Generating an improved Sentinel-2 GPP dataset using AI

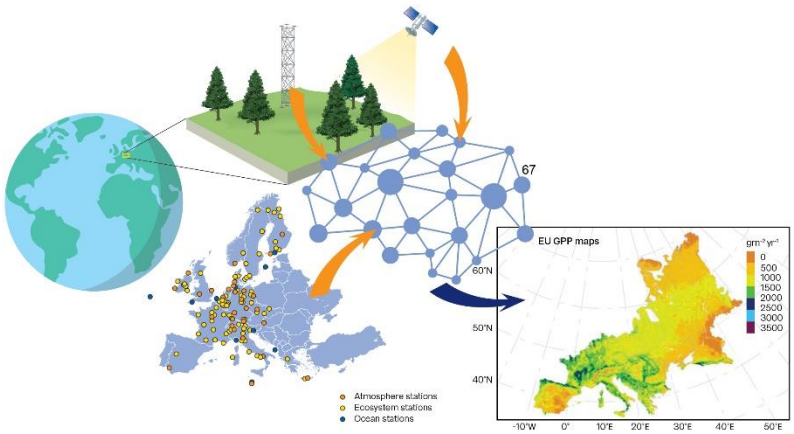
Supervision Provided by Anna Spinosa, Dr. Marieke Eleveld, and Dr. Ghada El Serafy
Collaboration between Deltares and TU Delft DIAM (EWI)

Research Context

We have an interesting opportunity for a MSc candidate to collaborate on our project aiming at exploring the use of machine learning methods for the generation of an improved European dataset of Gross Primary Productivity (GPP). GPP, defined as the total carbon fixed by the ecosystems through vegetation photosynthesis, is a key element for the assessment of ecosystem health. Healthy ecosystems regulate the impacts of natural hazards, thus playing a key role in mitigating climate change effects (Orradóttir, 2015). Satellite imageries offer an ideal platform to monitor the dynamics of the ecosystems in time and space and with that improve the understanding of the relevant processes and drivers of ecosystem changes at a system level.

Research Gap

Yet, remote sensing based global products of GPP are solely provided by MODIS sensor at spatial resolutions of 500 to 1000 m (MOD17 products). The coarse resolution of these products highly affects the accuracy of local estimates of GPP, especially in heterogeneous landscapes (Balzarolo et al., 2019; Wang et al., 2019). Sentinel-2 (S2) offers a promising alternative capable of improving the accuracy and precision of the estimated GPP at a local scale, given its significantly higher spatial resolution (10-60m) (Pettorelli et al. 2018).



Task Description

The scope of this MSc is to develop integrative algorithms for spatial distribution maps and indicators for habitats. The aim is to develop a generic method that allows for upscaling (temporally and spatially) at a European scale the approached tested by Spinosa et al. (2023).

The following aspects can be looked at:

a) we would like to create a data pipeline that enables fast reconstruction and monitoring of GPP at high-spatial resolution across targeted ecosystems. This will raise several challenges, including the acquisition, storage, searching, sharing, transferring, analysis, and visualization of big data. Thus, effective methods for estimating the GPP should be assessed and compared in terms of their advantages and limitations (assumptions, accuracy, understandability, speed). The identified methods will be applied to the high-resolution data set at hand. The workflow includes discovery and download of datasets, calibration and validation of algorithms.
b) newly available very high resolution (~2 m) optical data can fill the current knowledge gap, however, these datasets only contain spectral reflectances (wave lengths of the reflected light in various bands). Consequently, new so-called “band-math” or “spectral-math” algorithms should be developed to map the reflectance values into indicators. These

algorithms may range from simpler linear statistical models to non-linear machine learning models.

Methodology

1) Literature review on state-of-the-art spectral math techniques (e.g. map existing methods, algorithms, tools) mainly, including AI methods.

2) Comprehensive assessment and comparison of identified techniques in terms of their advantages and limitations.

3) Collection and sorting of various relevant remote sensing data sources (e.g. Sentinel-2, MODIS, Landsat-8, Pleiades) and in situ measurements of flux tower and meteorological products within the selected ecosystem. Investigation into the spatial and temporal variability within collected data sets including its quality. Preparation of an inventory of training, and validation datasets including indicators.

4) Application or development of selected techniques on the relevant data. Assessment of the accuracy level of the methods and computational efficiency, limitations and future trends associated with the methods.

The Project

The findings of this MSc thesis will contribute to the OBAMA-NEXT project (https://obama-next.eu/), an EU-funded project aiming at creating a toolbox capable of producing and disseminating precise, user-friendly, and cost-effective information to assess conditions and changes in marine ecosystems and their biodiversity.

Student Profile

You will program in python. You will deal with geo-big data and work on cloud-based geospatial processing platform. You are motivated to work at the intersection of multiple fields, including earth sciences, machine learning and biogeochemistry.

Application

Interested applicants can submit a Cover letter stating interest and relevant background experience with a Resume attached and Subject “GPP project internship opportunity” to Anna Spinosa: Anna.Spinosa@deltares.nl and Ghada El Serafy Ghada.elserafy@deltares.nl.

Relevant Literature

Spinosa, A., Fuentes-Monjaraz, M.A., El Serafy, G., 2023. Assessing the Use of Sentinel-2 Data for Spatio-Temporal Upscaling of Flux Tower Gross Primary Productivity Measurements. *Remote Sensing*. 15(3):562. https://doi.org/10.3390/rs15030562

Pereira, H.M., Ferrier, S., Walters, M., Geller, G.N., Jongman, R.H., Scholes, R.J., Bruford, M.W., Brummitt, N., Butchart, S.H., Cardoso, A.C. and Coops, N.C., 2013. Essential biodiversity variables. *Science*, *339*(6117), pp.277-278. DOI:10.1126/science.1229931