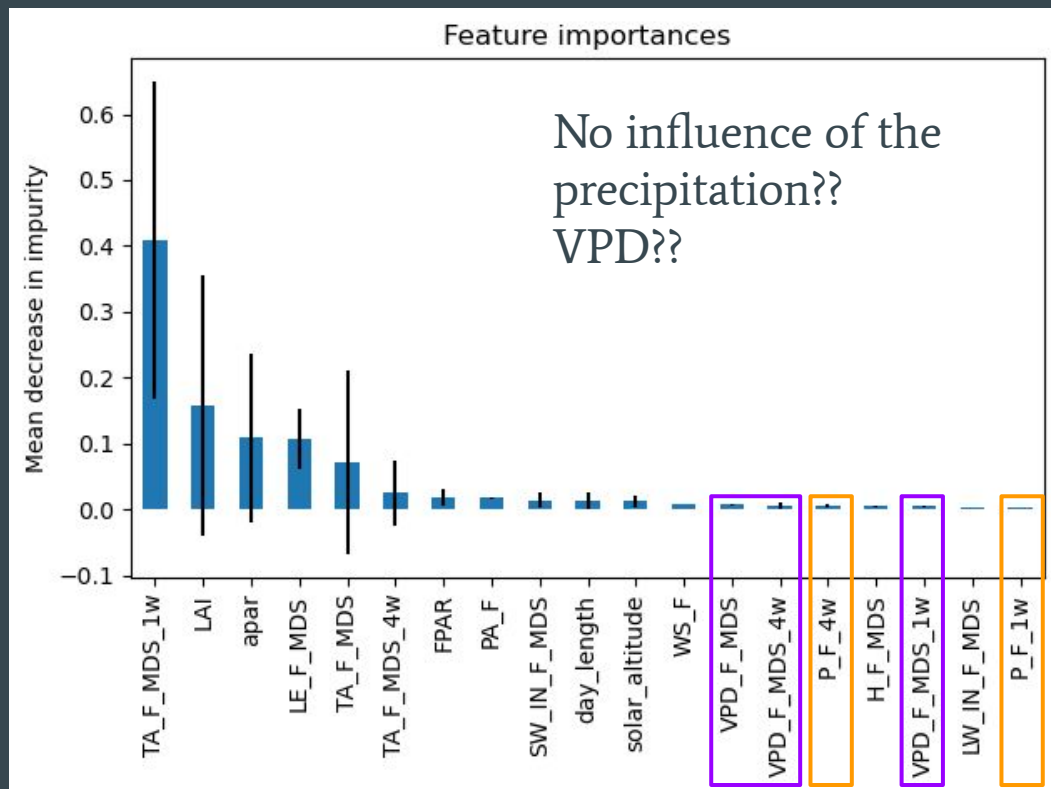
Results - Random forest



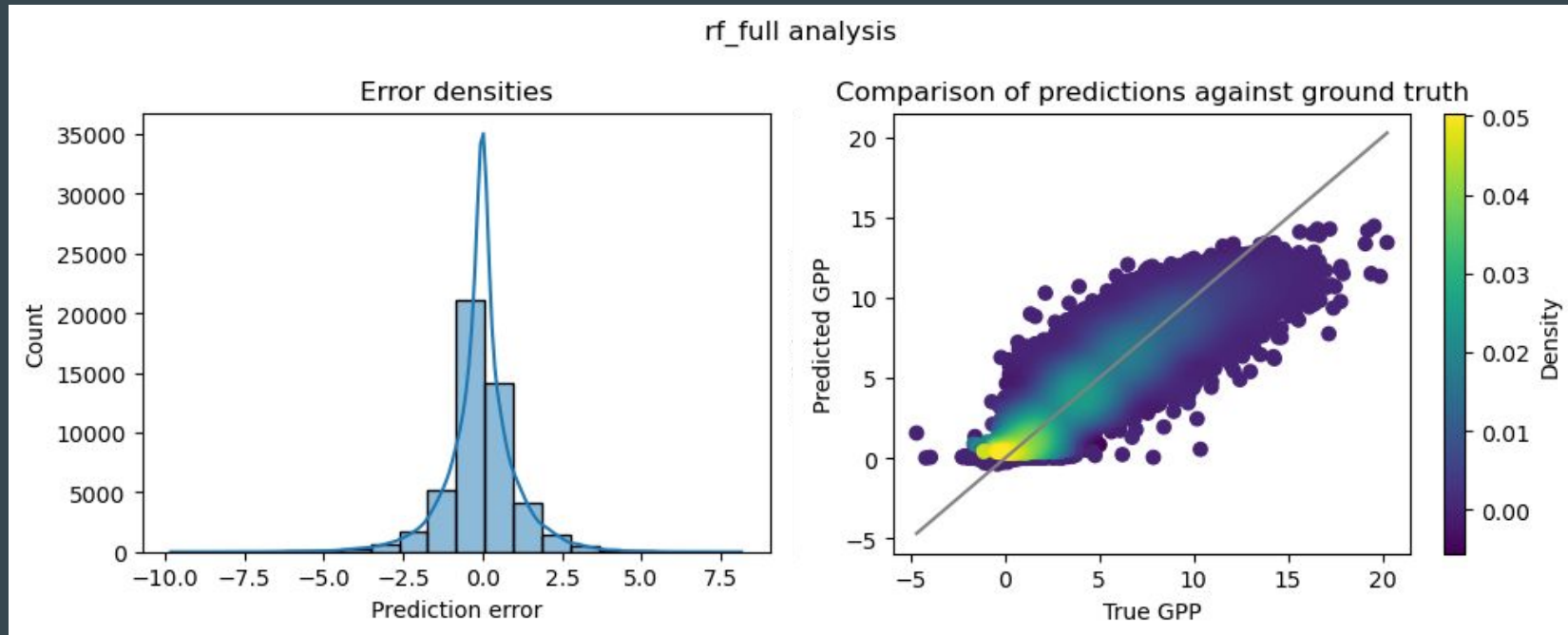
Results - Random forest



Results - Random forest

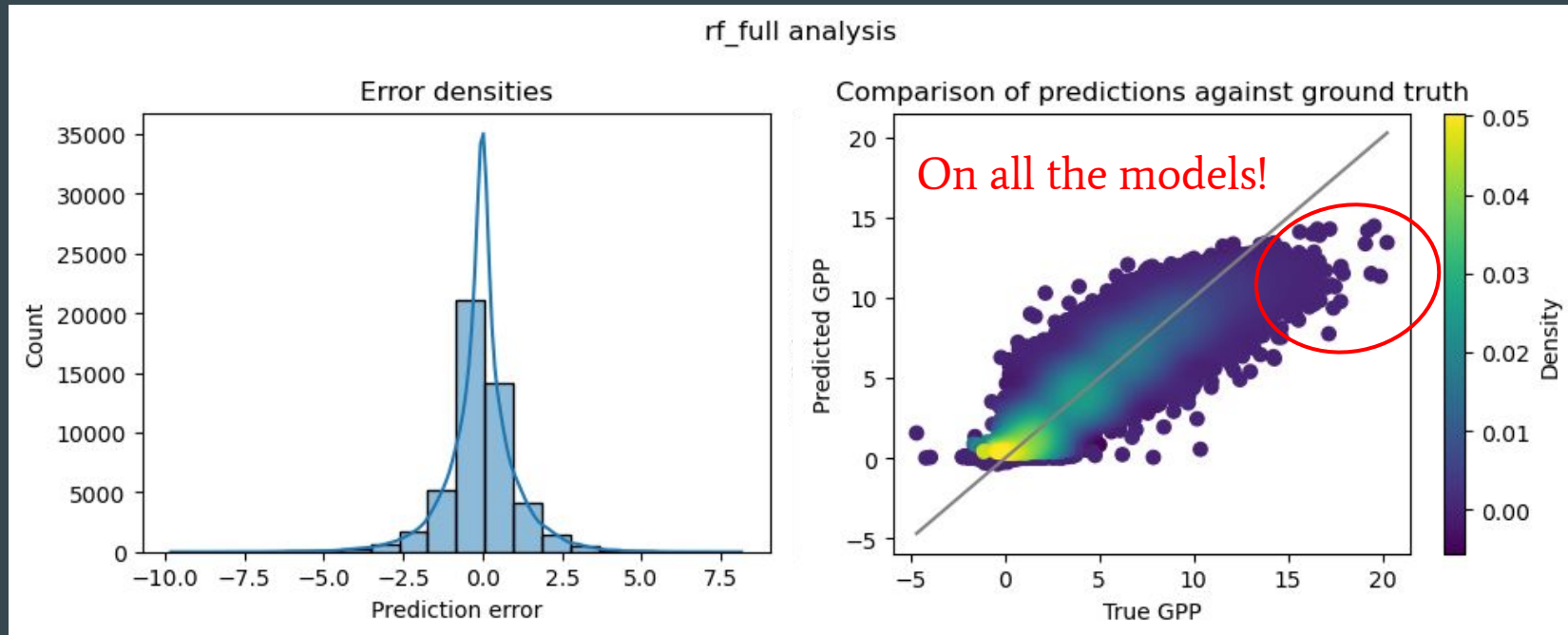


Results - Random forest



(Out of sample)

Results - Random forest

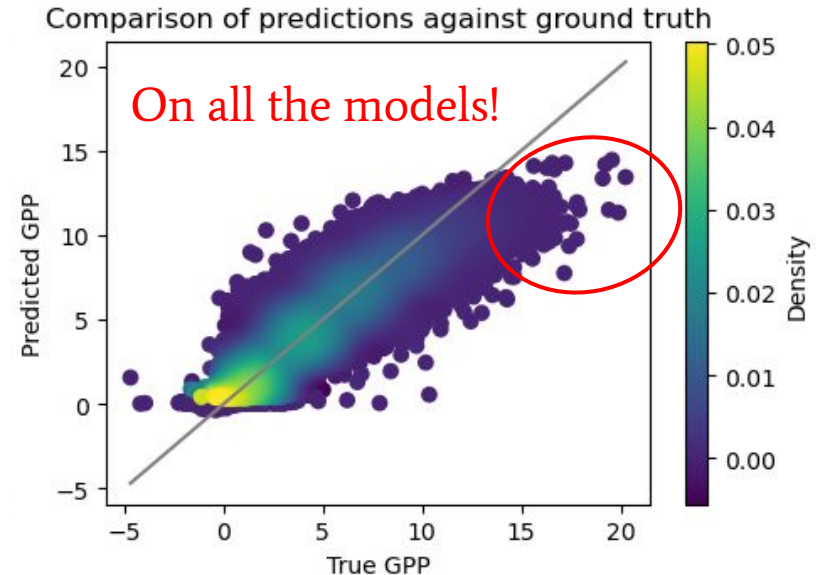


(Out of sample)

Results - Random forest

- Different meteorological patterns??
- Hidden variable??
 - Different species? Age, density?
 - Different altitude (photorespiration)?
 - Soil composition, pH, temperature, ...
- Time-dependent behavior?

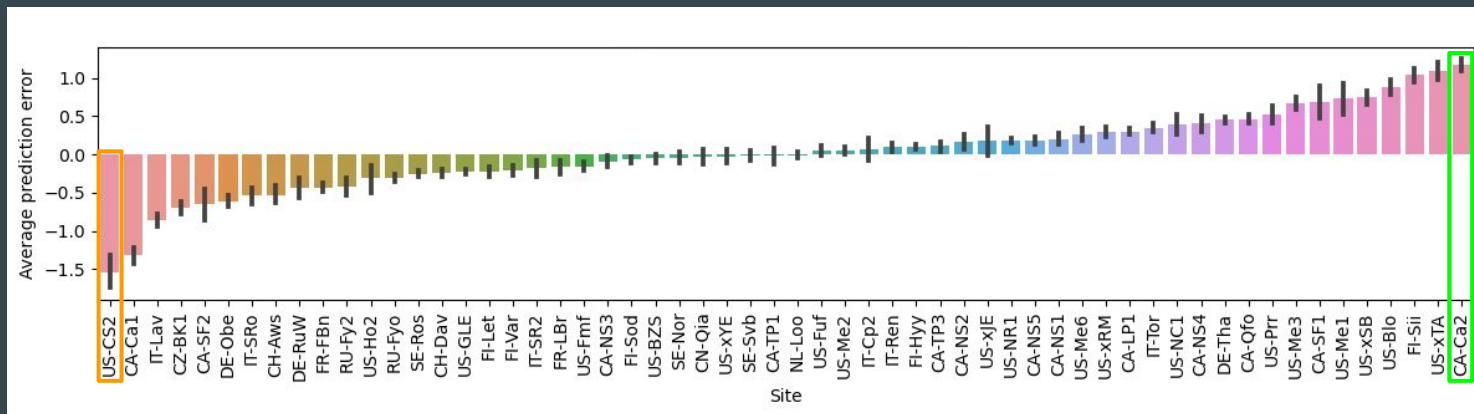
rf_full analysis



(Out of sample)

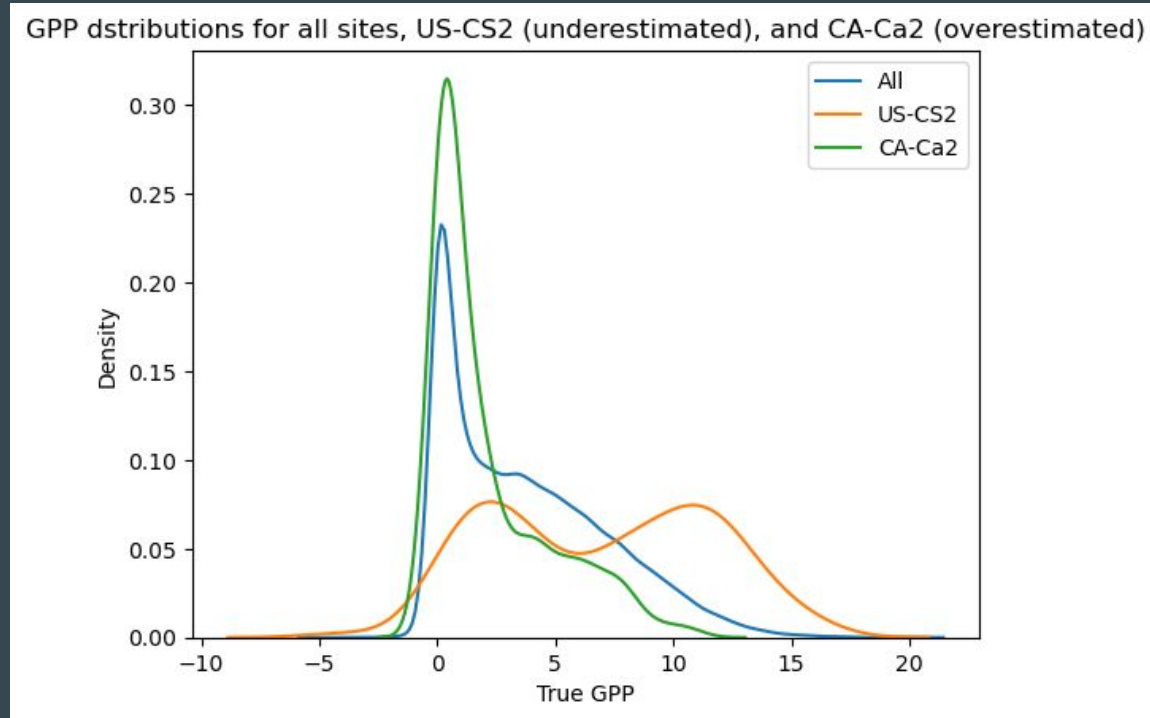
Results - Random forest

Error per site

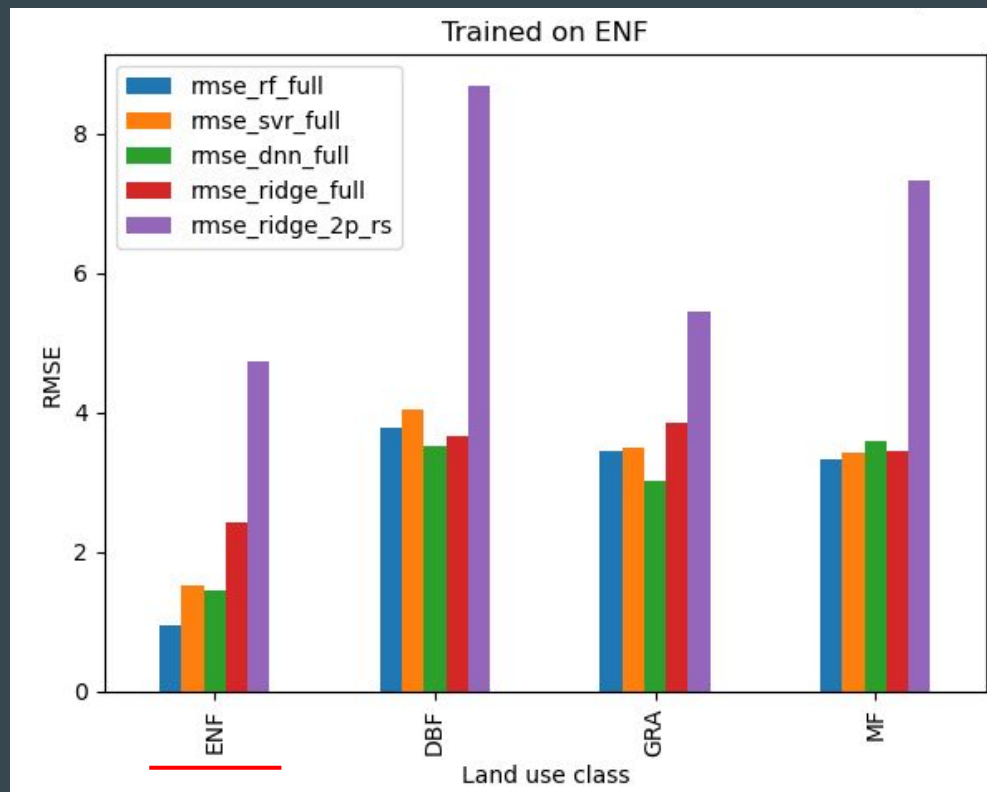


(Out of sample)

Results - Random forest

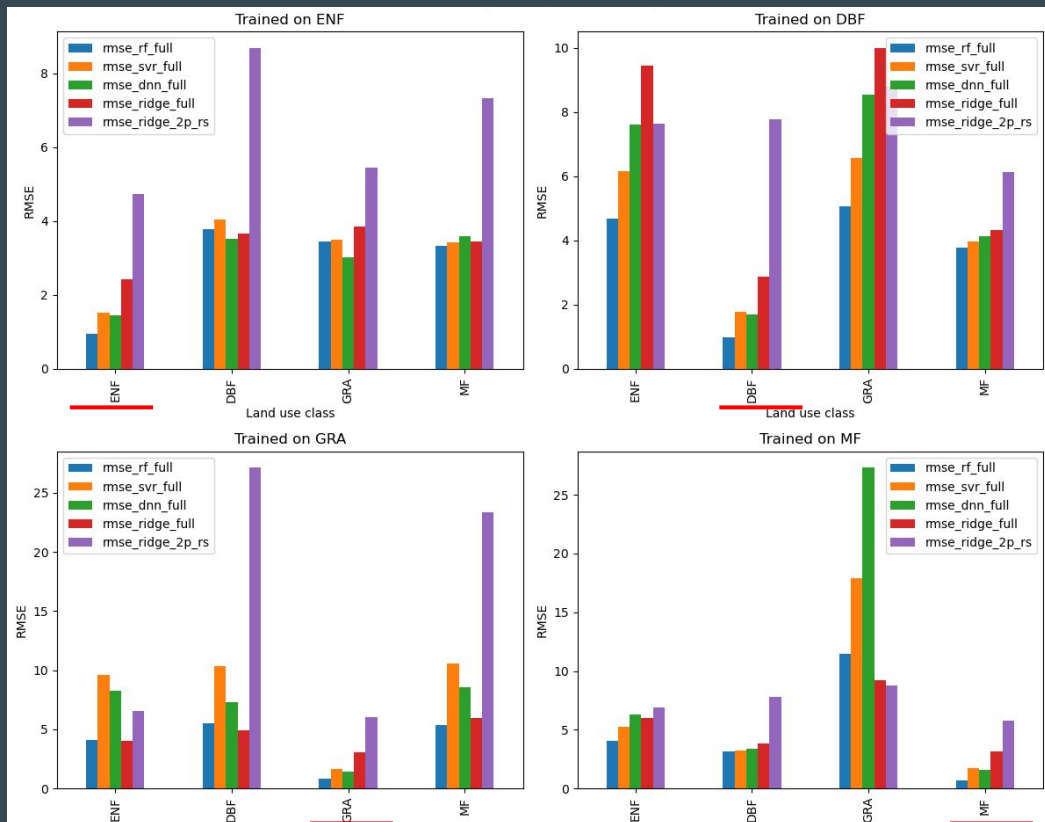


Results - Cross type generalization



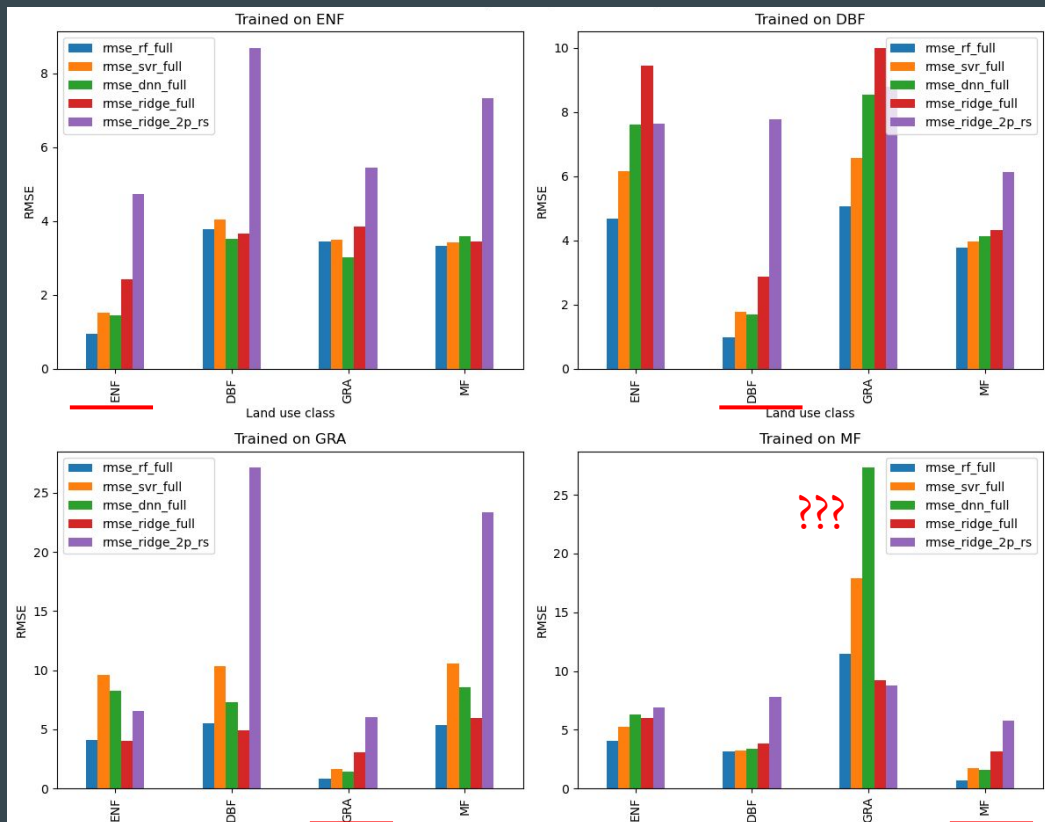
(Out of sample)

Results - Cross type generalization



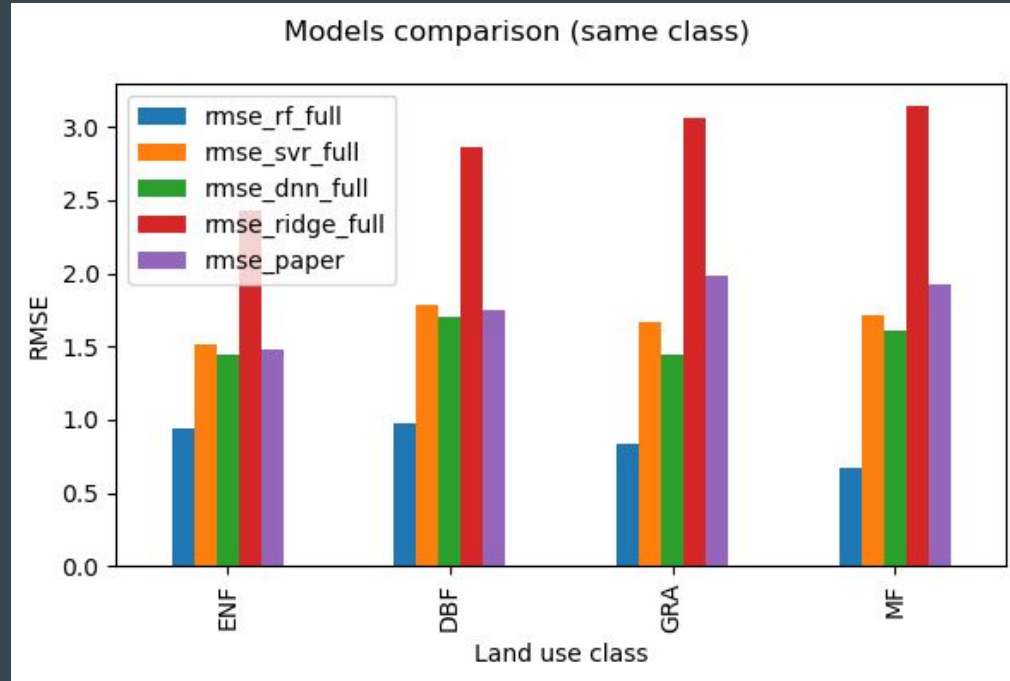
(Out of sample)

Results - Cross type generalization



(Out of sample)

Results - Unfair comparison with the paper

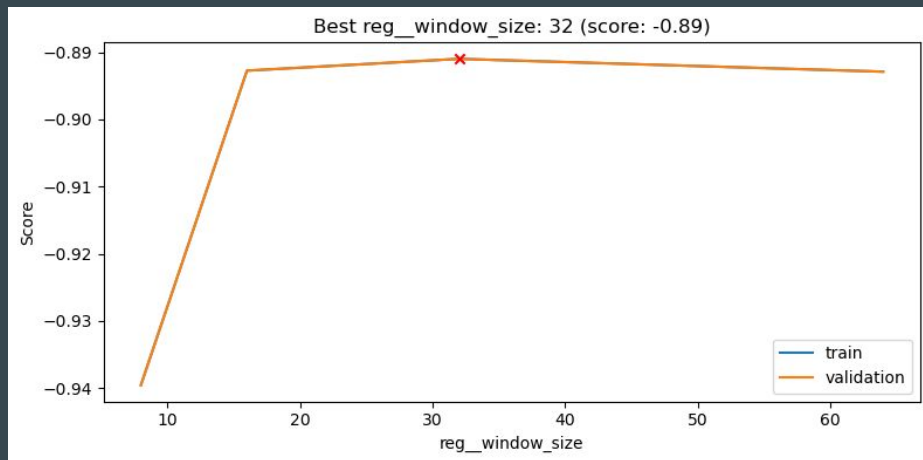


(Out of sample)

My two cents

- For DNN, 700 years = 250k samples
- For LSTM, 700 years = 700 (1-year) samples

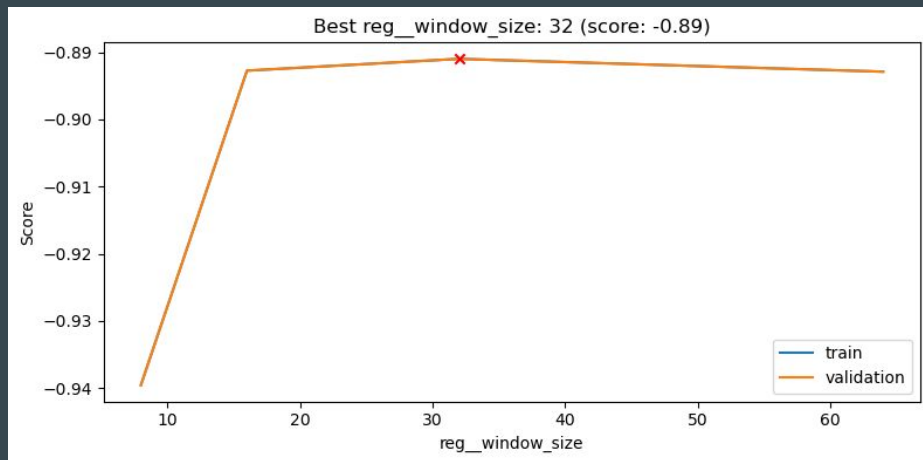
→ Not enough data to train LSTM on long time windows.



My two cents

- For DNN, 700 years = 250k samples
- For LSTM, 700 years = 700 (1-year) samples

→ Not enough data to train LSTM on long time windows.



→ Try 1d-CNN on half-hourly data?

A nice read

Partitioning net carbon dioxide fluxes into photosynthesis and respiration using neural networks

2020, Tramontana et al.

<https://doi.org/10.1111/gcb.15203>

Thank you



Notebooks: <https://github.com/marcadella/GPP-prediction>