Automated Gross Primary Production Application for Monitoring Ecosystem Health Within GEOSS

Spinosa, Anna; Eleveld, Marieke; Mallast, Ulf; Peterseil, Johannes; Mobilia, Valeria; Karisma, Karisma; Fuentes-Monjaraz, Mario Alberto; El Serafy, Ghada

DOI
10.1109/IGARSS53475.2024.10642481

Publication date
2024

Document Version
Final published version

Citation (APA)
Spinosa, A., Eleveld, M., Mallast, U., Peterseil, J., Mobilia, V., Karisma, K., Fuentes-Monjaraz, M. A., & El Serafy, G. (2024). *Automated Gross Primary Production Application for Monitoring Ecosystem Health Within GEOSS*. 4544-4547. Paper presented at 2024 IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2024, Athens, Greece. https://doi.org/10.1109/IGARSS53475.2024.10642481

IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium | 979-8-3503-6032-5/24/$31.00 ©2024 IEEE | DOI: 10.1109/IGARSS53475.2024.10642481

AUTOMATED GROSS PRIMARY PRODUCTION APPLICATION FOR MONITORING ECOSYSTEM HEALTH WITHIN GEOSS

Anna Spinosa[1,2], Marieke Eleveld[1,2], Ulf Mallast[3], Johannes Peterseil[4], Valeria Mobilia[1], Karisma Karisma[5], Mario Alberto Fuentes-Monjaraz[1], Ghada El Serafy[1,2]

[1] Deltares, Delft, The Netherlands
[2] Delft University of Technology, Delft, The Netherlands
[3] Environmental Research Centre Leipzig-Halle (UFZ), Leipzig, Germany
[4] Umweltbundesamt GmbH, Vienna, Austria
[5]University of Twente, Enschede, The Netherlands

ABSTRACT

This study addresses the challenges posed by climatic changes and biodiversity loss to ecosystem stability, by quantifying gross primary production (GPP) changes. An improved earth observation product is obtained by integrating in-situ and remote sensing data via data-driven models. Employing a user-centered strategy, our methodology builds on users' engagement, ensuring both the identification of user needs and practical product demonstrations. With GEOSS as central and integrated stakeholder, we strive for a broad interoperability and accessibility of generated outcomes. The project outcomes include a curated dataset with FAIR metadata, openly available code, and reports for reproducibility, contributing to the broader Earth Intelligence supply chain.

Index Terms— Gross Primary Production, Ecosystems, Earth Observation, Earth Intelligence, GEOSS

1. INTRODUCTION

Changes in climatic conditions and loss of biodiversity pose significant challenges to the stability of ecosystem biomass production, ultimately affecting its role as carbon sing and the ecosystem services it provides [1, 2]. The accurate quantification of gross primary production (GPP), defined as the total carbon fixed by the ecosystems through vegetation photosynthesis, and its dynamic spatiotemporal changes is an important prerequisite for ecosystem function assessment and carbon balance research [3].

The request for an improved remote sensing -- in-situ nexus has been repeatedly raised as Group of Earth Observation (GEO) Work Programme priority. This emphasis reflects the recognition that a more robust integration of remote sensing technologies and ground-based observations is essential for advancing our understanding of Earth's complex systems. This study aims at integrating in-situ measurements with Sentinel-2 data, enhancing our capacity to discern changes and trends in GPP production at a European scale. Fostering the provision of Earth Intelligence [4], collaboration with the end-users guides the co-design of these products, ensuring the EO chain is fully exploited for the creation of actionable knowledge.

2. MATERIAL AND METHODS

The approach for the derivation of GPP builds upon our previous research [5] on spatio-temporal upscaling of flux tower GPP measurements using open access Copernicus Sentinel-2 data. The high, 10-60 m spatial resolution of the Sentinel-2 multi-spectral instrument (MSI) promises to enhance the estimation of GPP at a local scale [6, 7].

The methodology suggested for the derivation of GPP beyond the flux footprint is depicted in Figure 1. The approach involves evaluating various satellite radiometric indices, particularly NDVI, EVI, EV2, CLr, MNDVI, MNDWI, LSWI, and NDII. To account for physiological changes in photosynthesis driven by meteorological conditions which cannot be discerned solely from satellite reflectance values, the methodology integrates time series of environmental variables drivers (EVs) of GPP as predictors. Those include vapour pressure deficit, air temperature, rainfall, relative humidity, global radiation, and precipitation. Forward feature selection generates models based on predictor combinations to find the most relevant predictors for GPP. This provides information on the vegetation indexes (VIs), spectral bands, and environmental variables most relevant for predicting gross primary prediction [8].

In situ 30-minute measurements of net ecosystem exchanges (NEE) and auxiliary data, including the EVs listed above, are collected from the Integrated Carbon Observation System (ICOS) portal, and the FLUX Network (FLUXNET) portal, both open repositories. In some cases, additional environmental variables data sources are provided by eLTER. The temporal coverage of the dataset is site-dependent, and minimum of eight months. The preliminary selected sites for the analysis are Torgnon (grassland, Italy), Hyytiälä (evergreen needleleaf forests, Finland), Tereno Harz/Central Lowland (deciduous forest, Germany), Doñana National Park (grassland, Spain), and the Wüstebach catchment (deciduous forest, Germany). Those sites are selected to encompass different types of ecosystems, at different latitudes.

The workflow adopts and compares the linear formulation suggested by [9] and used by [6] for the combination of VIs and EVs, and state-of-the-art machine learning-based models. In this study, the models used are gradient-boosted decision trees (GBDT), and long short-term memory (LSTM) neural networks. At first, data collected for the three sites of Torgnon, Hyytiälä, and Tereno Harz are used to train and test the GBDT and LSTM models. Initially, both GBDT and LSTM were applied independently to each one of the three sites. Following this step, a preliminary training phase was conducted, encompassing three different sites. In the latter phase, both models underwent three training cycles, with each cycle integrating an additional site. GPP is derived at different European long-term observation (eLTER) sites. The workflow produces high resolution maps with 5-day temporal resolution (same resolution as Sentinel-2 data in cloud-free conditions). In places with high cloud coverage monthly maps are calculated with monthly mean composite maps.

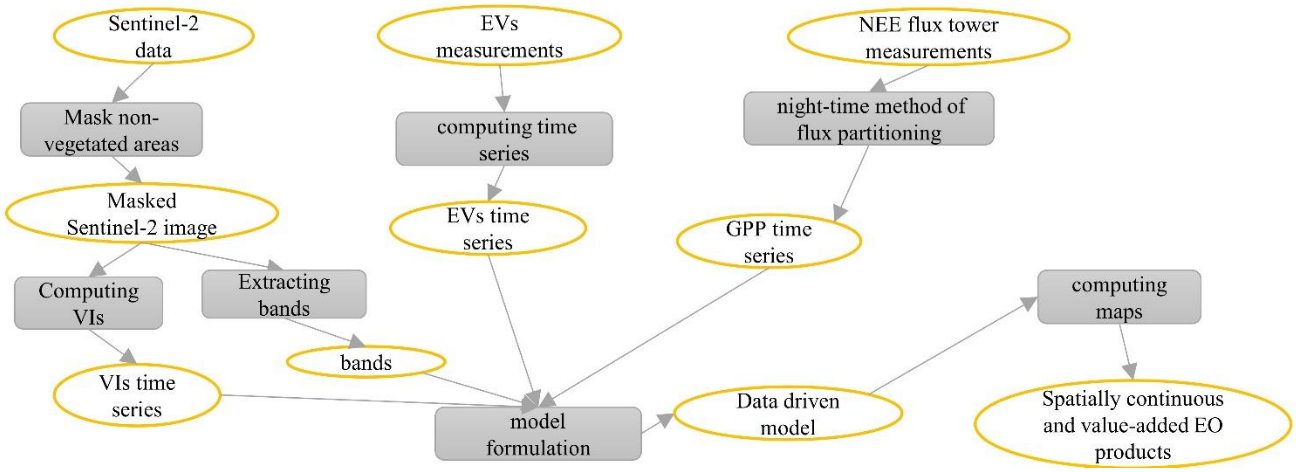

Figure 1: Schematization of the methodology for the derivation of Gross Primary Production (GPP). Abbreviations: VI: Vegetation Index, NEE: Net Ecosystem Exchange, EV: Environmental Variable, EO: Earth Observation.

2.1. AGAME co-design and validation

Harnessing the potential of EO data to offer considerable benefits to a broad spectrum of socio-economic stakeholders is a complex task [9]. Especially in the context of "big data", where new data sources emerge continuously, transforming raw data into valuable and usable information represents a recurring challenge.

The approach we follow in the AGAME project (funded by ESA) builds on an end-user centered approach, ensuring the generation and demonstration of the added value of EO based products stepping into actionable knowledge. We are engaging with different groups of users, each playing a critical role within the GEO community [10]. Our user engagement strategy is structured into two distinct phases: a) co-define/co-design phase, focusing on mapping users' needs and requirements to identify added value products beneficial to the users, and b) user validation phase, allowing for practical demonstrations and real-time feedback on the developed products.

This structured approach ensures not only the dissemination of information about our product, but also the active consultation and involvement of all user groups through a dedicated engagement process.

3. PRELIMINARY RESULTS

When applying the linear formulations to different ecosystem types, different combinations of EV and VI were discovered to exhibit stronger correlations with GPP. For instance, when applied to the wetlands of the Protected Area of Doñana National Park, the workflow resulted in a regression model driven by the Red-edge Index (CLr), affirming the sensitivity of the red-edge to canopy biomass, chlorophyll content, and photosynthesis activity [5, 11-13]. However, when the workflow was applied to the deciduous forest of the Wüstebach catchment, Germany, and the Hyytiälä evergreen forests, Finland, it revealed a higher correlation between Enhanced Vegetation Index (EVI) and GPP. This finding aligns with previous studies conducted in high-biomass vegetated areas, such as dense grass or forest ecosystems [14, 15].

When the GBDT and LSTM models were applied independently to each site, the results indicated the best performance at the Torgnon site, with Mean Squared Error (MSE) values of 0.757 and 0.928 for the GBDT and LSTM models, respectively. This site had the longest data span of 2 years. The Hyytiälä site, with a data span of 1.5 years, followed closely with MSE values of 1.474 and 1.709 for the

GBDT and LSTM models, respectively. The Tereno Harz site in Germany, which had a shorter data span of 11 months, demonstrated lower performance with MSE values of 10.338 and 3.463 for the GBDT and LSTM models, respectively. These preliminary results suggest that a minimum data span of 1.5 years is required for optimal performance of both the GBDT and LSTM models.

These preliminary results demonstrate the significance of incorporating various VIs, however, still further investigation is needed and research in AGAME is aimed at improving the model formulation for a better representation of different ecosystem types. Particularly, more sites are planned to be included in the training cycle of GBDT and LSTM to improve the performance of the algorithms and have more robust models that account for the differences in ecosystems.

4. CONCLUSIONS

We propose a method for bridging the gap between data acquired from satellites and real-time, in-situ measurements, to enhance the accuracy and reliability of environmental monitoring. The paper presents a high-resolution product of GPP. These outputs, enhancing our understanding of the ecosystem's capability to fulfill its ecosystem role, may serve as a cornerstone for evidence-based decision-making and sustainable resource management. The model implementation in the different areas helps us to construct robust solutions for all the different user communities.

We facilitate the supply of Earth Intelligence by engaging with users and completing the EO value-chain to generate actionable knowledge and enable the connections with the data. The project outcome extends beyond a mere data product. It includes a curated dataset and Findable, Accessible, Interoperable, Re-useable (FAIR) metadata, along with openly available code and reports for reproducibility. To ensure the generated products are easily accessible to a wide range of stakeholders and researchers, results are disseminated via the Global Earth Observation System of Systems (GEOSS) portal (https://www.geoportal.org/). The GEOSS platform architecture is realized as a system of systems that collects data from diverse sources and shares independent and open EO information and processing services. Integration of the data products within the GEOSS data infrastructure ensures the findability and interoperability of the results [16].

4. REFERENCES

[1] M. Fernández-Martínez *et al.*, "The role of climate, foliar stoichiometry and plant diversity on ecosystem carbon balance," *Global Change Biology,* vol. 26, no. 12, pp. 7067-7078, 2020, doi: https://doi.org/10.1111/gcb.15385.

[2] S. Piao *et al.*, "Characteristics, drivers and feedbacks of global greening," *Nature Reviews Earth & Environment,* vol. 1, no. 1, pp. 14-27, 2020/01/01 2020, doi: 10.1038/s43017-019-0001-x.

[3] M. Reichstein *et al.*, "Climate extremes and the carbon cycle," *Nature,* vol. 500, no. 7462, pp. 287-295, 2013/08/01 2013, doi: 10.1038/nature12350.

[4] GEO, "GEO Post 2025 strategy: Earth Intelligence for all " 2023. [Online]. Available: https://earthobservations.org/storage/app/media/documents/Key-Documents/GEO%20Post%202025%20Strategy%20Full%20Document.pdf.

[5] A. Spinosa, M. A. Fuentes-Monjaraz, and G. El Serafy, "Assessing the Use of Sentinel-2 Data for Spatio-Temporal Upscaling of Flux Tower Gross Primary Productivity Measurements," *Remote Sensing,* vol. 15, no. 3, p. 562, 2023, doi: https://doi.org/10.3390/rs15030562.

[6] Z. Cai *et al.*, "Modelling Daily Gross Primary Productivity with Sentinel-2 Data in the Nordic Region–Comparison with Data from MODIS," *Remote Sensing,* vol. 13, no. 3, p. 469, 2021, doi: https://doi.org/10.3390/rs13030469.

[7] N. Pettorelli *et al.*, "Satellite remote sensing of ecosystem functions: opportunities, challenges and way forward," *Remote Sensing in Ecology and Conservation,* vol. 4, no. 2, pp. 71-93, 2018, doi: https://doi.org/10.1002/rse2.59.

[8] D. E. Pabon-Moreno, M. Migliavacca, M. Reichstein, and M. D. Mahecha, "On the Potential of Sentinel-2 for Estimating Gross Primary Production," *IEEE Transactions on Geoscience and Remote Sensing,* vol. 60, pp. 1-12, 2022, doi: 10.1109/TGRS.2022.3152272.

[9] R. Barbier, S. B. Yahia, P. Le Masson, and B. Weil, "Expanding usages of Earth Observation data: a co-design approach to grow an ecosystem of efficient service designers," in *2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS*, 2021: IEEE, pp. 296-299, doi: 10.1109/IGARSS47720.2021.9553914.

[10] GEO, "GEO Strategic Plan 2016-2025: Implementing GEOSS," 2021. [Online]. Available: https://earthobservations.org/resources#key.

[11] G. A. Blackburn and J. Pitman, "Biophysical controls on the directional spectral reflectance properties of bracken (Pteridium aquilinum) canopies: results of a field experiment," *International Journal of Remote Sensing,* vol. 20, no. 11, pp. 2265-2282, 1999, doi: https://doi.org/10.1080/014311699212245.

[12] O. Mutanga, A. K. Skidmore, and S. Van Wieren, "Discriminating tropical grass (Cenchrus ciliaris) canopies grown under different nitrogen treatments using spectroradiometry," *ISPRS Journal of Photogrammetry and Remote Sensing,* vol. 57, no. 4, pp. 263-272, 2003, doi: https://doi.org/10.1016/S0924-2716(02)00158-2.

[13] P. S. Thenkabail, R. B. Smith, and E. De Pauw, "Hyperspectral vegetation indices and their relationships with agricultural crop characteristics," *Remote sensing of Environment,* vol. 71, no. 2, pp. 158-182, 2000, doi: https://doi.org/10.1016/S0034-4257(99)00067-X.

[14] P. Olofsson *et al.*, "Towards operational remote sensing of forest carbon balance across Northern Europe," *Biogeosciences,* vol. 5, no. 3, pp. 817-832, 2008, doi: https://doi.org/10.5194/bg-5-817-2008.

[15] C. Xu, Y. Li, J. Hu, X. Yang, S. Sheng, and M. Liu, "Evaluating the difference between the normalized difference vegetation index and net primary productivity as the indicators of vegetation vigor assessment at landscape scale," *Environmental monitoring and assessment,* vol. 184, pp. 1275-1286, 2012, doi: https://doi.org/10.1007/s10661-011-2039-1.

[16] R. Roncella *et al.*, "Publishing NextGEOSS data on the GEOSS Platform," *Big Earth Data,* vol. 7, no. 2, pp. 413-427, 2023/04/03 2023, doi: 10.1080/20964471.2022.2135234.