High-spatial-resolution gross primary production estimation from Sentinel-2: A baseline for future SIF integration

Emma De Clerck[1], Pablo Reyes-Muñoz[1], Egor Prikaziuk[2], Dávid D.Kovács[3], Jochem Verrelst[1]

[1] Image Processing Laboratory, Universitat de València
[2] Faculty of Geo-Information Science and Earth Observation, Universiteit Twente
[3] Department of Geodesy and Geoinformation, Technische Universität Wien

HIGHLIGHTS

- **Hybrid vs. data-driven** models for high-spatial-resolution (20m) **Gross Primary Productivity** from **Sentinel-2** reflectance
- Optimizing **Gaussian processes regression** models with **Active Learning** for 10 **Plant Functional Types**
- Incorporation of **meteorological, soil and DEM** variables
- Validation against **eddy-covariance** flux towers in Europe and US
- Scalable workflow in **Google Earth Engine / openEO** for accessible and reproducible carbon monitoring

1 INTRODUCTION

Problem: Accurate, high-resolution GPP is crucial for local carbon monitoring, especially in heterogeneous landscapes where global products lack spatial detail.

Challenge: Developing scalable, fine-resolution GPP models that capture ecosystem-specific dynamics is complex.

Objective: Develop and validate novel, scalable, PFT-specific high-resolution GPP models from Sentinel-2 and auxiliary variables using both a hybrid and a data-driven approach

2 MATERIAL AND METHODS

HYBRID MODEL

Training data:
- SCOPE radiative transfer model

Validation data:
- EC flux tower data from ICOS
- Sentinel-2 surface reflectance from GEE

Model configuration:
- Gaussian Processes Regression
- 10 PFT-optimized models with active learning

DATA-DRIVEN MODEL

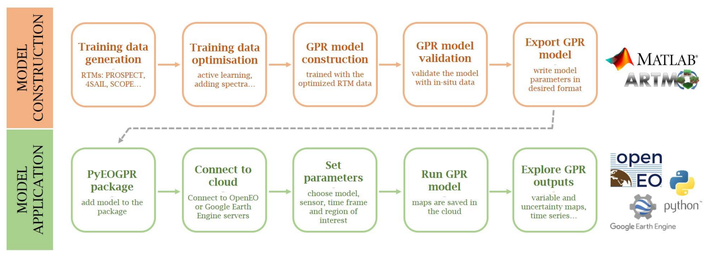
Training & validation data:
- EC flux tower data from ICOS, AmeriFlux and ChinaFlux
- Sentinel-2 SR, Sentinel-1 SAR, DEM, SoilGrids, ERA-5 Land from GEE

Model configuration:
- Gaussian Processes Regression
- Dimensionality reduction: GPR-PCA and GPR-BAT

IMPLEMENTATION AND APPLICATION

PyEOGPR

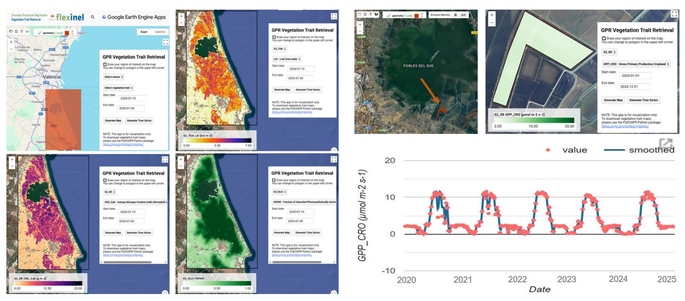
https://pypi.org/project/pyeogpr/



GEE Web Application

Gaussian Processes Regression
Vegetation Trait Retrieval
Google Earth Engine Apps

https://declerckemma.users.earthengine.app/view/gpr-vegetation-trait-retrieval



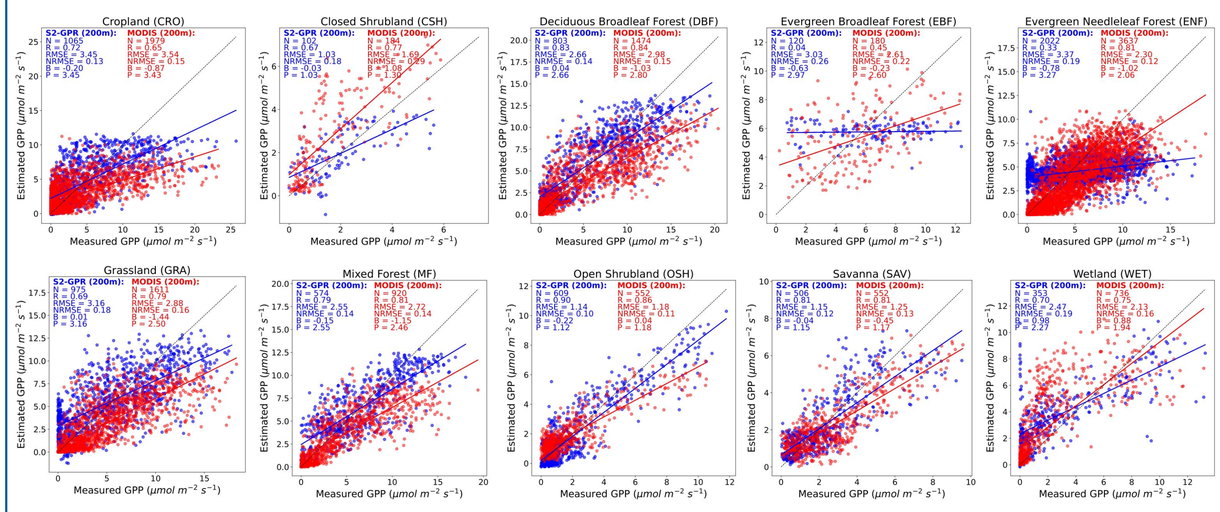
3 RESULTS AND DISCUSSION

Model performance – hybrid models

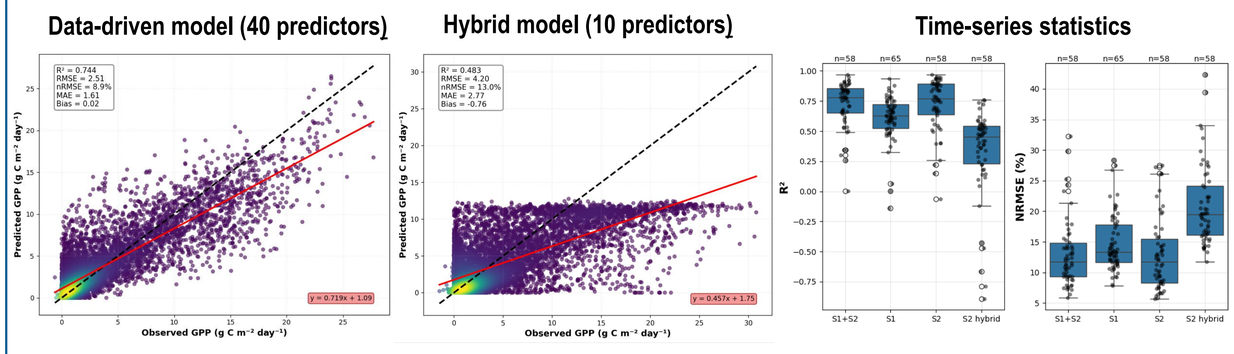
Strong agreement:
- deciduous broadleaf forests, savannas, croplands, open shrublands and mixed forests

Intermediate accuracy:
- closed shrublands, grasslands and wetlands

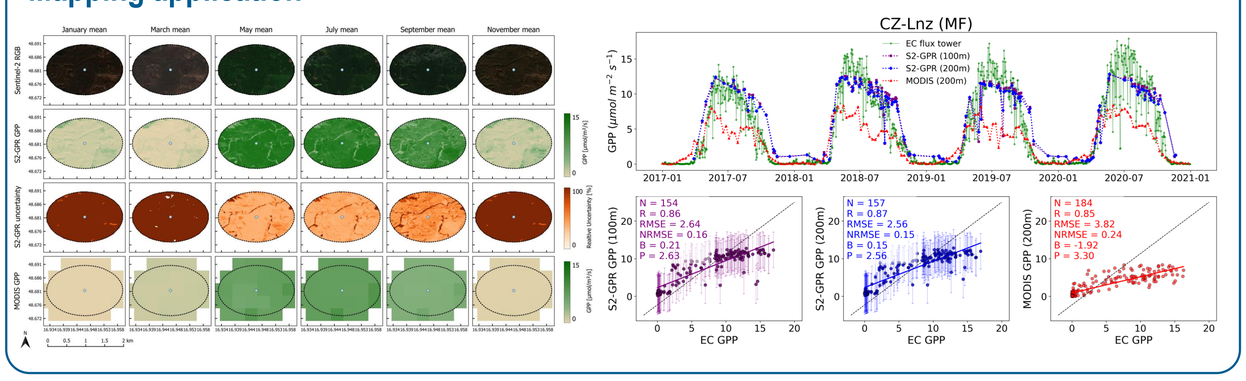
Lower performance:
- evergreen needleleaf forests and evergreen broadleaf forests



Model performance – data-driven models (cropland)



Mapping application



4 CONCLUSION

- PFT-specific hybrid modelling based on Sentinel-2 reflectance and GPR with AL can estimate 20 m GPP with strong agreement to eddy-covariance flux tower measurements
- For cropland, data-driven approach outperforms hybrid model
- Incorporating meteorological, soil, and DEM predictors significantly improved model accuracy and reduced bias.
- The workflow implemented in Google Earth Engine / openEO enables scalable, reproducible, and accessible high-resolution carbon monitoring.

Future Work

- Integrate SIF from FLEX mission for improved GPP monitoring
- Expand data-driven approach to other vegetation types

emma.clerck@uv.es

De Clerck, E., & Verrelst, J. (2026). A user-friendly Google Earth Engine Web Application for vegetation trait mapping and time series analysis from Sentinel-2 and Sentinel-3 imagery. *IEEE Geoscience and Remote Sensing Letters.*
De Clerck, E., Reyes-Muñoz, P., Prikaziuk, E., Kovács, D. D., & Verrelst, J. (2026). High-spatial-resolution gross primary production estimation from Sentinel-2 reflectance using hybrid Gaussian processes modeling. *ISPRS Journal of Photogrammetry and Remote Sensing.*
D.Kovács, D., De Clerck, E., Verrelst, J. (2025) PyEOGPR: A Python package for vegetation trait mapping with Gaussian Process Regression on Earth observation cloud platforms. *Ecological Informatics.*