# Estimating Gross Primary Productivity in Crops with Satellite Data, Radiative Transfer Modeling and Machine Learning

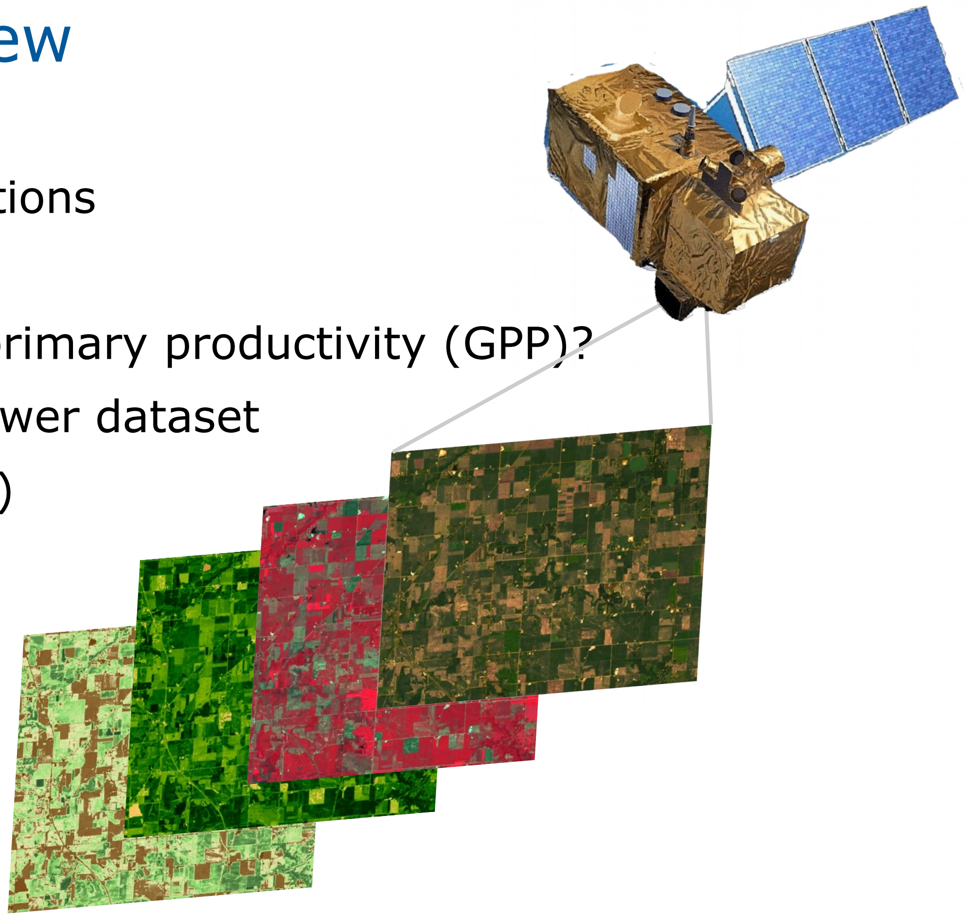
Aleksandra Wolanin[a,*], Luis Guanter[a], Gustau Camps-Valls[b], Luis Gómez-Chova[b], Gonzalo Mateo García[b], Christiaan van der Tol[c], Yongguang Zhang[d]

[a] GFZ German Research Centre for Geosciences, Potsdam, Germany
[b] Universitat de Valncia, Valncia, Spain
[c] University of Twente, Enschede, The Netherlands
[d] Nanjing University, Nanjing, China

*ola@gfz-potsdam.de

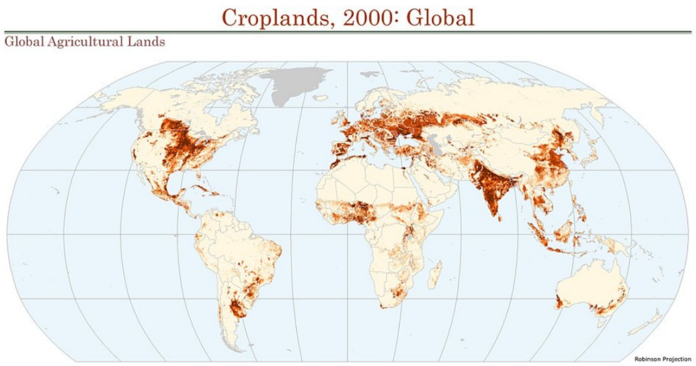
# Overview

1) Introduction: croplands and satellite observations
2) SCOPE model and satellite data...
3) … and how to apply them to estimate gross primary productivity (GPP)?
4) Feasibility test: a comparison with the flux tower dataset
5) Implementation in Google Earth Engine (GEE)
6) Conclusions

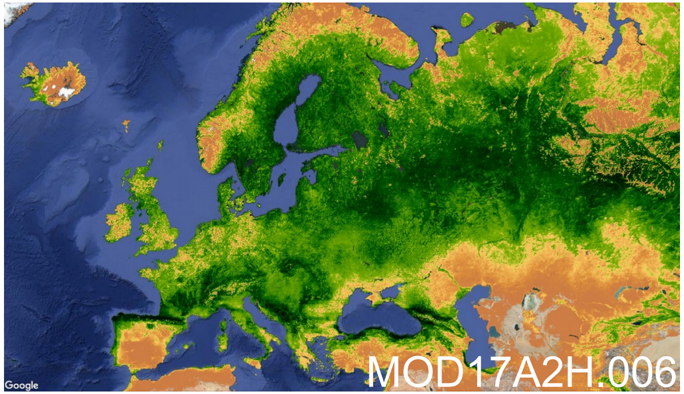


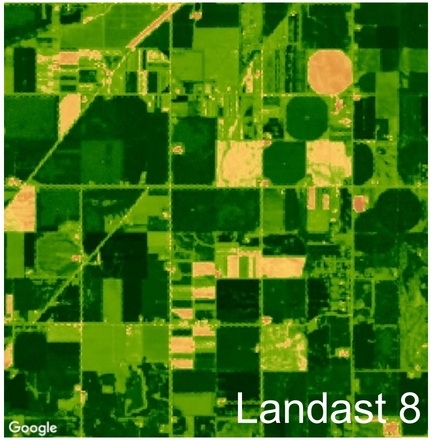
# Introduction

- Croplands cover ~ 12% of the Earth's ice-free land surface
- Satellite data - global and continues characterization
- Gross primary productivity (GPP) - amount of carbon fixed by plants through photosynthesis




Ramankutty et al., 2000



# Modeling GPP

- Process-based model SCOPE (Van der Tol et al. 2009)
- Simulates photosynthesis, radiative transfer in the leaf and canopy, and surface energy balance
- Based on PROSPECT-D (leaf, Féret et al., 2017) and 4SAIL (canopy, Verhoef et al., 2007)
- Output: reflectance, GPP, etc.
- leaf maximum carboxylation capacity Vcmax = f(Cab) (Houborg et al. 2013)

GFZ Helmholtz-Zentrum POTSDAM

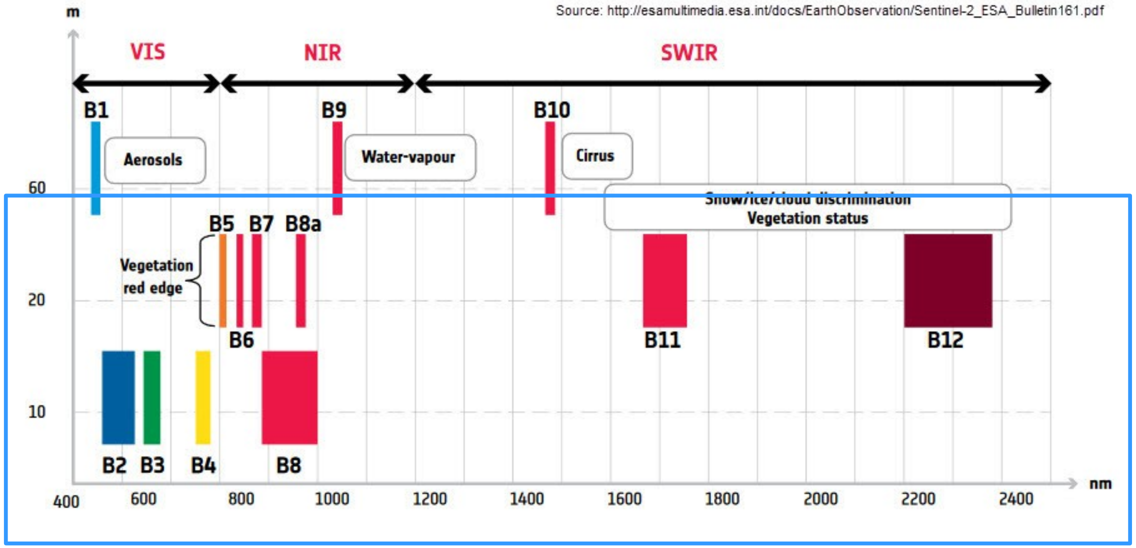
| Leaf & canopy & soil | | Unit | min | max |
|---|---|---|---|---|
| Cab | Chlorophyll AB content | ug cm-2 | 11 | 90 |
| Cca | Carotenoid content. Usually 25% of Cab | ug cm-2 | 0 | 40 |
| Cdm | Dry matter content | g cm-2 | 0 | 0.1 |
| Cw | leaf water equivalent layer | cm | 0 | 0.05 |
| Cs | scenecent material fraction | fraction | 0 | 0.9 |
| Cant | Anthocyanins | ug cm-2 | 0 | 40 |
| N | leaf thickness parameters | [] | 1 | 2.5 |
| LIDFa | leaf inclination | | -1 | 1 |
| LIDFb | variation in leaf inclination | | -1 | 1 |
| LAI | Leaf area index | m2 m-2 | 0 | 9 |
| hc | vegetation height | m | 0.1 | 2 |
| SMC | volumetric soil moisture content in the root zone | | 0.01 | 0.7 |
| BSMBrightness | BSM model parameter for soil brightness | | 0.01 | 0.9 |
| BSMlat | BSM model parameter 'lat' | | 20 | 40 |
| BSMlon | BSM model parameter 'long' | | 45 | 65 |

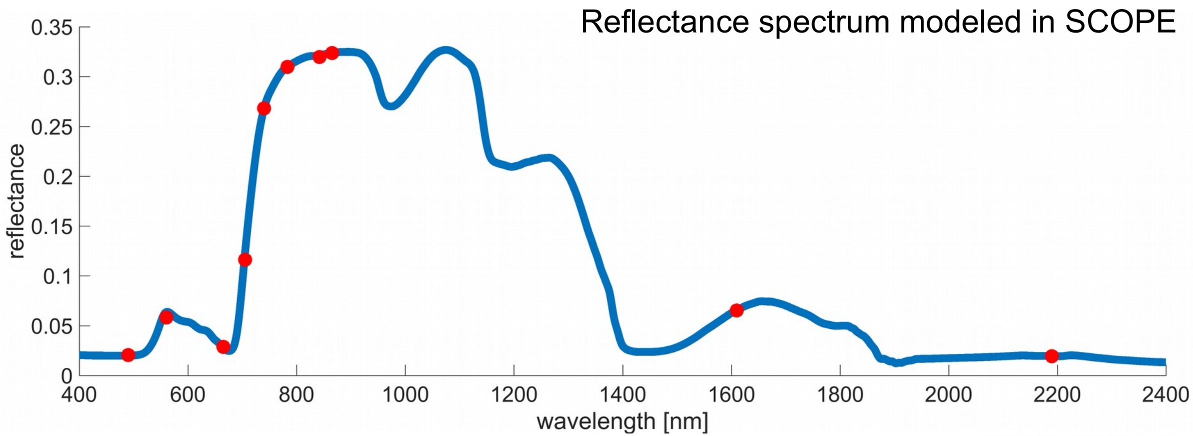
meteo

| | | | | |
|---|---|---|---|---|
| Rin | broadband incoming shortwave radiation (0.4-2.5 um) | W m-2 | 0 | 1400 |
| Ta | air temperature | T | -10 | 50 |
| Rli | broadband incoming longwave radiation (2.5-50 um) | W m-2 | 0 | 400 |
| p | air pressure | hPa | 500 | 1030 |
| ea | atmospheric vapour pressure | hPa | 0 | 125 |
| u | wind speed | m s-1 | 0 | 25 |

# Sentinel-2

- Good spatial resolution (10 & 20 m)
- Wide spectral coverage 400-2500nm (10 bands)
- Sentinel-2 A & B (launched 06.2015 & 03.2017)



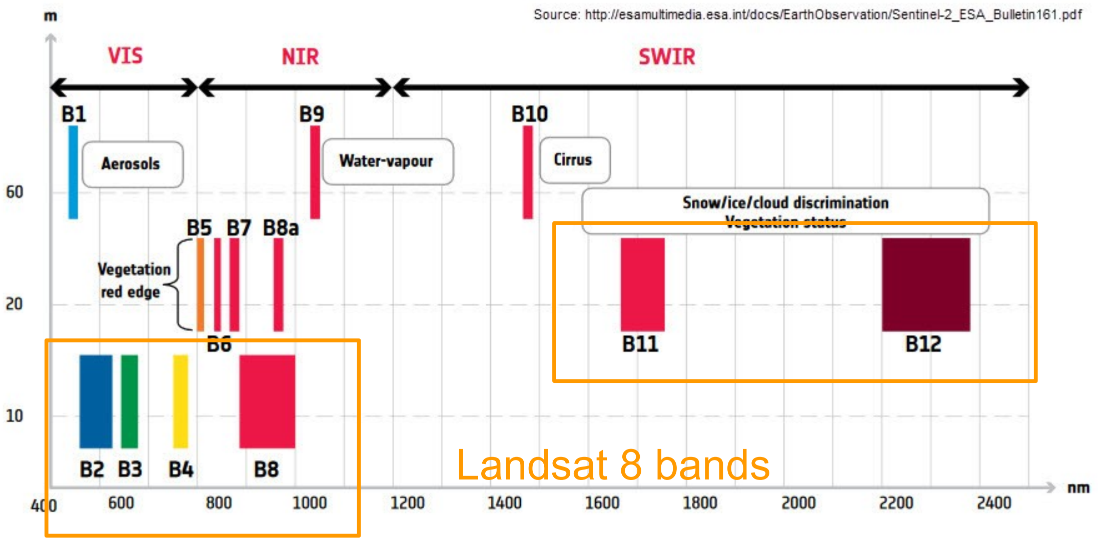



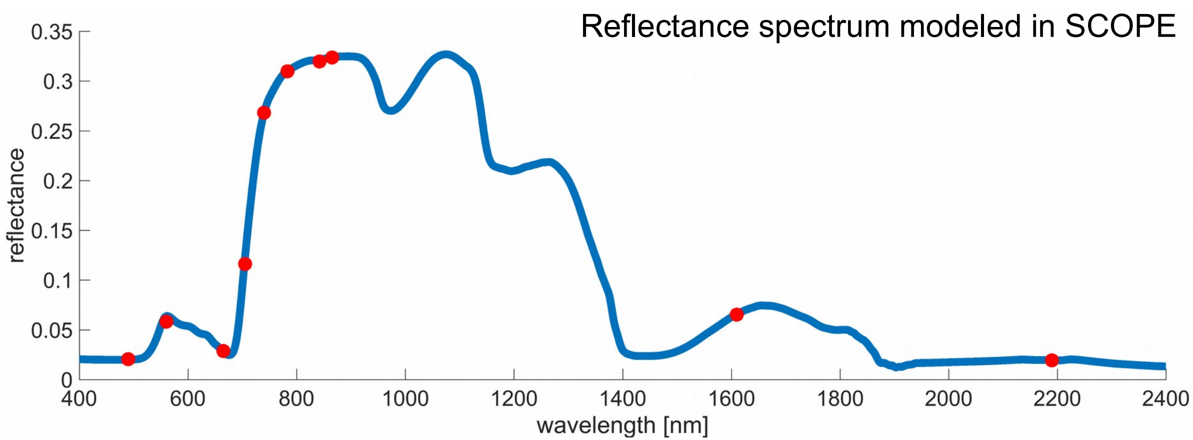
# Sentinel-2

- Good spatial resolution (10 & 20 m)
- Wide spectral coverage 400-2500nm (10 bands)
- Sentinel-2 A & B (launched 06.2015 & 03.2017)

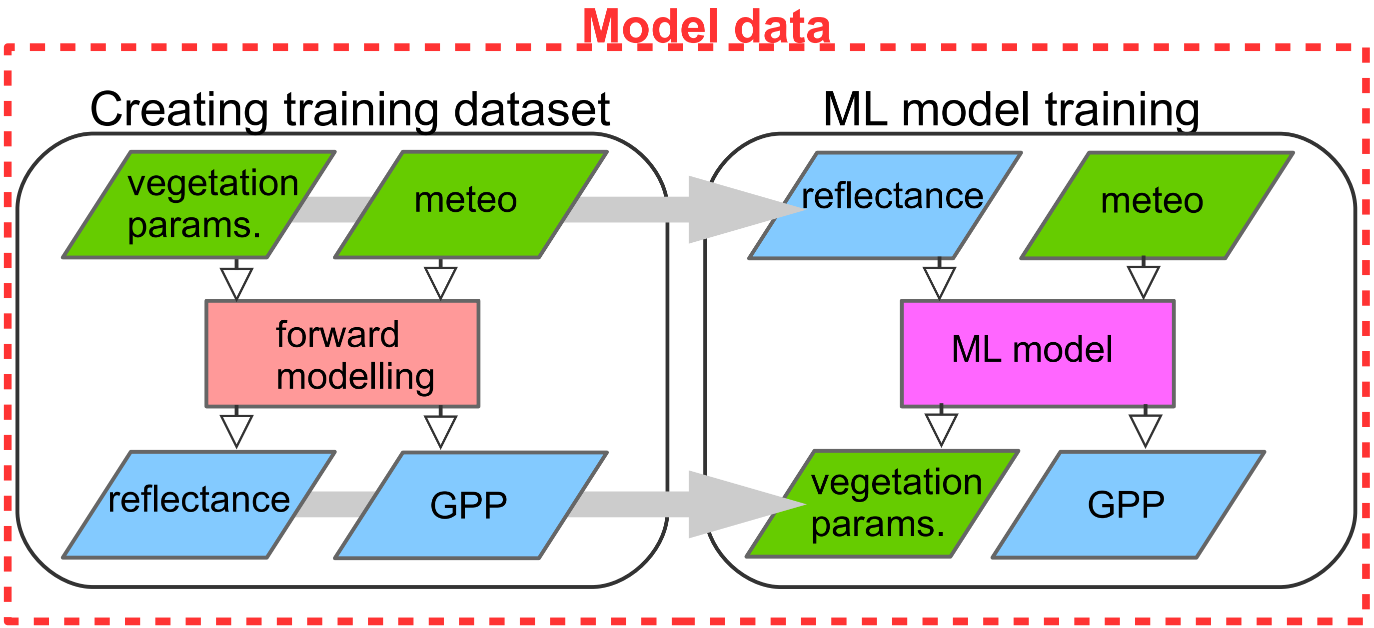
# Landsat 8

- Spatial resolution 30 m
- 6 bands common with Sentinel-2
- Launched 02.2013

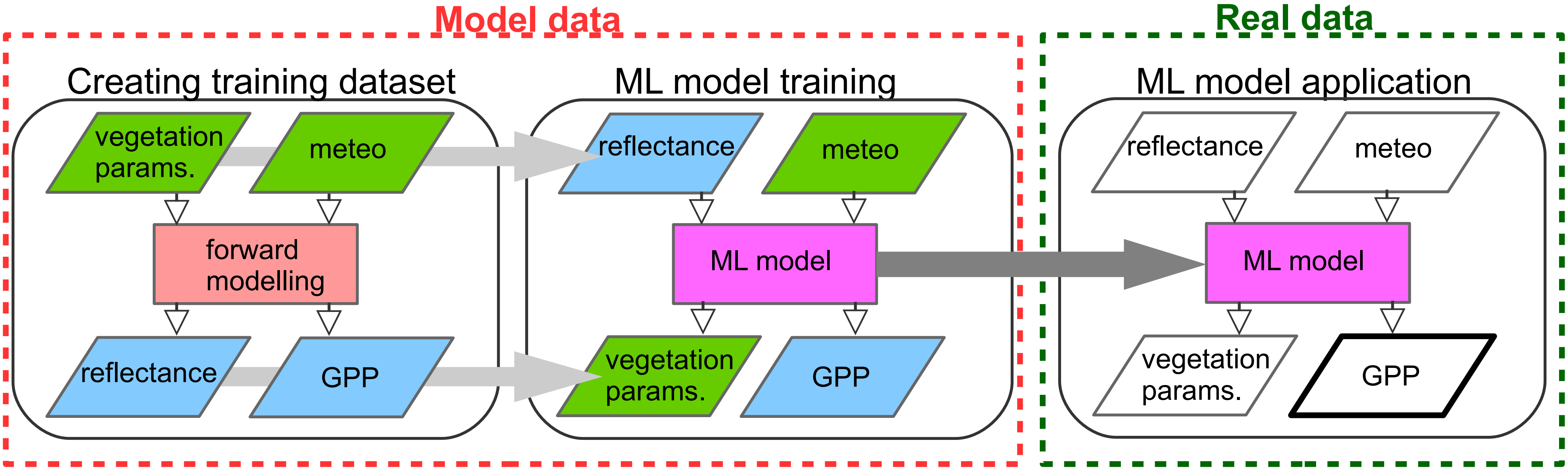


# How to apply SCOPE to satellite data?

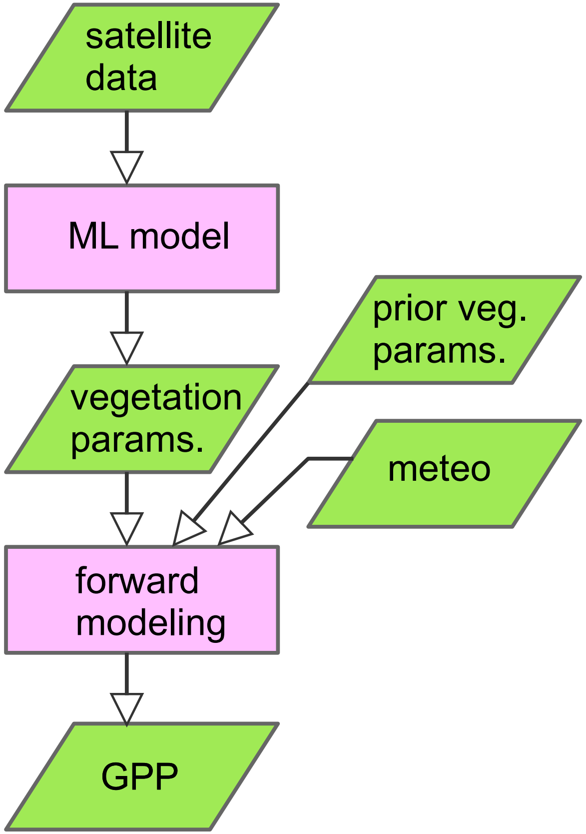


# How to apply SCOPE to satellite data?



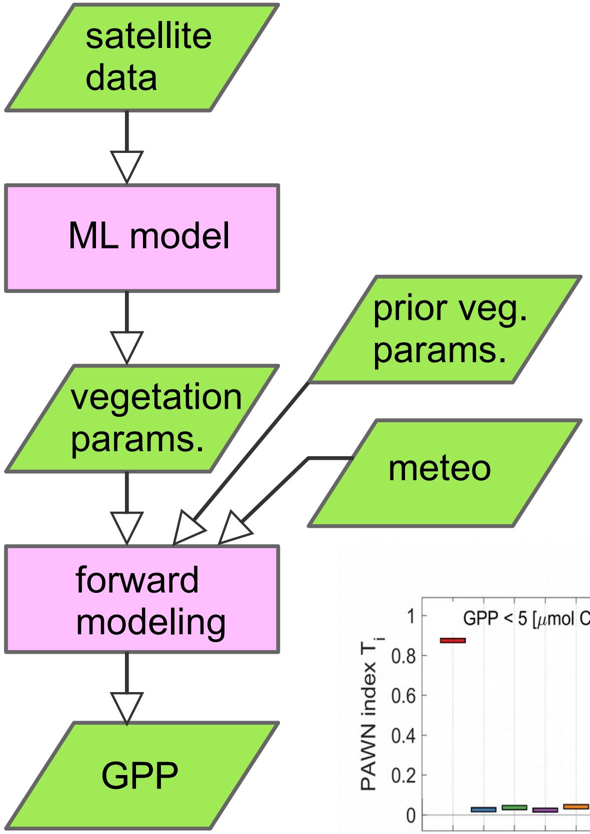
# How to apply SCOPE to satellite data?

1°



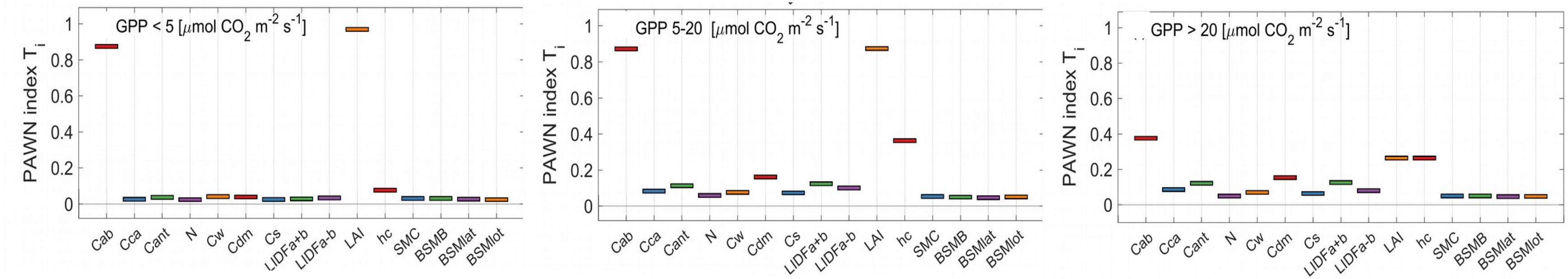
# How to apply SCOPE to satellite data?

1°

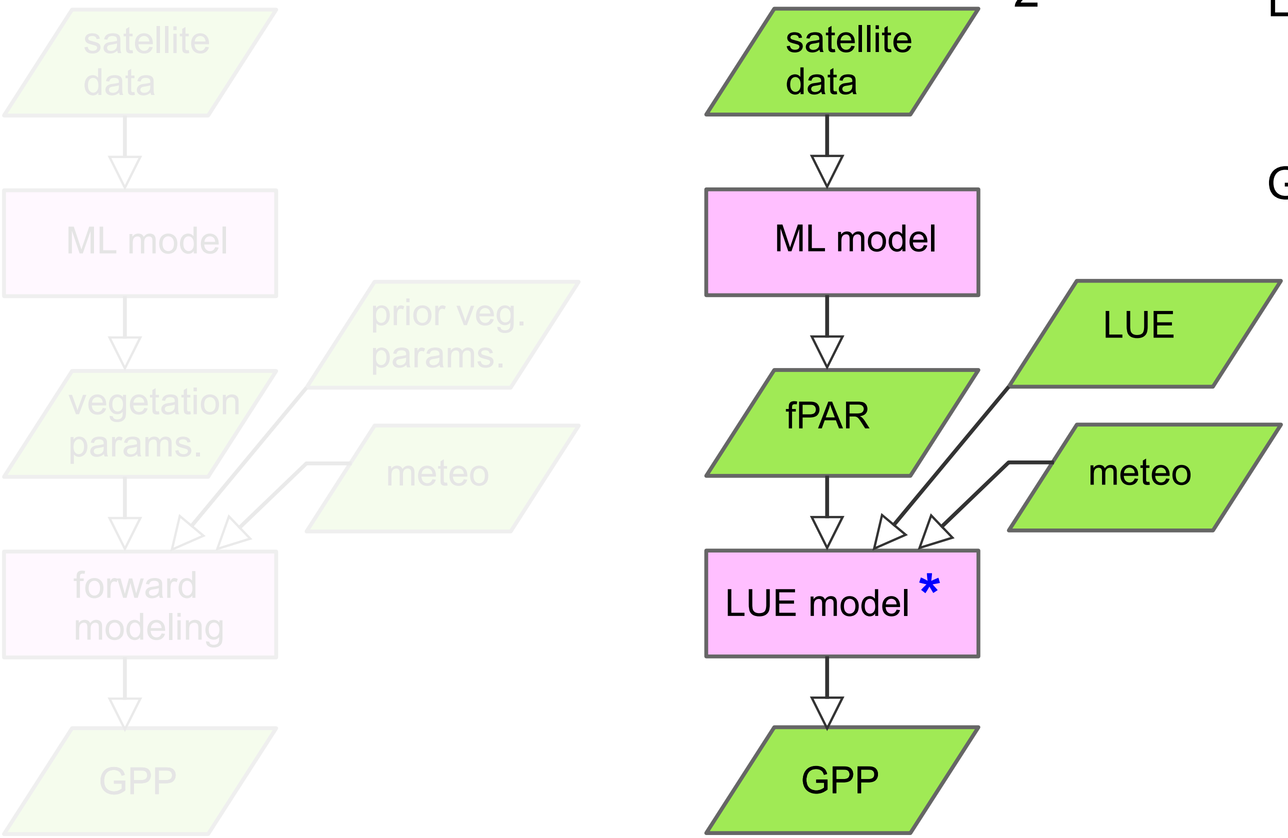



- Sensitivity analysis of parameters driving GPP (with PAWN method, Pianosi and Wagener, 2015)
- Evaluate the relative importance of input variables and identify the most influential variables affecting model outputs

For different sub-ranges of GPP



# How to apply SCOPE to satellite data?




2°

- LUE model*

$$GPP = \overbrace{fPAR * PAR}^{APAR} * LUE * f(meteo)$$

fraction absorbed by vegetation

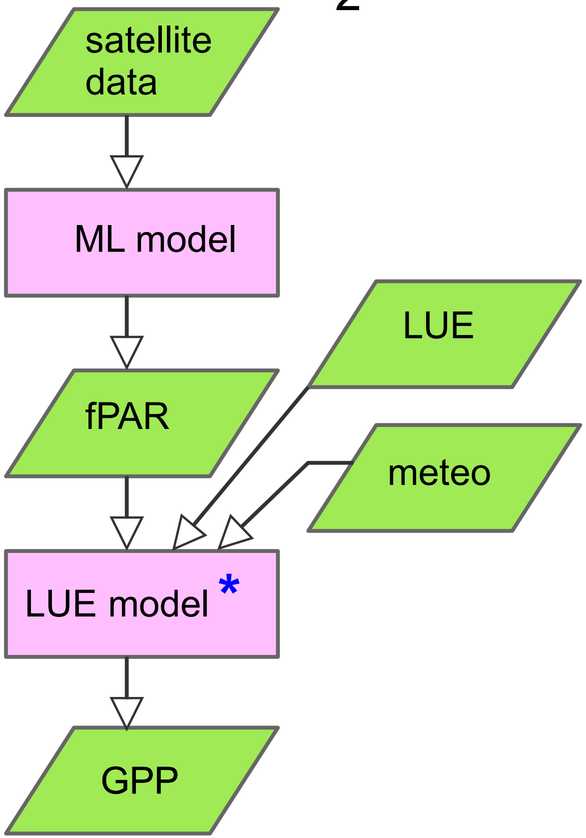
photosynthetically active radiation

light use efficiency

stress due to meteorological conditions

# How to apply SCOPE to satellite data?

2°




- LUE model*: $GPP = \overbrace{fPAR * PAR}^{APAR} * LUE * f(meteo)$

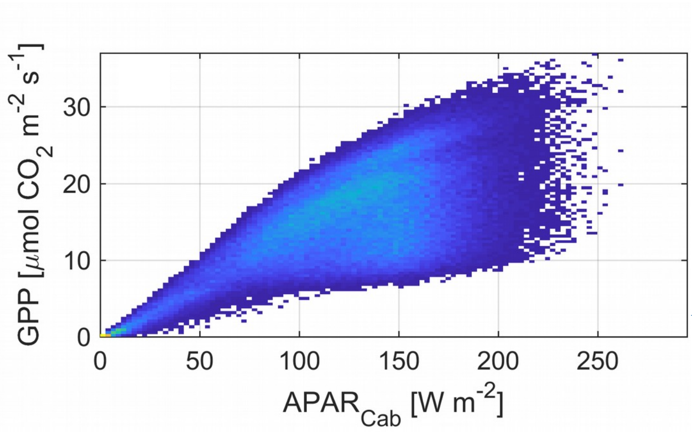
fraction absorbed by vegetation

photosynthetically active radiation

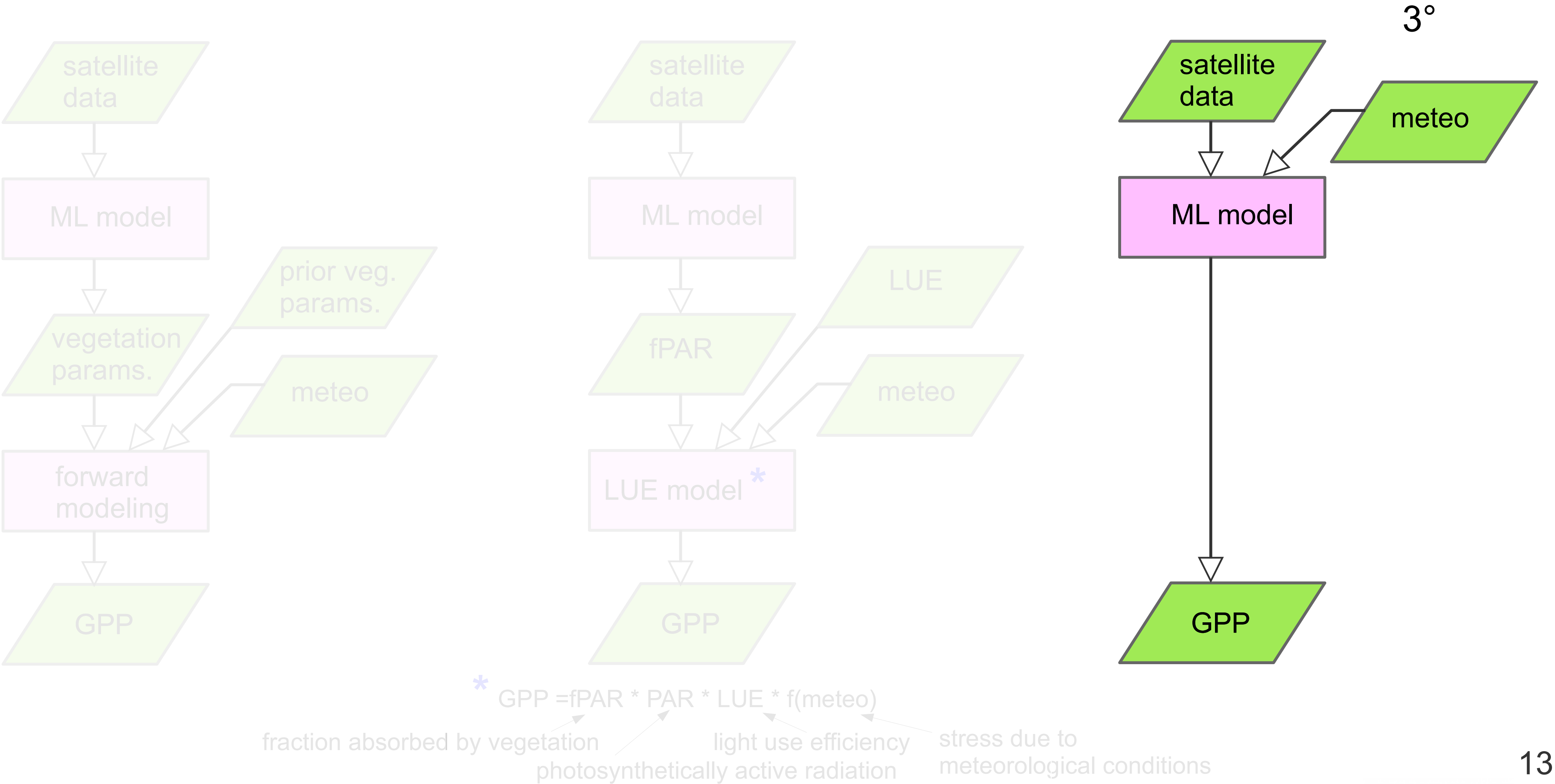
light use efficiency

stress due to meteorological conditions

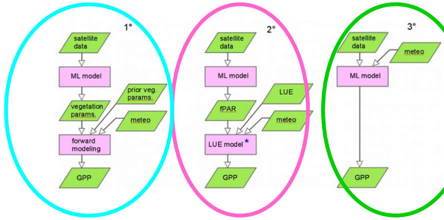
- LUE (light use efficiency) not constant
- All variability is due to changes in leaf and canopy properties (meteo constant)



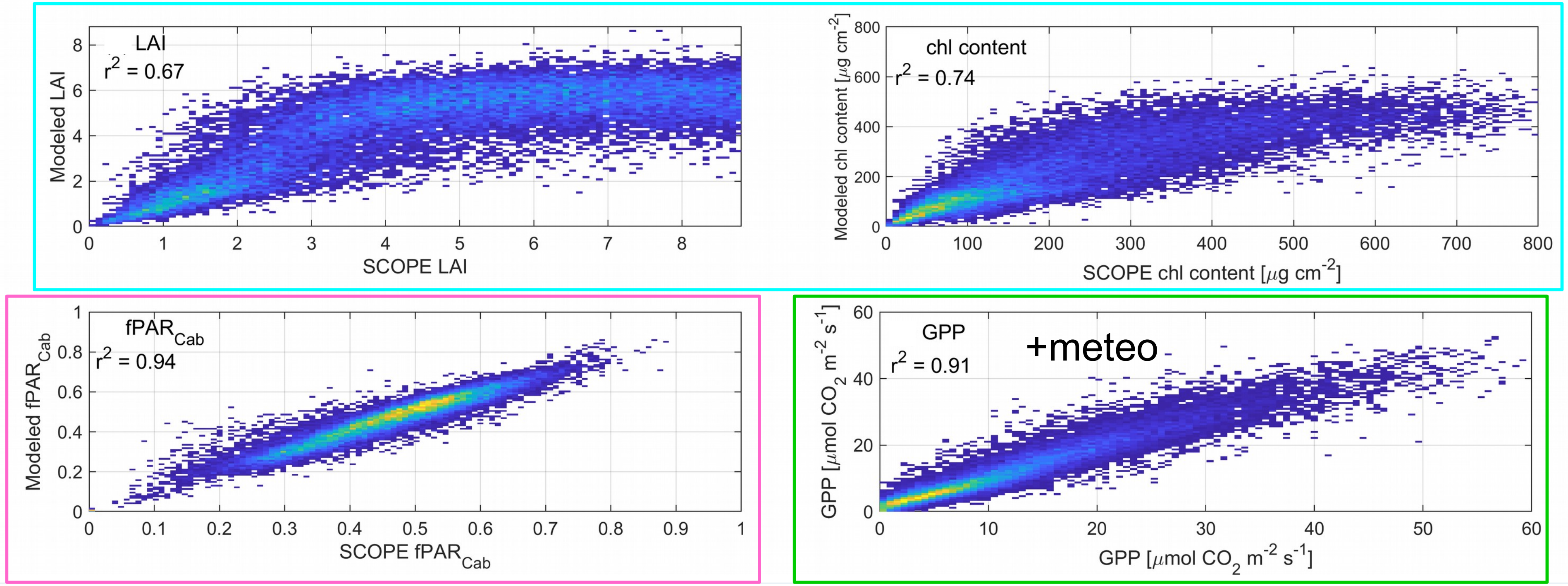
# How to apply SCOPE to satellite data?

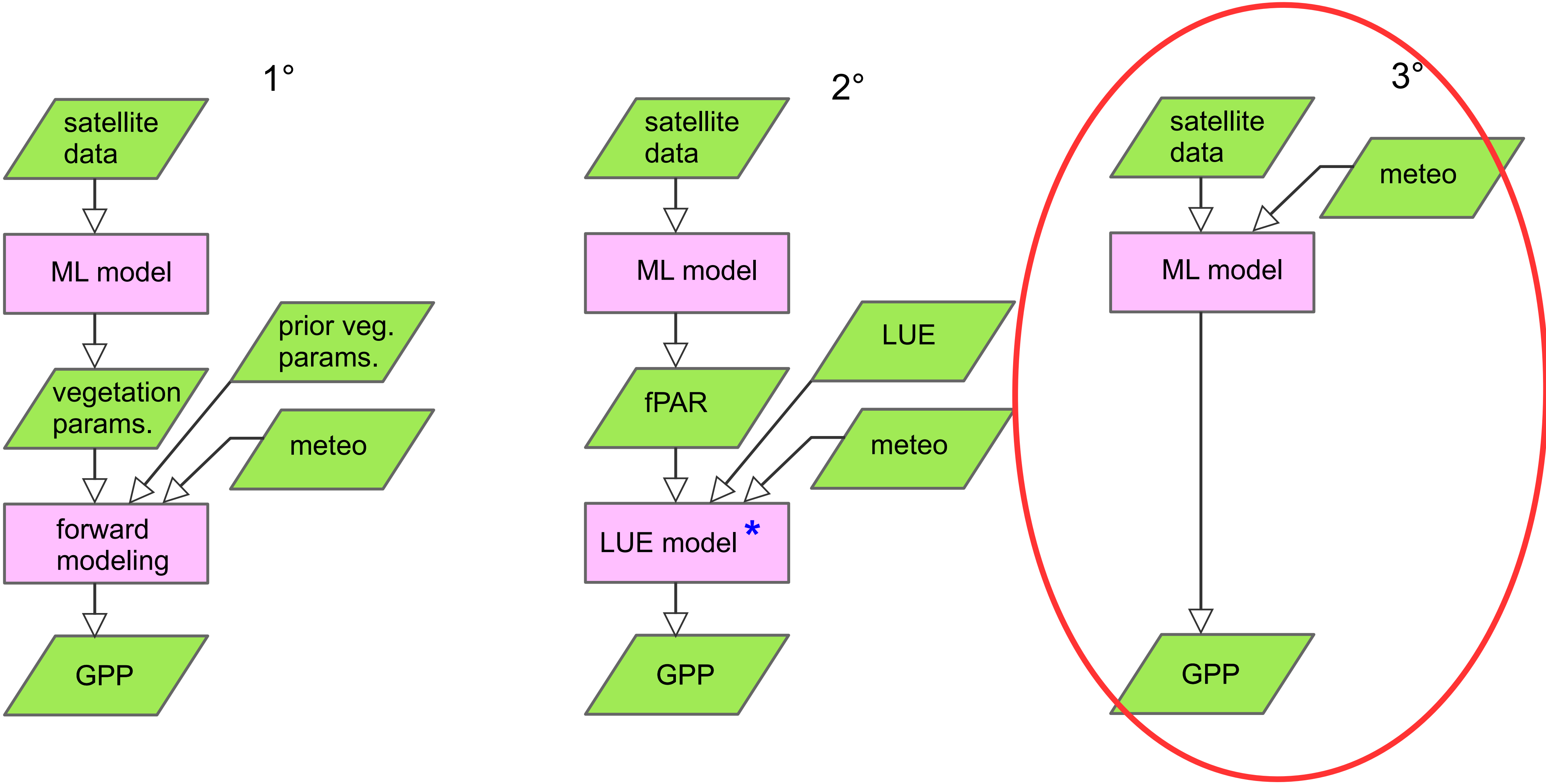


# Testing performance on synthetic dataset



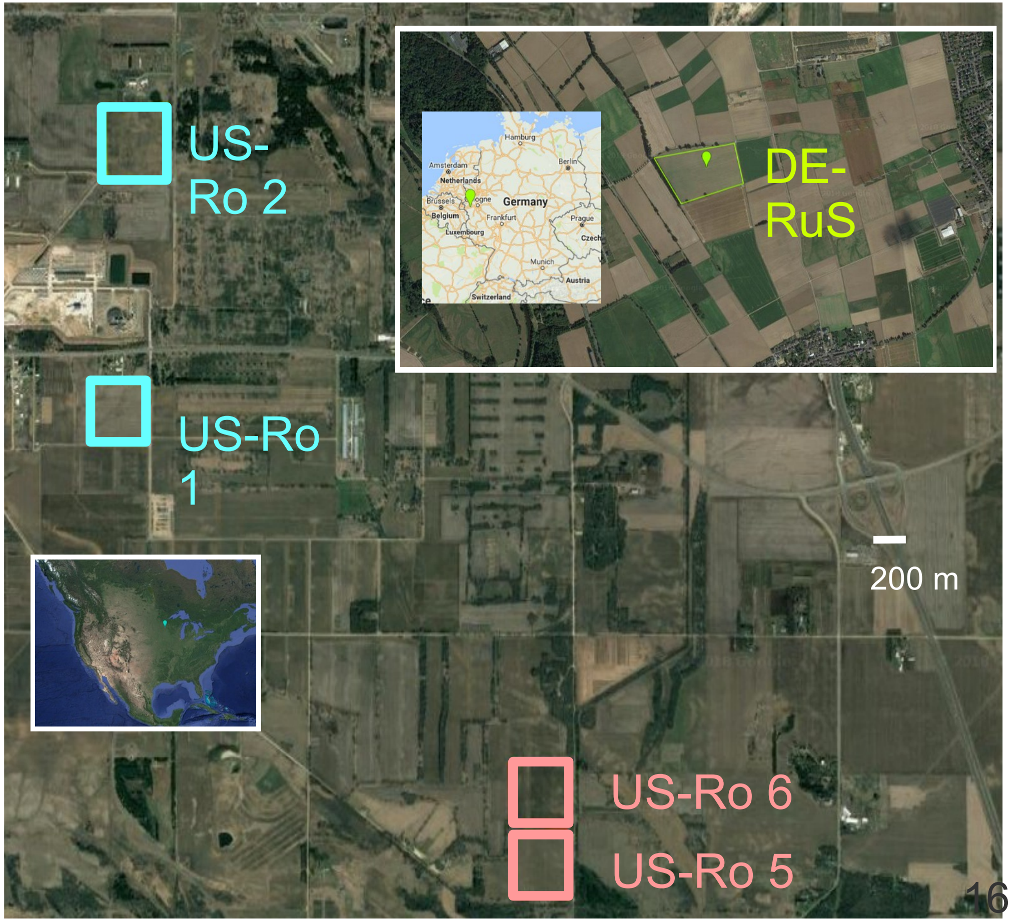

- Performance of the ML models (NN) for the synthetic datasets (at Sentinel-2 bands)




1°
satellite data
ML model
vegetation params.
prior veg. params.
meteo
forward modeling
GPP
2°
satellite data
ML model
fPAR
LUE
meteo
LUE model *
GPP
3°
satellite data
meteo
ML model
GPP

# Feasibility test with the flux tower dataset

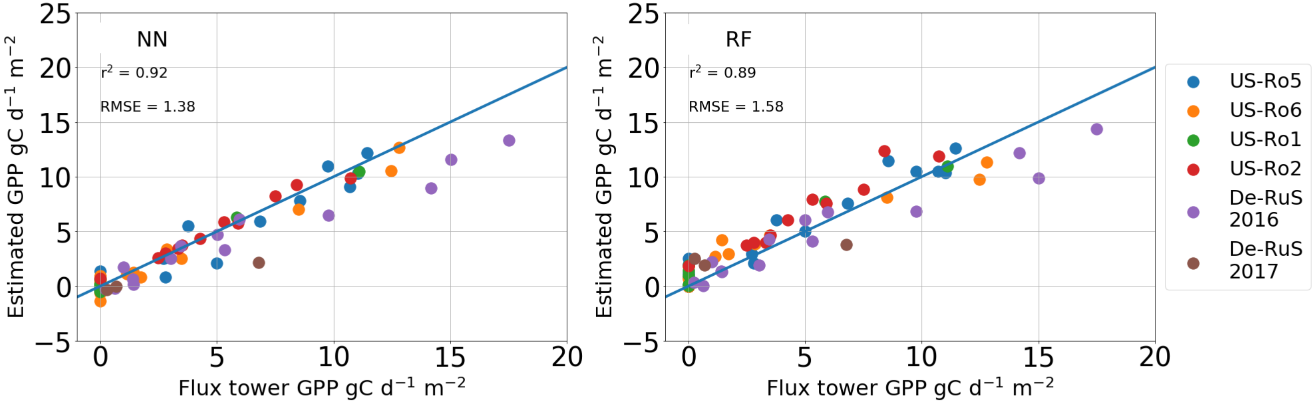
- Sentinel-2, atmospheric correction with Sen2Cor (version 2.4)
- Landsat 8, exported from Google Earth Engine (GEE)
- 2016 & 2017
- Meteorology from GLDAS 2.1 (Rodell et al., 2004), exported from GEE
  - 3 h resolution
- Integrated to daily values




Great thanks to PIs and Data Managers of the flux tower sites: Timothy Griffis (University of Minnesota), Cody Winker (USDA-ARS), John Baker (USDA–ARS), Alexander Graf (Forschungszentrum Julich) and Marius Schmidt (Forschungszentrum Julich)

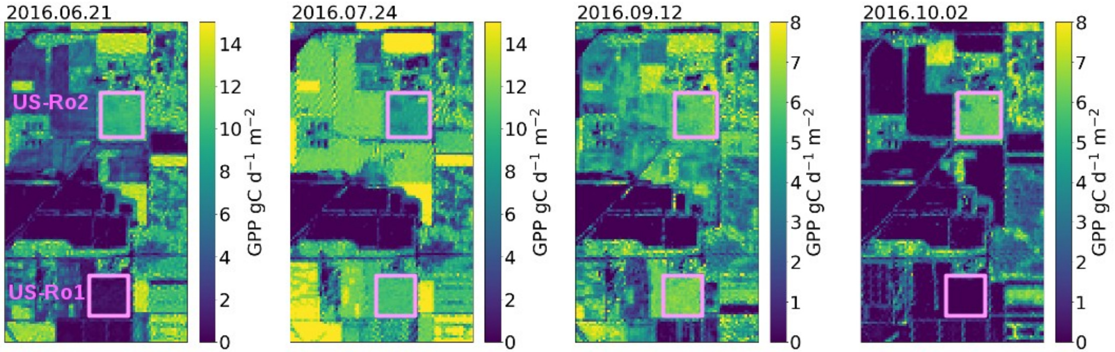
# Results: Sentinel-2

- Good performance of ML models (Neural Network and Random Forest)



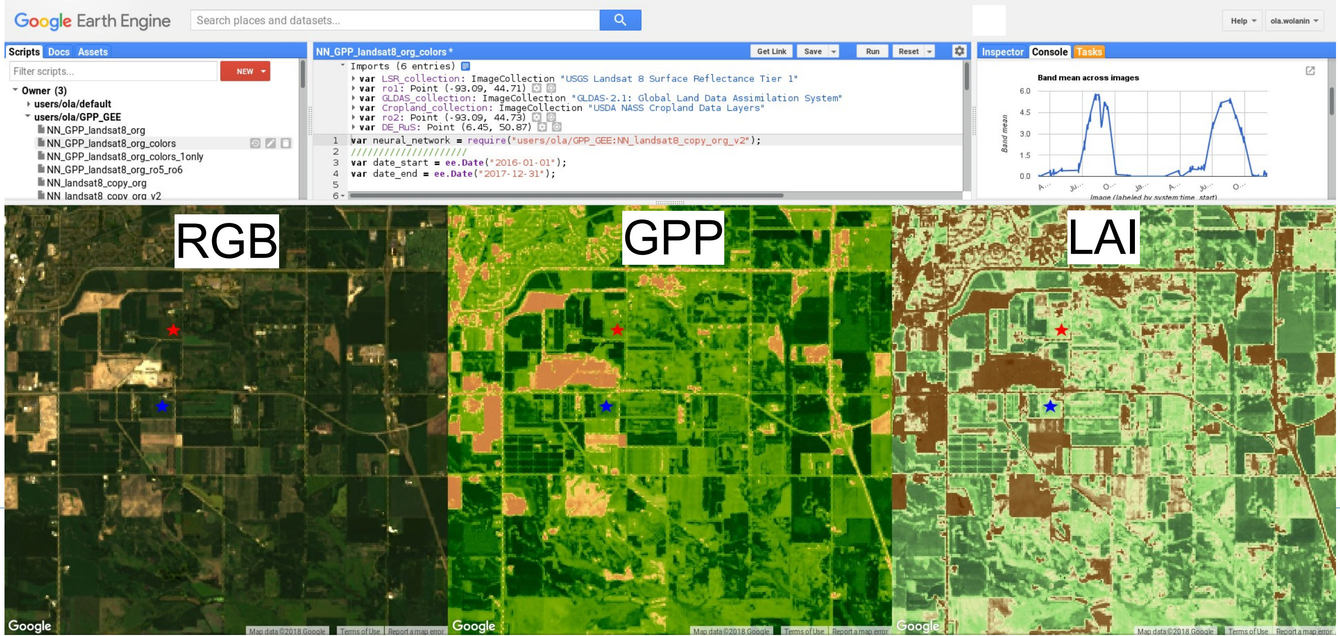

Kura clover

soybeans



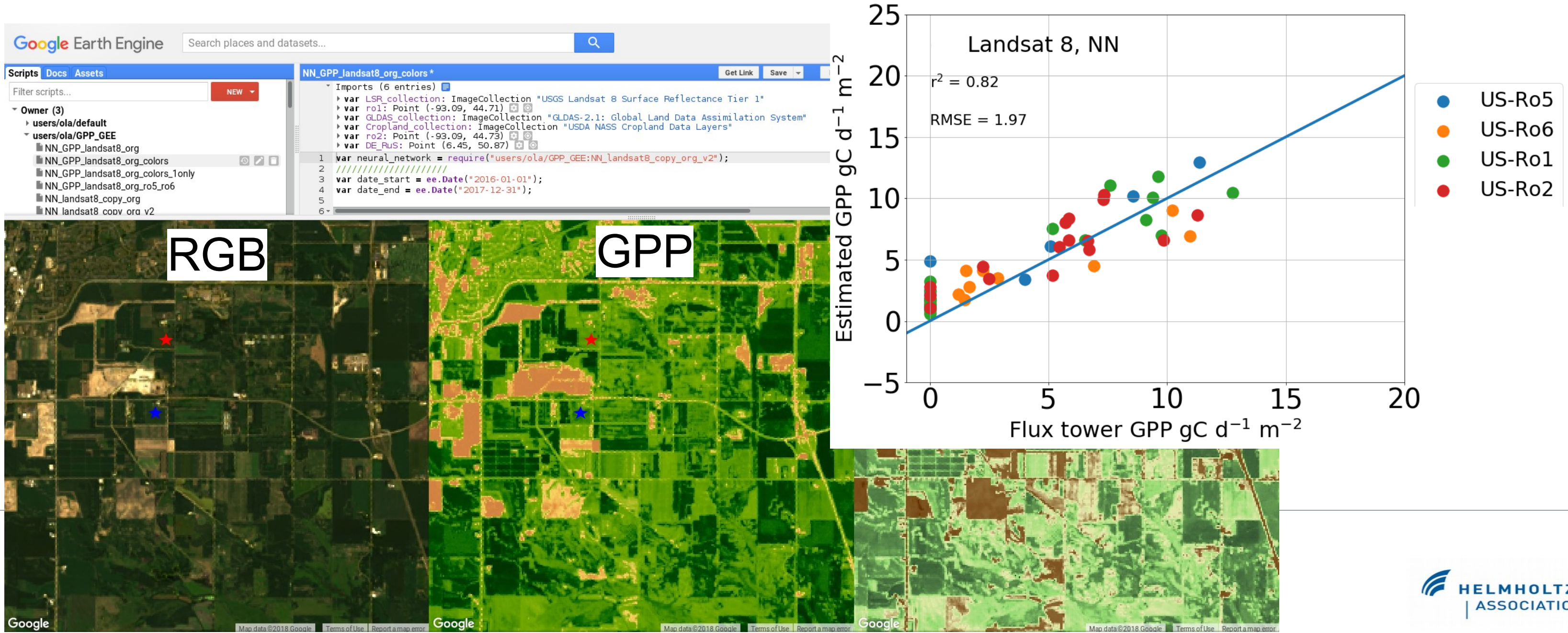
# Towards a global application: Landsat 8

- Looking for cloud-based solution: e.g., Google Earth Engine
  - Data archived in one location
  - Computational power of GEE
  - Easily available to others



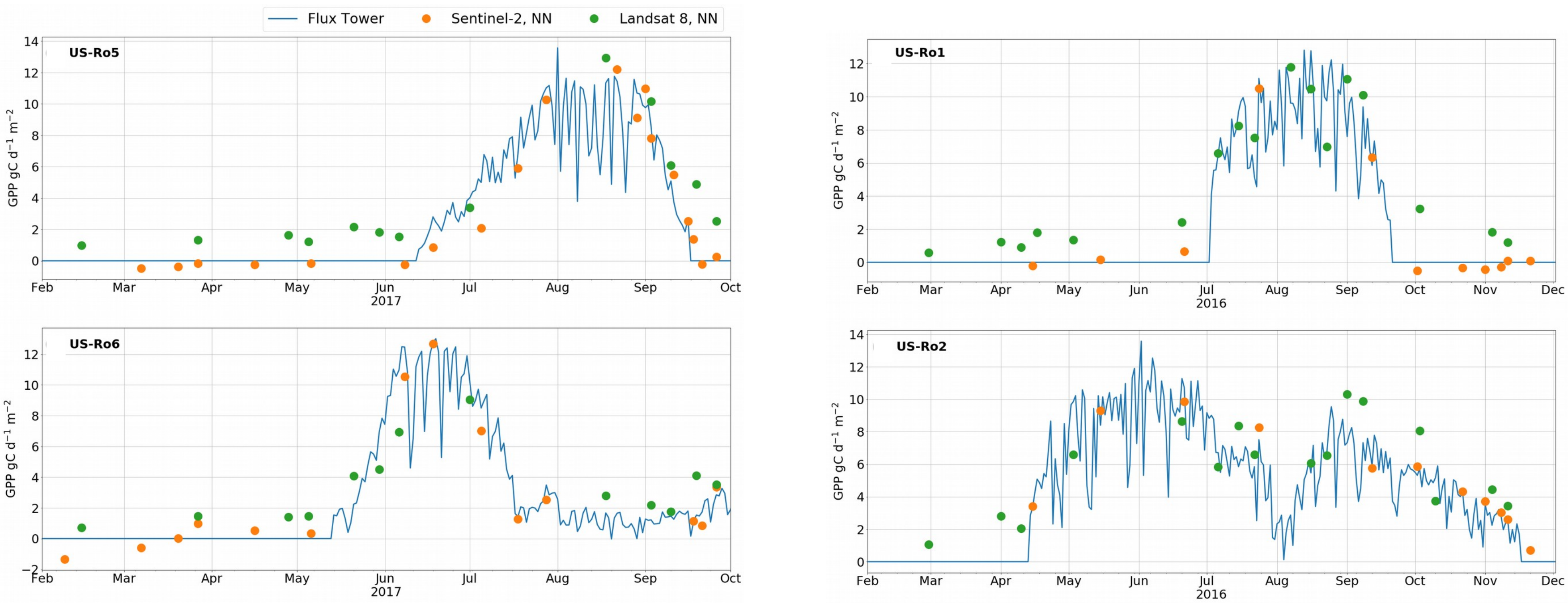
# Towards a global application: Landsat 8

- Looking for cloud-based solution: e.g., Google Earth Engine
  - Data archived in one location
  - Computational power of GEE
  - Easily available to others



# Time series for Sentinel-2 and Landsat 8

- Similar values for the growing seasons
- Overestimation outside of the growing season, especially for Landsat 8
- Significant increase in available datapoints
- Shows potential for using Sentinel-2 and Landsat 8
  - as well as promising application to other satellites



# Conclusions

- Developed ML model for GPP trained on data modeled with SCOPE
  - applied across a range of satellite instruments
  - using all available bands
  - not need to wait for a sufficient dataset to train the model (more data always good for validation and improvement!)
  - if model improved or changes – can be easily adapted
  - fast and can be easily applied globally in GEE
  - no physiological stress simulation (e.g., soil moisture limitation)
  - noise associated to the data
- Demonstrated the feasibility of this method for Sentinel-2 and Landsat 8
- Outlook:

  - further implementation and validation

  - model improvement: testing the model architecture; selecting the training dataset


Aleksandra Wolanin, ola@gfz-potsdam.de

# Conclusions

- Developed ML model for GPP trained on data modeled with SCOPE
  - applied across a range of satellite instruments
  - using all available bands
  - not need to wait for a sufficient dataset to train the model (more data always good for validation and improvement!)
  - if model improved or changes – can be easily adapted
  - fast and can be easily applied globally in GEE
  - no physiological stress simulation (e.g., soil moisture limitation)
  - noise associated to the data
- Demonstrated the feasibility of this method for Sentinel-2 and Landsat 8
- Outlook:

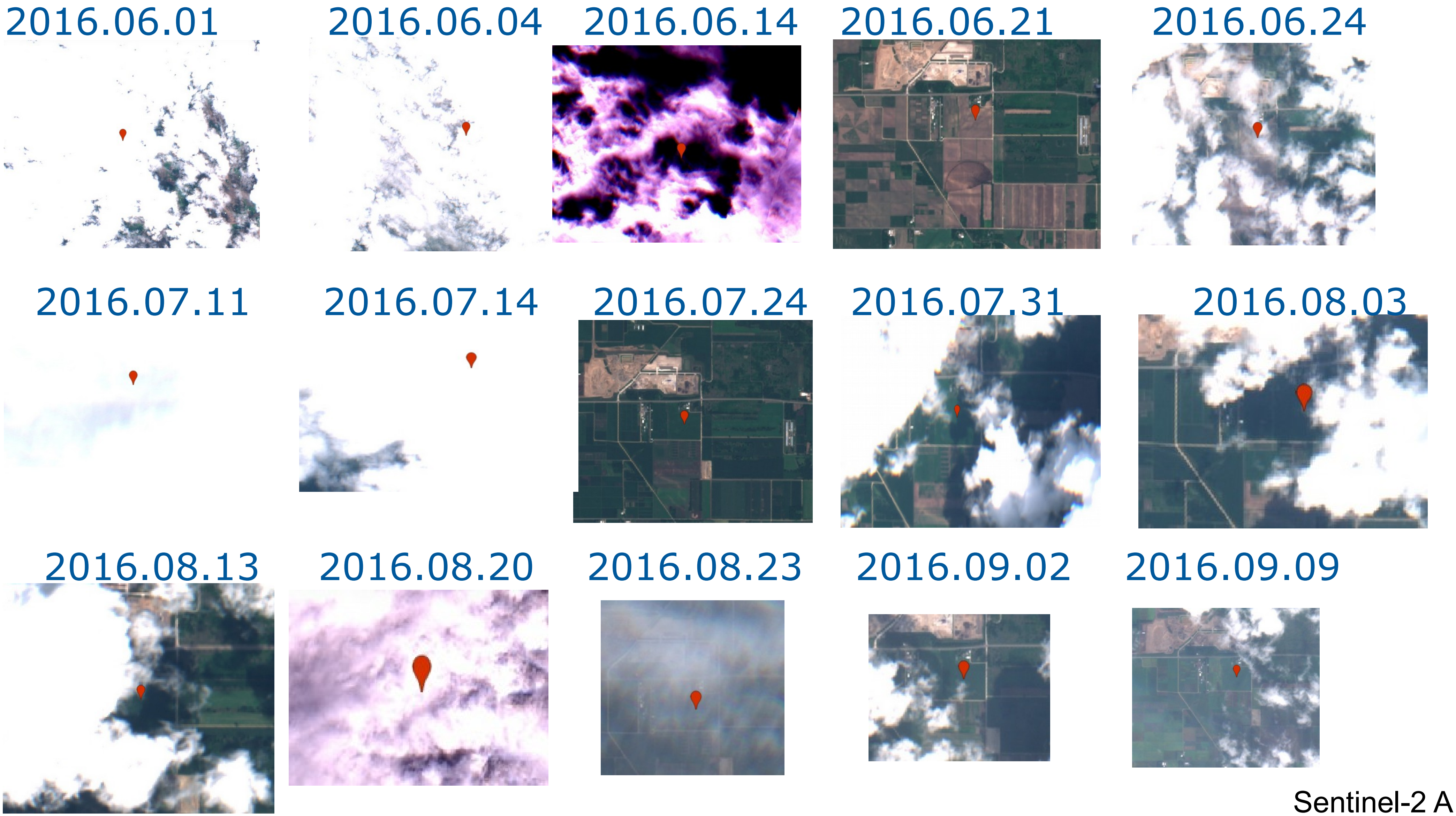
  - further implementation and validation

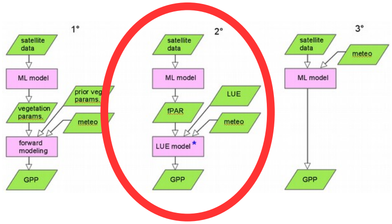
  - model improvement: testing the model architecture; selecting the training dataset

Thank you!



Aleksandra Wolanin, ola@gfz-potsdam.de

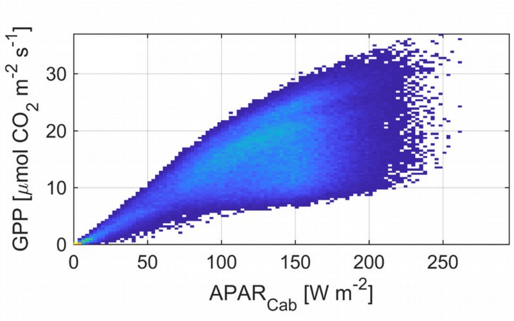

2016.06.01
2016.06.04
2016.06.14
2016.06.21
2016.06.24
2016.07.11
2016.07.14
2016.07.24
2016.07.31
2016.08.03
2016.08.13
2016.08.20
2016.08.23
2016.09.02
2016.09.09
Sentinel-2 A

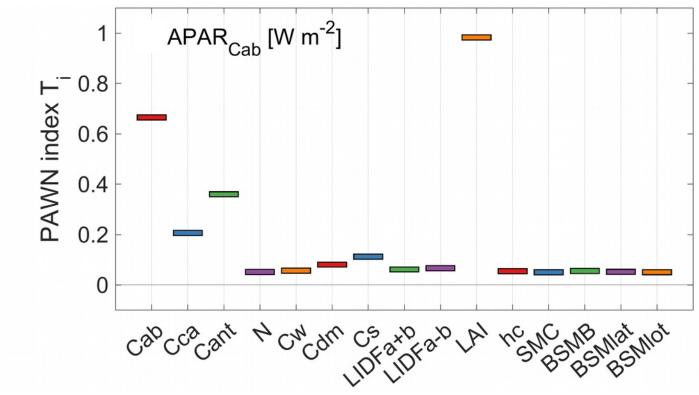


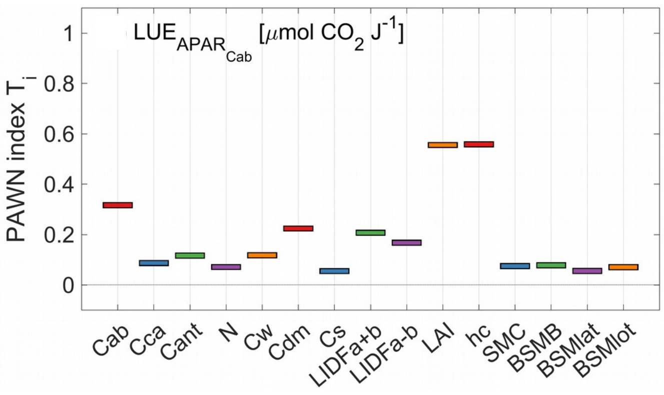

# LUE model based on SCOPE

- GPP = APAR * LUE * f(meteo)
- LUE (light use efficiency) not constant
- All variability is due to changes in leaf and canopy properties

- More parameter driving variability in LUE than APAR




GFZ
Helmholtz-Zentrum
Potsdam

Helmholtz Association

Table 4: Details about the flux tower sites used in this study (Griffis et al., 2004; Eder et al., 2015).

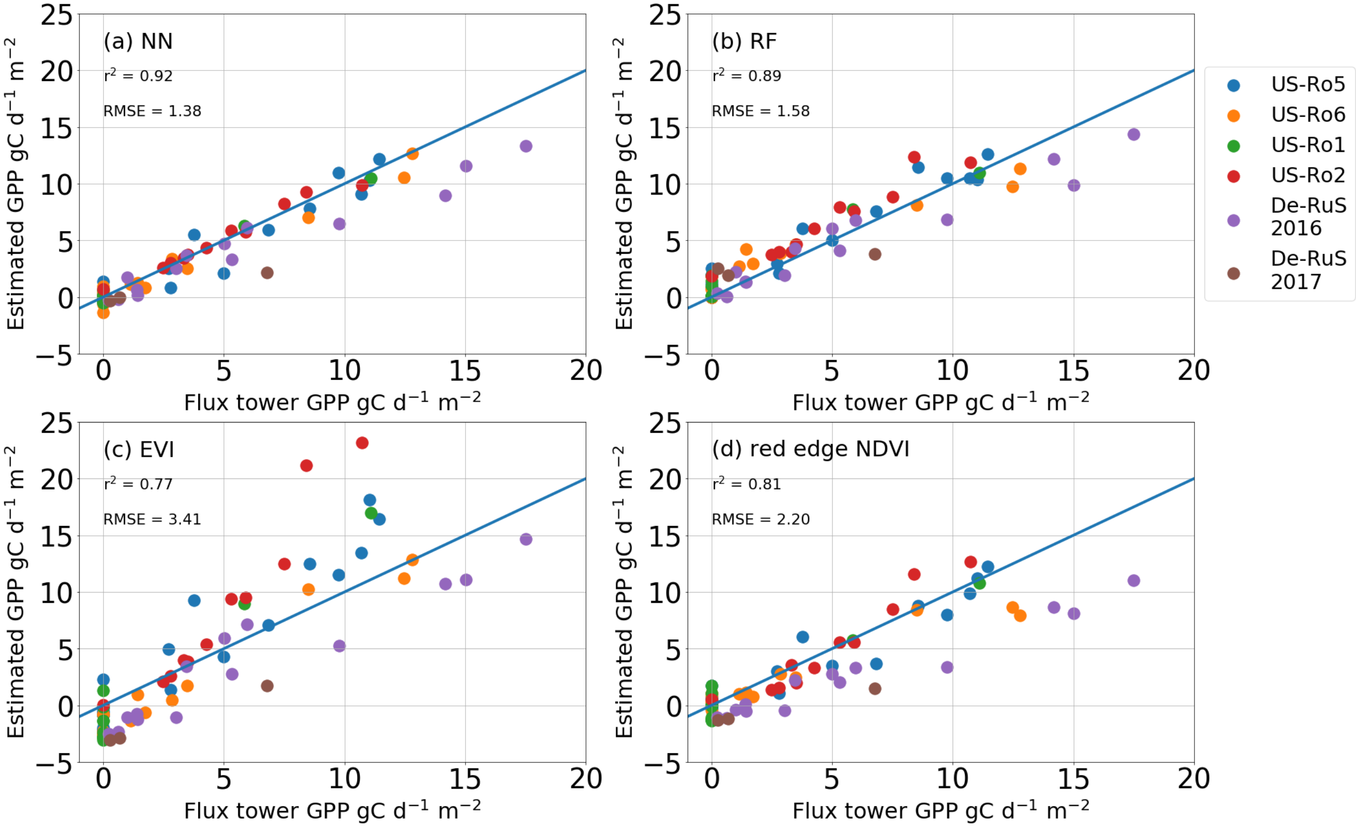
| Site ID | Lon (°W) | Lat (°N) | Period | Crops |
|---|---|---|---|---|
| US-Ro1 | -93.0898 | 44.7143 | 2016 | soybeans |
| US-Ro2 | -93.0888 | 44.7288 | 2016 | Kura clover only |
| US-Ro5 | -93.0576 | 44.6910 | 2017 | soybean |
| US-Ro6 | -93.0578 | 44.6946 | 2017 | wheat/Kura Clover |
| DE-RuS | 6.4472 | 50.8659 | 2016 & 2017 | winter barley in spring 2016, a catch crop mixture in fall 2016 and sugarbeet in 2017 |

Table 1: Sentinel-2A spectral specifications and spatial resolution. The bands written in bold are those that overlap with Landsat 8 bands. Source: https://earth.esa.int/web/sentinel/user-guides/sentinel-2-msi/resolutions/radiometric

| Bands | B1 | **B2** | **B3** | **B4** | B5 | B6 | B7 | B8 | **B8a** | B9 | B10 | **B11** | **B12** |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Center wavelength (nm) | 443.9 | **496.6** | **560.0** | **664.5** | 703.9 | 740.2 | 782.5 | 835.1 | **864.8** | 945.0 | 1375 | **1613.7** | **2202.4** |
| Bandwidth (nm) | 27 | **98** | **45** | **38** | 19 | 18 | 28 | 145 | **33** | 26 | 75 | **143** | **242** |
| Spatial resolution (m) | 60 | **10** | **10** | **10** | 20 | 20 | 20 | 10 | **20** | 60 | 60 | **20** | **20** |

Table 3: Summary of vegetation indices used in this study. $\rho_{\text{NIR}}$, $\rho_{\text{red edge}}$, $\rho_{\text{green}}$ and $\rho_{\text{red}}$ are reflectance in spectral bands of NIR, red edge, red, green regions and the refer to Sentinel-2 bands B8 (842 nm), B5 (705 nm), B4 (665 nm), B3 (560 nm) respectively.

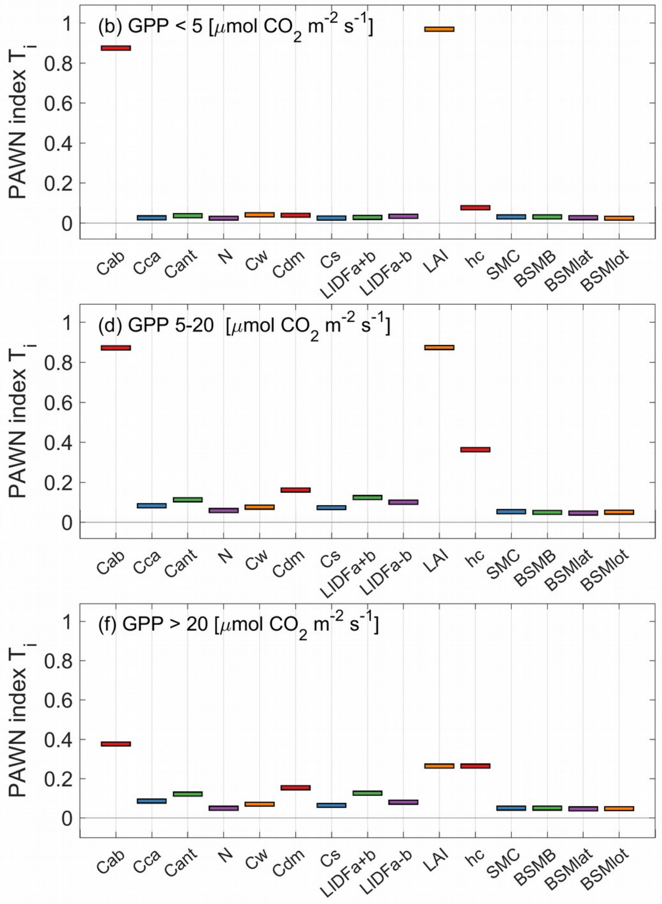
| Vegetation index (VI) | VI abbreviation | VI formula | GPP ($x = VI \times PAR_{in}$) | Reference |
|---|---|---|---|---|
| Red edge chlorophyll index | $\text{CI}_{\text{red edge}}$ | $\rho_{\text{NIR}}/\rho_{\text{red edge}} - 1$ | $4.80ln(x) - 37.93$ | Peng and Gitelson (2012) |
| Green chlorophyll index | $\text{CI}_{\text{red edge}}$ | $\rho_{\text{NIR}}/\rho_{\text{green}} - 1$ | $5.13ln(x) - 46.92$ | Peng and Gitelson (2012) |
| Normalized difference vegetation index | NDVI | $(\rho_{\text{NIR}} - \rho_{\text{red}})/(\rho_{\text{NIR}} + \rho_{\text{red}})$ | $2.07x - 6.19$ | Gitelson et al. (2012) |
| Green normalized difference vegetation index | greenNDVI | $(\rho_{\text{NIR}} - \rho_{\text{green}})/(\rho_{\text{green}} + \rho_{\text{green}})$ | $2.86x - 11.9$ | Gitelson et al. (2012) |
| Enhanced vegetation index | EVI | | $2.26x - 3.73$ | Peng et al. (2013) |
| Red edge normalized difference vegetation index | reNDVI | $(\rho_{\text{NIR}} - \rho_{\text{red edge}})/(\rho_{\text{NIR}} + \rho_{\text{red edge}})$ | | |



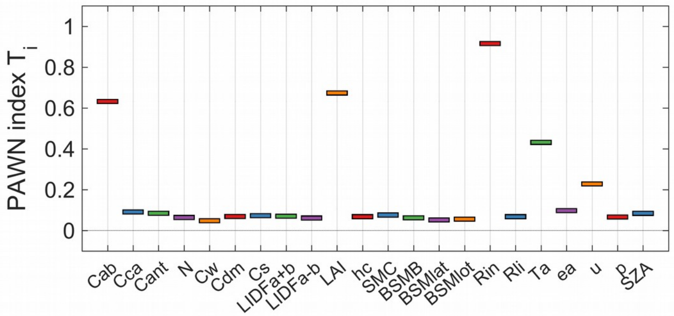
# Sensitivity analysis of parameters driving GP

- PAWN sensitivity analysis (Pianosi and Wagener, 2015)
- Relative importance of each input variable

For different subranges of GPP

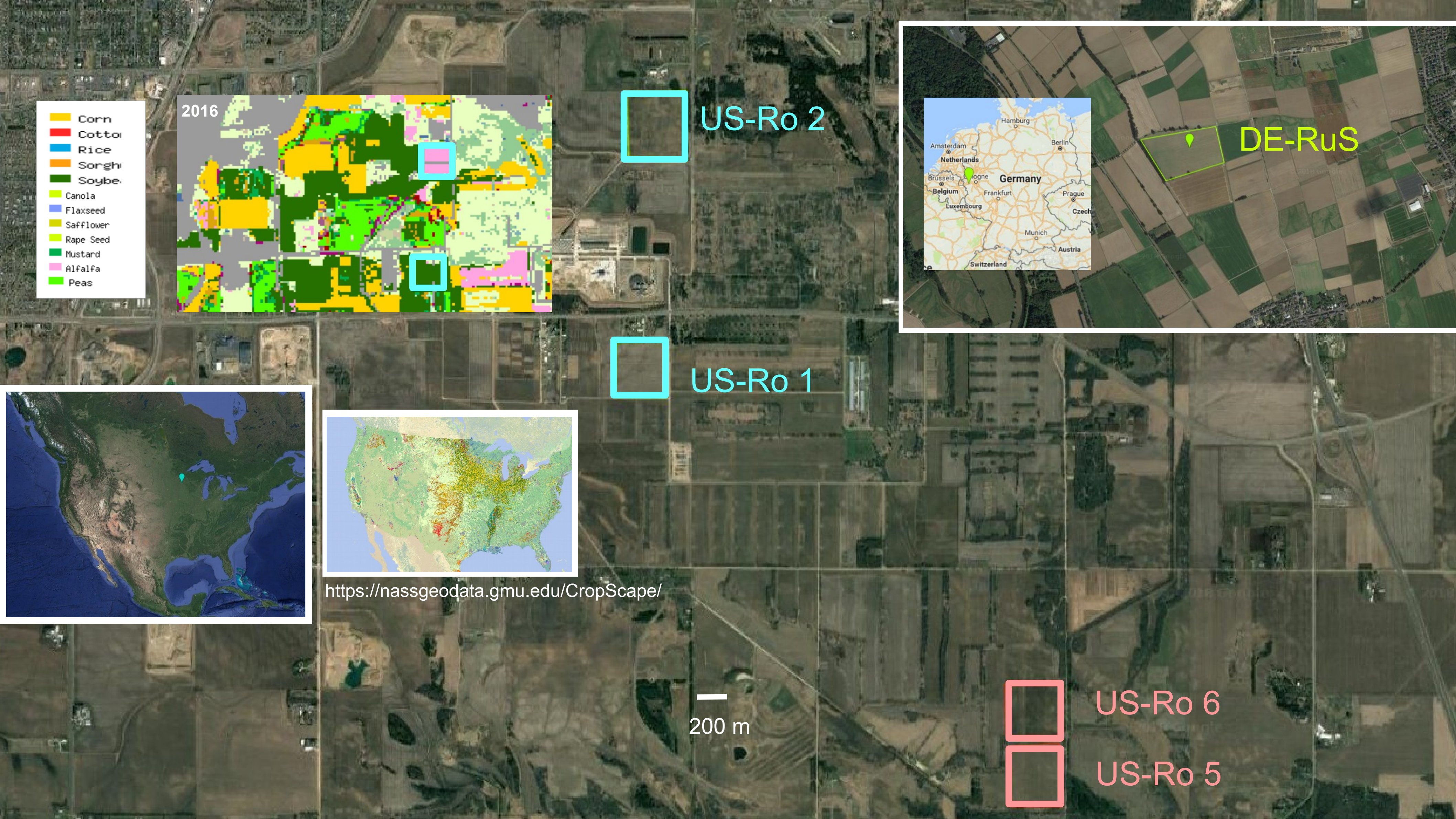



Including meteo data




GFZ Helmholtz-Zentrum Potsdam

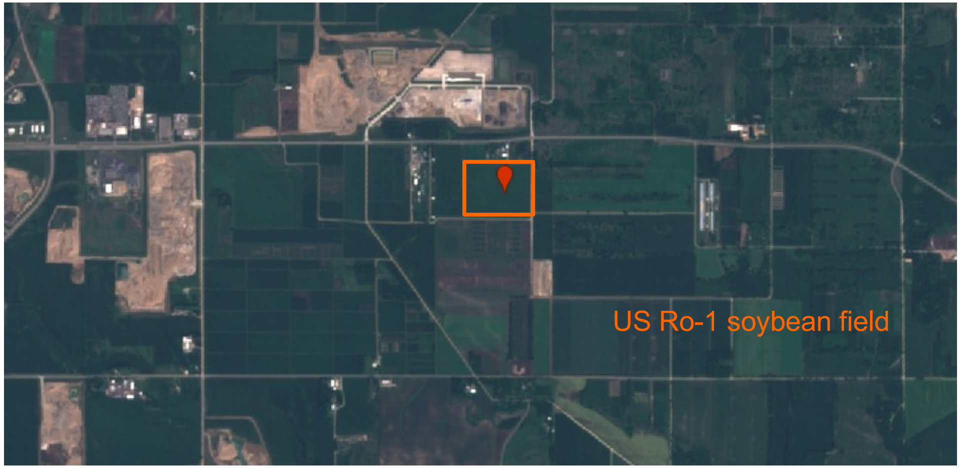
Helmholtz Association



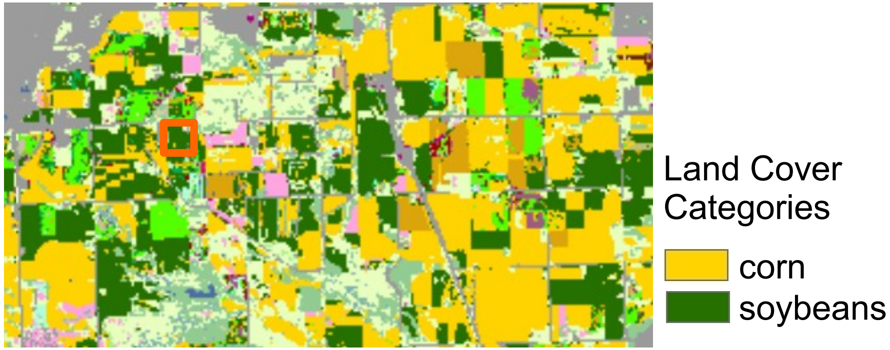
2016
Corn
Rice
Canola
Flaxseed
Safflower
Rape Seed
Mustard
Alfalfa
Peas
US-Ro 2
US-Ro 1
US-Ro 6
US-Ro 5
DE-RuS
Hamburg
Amsterdam
Netherlands
Berlin
Brussels
Belgium
Germany
Frankfurt
Prague
Luxembourg
Czech
Munich
Austria
Switzerland
200 m
https://nassgeodata.gmu.edu/CropScape/

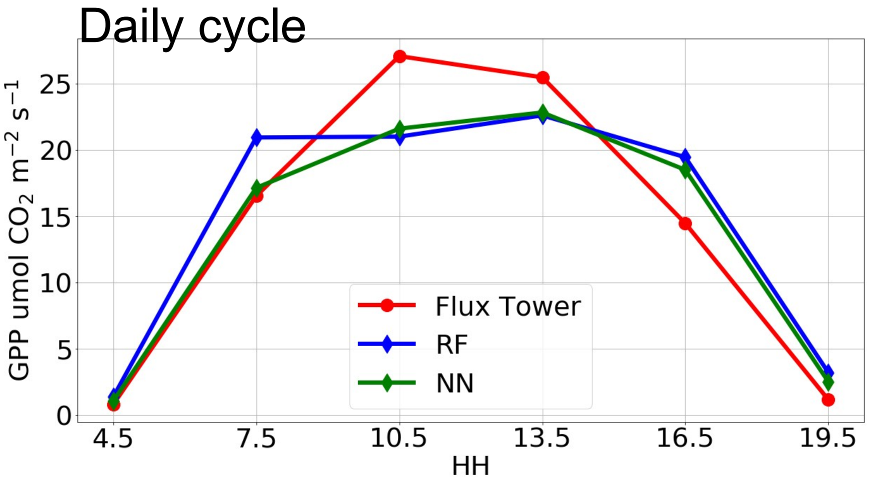
# Application: daily cycle

- US-Ro1 flux tower
- Clear sky Sentinel-2 image 24.07.2016

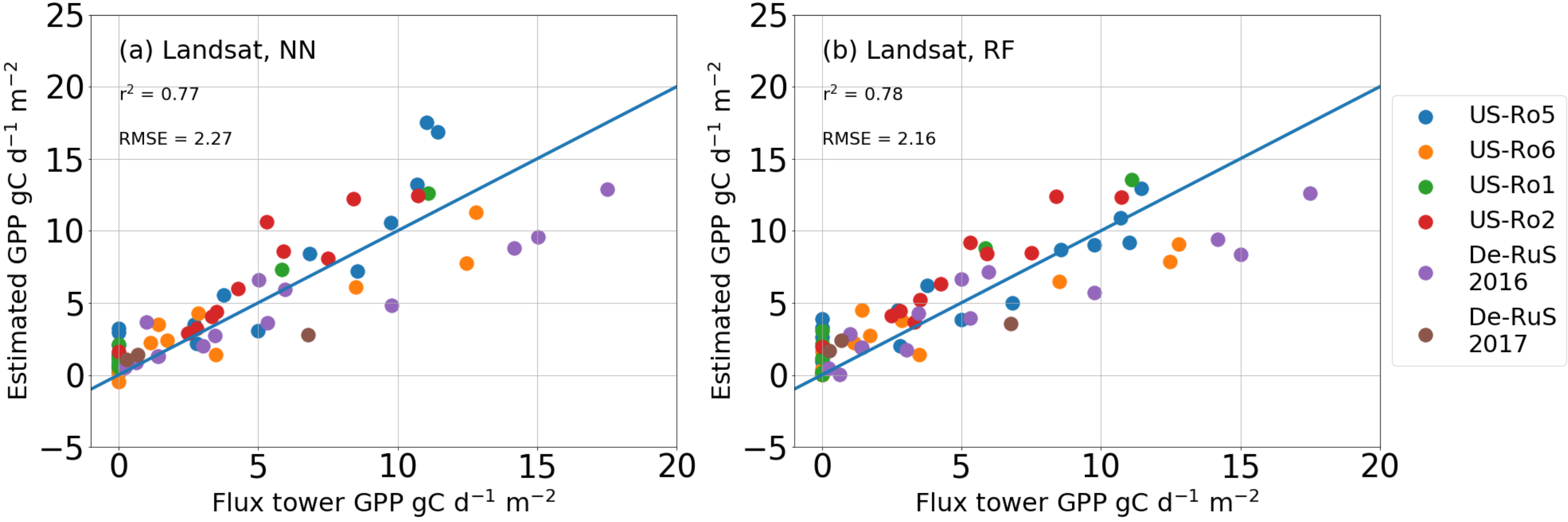




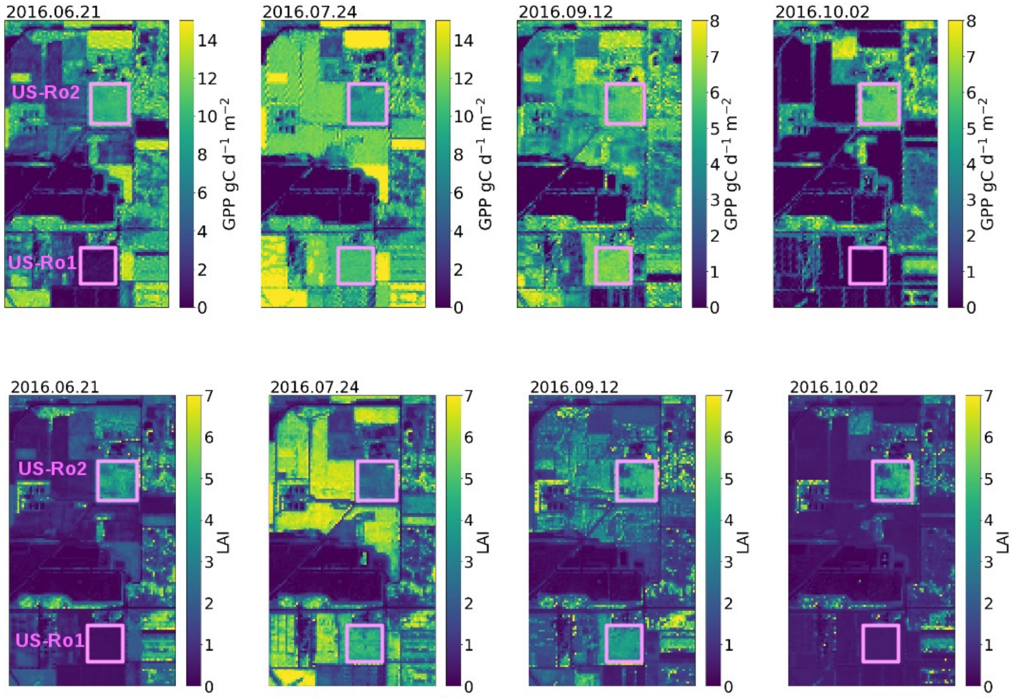


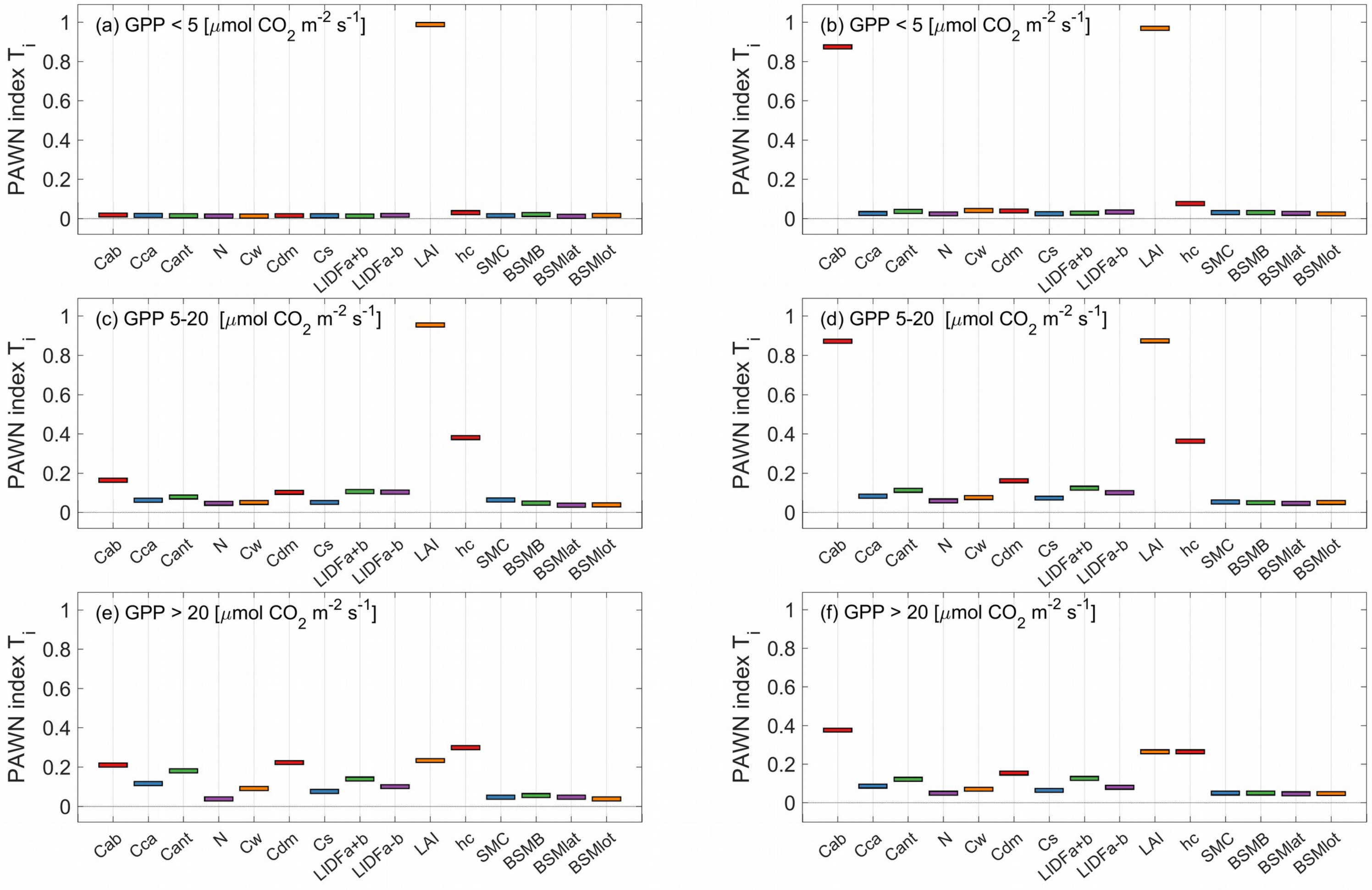



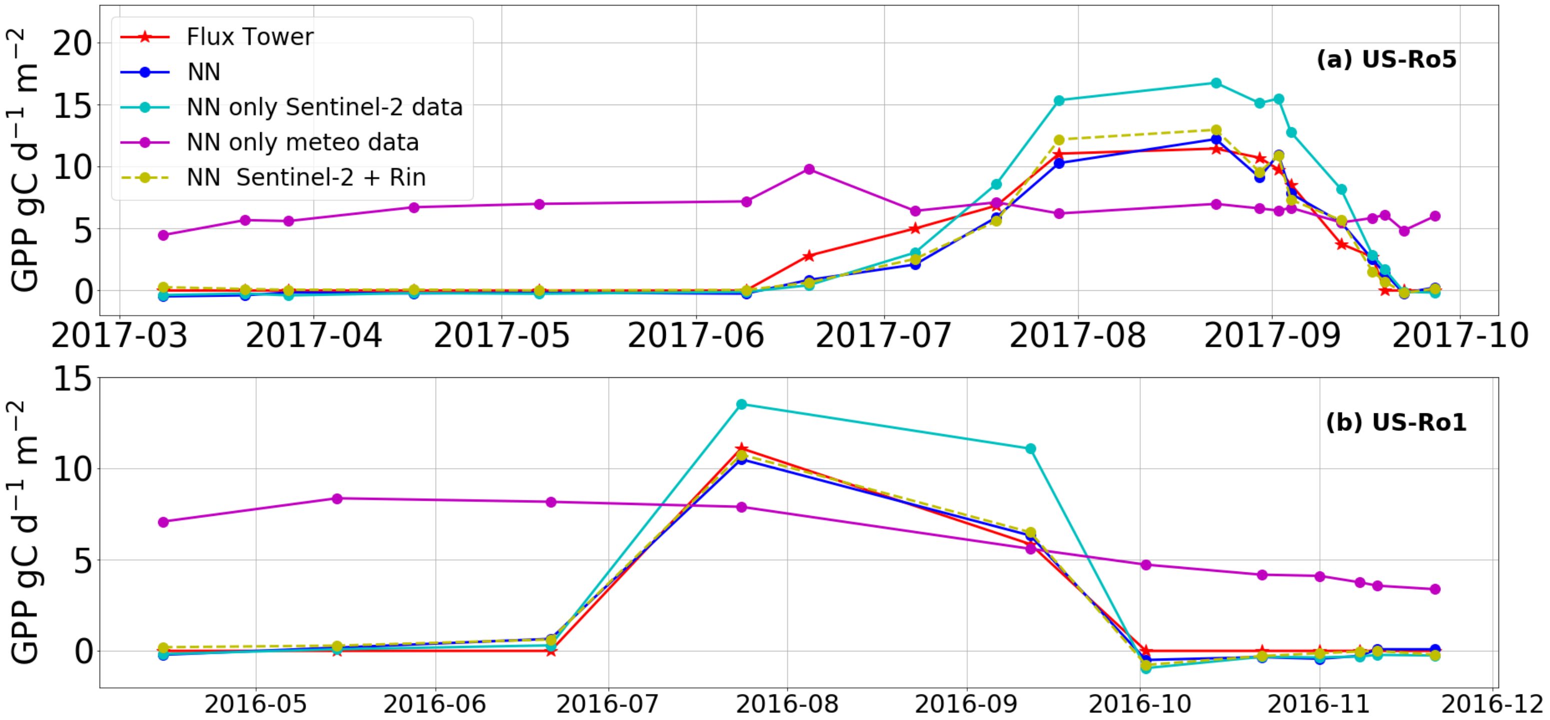

Flux data: curtesy of the Biometeorology Group at the University of Minnesota-Twin Cities



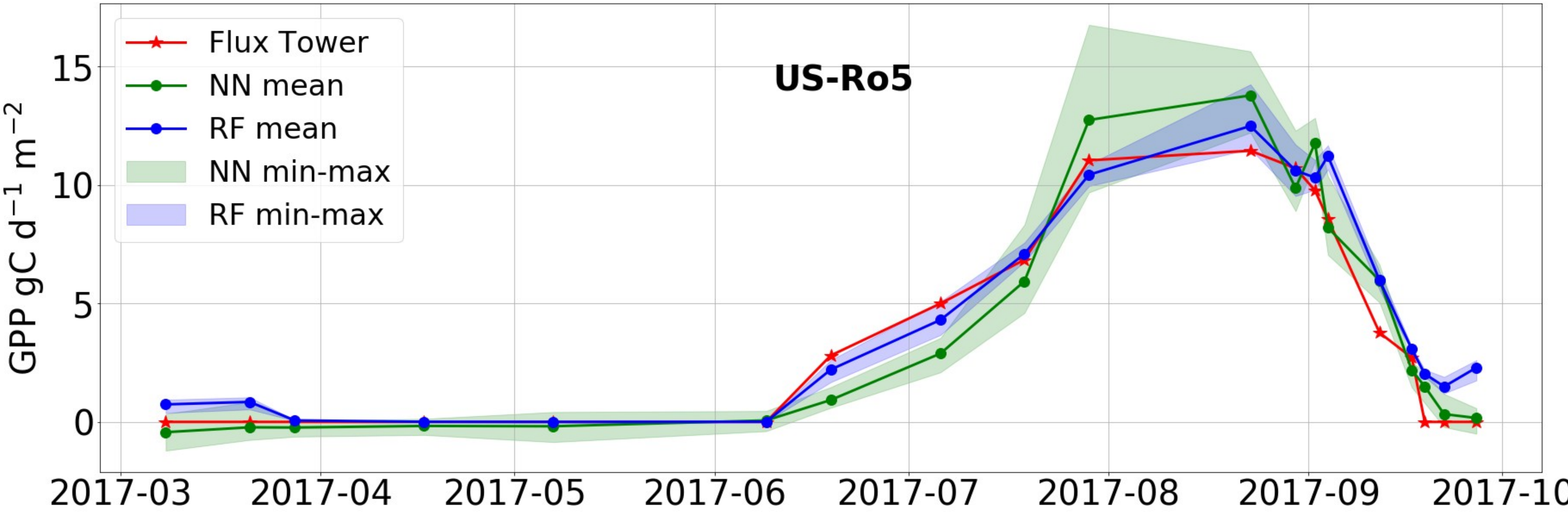

Sentinel-2 data, bands shared with Landsat 8

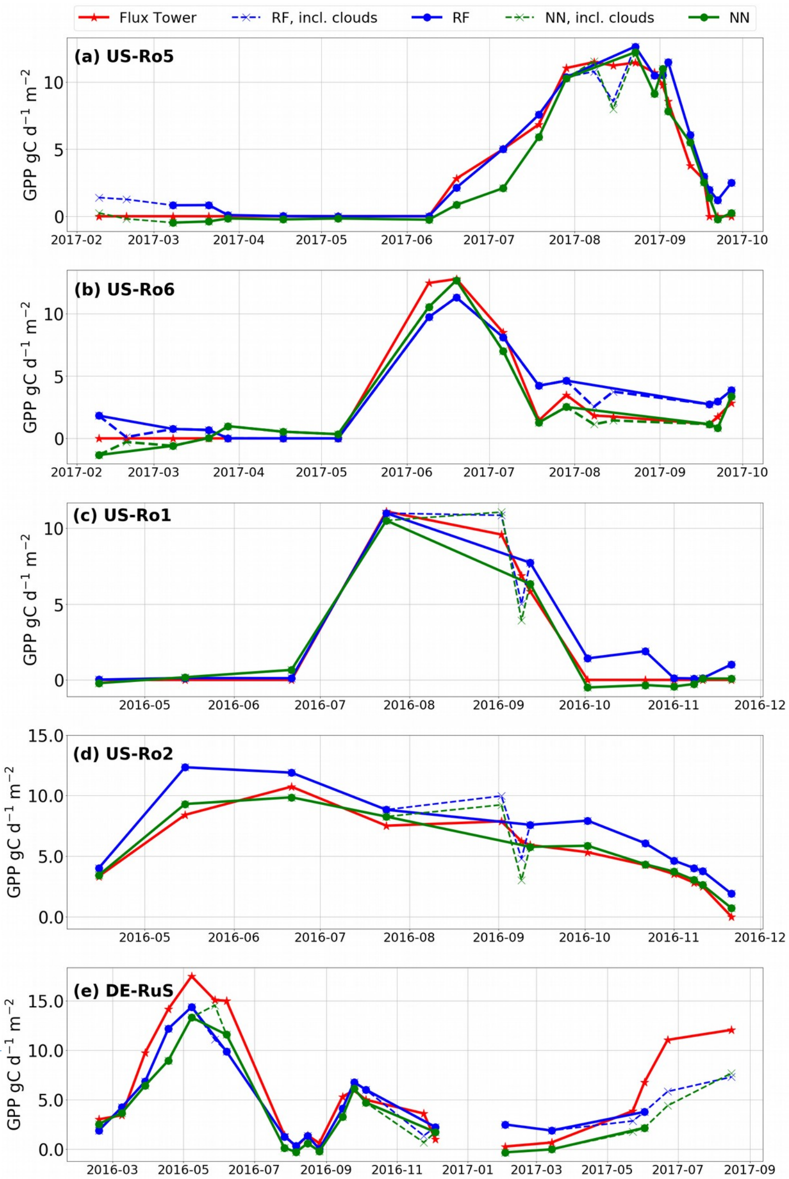

2016.06.21
2016.07.24
2016.09.12
2016.10.02
US-Ro2
US-Ro1
GPP gC d$^{-1}$ m$^{-2}$
LAI
0 2 4 6 8 10 12 14
0 1 2 3 4 5 6 7 8
0 1 2 3 4 5 6 7

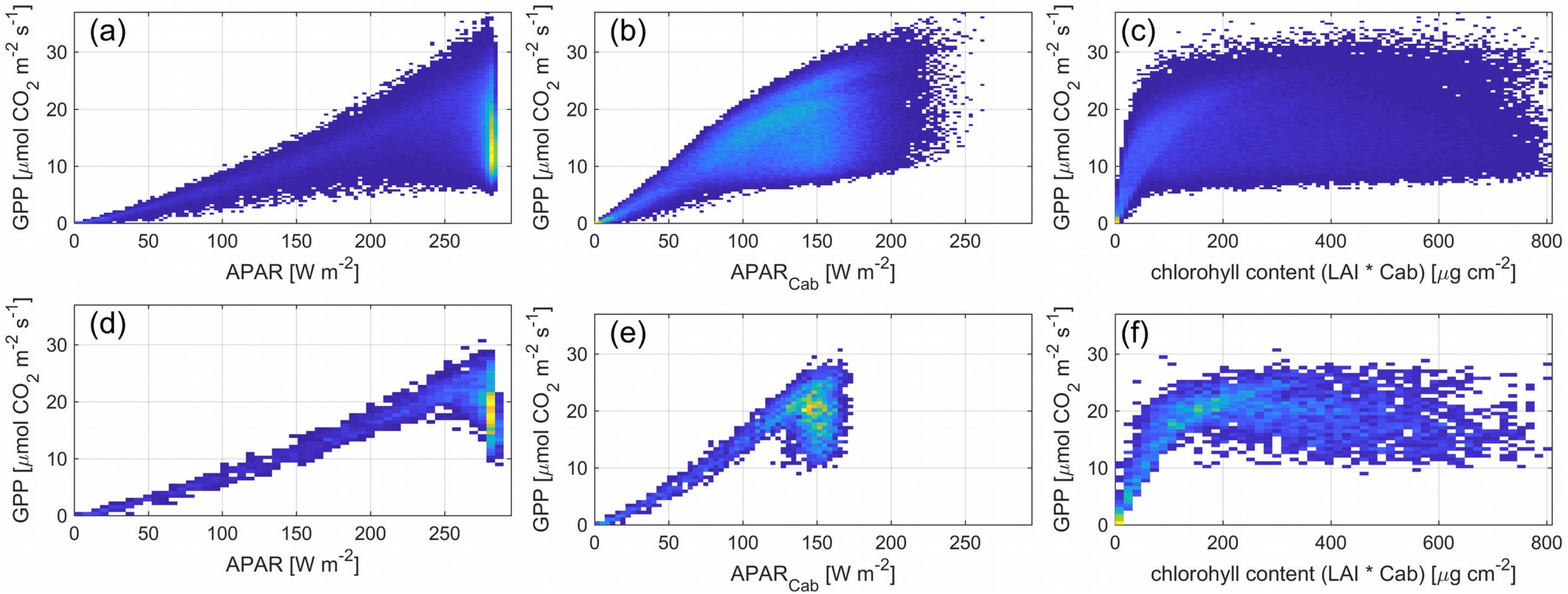

(a) GPP < 5 [μmol CO2 m-2 s-1]
(b) GPP < 5 [μmol CO2 m-2 s-1]
(c) GPP 5-20 [μmol CO2 m-2 s-1]
(d) GPP 5-20 [μmol CO2 m-2 s-1]
(e) GPP > 20 [μmol CO2 m-2 s-1]
(f) GPP > 20 [μmol CO2 m-2 s-1]
PAWN index Ti
0
0.2
0.4
0.6
0.8
1
Cab
Cca
Cant
N
Cw
Cdm
Cs
LIDFa+b
LIDFa-b
LAI
hc
SMC
BSMB
BSMlat
BSMlot



Flux Tower
NN
NN only Sentinel-2 data
NN only meteo data
NN Sentinel-2 + Rin
(a) US-Ro5
GPP gC d$^{-1}$ m$^{-2}$
20
15
10
5
0
2017-03
2017-04
2017-05
2017-06
2017-07
2017-08
2017-09
2017-10
(b) US-Ro1
GPP gC d$^{-1}$ m$^{-2}$
15
10
5
0
2016-05
2016-06
2016-07
2016-08
2016-09
2016-10
2016-11
2016-12


US-Ro5
Flux Tower
NN mean
RF mean
NN min-max
RF min-max
GPP gC d$^{-1}$ m$^{-2}$
0
5
10
15
2017-03
2017-04
2017-05
2017-06
2017-07
2017-08
2017-09
2017-10


Flux Tower
RF, incl. clouds
RF
NN, incl. clouds
NN
(a) US-Ro5
(b) US-Ro6
(c) US-Ro1
(d) US-Ro2
(e) DE-RuS
GPP gC d$^{-1}$ m$^{-2}$


(a)
(b)
(c)
(d)
(e)
(f)
GPP [μmol CO2 m-2 s-1]
APAR [W m-2]
APARCab [W m-2]
chlorohyll content (LAI * Cab) [μg cm-2]
0
10
20
30
50
100
150
200
250
400
600
800