

# StatEO

5-7 May 2026 | ESA-ESRIN | Frascati (Rome), Italy

## Towards integration of Eddy covariance and EO data to support CO<sub>2</sub> flux monitoring for Irish grasslands

**Wahaj Habib**  
Postdoctoral Research Fellow  
Trinity College Dublin, Ireland

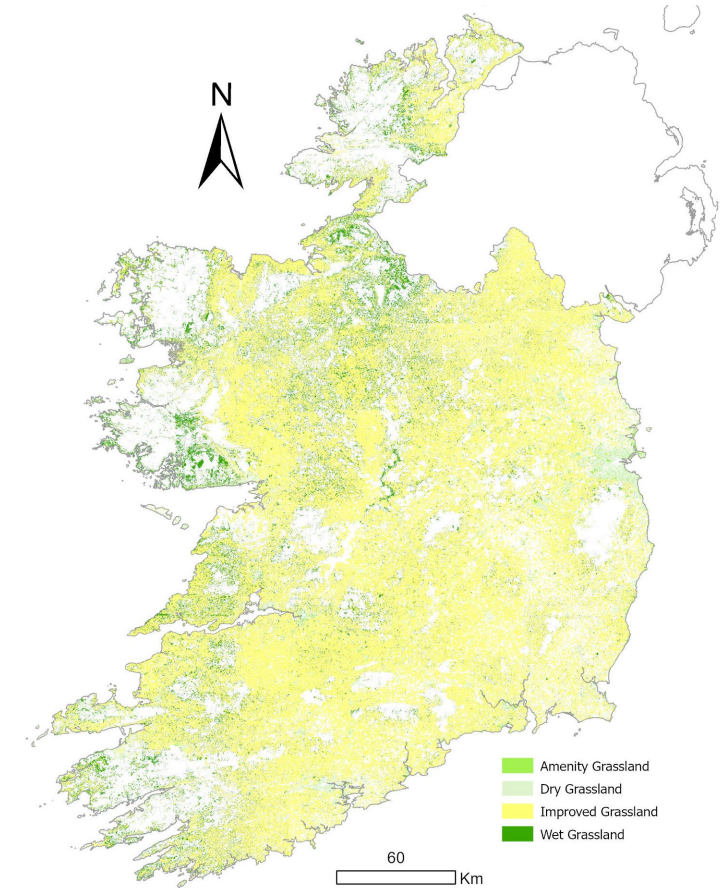


ESA UNCLASSIFIED - For ESA Official Use Only



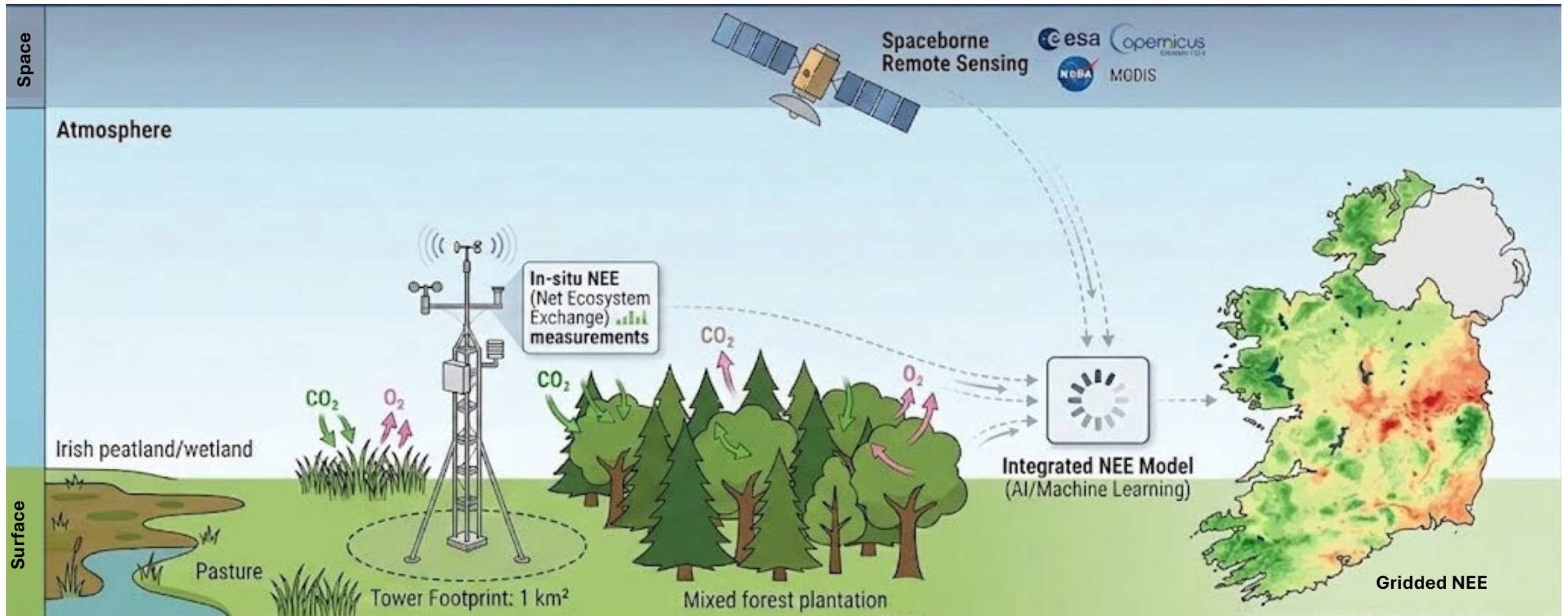
# Grasslands in Ireland - overview

- Covers ~50-60 % of the land cover of Ireland
  - Improved grassland: 77%
  - Wet grassland: 18%
  - Dry grassland: 2%
  - Amenity grassland: 3%
- Improved grassland:
  - Includes agricultural grasslands with moderate to high biomass intensity, 41.53% of the total land area of Ireland



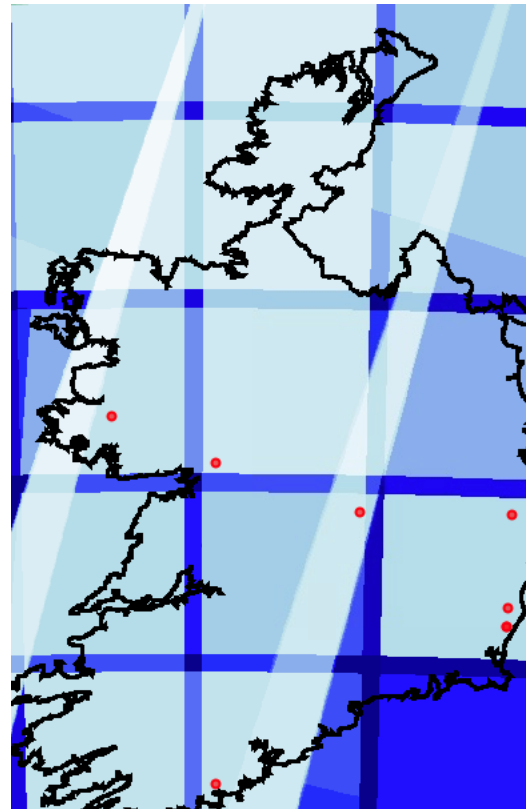
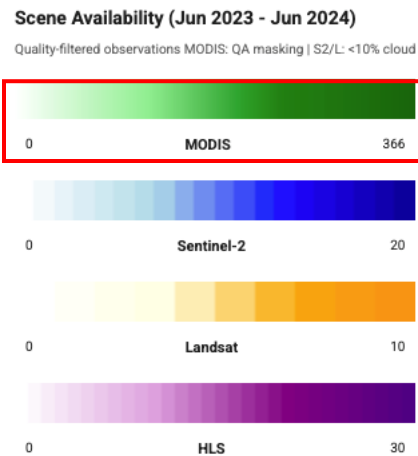
Map source: NLCM

# Project overview

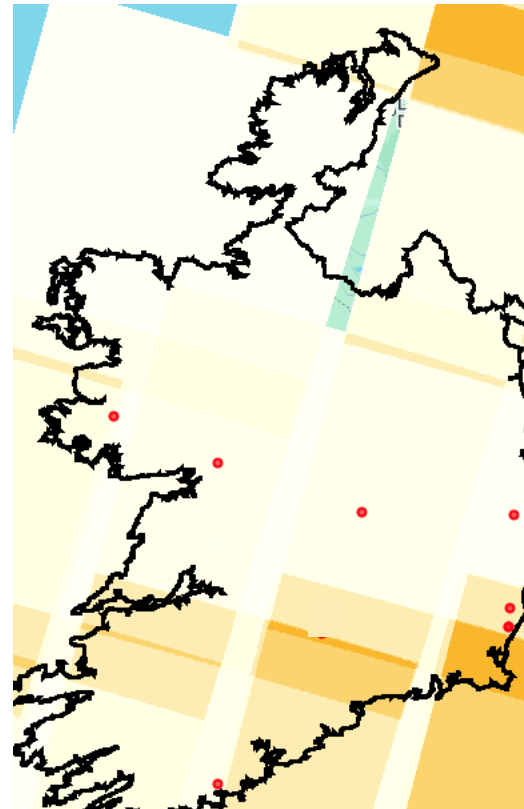


# Data availability – Optical Remote Sensing

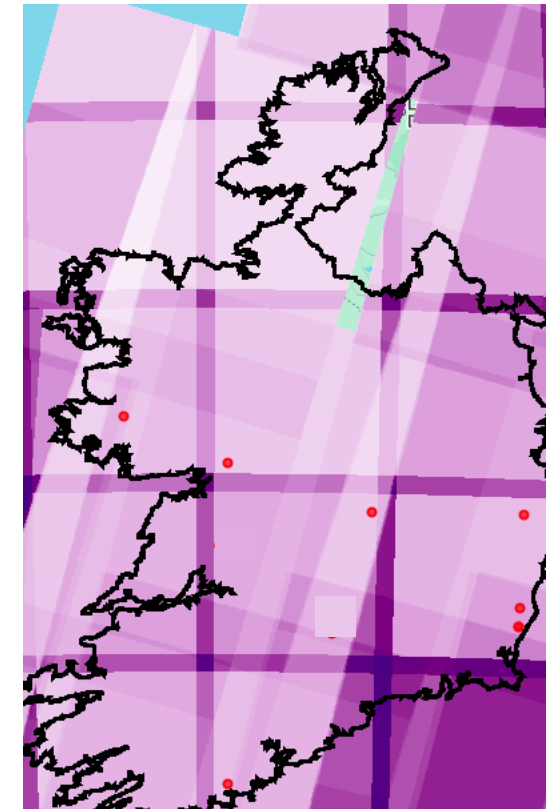
Lighter colour = close to zero cloud-free (<10% cloud) images



Sentinel-2 (10m)



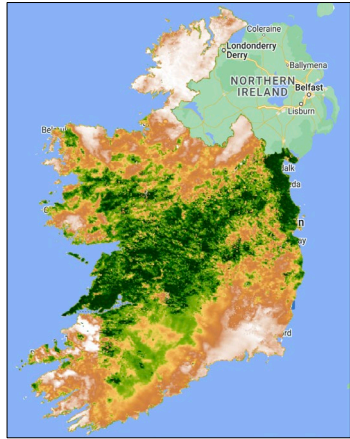
Landsat (30m)



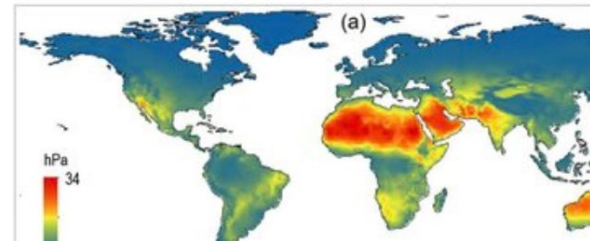
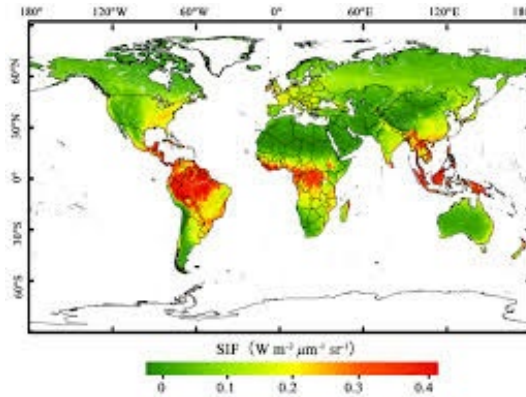
Harmonized  
Landsat/Sentinel (30m)

Bottom line: Not useful for high spatial/temporal NEE prediction

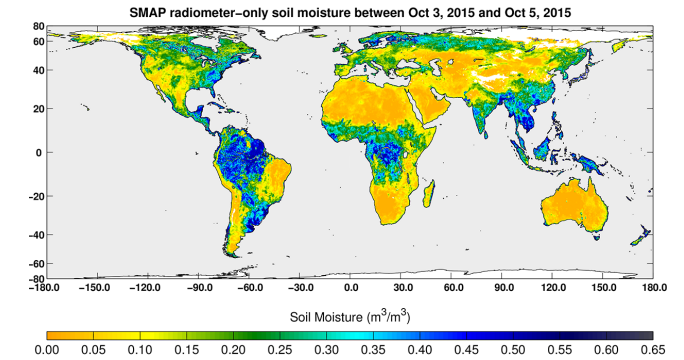
# Remote sensing – four main sources



MODIS NDVI



ERA5 e.g., VPD



variables

## MODIS Vegetation

- NDVI & EVI (16-day, 250m)
- LAI & FPAR (8-day, 500m)
- BRDF reflectance bands:
  - NIR (858nm), Green (545nm), SWIR (1640/2105nm)
- Surface albedo (BSA/WSA)
- ET/PET (evapotranspiration)

*Captures: canopy greenness, structure, phenology, water use*

## Solar-Induced Fluorescence

- GOSIF v2 (8-day, 0.05deg)
- Derived from OCO-2 + MODIS
- Faint 757nm glow from active chlorophyll

### Direct proxy for photosynthetic activity

Best single predictor for GPP (r=0.52)

**2025 gap-filled using transfer model (R<sup>2</sup>=0.89)**

*Captures: real-time photosynthesis rate*

## ERA5-Land Reanalysis

- 2m air temperature (hourly, 9km)
- Vapour pressure deficit (VPD)
- Surface shortwave radiation (R<sub>g</sub>)
- Precipitation, wind, pressure

### Key respiration driver: temperature

*Captures: climate forcing, energy balance, atmospheric demand*

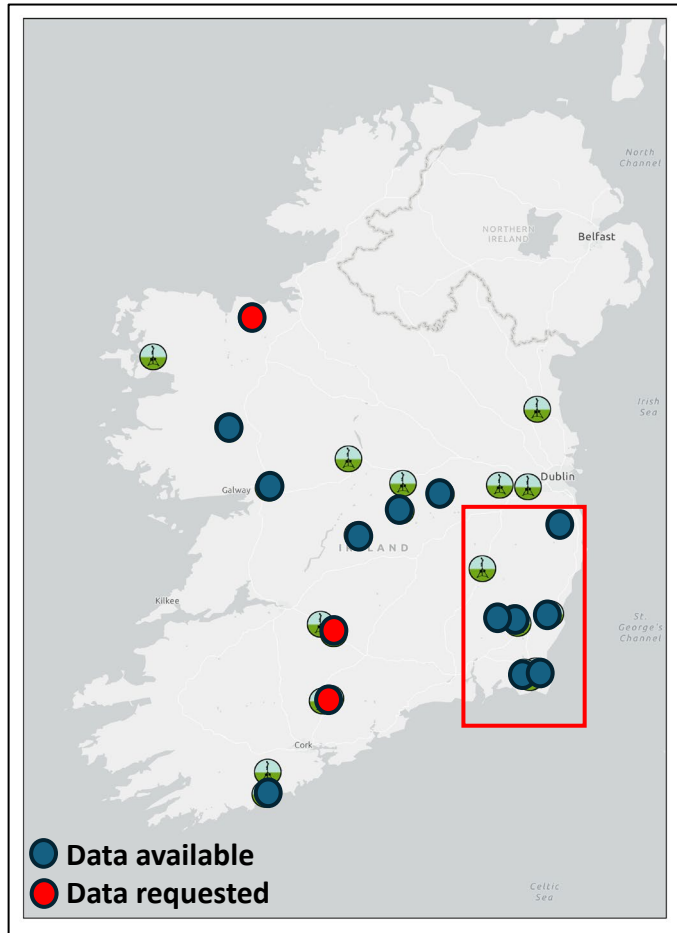
**Combined: 29 features at 8-day resolution, aggregated to monthly and seasonal for multi-scale analysis**

## Soil Moisture & PAR

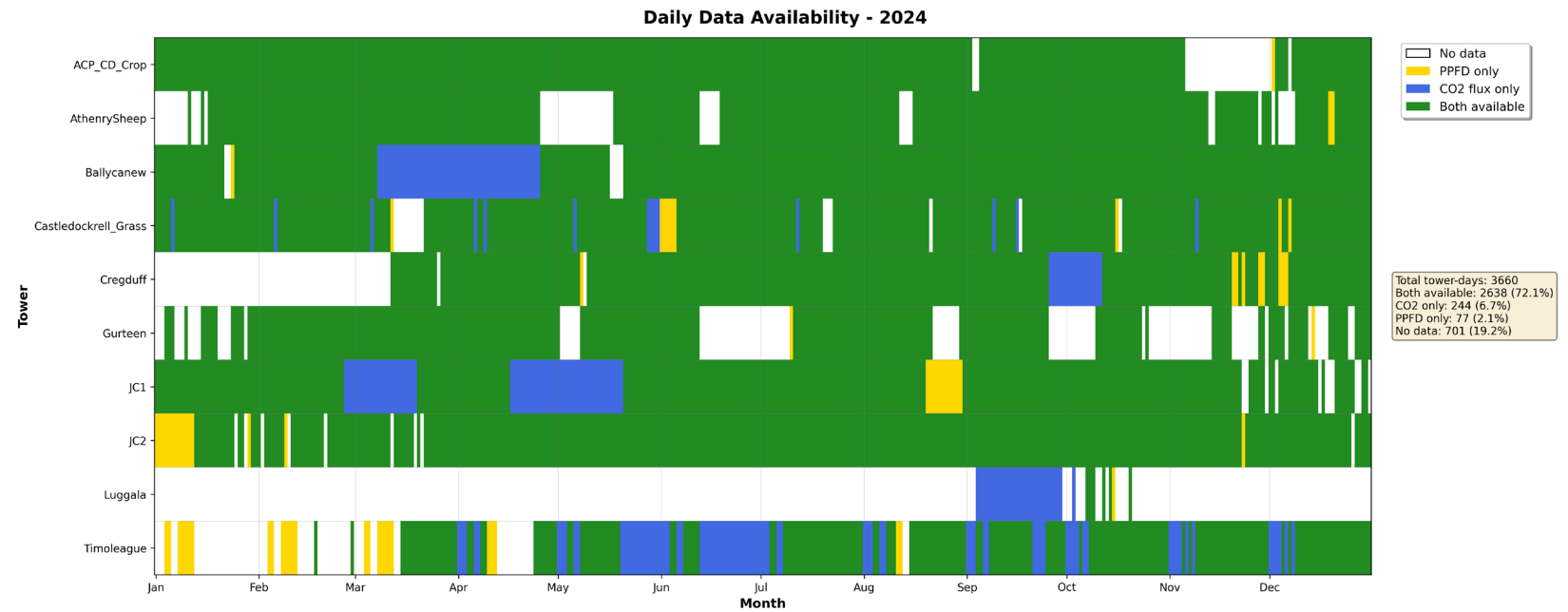
- SMAP v008 (surface + root-zone)
- L-band microwave, top 5cm
- MODIS MCD18C2 PAR/PPFD
- Photosynthetically active radiation
- NDWI moisture index from BRDF

*Captures: water availability, light energy for photosynthesis*

# NASCO Tower Network –Data availability



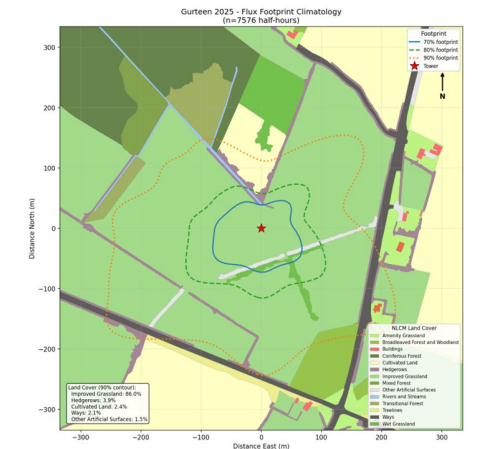
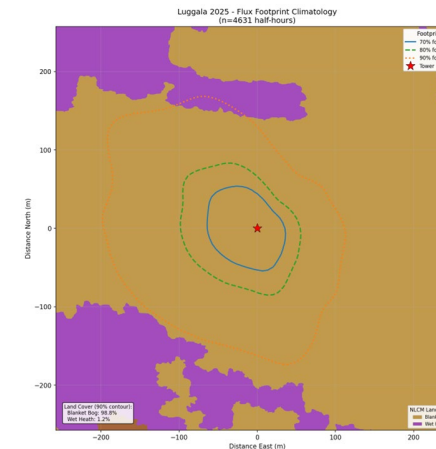
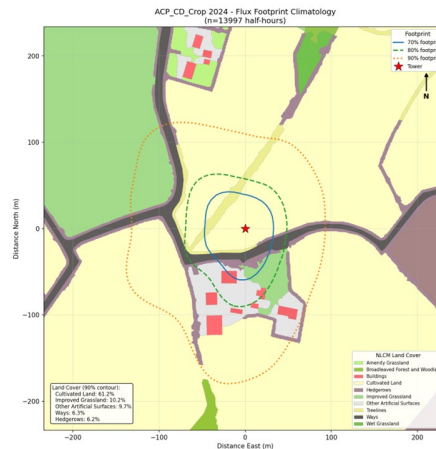
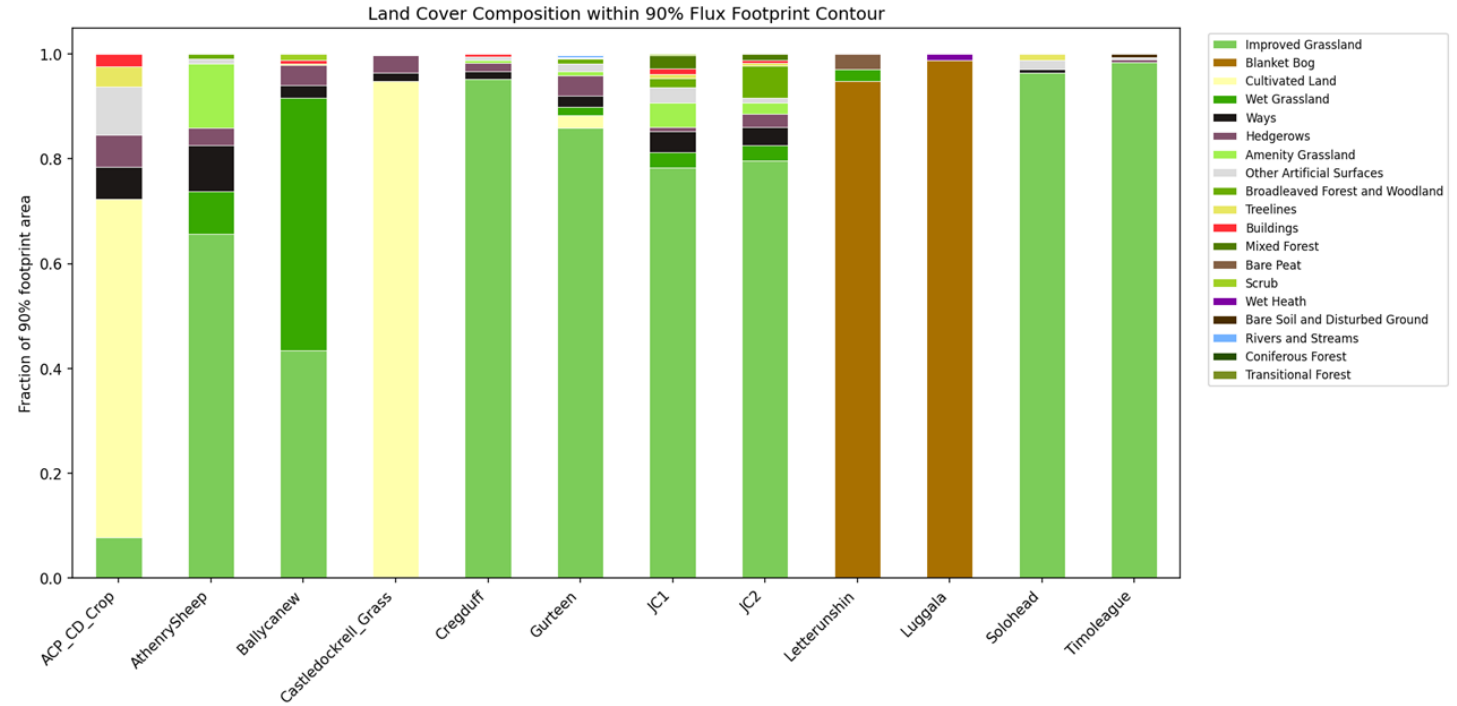
NASCO ECFT Locations (Teagasc)



LRC parametrization model struggles to predict NEE

# ECFT Data - footprint

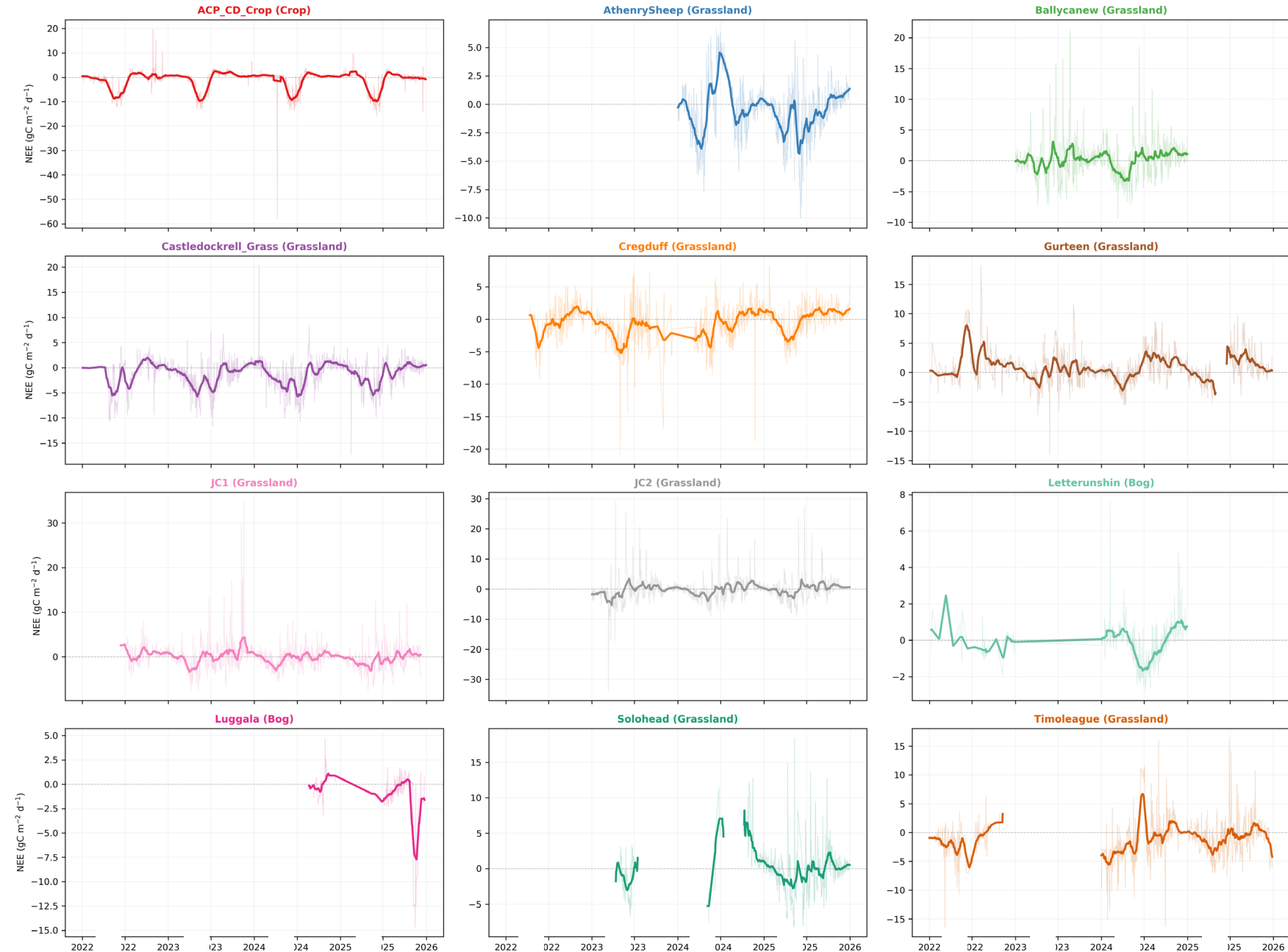
- 12 eddy covariance flux towers across Ireland
- 10 grassland sites (improved grassland -- dominant Irish land cover)
- 1 cropland (ACP\_CD\_Crop -- winter wheat/barley rotation)
- 2 bog sites (Letterunshin blanket bog, Luggala upland blanket bog)
- Continuous 30-min eddy covariance measurements
- Reichstein flux partitioning: NEE --> GPP + Reco
- Kljun FFP for footprint
- 11,922 daily observations after quality control
- Training window: 2023-2025 (3 years)
- Latitude range: 52.2N - 54.2N
- Elevation: 10 - 450 m



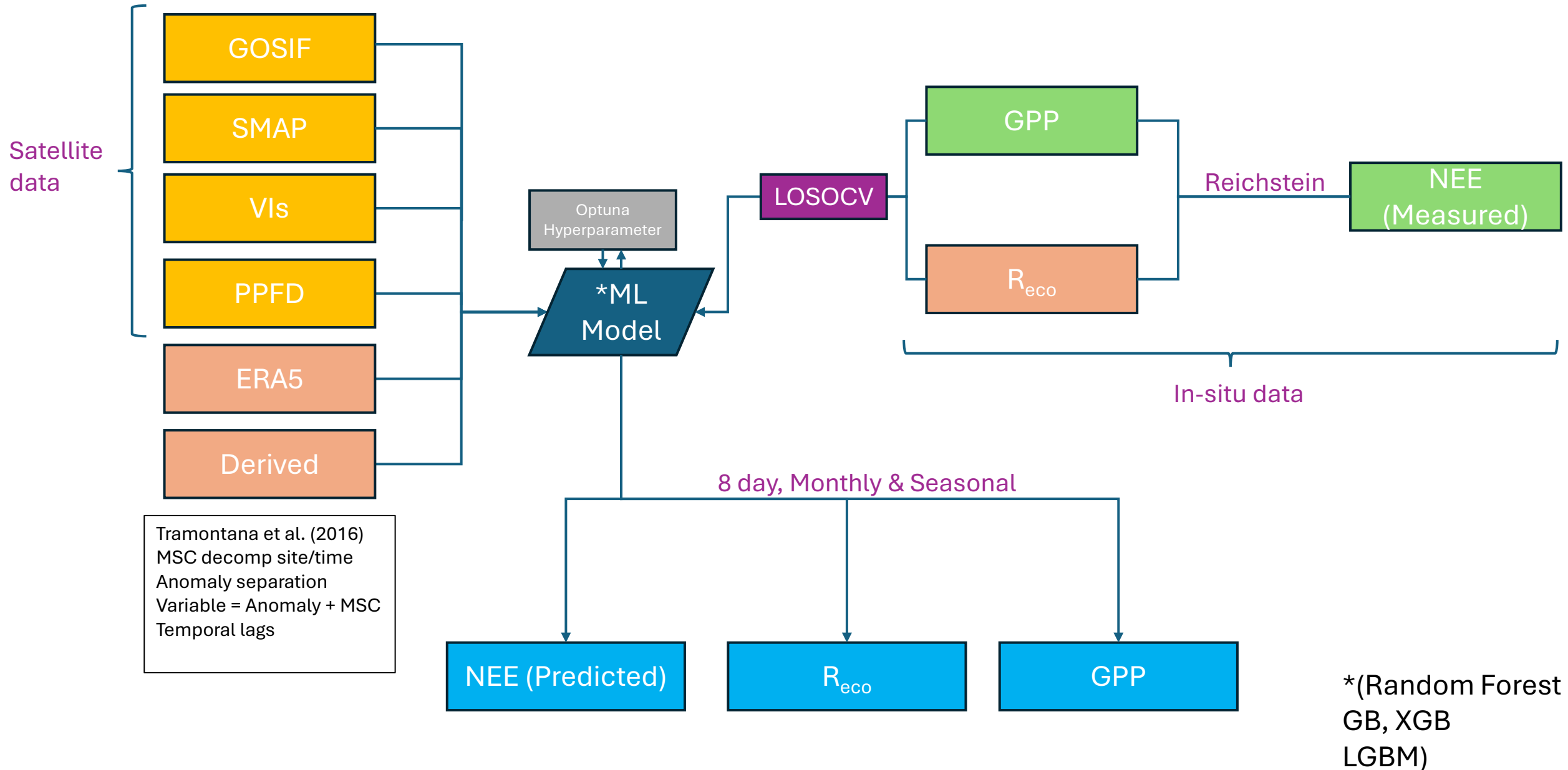
# NEE - Timeseries

- Strong interannual variability (2024 anomalous across multiple sites)
- Sharp daily spikes, management events?
- Daily noise vs 30-day mean, satellites only sample snapshots of this variance
- Several sites are data-limited (from 2024 only), constraining model training
- Cregduff persistent source; bogs near-neutral

Daily NEE Time Series (tower, gap-filled)  
Thin = daily, Bold = 30-day rolling mean



# Machine learning approach (Spatial Upscaling)

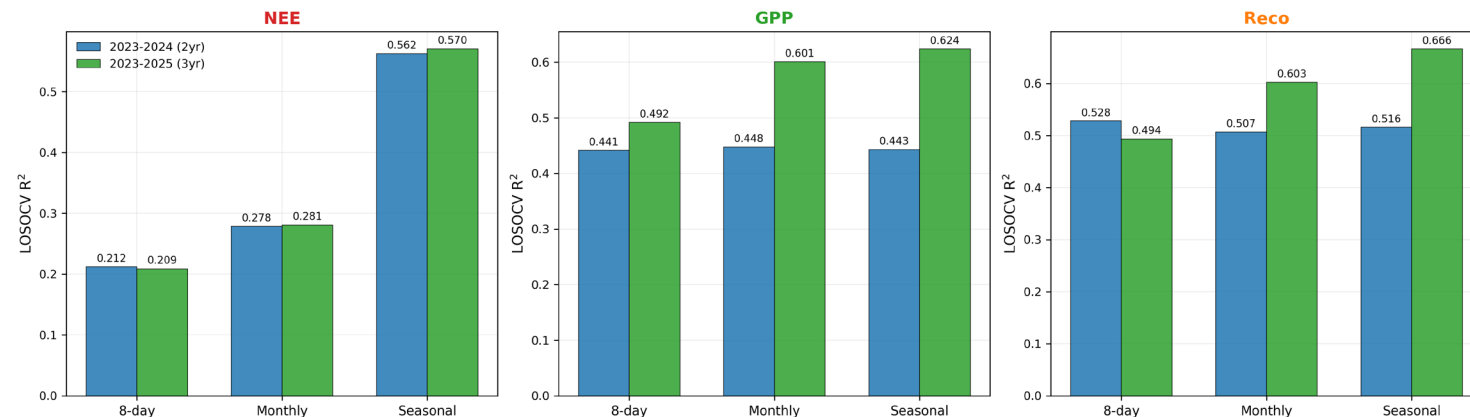


# Results Overview: LOSOCV $r^2$ Across All Targets and Resolutions

	8-day	Monthly	Seasonal	Best Model
<b>GPP</b>	0.492	0.601	<b>0.624</b>	RF / GBR
<b>Reco</b>	0.494	0.603	<b>0.666</b>	XGB / RF
<b>NEE</b>	0.209	0.281	0.570	RF / XGB

Training: 2023-2025 (3 years) | 17 RS-only features (Tramontana-style) | Optuna-tuned tree ensembles

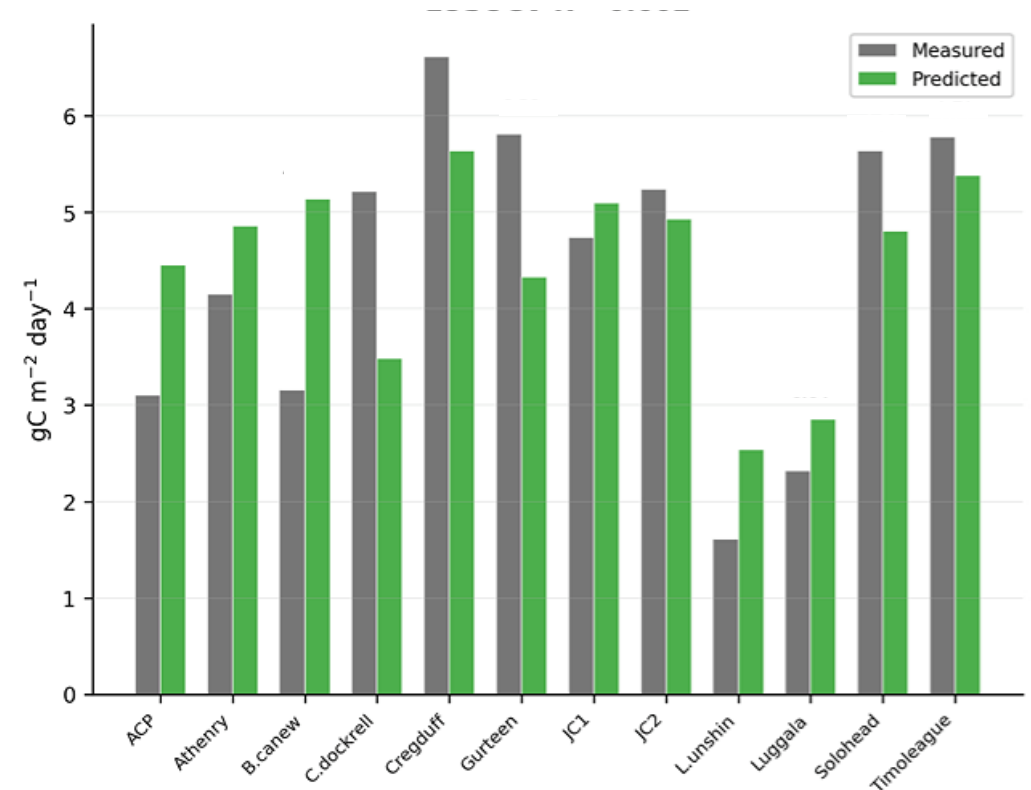
Effect of Adding 2025 Data (with Predicted SIF)



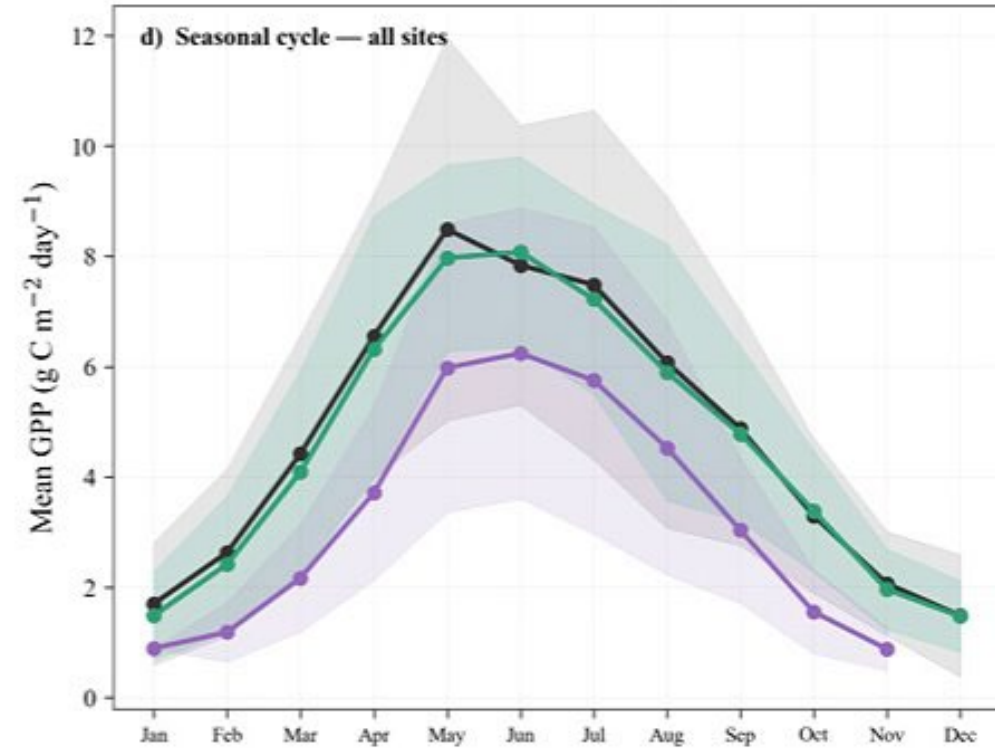
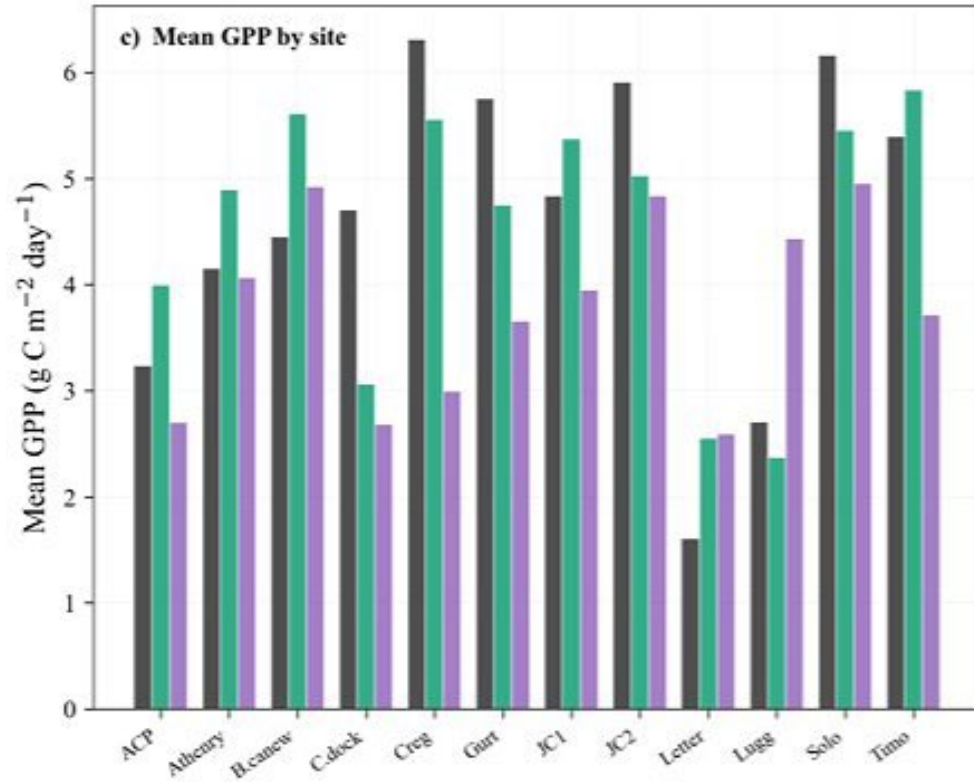
# GPP: Results

## Best LOSOCV $R^2 = 0.60$ (monthly, RF)

- 8-day: 0.49 | Monthly: 0.60 | Seasonal: 0.62
- Key drivers: SIF, EVI, temperature, APAR
- SIF is the single strongest predictor ( $r=0.52$ )
- Tramontana interactions (NDVI x Rg) critical
- **Adding 2025 data: +0.15  $R^2$  at monthly scale**



# GPP product comparison: Tower - ML - MODIS



— Tower (hesseflux) — ML model (LOSOCV) — MODIS MOD17A2H

# Ecosystem Respiration: Results

**Best LOSOCV  $R^2 = 0.67$  (seasonal, XGB)**

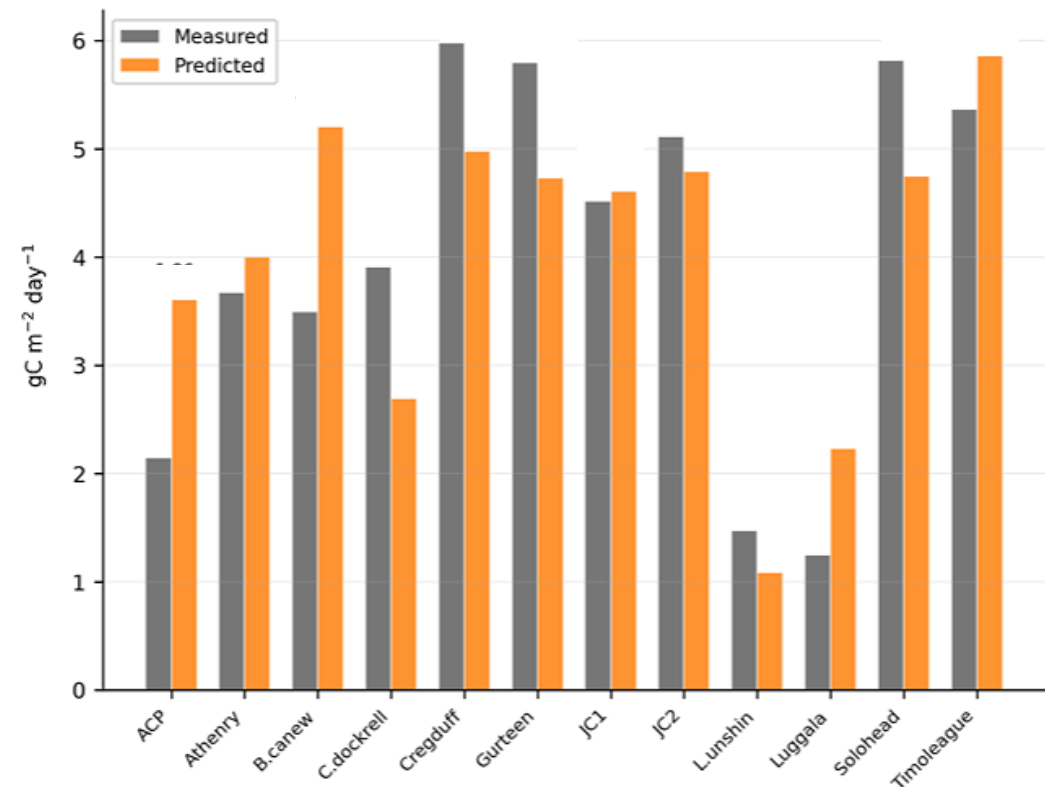
8-day: 0.49 | Monthly: 0.60 | Seasonal: 0.67

Strongest single-flux prediction in the project

Dominant drivers: ERA5 temperature, LST, soil moisture

Exponential temperature response ( $Q_{10} \sim 1.5-3.0$ )

**Adding 2025: +0.15  $R^2$  at seasonal scale**



# NEE: Results

**NEE = Reco - GPP: a small residual of two large co-varying fluxes**

8-day  $R^2 = 0.22$  | Monthly = 0.28 | Seasonal = 0.57

## **Three fundamental barriers:**

1. Signal-to-noise: measurement uncertainty 4x larger than the NEE signal
2. Cross-site sign reversals: 13/17 features flip sign between sites
3. Satellite singles struggle to capture management

## **Improvement strategies tested**

Direct, Indirect (Reco-GPP), Stacked, Residual correction, Stratified, NEE-tuned features

NEE-tuned features best at 8-day (+0.014), but gains marginal overall

Consistent with literature: RS-only NEE prediction hits a wall at  $R^2 \sim 0.2-0.3$

# NEE: Improvement strategies tested

## Strategies:

1. Direct: standard RS features --> NEE
2. Indirect: predict GPP + Reco separately, compute  $NEE = Reco - GPP$
3. Stacked: add predicted GPP/Reco as features for NEE model
4. Residual correction: indirect NEE + model on residuals
5. Ecosystem-stratified: separate grassland vs bog/crop models
6. NEE-tuned features: expanded set with anomalies + respiration drivers

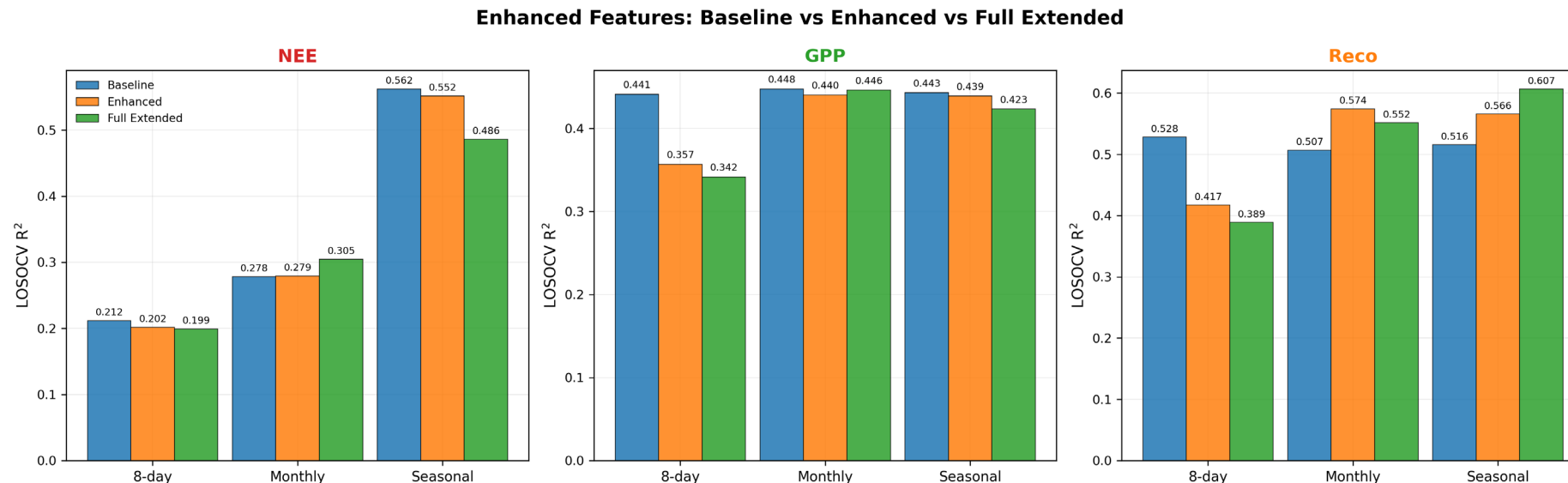
## Results:

8-day: NEE-tuned features best ( $R^2=0.22$ , +0.014)

Monthly/Seasonal: direct baseline best -- no strategy meaningfully helped

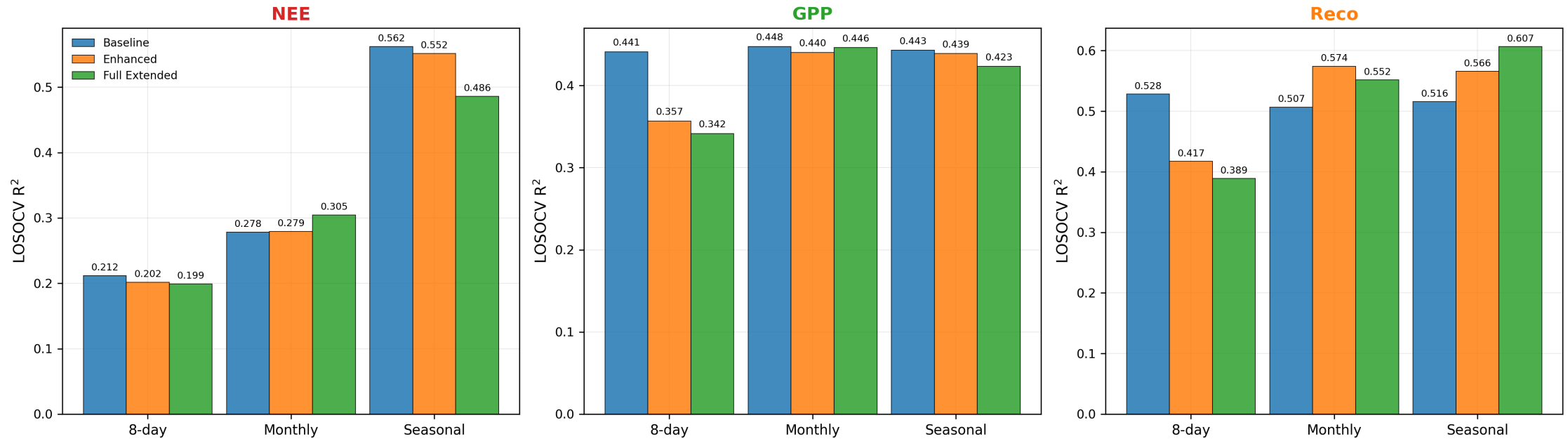
Indirect (Reco-GPP) WORSE: error propagation compounds

**Conclusion: NEE  $R^2 \sim 0.2-0.3$  appears to be an RS-only ceiling**



# Temporal Resolution Trade-offs

Enhanced Features: Baseline vs Enhanced vs Full Extended



# Conclusions and future directions

## What we showed:

- Satellite RS can predict GPP ( $R^2=0.60$ ) and Reco ( $R^2=0.67$ ) at monthly/seasonal scales
- NEE prediction limited to  $R^2\sim 0.2-0.3$  sub-seasonally -- fundamental RS constraint.
- 3 years > 2 years: more data directly improves generalisation
- SIF is gap-fillable at high accuracy, enabling temporal extension

## Next steps:

- Sentinel-1/2 integration (5-day revisit, 10m) for management event detection
- National upscaling: apply GPP/Reco models to wall-to-wall satellite grids
- Deep learning approaches (temporal CNNs, transformers) for sequence modelling

## Project by the numbers:

- 12 eddy covariance flux towers
- 29 satellite remote sensing features
- 3 years of data (2023-2025)
- 3 temporal resolutions (8-day, monthly, seasonal)
- 4 ML algorithms, Optuna-tuned
- 17 features in the best model (Tramontana-style)
- 152 analysis scripts

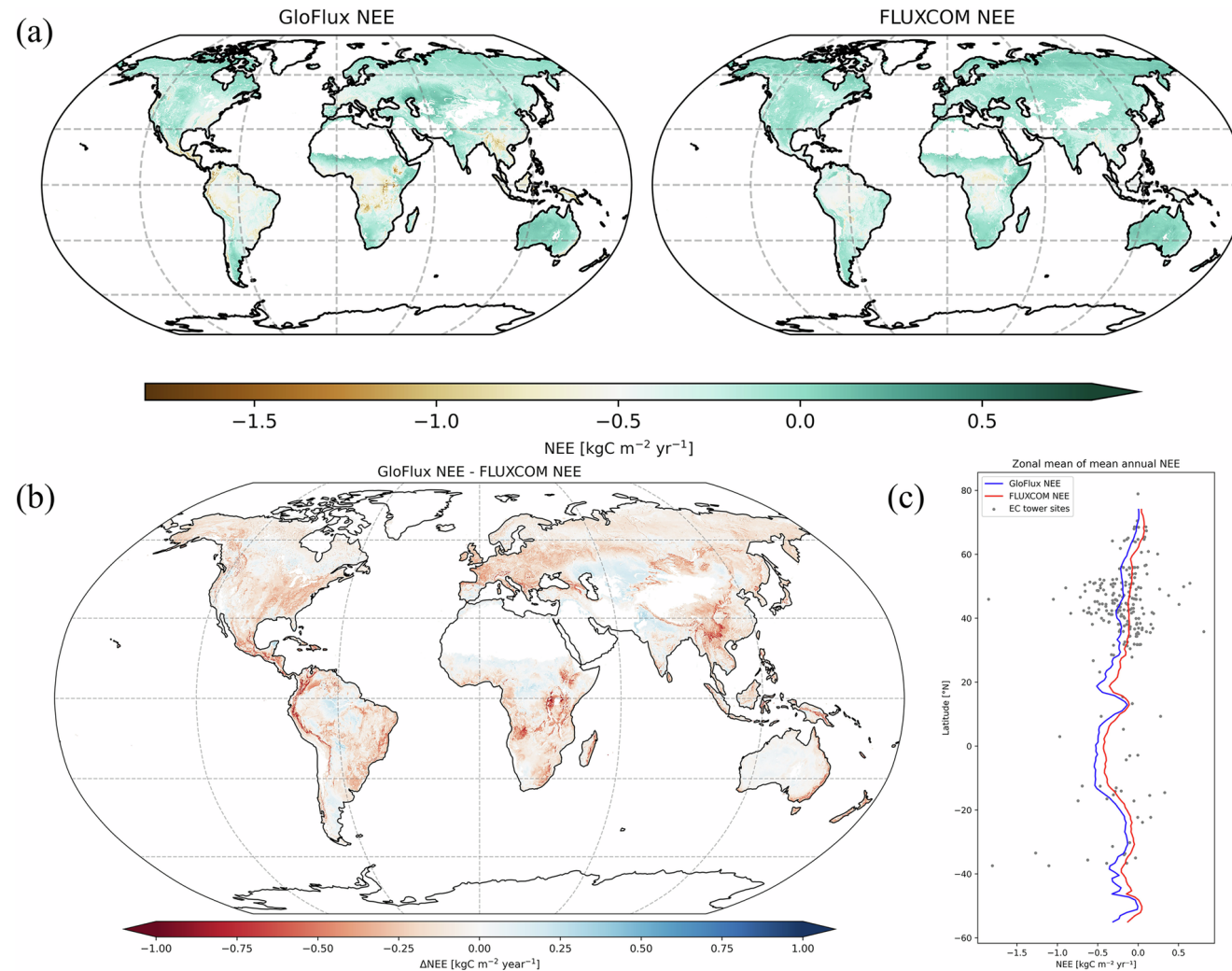
Best results (LOSOCV  $R^2$ ):

**GPP: 0.60 (monthly RF)**

**Reco: 0.67 (seasonal XGB)**

**NEE: 0.57 (seasonal XGB)**

# What's next? **ÉireFlux**: Ireland Carbon Flux Gridded Dataset



Thank you