Internship Report

# VERIFYING THE PRIMARY PRODUCTIVITY DATA FROM BETHY/DLR MODEL WITH OTHER MODEL OUTPUTS

**Intern/Student:** Md Saiful Islam

**Student no.:** s6048943

**Name of the Company:** German Aerospace Center (DLR)

**Address of the Company:** Münchener Str. 20, Oberpfaffenhofen, 82234 Weßling, Germany

**Period of the Internship:** 24th April to 16th June 2023 (08 weeks)

**Host Supervisor:** dr. Andreas Hirner, Agriculture and Forest Ecosystems Team, Land Surface Dynamics

**UT Supervisor:** dr. Roshanak Darvishzadeh, Natural Resources Department, Faculty ITC

## ACKNOWLEDGEMENT

I express my sincere thanks to dr. Andreas Hirner, the host supervisor at DLR and dr. Roshanak Darvishzadeh, the university supervisor, for their help and guidance throughout my internship. I also sincerely thank Stefanie Holzwarth for all her support and effort from the recruiting to the end of this internship. My sincere thanks also go to Belinda Jaarsma-Knol, internship coordinator at ITC, for all the support and help she provided at different stages of this internship. I also thank GEM Management, especially drs. Raymond Nijmeijer and Laura Windig for their support in solving the issue of the internship period. Finally, I thank my friends and colleagues at DLR, who gave me company, inspiration, and perspectives during my work.

# Table of Contents

1 Introduction ..... 5
1.1 Background ..... 5
1.2 Objectives ..... 5
1.3 Deliverables ..... 6
2 Data and Methods ..... 6
2.1 Data ..... 6
2.2 Methods ..... 7
2.2.1 Data pre-processing and harmonization ..... 7
2.2.2 Sample point generation ..... 7
2.2.3 Evaluation statistics and pattern identification ..... 7
2.2.4 Computing environment ..... 8
3 Results ..... 8
3.1.1 Comparison of BETHY with other models ..... 8
3.1.2 Patterns of the relations between BETHY and other models ..... 9
4 Discussion and Conclusion ..... 13
5 Reflection ..... 13
References ..... 14
Annex I ..... 15
Annex II ..... 16

# List of Figures

**Figure 1.** The flowchart shows the workflow of the assignment. ........ 8

**Figure 2.** Frequency distribution of the evaluation statistics between BETHY and other models; a. $R^2$ of BETHY and LPJ GPP (30 sample points in Europe); b. RMSE of BETHY and LPJ GPP (30 points in Europe); c. $R^2$ of BETHY and CASA NPP (20 and 12 points from West Africa and China); d. RMSE of BETHY and CASA NPP (20 and 12 points from West Africa and China). The x-axis shows the different range (bins) of $R^2$ and RMSE. The y-axis shows the no. of sample points in each range/bin. The unit for RMSE on x-axis is gC/m$^2$/timestep. ........ 9

**Figure 3.** Frequency distribution of the evaluation statistics between BETHY and other models. a. $R^2$ of BETHY and MODIS monthly GPP (11 points in Europe, 4 points were omitted due to no data); b. RMSE of BETHY and monthly MODIS GPP (11 points in Europe); c. $R^2$ of BETHY and MODIS annual NPP (20 points in Europe); d. RMSE of BETHY and MODIS annual NPP (20 points in Europe). The x-axis shows the different range (bins) of $R^2$ and RMSE. The y-axis shows the no. of sample points in each range/bin. The unit for RMSE on x-axis is gC/m$^2$/timestep. ........ 10

**Figure 4.** An example of good fit between BETHY and LPJ-GUESS decadal GPP with high $R^2$ and low RMSE. ........ 10

**Figure 5.** An example plot of good fit between monthly NPP from BETHY and CASA with high $R^2$ and low RMSE. ........ 11

**Figure 6.** An example of overestimation of BETHY compared to CASA monthly NPP. $R^2$ is higher since both models follow the same pattern, but higher RMSE indicates the difference in values. ........ 11

**Figure 7.** An example of underestimation of BETHY compared to monthly NPP of CASA. ........ 11

**Figure 8.** An example plot of leftward bias of BETHY decadal GPP compared to LPJ-GUESS. ........ 12

**Figure 9.** An example of irregular pattern between BETHY and LPJ-GUESS decadal GPP. ........ 12

**Figure 10.** Another example of irregular pattern between BETHY and CASA monthly NPP. ........ 12

# List of Tables

**Table 1.** The properties of datasets used to compare with the BETHY data. ........ 6

**Table 2.** The properties of BETHY data used to compare with other models' data. ........ 6

# 1 Introduction

## 1.1 Background

The carbon exchange in the terrestrial biosphere plays a significant role in the global carbon balance and climate system. Among the different elements of the biosphere, vegetation is a significant interface of carbon exchange between the atmosphere and soil. Vegetation produces organic matter through photosynthesis using solar radiation, carbon dioxide, and water as input variables, known as gross primary productivity (GPP). In addition to carbon production, vegetations release carbon through autotrophic respiration, a combination of maintenance and growth respiration. The remaining organic matter after respiration is net primary productivity (NPP).

Models are required at different spatial scales to monitor the function of vegetation in terms of their productivity. However, the existing models show inconsistency in their outputs due to their theoretical uncertainties and variations in the representation of vegetation functions on a global scale (Knorr, 1997). The long-term remote sensing observation made it possible to monitor global ecosystem productivity where different indices and biophysical parameters extracted from the reflectance values are used as model inputs. Nevertheless, there are distortions in these measurements that originated from viewing angles, soil background, and atmospheric composition.

The Biosphere Energy Transfer Hydrology (BETHY/DLR) model, a process model operated by German Aerospace Center(DLR), took a different approach to deal with these problems by simulating satellite signals by vegetation and radiative transfer models instead of using satellite images (Knorr, 1997). The model uses leaf area index and land cover data derived from satellite reflectance and meteorological data as input variables. The meteorological data includes air temperature, precipitation, cloud cover, wind speed, and soil water content (Tum et al., 2012). The model can produce GPP and NPP at daily timesteps.

Since the model involves many input variables and can be extracted from different sources, including local and global scales, the model outputs are sensitive to input data. Tum et al. (2012) found that the uncertainty level of NPP reaches 36% when changing major input datasets. For instance, this model showed sensitivity to temperature and the selection and accuracy of land cover products. Moreover, the work is undergoing at DLR to convert and adapt the BETHY/DLR model to use a new input data set. The new input dataset includes land cover map from ESA-CCI (https://www.esa-landcover-cci.org/), gap filled Leaf Area Index (LAI) using linear least square regression-based smoothing method, i.e., Savitzky–Golay filter.

Therefore, further investigation is needed to evaluate and compare the model output for these new input datasets to understand the performance of the model at this stage.

## 1.2 Objectives

The internship task aims to evaluate the BETHY/DLR model output from new input variables by comparing them to the output of the other ecosystem models.

The specific tasks involved to achieve the objective includes:

I. Literature review to identify the available GPP and NPP datasets produced from different ecosystem models and field measurements.
II. Selecting suitable evaluation datasets and processing the data to make the analysis ready.
III. Generating sample points in different geographic regions and land cover types.
IV. Preparing scripts using Python to create the data analysis environment. Then, data analysis using the scripts to estimate the evaluation statistics.
V. Preparing the deliverables.

## 1.3 Deliverables

The deliverables of this internship project includes:

I. Jupyter notebooks of the Python scripts used for data analysis.
II. The data derived from other ecosystem models were used to evaluate the BETHY outputs.
III. The sample points to extract data from BETHY and reference models covering different geographic regions and land covers.
IV. A presentation file containing case studies of the patterns of the relations between BETHY and other models. Also, a report will be prepared to submit to the university.

# 2 Data and Methods

## 2.1 Data

Several potential data sources have been identified, which can be used to compare with the BETHY model outputs. These datasets included estimated GPP and NPP from other models and measured GPP and NPP from eddy covariance flux tower and forest inventory. The datasets can be classified into four groups: remote sensing-based datasets, vegetation modelling-based data, field-measurement-based data, and eddy covariance flux tower-based data. The remote sensing-based datasets include different Light Use Efficiency (LUE)-based model-generated data (i.e., MODIS products). On the other hand, vegetation model-based data means the NPP estimated by process-based dynamic global vegetation models (i.e., LPJ-GUESS data). However, the process-based model may also use input variables from remote sensing data. The field-measured data are mainly the forest inventory data collected from the sample plots in the forest (i.e., Harvard forest NPP dataset). The eddy covariance flux tower dataset contains the flux tower-measured GPP, ecosystem respiration, climate data, etc. (i.e., FLUXNET2015 dataset). An overview of some datasets from all the types mentioned above is available in Annex I.

**Table 1.** The properties of datasets used to compare with the BETHY data.

| No. | Model | Variables | Timestep | Period | Resolution | Spatial coverage | Source |
|---|---|---|---|---|---|---|---|
| 1. | LPJ-GUESS | GPP | Hourly | 2017-2021 | 0.5 degree | Europe | ICOS |
| 2. | CASA (N18) | NPP | Monthly | 1985-2015 | 1 km | China | (P F, 2019) |
| 3. | CASA (Global) | NPP | Monthly | 2003-2017 | 0.5 degree | Global | NASA-GES DISC |
| 4. | MOD17 (MODIS) | GPP | 8-day | 2016-2021 | 500 m | Europe | NASA-LPDAAC |
| 5. | MOD17 (MODIS) | NPP | Annual | 2014-2022 | 500 m | West Africa, Europe | NASA-LPDAAC |

Note: ICOS = Integrated Carbon Observation System; NASA-GES = NASA Goddard Earth Sciences (GES) Data and Information Services Center (DISC); NASA-LPDAAC = Land Processes Distributed Active Archive Center.

**Table 2.** The properties of BETHY data used to compare with other models' data.

| No. | Model | Variables | Timestep | Period | Resolution | Spatial coverage | Source |
|---|---|---|---|---|---|---|---|
| 1. | BETHY/DLR | NPP | 10-day, monthly, annual | 2014-2022 | 1 km | Europe, West Africa, China | DLR |
| 2. | BETHY/DLR | GPP | 10-day, monthly, annual | 2014-2022 | 1 km | Europe, West Africa, China | DLR |

However, a few datasets were finally collected to evaluate the BETHY output (Table 1). These datasets contain GPP and NPP from different ecosystem models, such as LPJ-GUESS ((Sitch et al., 2003), CASA (Potter et al., 1993), and MODIS Products (Running & Zhao, 2019). The datasets vary in timesteps, units, spatial resolution, spatial coverage, and temporal extent. Apart from the evaluation data, two sets of data, i.e., GPP and NPP from BETHY, have been generated to perform the evaluation. The BETHY data contains decadal, monthly, and annual timesteps with a 1 km spatial resolution (Table 2). Although flux tower data was collected, we could not use it for evaluation as BETHY data started from 2014, whereas flux tower data was mainly before 2014.

## 2.2 Methods

### 2.2.1 Data pre-processing and harmonization

The collected data comes with different units and timesteps. The data format is also different; for instance, LPJ and CASA data are in NetCDF format, MODIS data are in HDF format, and CASA N18 and BETHY data are in GeoTIFF format. Moreover, the spatial reference system is different in all the datasets. The BETHY, CASA Global, and LPJ datasets have WGS 1984 spatial reference system. On the other hand, the MODIS dataset has MODIS sinusoidal reference system. Therefore, the data has undergone pre-processing work to make them compatible to compare with the BETHY data.

The LPJ dataset provides hourly average GPP per second timestep μmolC/m² unit. This data was converted to 10-day or decadal total GPP in gC/m² unit to match the BETHY 10-day data. The MODIS 8-day GPP has been converted from the monthly GPP to the monthly GPP of BETHY. MODIS annual NPP and BETHY annual timestep are the same, but the unit was harmonized for MODIS 8-day GPP and MODIS annual NPP. The CASA Global and CASA N18 datasets were in monthly timestep and kgC/m² units.

### 2.2.2 Sample point generation

Sample points have been generated from Europe, West Africa, and China to perform the time series comparison between BETHY and reference model outputs. Several sets of sample points have been generated for each reference model dataset since the spatial resolution varies with the datasets. A simple random sampling technique has been applied with the minimum distance between two points more than the pixel size to generate the sample points in GIS software. Then, manual adjustment was performed to relocate the point in a homogenous area to minimize the mixed land cover originated error. The number of sample points is 30 for LPJ in Europe, 20 for CASA Global and MODIS in Africa, 11 for MODIS in Europe, and 12 for CASA Global and CASA N18 in China. The coordinates of the sample points are available in Annex II. Although the coordinates are provided in WGS 1984 reference system they have been transformed to MODIS Sinusoidal and Krasovsky 1940 Albers when applied with MODIS and CASA (N18 in China) datasets.

### 2.2.3 Evaluation statistics and pattern identification

The time series extracted from BETHY and reference models in different sample points have been compared to estimate the Coefficient of Determination ($R^2$) and Root Mean Square Error (RMSE). The $R^2$ is a measure of correlation, which shows the dependency of dependent variables on the independent variables. In other words, it expresses how much of one variable can be explained by another. It can also be expressed as a percentage. On the other hand, by computing the square root of the mean of the squared differences between estimated and reference values, the RMSE expresses the level of accuracy or inaccuracy in a predictive model or estimator. The GPP and NPP from other models serve as the reference values in this situation. A smaller RMSE suggests that the estimates are nearer the reference values, indicating a more accurate model. A greater RMSE, on the other hand, denotes less accuracy and more serious estimating errors.

### 2.2.4 Computing environment

GIS software, such as QGIS and ArcGIS Desktop, have been used to generate the sample points and some of the image processing. Several Python scripts have been prepared and used to read, subset, pre-process, harmonize, and analyze the raster data of different models in different formats, i.e., GeoTIFF, NetCDF, and HDF. The flowchart in Figure 1 shows the general workflow of the assignment, including the data sources, data, and analysis steps.

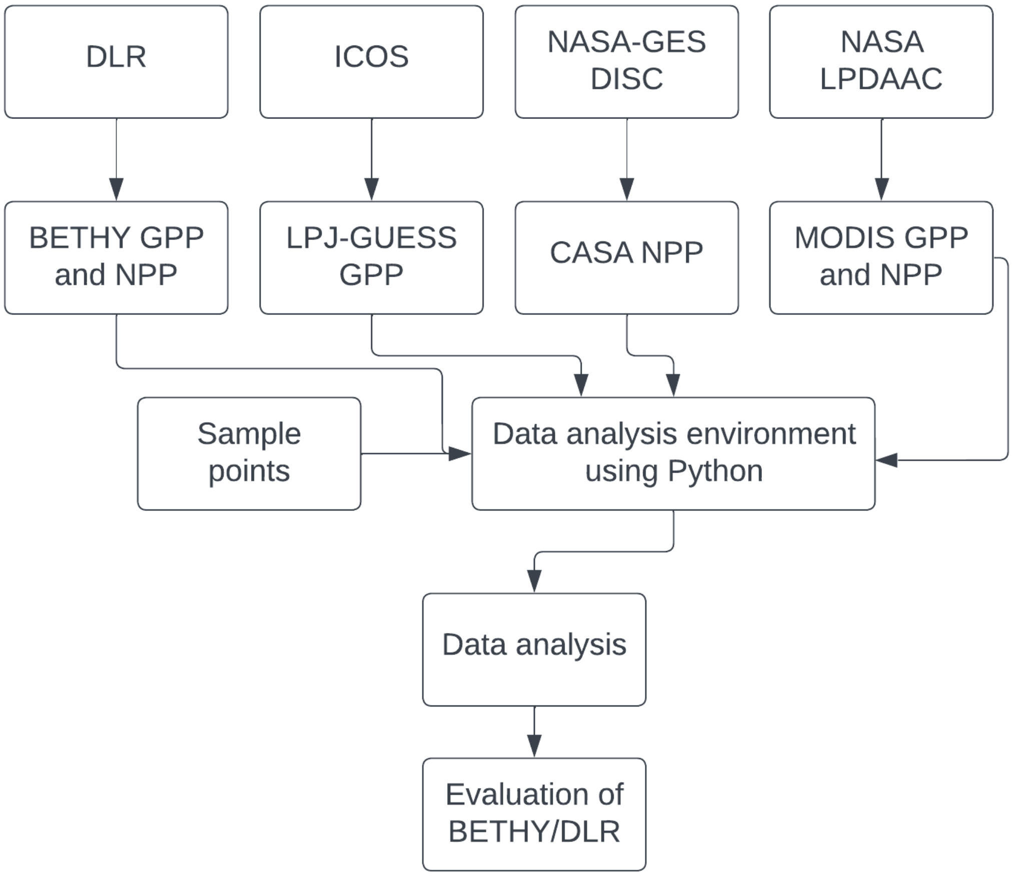


**Figure 1.** The flowchart shows the workflow of the assignment.

# 3 Results

### 3.1.1 Comparison of BETHY with other models

The frequency distribution of $R^2$ and RMSE values between BETHY and other models is given in Figure 2 and Figure 3. The results show that most of the $R^2$ values of BETHY with LPJ and CASA are above 0.5, suggesting a strong relation. The RMSE between BETHY and LPJ decadal GPP shows that the error ranges from 20 to 63 gC/m$^2$/10-day; however, most values are between 25 and 36 gC/m$^2$/10-day. As for the monthly NPP of BETHY and CASA, most values range from 22 to 57 gC/m$^2$/month. In the monthly GPP of BETHY and MODIS, very good $R^2$ is found as the values range from 0.7 to 0.94. However, the RMSE is higher in this case. The annual NPP of BETHY matches poorly with the annual NPP of MODIS, as the $R^2$ values are mostly at the lower quartile in the histogram. The RMSE also shows an irregular distribution pattern.

### 3.1.2 Patterns of the relations between BETHY and other models

Several patterns have been identified in the relationship between BETHY and other models. Overall, five patterns are found in every model comparison: good fit, BETHY overestimated, BETHY underestimated, leftward bias, and irregular pattern. These patterns are present in every comparison, such as BETHY vs LPJ-GUESS, BETHY vs CASA, and BETHY vs MODIS. The patterns are identified from the visual interpretation of the time series plot, $R^2$, and RMSE values. The good fit has been determined based on the lowest RMSE value in each comparison. On the other hand, the RMSE with lowest $R^2$ indicates irregular pattern. Figure 4 to Figure 10 shows a few example plots of different patterns.

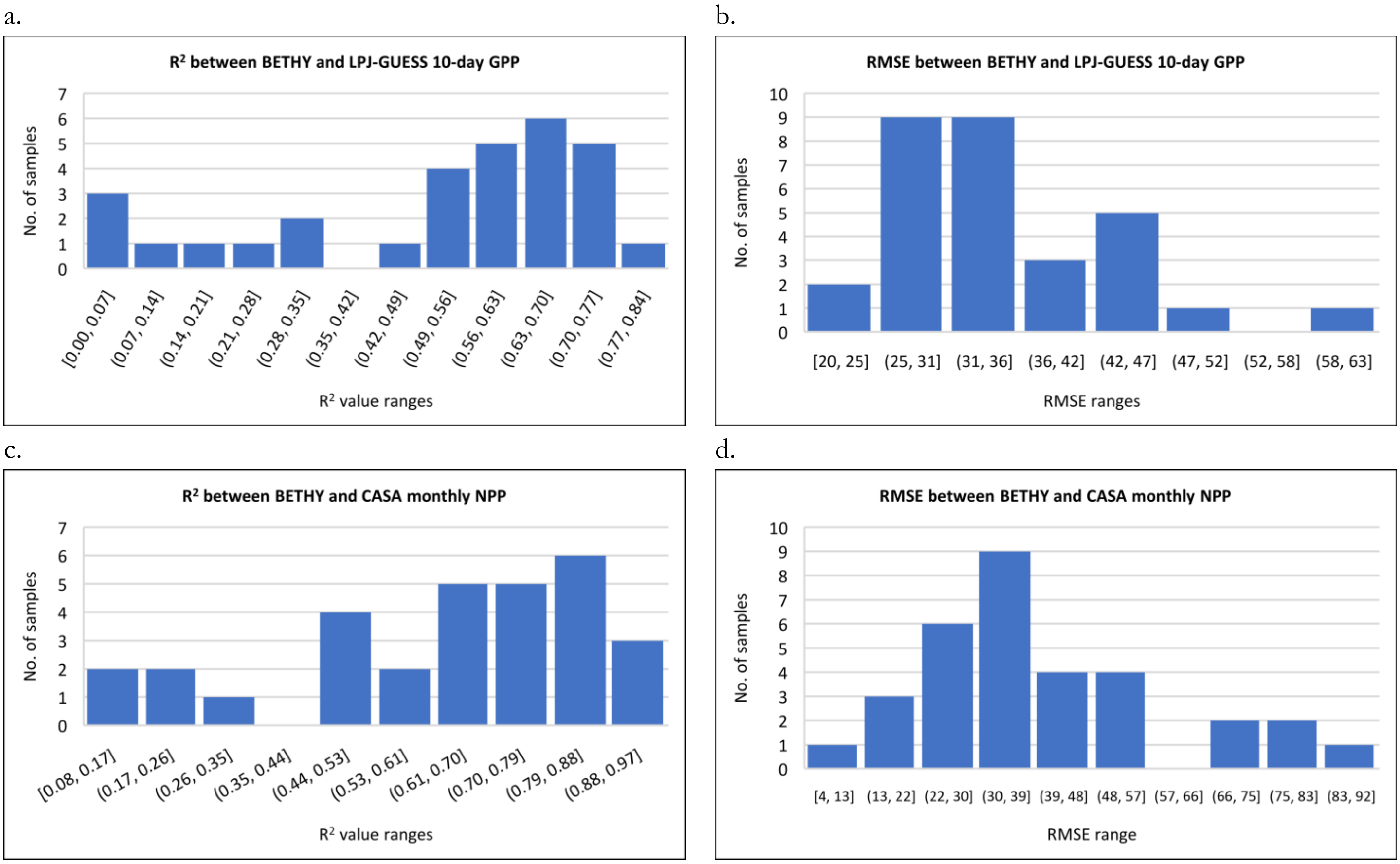


**Figure 2.** Frequency distribution of the evaluation statistics between BETHY and other models; a. $R^2$ of BETHY and LPJ GPP (30 sample points in Europe); b. RMSE of BETHY and LPJ GPP (30 points in Europe); c. $R^2$ of BETHY and CASA NPP (20 and 12 points from West Africa and China); d. RMSE of BETHY and CASA NPP (20 and 12 points from West Africa and China). The x-axis shows the different range (bins) of $R^2$ and RMSE. The y-axis shows the no. of sample points in each range/bin. The unit for RMSE on x-axis is $gC/m^2$/timestep.

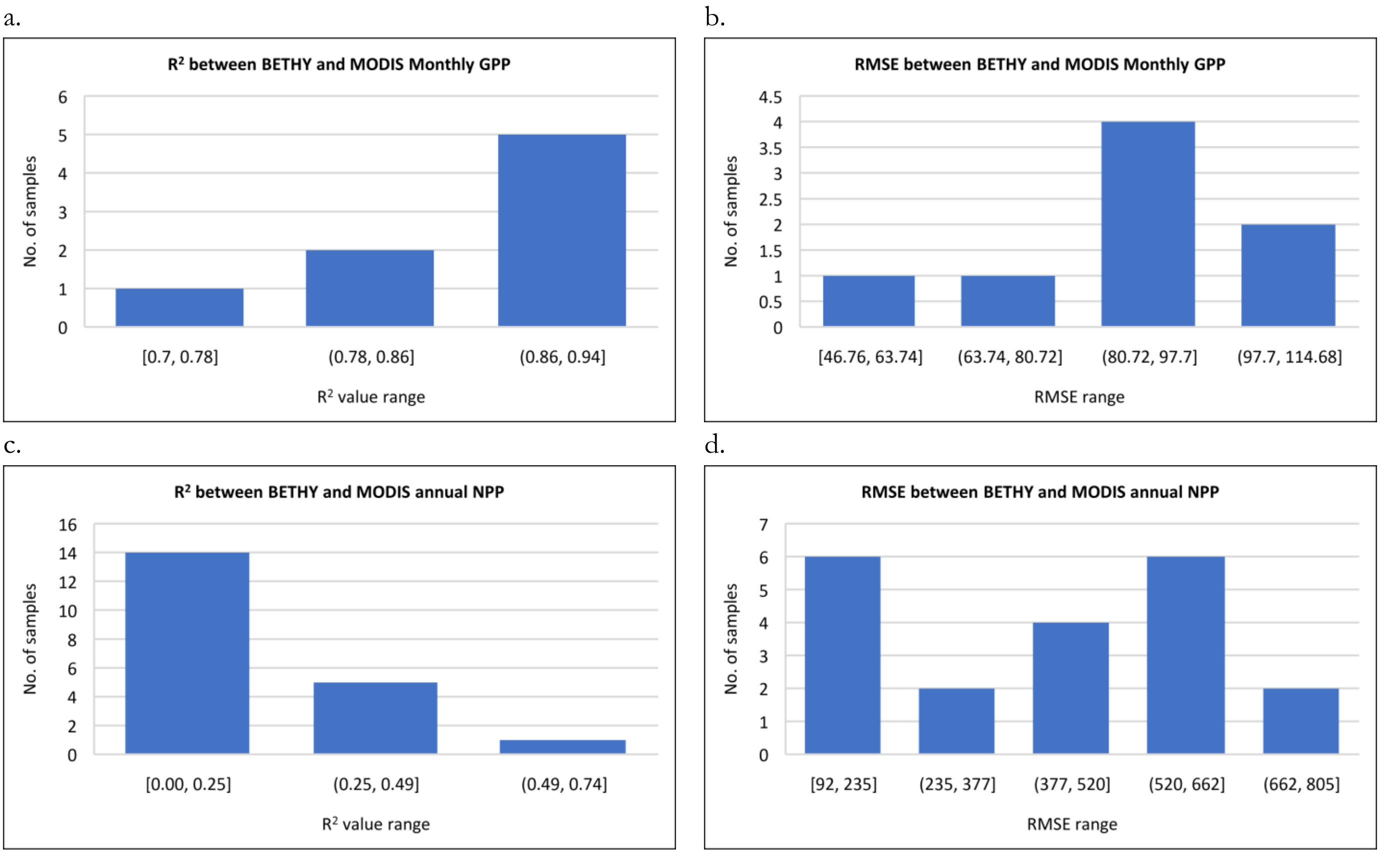


**Figure 3.** Frequency distribution of the evaluation statistics between BETHY and other models. a. $R^2$ of BETHY and MODIS monthly GPP (11 points in Europe, 4 points were omitted due to no data); b. RMSE of BETHY and monthly MODIS GPP (11 points in Europe); c. $R^2$ of BETHY and MODIS annual NPP (20 points in Europe); d. RMSE of BETHY and MODIS annual NPP (20 points in Europe). The x-axis shows the different range (bins) of $R^2$ and RMSE. The y-axis shows the no. of sample points in each range/bin. The unit for RMSE on x-axis is gC/m²/timestep.

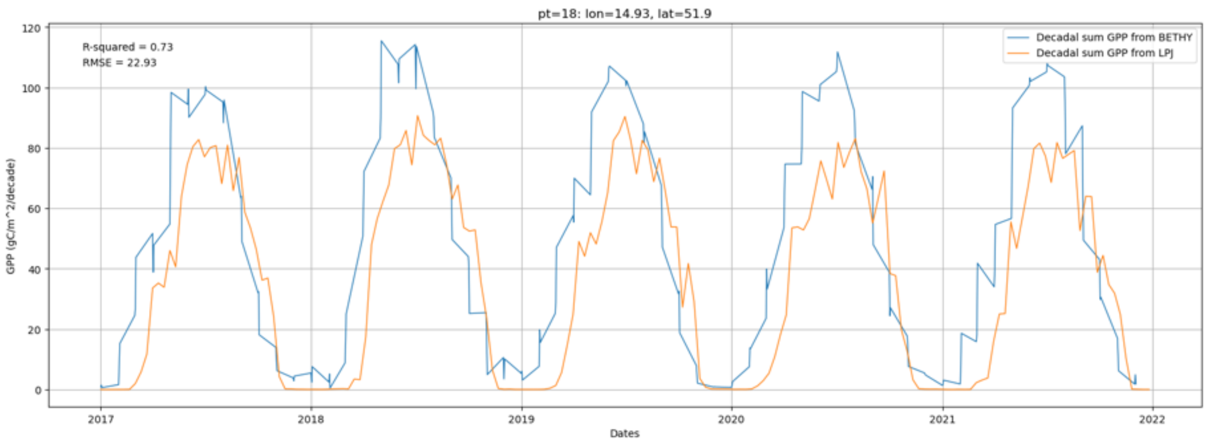


**Figure 4.** An example of good fit between BETHY and LPJ-GUESS decadal GPP with high $R^2$ and low RMSE.

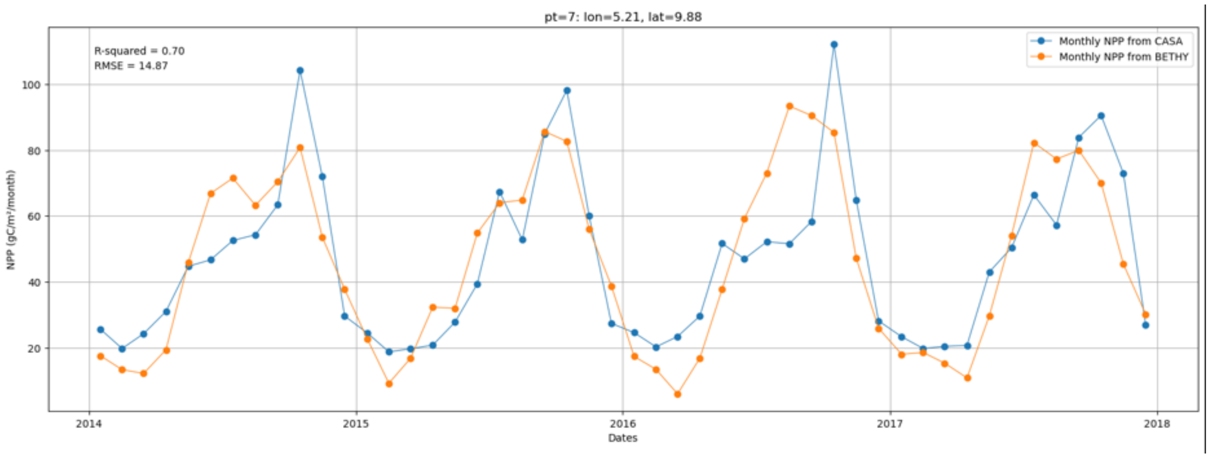


**Figure 5.** An example plot of good fit between monthly NPP from BETHY and CASA with high $R^2$ and low RMSE.

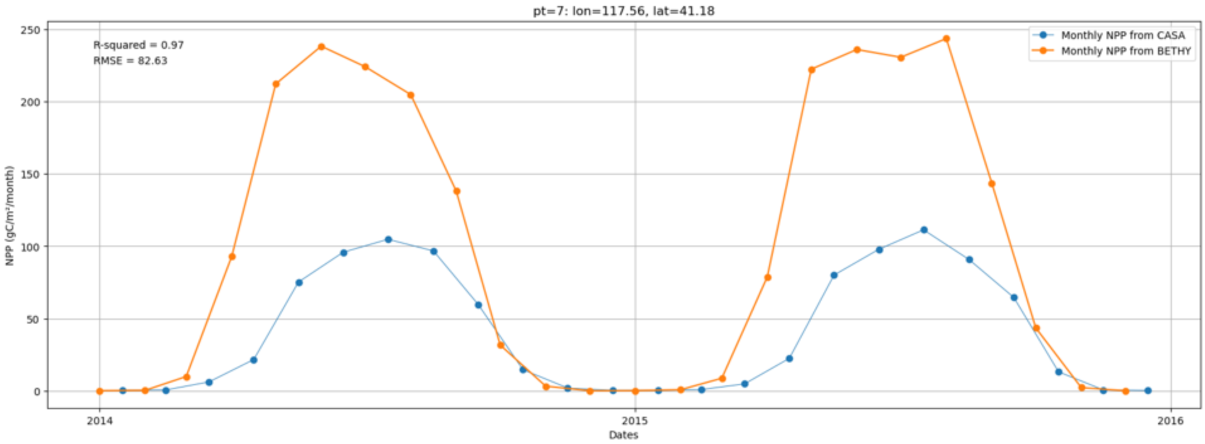


**Figure 6.** An example of overestimation of BETHY compared to CASA monthly NPP. $R^2$ is higher since both models follow the same pattern, but higher RMSE indicates the difference in values.

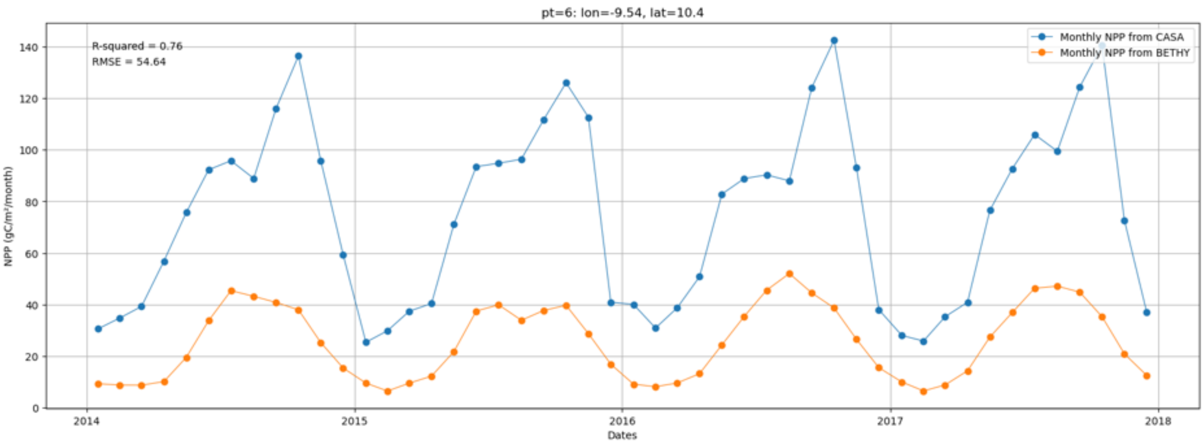


**Figure 7.** An example of underestimation of BETHY compared to monthly NPP of CASA.

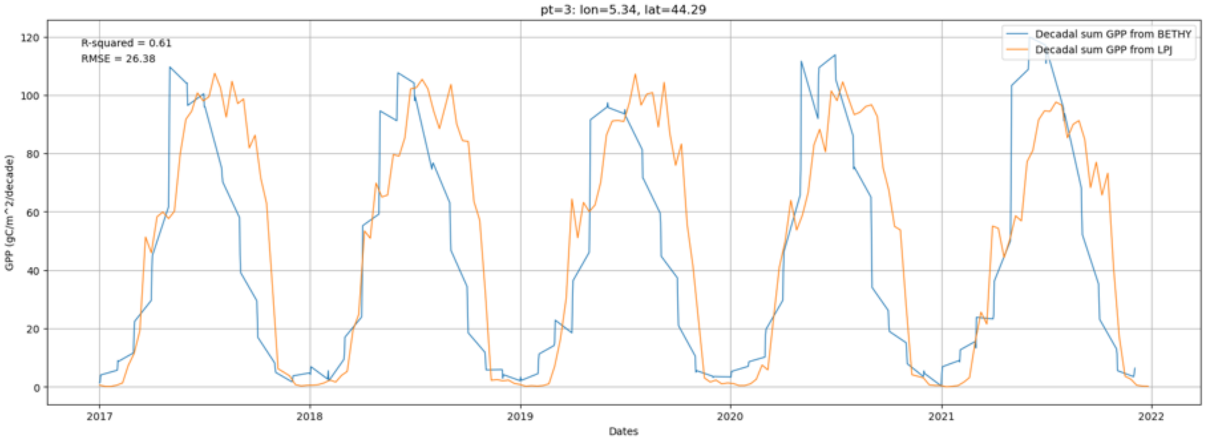


**Figure 8.** An example plot of leftward bias of BETHY decadal GPP compared to LPJ-GUESS.

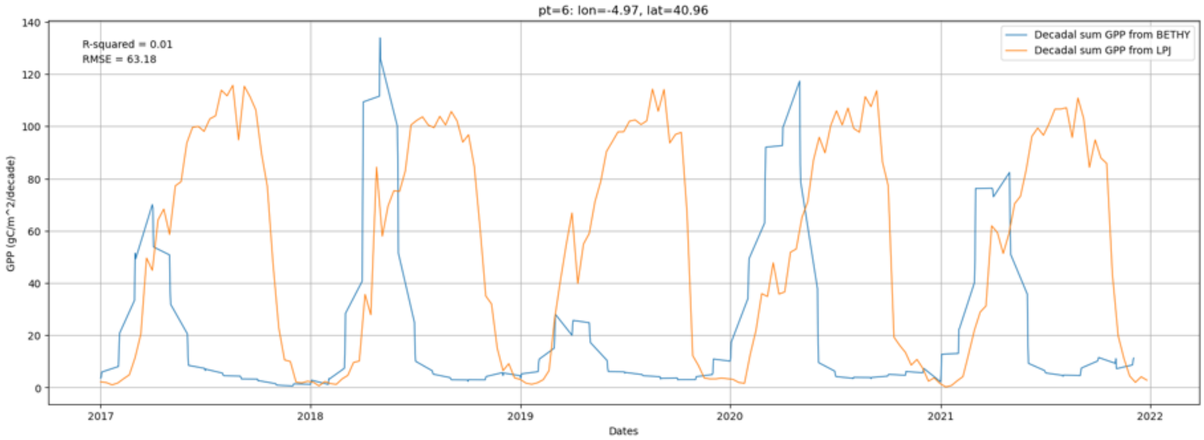


**Figure 9.** An example of irregular pattern between BETHY and LPJ-GUESS decadal GPP.

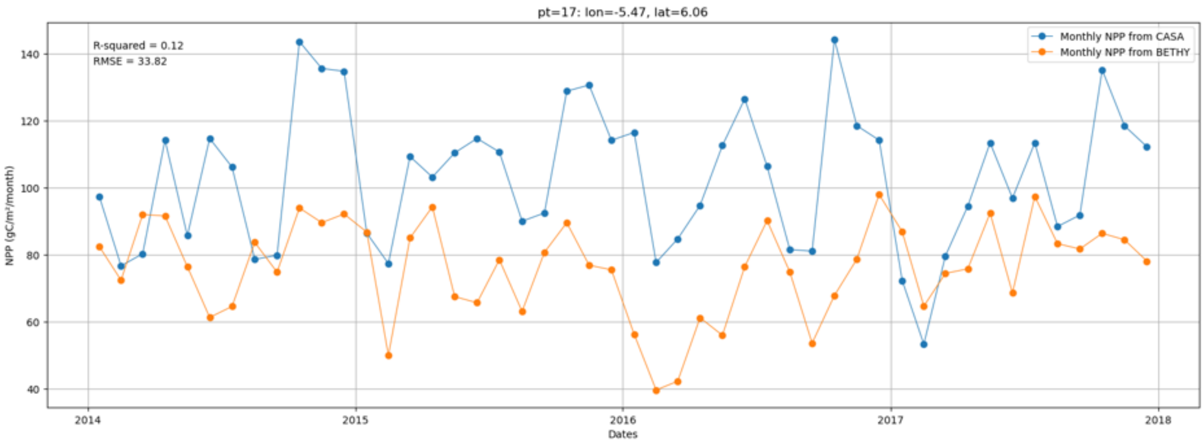


**Figure 10.** Another example of irregular pattern between BETHY and CASA monthly NPP.

# 4 Discussion and Conclusion

Since LPJ-GUESS and CASA are process-based models, a good fit of BETHY with these models indicates its compatibility. On the other hand, poor relations with MODIS annual NPP require further investigations. The LPJ-GUESS and CASA have a very coarse spatial resolution (i.e., about 55 km). Therefore, the pixel-wise GPP and NPP values represent a generalized picture containing different land cover classes. On the other hand, BETHY has a 1 km spatial resolution, which can simulate NPP preserving the homogenous characteristics of a land cover class more than LPJ and CASA. Moreover, there are differences in input variables and algorithms across the models. For instance, CASA uses the Normalized Difference Vegetation Index (NDVI) to estimate LUE, whereas BETHY uses Leaf Area Index (LAI). These differences contribute to the distortion in model outputs. Furthermore, the model's accuracy largely depends on the accuracy of the input variables. Notably, the accuracy of the land cover map is critical since it determines the use of vegetation-specific parameters in the model algorithm. Since the ESA-CCI land cover map used in this model was resampled to 1 km spatial resolution from 300 m, it may have generated the error. This error affects the comparison result with another model. Therefore, to get a better picture of the model's performance, it is required to compare with the field-measured data, like eddy covariance flux tower data and forest inventory data. GPP and NPP can also be compared with crop production data, especially for the agricultural land cover type.

In conclusion, BETHY generates promising outputs considering the evaluation results against LPJ-GUESS GPP, CASA NPP, and MODIS GPP. Although further verifications are required using the field-measured data, we can understand the compatibility of BETHY at this stage based on this analysis. The case studies help understand what land cover BETHY is performing well on the other and where it is performing poorly. Based on this analysis, further calibration of the parameters and validation of the input variables may reduce the error and produce a more accurate output.

# 5 Reflection

Working and learning at DLR was an incredible experience, with many exciting works around the campus. The work environment at DLR was very exciting and encouraging. All my co-interns and other colleagues were very welcoming and supportive. I enjoyed my time there working at the office, going to lunch together, after-lunch walks, occasional coffee breaks and parties. I have learned many things from my work and the environment there. Overall, working at DLR was a great experience that opened many horizons in front of me. In this section, I discussed some reflections on my internship experience.

### Learning process

The internship project involved learning new technical and soft skills in a multinational and multicultural environment. The task required geospatial data processing and analysis with Python programming. This work helped me to learn coding with appropriate Python libraries. Moreover, GIS software has also been used. As for the soft skills, the internship provided opportunities for learning through working in an official environment, professional communication with supervisors and colleagues, time management for deadline meet up, and attending seminars and office parties.

Additionally, the internship improved my ability to solve problems. Throughout the internship, I had a variety of difficulties and impediments that forced me to use my critical thinking and come up with creative solutions. Working with my colleagues and asking for advice from mentors improved my problem-solving skills and helped me approach problems from several angles.

### Realization process

My internship involves many duties and deliverables. Understanding needs, dividing large projects into simple parts, and systematically executing them were the realization process. I set reasonable goals and milestones to keep on track and deliver results. I also learned about the tools and resources available through the organization during the realization process. My effectiveness increased as a result of adopting their procedures and making use of their tools and resources. My supervisor's feedback helped me find areas for improvement and gave me helpful advice for getting better results.

### Communication process

Meetings, emails, and presentations conducted the communications. Two presentation sessions were organized, in which DLR and ITC supervisors participated. One presentation was at the internship's beginning, where the task's objective and process were communicated. At the final presentation after the internship's end, the work's final output and its implications were presented. Furthermore, regular emailing and in-person meetings were the primary mediums of communication.

### Self-assessment reflection

Throughout the internship, I had the opportunity to reflect on my strengths, values, and interests in the professional work environment. I understood my capability in critical thinking and problem-solving. I thoroughly enjoyed analyzing the problems, identifying potential solutions, and implementing effective strategies. Additionally, I value the mentorship approach and the opportunity to collaborate with colleagues, leveraging their diverse perspectives and knowledge to achieve common goals.

My professional interests focus on learning new things and developing new ways to do things. The internship experience made me even more interested in keeping up with trends and new developments. I enjoy trying new tools and looking for ways to grow and improve myself.

## References

Knorr, W. (1997). *Remote Sensing and Modelling of the Global CO2 Exchange of Land Vegetation: A Synthesis Study*. Max-Planck-Institut für Meteorologie.

P F, C. (2019). Monthly NPP Dataset Covering China's Terrestrial Ecosystems at North of 18°N (1985–2015). *Journal of Global Change Data & Discovery*, *3*(1), 34–41. https://doi.org/10.3974/geodp.2019.01.05

Potter, C. S., Randerson, J. T., Field, C. B., Matson, P. A., Vitousek, P. M., Mooney, H. A., & Klooster, S. A. (1993). Terrestrial ecosystem production: A process model based on global satellite and surface data. *Global Biogeochemical Cycles*, *7*(4), 811–841. https://doi.org/10.1029/93GB02725

Running, S. W., & Zhao, M. (2019). *User's Guide Daily GPP and Annual NPP (MOD17A2H/A3H) and Year-end Gap-Filled (MOD17A2HGF/A3HGF) Products NASA Earth Observing System MODIS Land Algorithm (For Collection 6)*.

Sitch, S., Smith, B., Prentice, I. C., Arneth, A., Bondeau, A., Cramer, W., Kaplan, J. O., Levis, S., Lucht, W., Sykes, M. T., Thonicke, K., & Venevsky, S. (2003). Evaluation of ecosystem dynamics, plant geography and terrestrial carbon cycling in the LPJ dynamic global vegetation model. *Global Change Biology*, *9*(2), 161–185. https://doi.org/10.1046/j.1365-2486.2003.00569.x

Tum, M., Strauss, F., McCallum, I., Günther, K., & Schmid, E. (2012). How sensitive are estimates of carbon fixation in agricultural models to input data? *Carbon Balance and Management*, *7*(1), 3. https://doi.org/10.1186/1750-0680-7-3

# Annex I

An overview of different GPP and NPP dataset reviewed during the literature review.

| No. | Dataset | Method | Period | Format | Source |
|---|---|---|---|---|---|
| 1. | MODIS NPP | Remote sensing-based models (MODIS) | 2001-2022 | Raster | NASA Earth Explorer / Google Earth Engine |
| 2. | GIMMS NPP | Remote sensing-based models (AVHRR) | 1982-2015 | Raster | GIMMS Website |
| 3. | VPM NPP | Remote sensing-based models | 2010-Present | Raster | NASA Earth Observation Website |
| 4. | N18 NPP | CASA model using MODIS and AVHRR dataset | 1985-2015 | Raster | https://doi.org/10.3974/geodb.2019.03.02.V1 |
| 5. | Harvard Forest NPP Dataset | Field measurement of biomass data | Various periods | TXT | https://harvardforest1.fas.harvard.edu/exist/apps/datasets/da-core.html?core=Primary%20Production |
| 6. | Saatchi NPP Dataset | Field measurement-based statistical estimation | 1990-2005 | TXT/ CSV | For cropland: https://cdiac.ess-dive.lbl.gov/carbonmanagement/cropcarbon/ |
| 7. | ORNL DAAC NPP | Field measurement and model outputs | Various years | TXT/ CSV/ GRID | https://daac.ornl.gov/ |
| 8. | CASA NPP Dataset | CASA Model Output | 2003-2018 | Raster | https://disc.gsfc.nasa.gov/datasets/GEOS_CASAGFED_M_FLUX_3/summary |
| 9. | LPJ NPP Dataset | LPJ-GUESS model output | 2017-2021 | Raster | https://www.icos-cp.eu/data-products/global-hourly-nee-gpp-and-total-respiration-2017-2021-based-lpj-guess-generated-2022 |
| 10. | DLEM NPP Dataset | Eddy covariance flux tower-based estimation | 1991-Present | TXT/CSV | https://fluxnet.org/data/fluxnet2015-dataset/ |

# Annex II

**Coordinates of 30 sample points in Europe for BETHY vs LPJ-GUESS decadal GPP:**

[[-5.75575452086703, 40.0080542529708, 1], [23.0736672527159, 52.4380396101478, 2], [5.33836956080819, 44.2870217296024, 3], [9.07495412722206, 50.6154094941613, 4], [7.83269818253244, 46.7281946887473, 5], [-4.97496947775681, 40.9596352322487, 6], [16.0824156190516, 53.1845674676702, 7], [22.1295333040348, 52.7160212534025, 8], [24.397939799647, 50.5051866827051, 9], [19.3433119990728, 49.0261763930267, 10], [30.4766429476738, 47.422969273992, 11], [17.1367090181793, 47.9896405710419, 12], [20.8123339577545, 49.4805377877087, 13], [3.39606317082162, 49.3321378562387, 14], [-0.167748595939941, 52.2797426890301, 15], [25.8081700035765, 47.6766160656794, 16], [-3.06005616404389, 42.6399522976834, 17], [14.9343141338962, 51.8953990318221, 18], [1.35931930878426, 47.8654403209983, 19], [1.5875203192093, 44.5245072398962, 20], [-3.22102691973168, 41.549969754164, 21], [9.92328200605477, 48.9559424247788, 22], [20.9823442174107, 48.018469830189, 23], [9.77329719848922, 45.8287724570193, 24], [20.7538792027791, 50.1206656146583, 25], [8.02444126831328, 48.1894516253384, 26], [17.3237845096236, 48.8197204792496, 27], [26.6417643357565, 55.9594206211683, 28], [24.9821451999522, 53.4318570381093, 29], [-2.27715908729653, 54.0076016316366, 30]]

**Coordinates of 20 sample points in West Africa for BETHY vs CASA-Global monthly NPP:**

[[4.5701272247065, 7.220695211232735, 1], [9.357573592279554, 5.073654343638358, 2], [-10.94496824232343, 11.61656891756962, 3], [1.739243453566037, 7.147101390054359, 4], [-9.637508356390425, 8.228369064226714, 5], [-9.54059449382756, 10.40335081372953, 6], [5.205059220908026, 9.875697479570679, 7], [9.29524121774936, 9.441090757882632, 8], [-6.528850135684388, 8.386483061610306, 9], [6.745777860699571, 6.690710728996129, 10], [-7.03559479766413, 9.722922117727133, 11], [9.54289959487523, 6.726958926303201, 12], [7.174084260636708, 7.223491849598336, 13], [-10.46360753345283, 6.589874348423368, 14], [-4.852543886473764, 9.909056951249003, 15], [6.347281371756131, 8.082578726883579, 16], [-5.468118559331866, 6.063670431948502, 17], [8.152364520933357, 9.206329248625801, 18], [-14.14543543627267, 10.65916249924607, 19], [0.2824384410115541, 9.20879215584487, 20]]

**Coordinates of 11 sample points in Europe for BETHY vs MODIS monthly GPP:**

[[0.820685449333405, 40.6866726857909, 1], [13.2595246143118, 48.9723894991341, 2], [2.66032824537647, 43.1348559827938, 3], [12.469295460795, 49.3597431193533, 4], [2.6093520062252, 47.285305172961, 5], [3.90630123026076, 44.0907950024944, 6], [11.4880118799735, 45.7763424133886, 7], [6.97391067018009, 48.4364575134944, 8], [1.14057159601772, 48.6425041937887, 9], [0.087512325172125, 44.624098473579, 10], [0.130834789874114, 43.2790690978156, 11]]

**Coordinates of 20 sample points in West Africa for BETHY vs MODIS annual NPP:**

[[-9.47281151679067, 6.232490322028184, 1], [-8.583595765440862, 8.81748224098692, 2], [-1.277576338316919, 7.559697246983262, 3], [-3.384457778877277, 6.850097208857221, 4], [-6.95629775528721, 4.657527079761337, 5], [-0.8241741579938093, 8.18200502955483, 6], [-1.278823111920815, 6.063312395682885, 7], [-3.618438474488471, 8.962415514767372, 8], [-6.628568188139752, 7.316886982320852, 9], [-5.579839073848461, 8.354121513597761, 10], [-4.518467893634806, 6.559856047677442, 11], [-6.987060920304202, 8.892772687744753, 12], [-4.477845926848488, 5.738317066654076, 13], [-4.764359204160542, 5.344498382521124, 14], [-7.930301688577458, 7.284811123734585, 15], [-7.793341903499879, 7.919603166496074, 16], [-8.443012485441955, 7.089629697020996, 17], [-4.230690508485964, 9.882168309426163, 18], [-2.112152524489943, 7.824982572192482, 19], [-6.664989764114921, 7.062191458098141, 20]]

**Coordinates of 12 sample points in China for BETHY vs CASA (Global and N18 dataset) monthly NPP:**

[[106.625803653902, 33.5699551390307, 1], [114.232288245745, 31.1528626271209, 2], [103.60805116092, 27.7607553750937, 3], [131.860530335124, 45.8344835525083, 4], [88.8478272315289, 31.2151124183236, 5], [79.3268532564864, 37.2858779194913, 6], [117.557000754284, 41.1823859423703, 7], [116.757531111798, 35.9353469271634, 8], [83.6076077059844, 43.7591290261723, 9], [110.220217248597, 26.040911393937, 10], [122.664836206656, 50.2197856686591, 11], [118.186564830561, 26.7485017960792, 12]]