Early fungal detection in pine forest by fusing PRISMA and Sentinel-2 with biophysical models

Jose Luis Pancorbo*
Consiglio Nazionale delle Ricerche
Institute of BioEconomy (CNR-IBE)
Florence, Italy
0000-0003-1837-7589

Paul Mille
Universidad de León
School of Agrarian and Forest Engineering
León, Spain
pmille00@estudiantes.unileon.es

Giandomenico De Luca
Consiglio Nazionale delle Ricerche
Institute of BioEconomy (CNR-IBE)
Florence, Italy
0000-0002-4740-6468

Beniamino Gioli
Consiglio Nazionale delle Ricerche
Institute of BioEconomy (CNR-IBE)
Florence, Italy
0000-0001-7631-2623

Lorenzo Genesio
Consiglio Nazionale delle Ricerche
Institute of BioEconomy (CNR-IBE)
Florence, Italy
0000-0001-9265-886X

Nicola Arriga
European Commission (EC)
Joint Research Centre (JRC)
Ispra, Italy
0000-0001-5321-3497

Flor Álvarez-Taboada
Universidad de León
School of Agrarian and Forest Engineering
León, Spain
0000-0002-1530-3309

Pieter S.A. Beck
European Commission (EC)
Joint Research Centre (JRC)
Ispra, Italy
0000-0002-0692-4779

Carlos Camino
Wageningen University.
Laboratory of Geo-Information Science and Remote Sensing,
Wageningen, Netherlands
0000-0001-5188-4406

Remote sensing of the spatio-temporal variability of plant traits provides an effective identification of biotic and abiotic stress in forest ecosystems. The early detection of biotic infections is essential to mitigate their ecological impacts. Hybrid inversion approaches that combine machine learning with biophysical radiative transfer models (RTMs) enable the retrieval of plant traits linked to ecosystem functioning and stress responses. In this study, we applied a hybrid Deep Learning–PROSAIL–PRO model to time series of Sentinel-2 (SE2) and PRISMA imagery acquired over the San Rossore pine forest ICOS site (IT-SR2), affected by a *Fomes fomentarius* infection. The objectives were to (i) assess the accuracy of satellite-derived plant traits and their capacity to estimate gross primary productivity (GPP), and (ii) evaluate their ability to detect fungal infection. Significant differences in leaf area index (LAI), total chlorophyll content (Chl-ab), and equivalent water thickness (EWT) were detected between infected and healthy areas using SE2-derived traits. In contrast, PRISMA-based retrievals did not allow a reliable discrimination of infected areas due to the sensor's coarser spatial resolution (30 m). These results highlight the potential of multispectral time series for early fungal detection and point towards future work focused on improving hyperspectral-based detection through spatial resolution enhancement with data fusion techniques.

Keywords—PROSAIL, inversion, Leaf area index, Gross Primary Productivity, ICOS

I. Introduction

Climate change is increasing forest biotic diseases caused by fungi, bacteria, viruses, and insects. These disturbances reduce tree growth and survival, with cascading effects on fauna, nutrient cycling, and forest carbon sequestration [1]. Ecosystem functioning and forest health can be assessed through continuous measurements of forest–atmosphere gas exchange provided by flux tower networks such as the Integrated Carbon Observation System (ICOS), where gross primary productivity (GPP) serves as a key indicator of photosynthetic carbon uptake.

Advances in data science and satellite remote sensing have become essential for large-scale, long-term forest health monitoring. Hyperspectral sensors such as PRISMA provide detailed spectral information for vegetation characterization, while multispectral sensors like Sentinel-2 offer fewer bands but higher spatial and temporal resolution for time-series analysis. Traditional empirical models linking reflectance or vegetation indices (VIs) to plant traits are computationally efficient but poorly transferable across conditions. In contrast, physically based radiative transfer models (RTMs), such as PROSAIL [2], simulate canopy reflectance from leaf optical properties and canopy structure, enabling the retrieval of structural (e.g., LAI), biochemical (e.g., chlorophyll content), and biophysical (e.g., water content) plant traits, supporting early detection of stress and disease [3].

This study compares the performance of multispectral (SE2), hyperspectral (PRISMA), and fused datasets to i) retrieve key plant traits, ii) estimate GPP, and iii) detect fungal outbreaks in a Mediterranean pine forest using a hybrid ANN–PROSAIL inversion framework.

II. Materials and Methods

A. Study site and ICOS ground measurements

The study was conducted in a *Pinus pinea* L. forest surrounding the ICOS IT-SR2 (San Rossore 2) flux tower, located in the Migliarino–San Rossore Natural Park, a Natura 2000 protected area ~8 km west of Pisa (Italy). The area has a Mediterranean climate (Köppen Csa), with a mean annual temperature of 15.3 °C and annual precipitation of ~950 mm.

Ground-based measurements were obtained from the ICOS data portal. Hourly gross primary productivity (GPP) was aggregated to daily values using the variable u* threshold (VUT) method [4] and expressed in µmol CO_2 m^{-2} s^{-1}. Additionally, 23 field-based LAI measurements collected around the flux tower between November 2019 and September 2023 were used for validation.

Two *Fomes fomentarius* outbreaks were identified and used to define three study zones (Fig. 1). Zone 1 corresponds to a 50 m radius around the ICOS tower that was not infected. Zone 2 (470 m^2) was affected in 2020 and cleared at the end of that year. Zone 3 (3820 m^2) experienced a new infection starting after 2021. *Fomes fomentarius* is a wood-decaying fungus causing white rot, weakening live trees and increasing their susceptibility to breakage, making early detection essential.

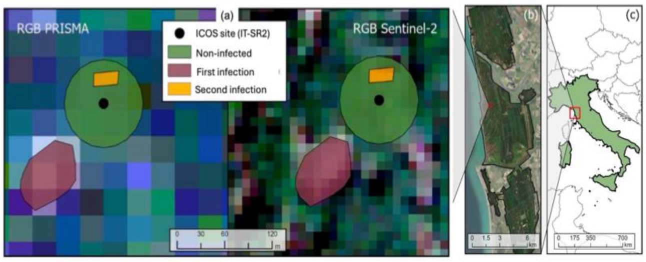

Fig. 1. a) RGB map of PRISMA and Sentinel-2 over the San Rossore ICOS site (IT-SR2), b) location of the site in San Rossore and c) location within Italy.

B. *Satellite imagery*

Multispectral SE2 and hyperspectral PRISMA imagery were used to retrieve plant traits, estimate GPP, and detect fungal outbreaks.

A total of 348 SE2 Level-2A images acquired between 2017 and 2024 were downloaded via Google Earth Engine. Images were filtered using the Scene Classification Layer (SCL) to retain cloud-free and radiometrically reliable observations. Surface reflectance from the 10 m and 20 m bands was extracted for pixels within the three study zones, forming the SE2 dataset (Dataset 1).

Twenty cloud-free PRISMA images acquired between 20 February 2021 and 28 October 2023 were used after being geometrically corrected and co-registered to temporally closest SE2 images [5]. From the full hyperspectral dataset, 167 bands with good radiometric quality were retained. Due to the lack of PRISMA acquisitions before 2021, reflectance values were extracted only for Zones 1 and 3.

A systematic spectral offset between PRISMA and SE2 reflectances was identified. To address this, several PRISMA-derived datasets were generated. First, PRISMA spectra were resampled to SE2 bands using spline interpolation, producing a multispectral PRISMA dataset (Dataset 2: PRISMA–SE2 original). Linear regression models were then derived for each SE2 band using reflectance values from spatially independent coincident observations and applied to all resampled PRISMA images, yielding a spectrally inferred dataset (Dataset 3: PRISMA–SE2 inferred). In parallel, a dataset using original PRISMA reflectance was created (Dataset 4: PRISMA original). To reduce collinearity, one out of every five spectral bands (34 bands) was retained. Finally, Dataset 3 was resampled back to hyperspectral resolution using spline interpolation, generating an inferred hyperspectral dataset (Dataset 5).

C. *Hybrid ANN-PROSAIL model*

Leaf area index (LAI), total chlorophyll content (Chl-ab), and equivalent water thickness (EWT) were retrieved using a hybrid ANN–PROSAIL framework implemented in the ToolsRTM R package [6]. First, a lookup table (LUT) of 20,000 PROSAIL-simulated spectra was generated using parameter ranges defined from literature for coniferous forests and iteratively refined by matching simulated and observed reflectance from the study zones (Table 1). Simulations were produced separately at SE2 and PRISMA spectral resolutions.

RTM inversion was performed using feedforward artificial neural networks (ANNs). LUT inputs were scaled prior to training. The ANNs were configured with three hidden layers (128 neurons), a batch size of 32, and up to 100 epochs with early stopping. The Adam optimizer was used. Models were trained with 80% of the LUT and validated with the remaining 20%, then retrained using the full LUT. Six inversion models were developed: three at SE2 resolution and three at PRISMA resolution, each targeting LAI, Chl-ab, or EWT. Predictor selection followed trait-specific spectral sensitivity, using all bands for LAI, VNIR bands for Chl-ab, and SWIR bands for EWT, complemented by selected VIs.

The trained models were applied to the five SE2 and PRISMA datasets. Trait estimates were aggregated per study zone using median values, and outliers were removed using an interquartile range filter. The final output consisted of five time series per dataset for LAI, Chl-ab, and EWT, enabling comparison across sensors, preprocessing strategies, and zones.

TABLE I. PLANT TRAITS RANGES SET IN PROSAIL TO CREATE THE LOOK-UP TABLE.

Trait	*Description*	*Range*	*Unit*
	PROSPECT		
N	Internal leaf structure parameter	1.5-3	[-]
Cab	Chlorophyll a+b content	5-70	µg/cm2
Car	Carotenoid content	0-15	µg/cm2
Anth	Anthocyanin content	0-5	µg/cm2
Cbrown	Brown pigments content	0.001-0.035	[-]
EWT	Equivalent water thickness	0.001-0.045	g/cm2
LMA	Leaf dry matter content	NO	g/cm2
Proteins	Protein content	0.001-0.01	mg/cm2
CBC	Carbon-based constituents	0.001-0.35	mg/cm2
	FourSAIL		
LAI	Leaf area index	1.5-4-5	[-]
LIDFa	Leaf angle distribution	50-90	deg
hspot	Hotspot parameter	0-1	deg
tts	Solar zenith angle	15-75	deg
tto	Observer zenith angle	0-15	deg
psi	Relative azimuth angle	0-150	deg
Soil factor	psoil	0-0-15	[-]

D. *Validation of satellite-derived plant traits*

The accuracy of retrieved plant traits (LAI, Chl-ab, and EWT) was validated using independent flux and field data. Daily GPP from the ICOS IT-SR2 flux tower served as a benchmark for ecosystem photosynthetic activity. A neural network regression model was trained with SE2-derived traits and seven vegetation indices as predictors, and ICOS GPP as the target. SE2 was selected due to its higher temporal resolution and coverage (2017–2024). Model performance was evaluated on a 20% independent test set using R^2, RMSE, and MAE.

LAI retrievals were validated against in situ measurements around the tower (Zone 1), allowing a maximum temporal mismatch of 20 days. This also enabled comparison of preprocessing strategies and satellite datasets. SE2 and spectrally inferred PRISMA traits were further compared for coincident acquisitions (within four days) using NRMSE.

Early fungal infection detection was assessed by analyzing temporal variations of LAI, Chl-ab, and EWT across study zones and periods (2019–2021 and 2021–2024). SE2 data were used for both periods, while PRISMA analyses were limited to Zones 1 and 3 post-2021 due to data gaps. Differences between zones and periods were evaluated using time series, boxplots, and ANOVA after confirming normality.

III. RESULTS

A. GPP estimation

The neural network model predicting GPP from SE2-derived plant traits and vegetation indices showed strong performance ($R^2 = 0.82$, RMSE = 1.66, MAE = 1.27 µmol CO_2 m^{-2} s^{-1}). The five-day averaged GPP time series reproduced the observed seasonal dynamics measured at the ICOS IT-SR2 flux tower (Fig. 2), with a slight underestimation of abrupt variations at the end of 2019 and early 2020. Overall, the results demonstrate the high capability of plant trait–based models to estimate ecosystem photosynthetic activity.

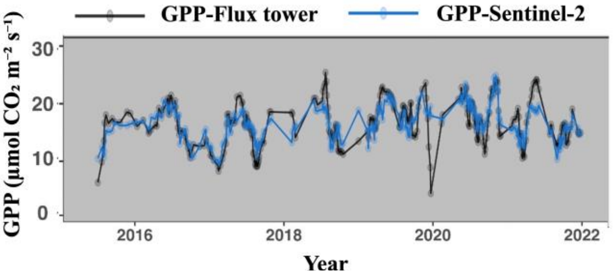

Fig. 2. Time series of the Gross Primary Productivity (GPP) at San Rossore ICOS site (IT-SR2) calculated with the flux tower and by combining the Sentinel-2 retrievals.

B. Plant traits retrieval

Among the retrieved plant traits, EWT showed the highest agreement between Sentinel-2 and PRISMA (NRMSE = 0.22), followed by Chl-ab (NRMSE = 0.25), while LAI exhibited the lowest consistency (NRMSE = 0.41).

Comparison with in situ LAI measurements revealed the best performance for Dataset 3 (PRISMA–SE2 inferred, RMSE = 0.41), closely followed by Dataset 1 (Sentinel-2, RMSE = 0.47). Dataset 4 (PRISMA original) showed moderate accuracy (RMSE = 0.74), Dataset 2 was similar (0.80), and Dataset 5 (PRISMA-inferred) performed poorly (RMSE = 4.93). These results highlight the critical role of spectral inference and preprocessing for accurate LAI retrieval from PRISMA data.

C. Early fungal infection detection

During the 2019–2021 period, Zone 2 (previous infection) consistently exhibited lower LAI, Chl-ab, and EWT than the other zones, indicating degraded vegetation conditions prior to tree removal. In the 2021–2024 period, larger differences between Zones 1 and 3 were observed for traits retrieved from SE2 compared to PRISMA. ANOVA results confirmed that SE2–based traits showed statistically significant differences between Zones 1 and 2 (2019–2021) and between Zones 1 and 3 (2021–2024), with p-values $< 2 \times 10^{-16}$. For LAI, a small but significant difference between Zones 1 and 3 was already detected in the first period, increasing markedly in the second period. In contrast, none of the traits retrieved from spectrally inferred PRISMA data showed significant differences between Zones 1 and 3, with p-values > 0.05 for all traits. This indicates that PRISMA-based trait retrievals did not enable detection of the new fungal infection.

During 2019–2021, Zone 2 (previous infection) consistently showed lower LAI, Chl-ab, and EWT than the other zones, reflecting vegetation degradation prior to tree removal. In 2021–2024, differences between Zones 1 and 3 were larger for traits retrieved from SE2 than from PRISMA. ANOVA confirmed that SE2-based traits differed significantly between Zones 1 and 2 (2019–2021) and Zones 1 and 3 (2021–2024), with p-values $< 2 \times 10^{-16}$. For LAI, a small but significant difference was already detected in the first period, increasing in the second. In contrast, traits from spectrally inferred PRISMA data did not show significant differences between Zones 1 and 3 ($p > 0.05$), indicating that PRISMA-based retrievals could not detect the new fungal infection.

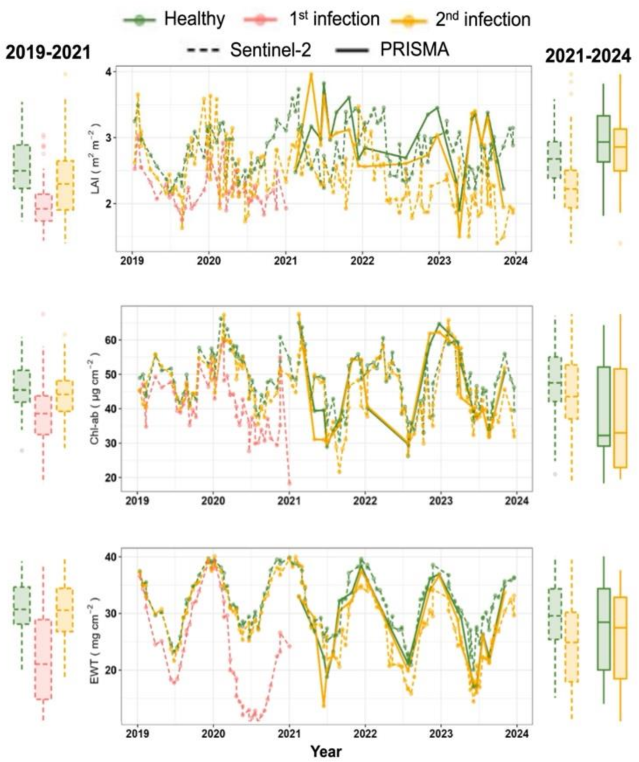

Fig. 3. Time series of LAI, Chl-ab and EWT calculated with the hybrid model applied to S2 and PRISMA reflectance of the non-infected area (green), the area infected in 2019 (red) and the area infected in 2021. Box plots describe the values retrieved during the first (left) and second (right) infection.

IV. Discussion

Linear regression approaches applied to satellite reflectance are widely used for atmospheric correction and inter-sensor harmonization. In this study, spectral inference effectively adjusted PRISMA reflectance to SE2, as confirmed by LAI field validation, which showed improved accuracy compared to using original PRISMA data. These results highlight the importance of preprocessing strategies and the need for algorithms that correct inter-sensor differences without degrading hyperspectral information.

Constraining RTM parameters was critical for reliable inversion. The resulting look-up table accurately represents the radiometric conditions captured by both satellites, providing a robust basis for training deep learning inversion models, and validated the parameter ranges used. Models using hyperspectral predictors consistently showed higher predictive power, reflecting the richer spectral information. Field LAI measurements provided the only direct validation of trait retrievals and demonstrated robust performance for both SE2 and PRISMA datasets, with errors consistent with similar forest studies [7]. Validation using ICOS GPP data confirmed that RTM-derived traits provide meaningful information on ecosystem photosynthetic activity, achieving accuracy comparable to state-of-the-art studies [8]. This supports the reliability of the methodology and its suitability for forest health analyses, disease detection and plant traits estimation.

Using SE2–derived traits, both fungal infection events were successfully detected, with significant reductions in LAI, Chl-ab, and EWT reflecting physiological stress responses. In contrast, PRISMA data did not detect the new infection due to its coarser 30 m spatial resolution, which mixed signals from affected and healthy trees. Previous disease detection studies relied on better spatial resolution [3], [9], suggesting that either larger infected areas or higher-resolution hyperspectral imagery are required to fully exploit PRISMA's spectral detail. Approaches such as pansharpening PRISMA images to 10–5 m resolution based on SE2 images [5], or integrating climatic and soil variables, could enhance early-detection capabilities for forest disease monitoring.

V. Conclusions

This study demonstrates the potential of combining SE2 and PRISMA imagery with PROSAIL-based RTM inversion to retrieve plant traits relevant for early forest disease detection. Statistically significant reductions in LAI, Chl-ab, and EWT were detected in areas affected by *Fomes fomentarius* using SE2 data, confirming its suitability for monitoring early infection dynamics. In contrast, PRISMA data did not enable reliable detection due to its coarser spatial resolution (30 m), which limited the discrimination of localized infection signals.

Despite this limitation, hyperspectral-based models showed higher predictive capability than multispectral ones. Validation with in situ LAI measurements indicated that spectrally inferred PRISMA data harmonized with SE2 achieved the highest accuracy (RMSE = 0.41), although this required resampling to multispectral resolution and resulted in a loss of hyperspectral detail. Gross primary productivity (GPP) was accurately estimated at the IT-SR2 ICOS site using SE2-derived plant traits and vegetation indices (R^2 = 0.82, RMSE = 1.66 μmol CO_2 m^{-2} s^{-1}), highlighting the value of RTM-based trait retrievals for assessing ecosystem functioning.

Future work should focus on improving the spatial resolution of hyperspectral observations through data fusion and pansharpening techniques, as well as on developing spectral harmonization methods that preserve hyperspectral information, to fully exploit the potential of upcoming hyperspectral missions for forest health monitoring.

Acknowledgment

This work is part of the project ITINERIS (Italian Integrated Environmental Research Infrastructures System), funded by the PNRR program (M4, C2, I3.1). This study made use of PRISMA Products of the Italian Space Agency (ASI), delivered under an ASI License to use, and was made possible by the funding support of: the agreement ASI-CNR, n. 20195HH0 "Attività scientifica di CAL/VAL della missione PRISMA - PRISCAV". Additional support was provided by the ESA-funded HYDRA-EO project (EXPRO+, Contract No. 1-12684.

References

[1] Guégan, B. De Thoisy, M. Gomez-gallego, and H. Jactel, "World forests, global change, and emerging pests and pathogens," Curr. Opin. Environ. Sustain., vol. 61, p. 101266, 2023, doi: 10.1016/j.cosust.2023.101266.

[2] J. B. Féret, K. Berger, F. de Boissieu, and Z. Malenovský, "PROSPECT-PRO for estimating content of nitrogen-containing leaf proteins and other carbon-based constituents," Remote Sens. Environ., vol. 252, no. October 2020, 2021, doi: 10.1016/j.rse.2020.112173.

[3] C. Camino, A. Valero-Jorge, E. G. Lima, R. Álvarez, P. S. A. Beck, and F. Álvarez-Taboada, "Detection and monitoring of Melampsora spp. Damage in multiclonal poplar plantations coupling biophysical models and Sentinel-2 time series," Int. J. Appl. Earth Obs. Geoinf., vol. 141, p. 104663, 2025, doi: https://doi.org/10.1016/j.jag.2025.104663.

[4] G. LASSLOP et al., "Separation of net ecosystem exchange into assimilation and respiration using a light response curve approach: critical issues and global evaluation," Glob. Chang. Biol., vol. 16, no. 1, pp. 187–208, 2010, doi: https://doi.org/10.1111/j.1365-2486.2009.02041.x.

[5] G. De Luca et al., "ISPRS Journal of Photogrammetry and Remote Sensing Improving PRISMA hyperspectral spatial resolution and geolocation by using Sentinel-2 : development and test of an operational procedure in urban and rural areas," ISPRS J. Photogramm. Remote Sens., vol. 215, no. February, pp. 112–135, 2024, doi: 10.1016/j.isprsjprs.2024.07.003.

[6] C. Camino, K. Ose, K. Araño, C. Bolyn, and P. S. A. Beck, "RT-Simulator: An Online Platform to Simulate Canopy Reflectance from Biochemical and Structural Plant Properties Using Radiative Transfer Models," in IGARSS 2024 - 2024 IEEE International Geoscience and Remote Sensing Symposium, 2024, pp. 2811–2814, doi: 10.1109/IGARSS53475.2024.10642442.

[7] J. L. Pancorbo, M. Quemada, M. D. Raya-sereno, B. Gioli, P. S. A. Beck, and C. Camino, "Integrating Artificial Neural Network-PROSAIL With Sentinel-2 to Monitor Crop Traits Dynamics and Nitrogen Status," vol. 18, pp. 17770–17786, 2025, doi: 10.1109/JSTARS.2025.3585080.

[8] R. Raj, B. Bayat, P. Lukeš, and L. Šigut, "Analyzing Daily Estimation of Forest Gross Primary Production Based on Harmonized Landsat-8 and Sentinel-2 Product Using SCOPE," pp. 1–22, 2020, doi: 10.3390/rs12223773.

[9] C. Camino et al., "Detecting Xylella fastidiosa in a machine learning framework using Vcmax and leaf biochemistry quantified with airborne hyperspectral imagery," Remote Sens. Environ., vol. 282, no. August, 2022, doi: 10.1016/j.rse.2022.113281.