

Influences of noise on paddy chlorophyll mapping and associated uncertainty across heterogeneous landscapes using Sentinel-2 and hybrid models[☆]

Giang Thi Phuong Thao^a, Nguyen An Binh^{a,*}, Tong Si Son^b, Le Thi Thu Ha^c

^a Department of Geography and Remote Sensing, Institute of Life Sciences, Vietnam Academy of Science and Technology, Ho Chi Minh City, Viet Nam

^b University of Science and Technology of Hanoi, Vietnam Academy of Science and Technology, Hanoi, Viet Nam

^c Hanoi University of Mining and Geology, Viet Nam

ARTICLE INFO

Keywords:

Remote sensing
Functional trait
Paddy chlorophyll
Hybrid model
Uncertainty
Noise
Vietnamese agriculture

ABSTRACT

Satellite-derived crop functional traits are increasingly pivotal for effective agricultural monitoring and informed decision-making, satisfying the demands of sustainable, smart, and precision agriculture. However, retrieving these traits leveraging multispectral derivatives is challenging due to inherent variability and uncertainties, highlighting the need for comprehensive quantification of the ill-posedness problem arising from both input data and predicted methods. This study investigated impacts of noise on paddy rice chlorophyll mapping and the associated pixel-level uncertainty across heterogeneous agricultural landscapes using Sentinel-2 imagery and hybrid models. Specifically, we employed Gaussian Process Regression (GPR) and Variational Heteroscedastic Gaussian Process Regression (VHGPR) to train models utilizing simulated data generated by the physically-based radiative transfer model PROSAIL. Subsequently, noise levels of 0 %, 5 %, and 10 % were incorporated to assess their effects, and pixel-level uncertainty was estimated across different land covers. Our findings revealed the influence of noise on model performance, band relevance, and associated uncertainty. The index of agreement (IOA) varying from the lowest 0.82 at 10 % noise to the highest 0.92 in the absence of noise. With relative root mean square errors (RRMSE) below 5 %, all models provided high-fidelity mapping of paddy chlorophyll content (mean relative uncertainties below 22 %), compared to the lower confidence levels observed for non-vegetated objects (mean relative uncertainties exceeding 50 %). This research contributes to the growing need for uncertainty mapping associated with plant trait products. Furthermore, it highlights the importance of noise management in satellite applications for achieving more robust and realistic retrievals across heterogeneous landscapes.

1. Introduction

Remote sensing in agriculture has become a mature science with streamlined products associated with categorical mapping, well-established spectral indices, and functional plant traits (Weiss et al., 2020b). The availability and accessibility of agricultural datasets facilitate the development of smart and precise agricultural applications such as crop phenology analysis, water and nutrient management, as well as pest and disease warning, contributing to accurate yield prediction (Berger et al., 2022; Sishodia et al., 2020). Generally, it is impossible to satisfy spatiotemporal monitoring objectives relying solely on field measurement and laboratory analysis, therefore highlighting

the capabilities of Earth observation technologies (Verrelst et al., 2019; Weiss et al., 2020b). However, in the context of remote sensing agricultural applications, the transferability of quantitative methods remains limited for truthful estimations across space and time dimensions, primarily due to data anomalies in field measurement combined with imaging spectroscopy data, and inherent model uncertainties (Brown et al., 2021; Martínez-Ferrer et al., 2022; Povey and Grainger, 2015). Consequently, this limits the broader agricultural application of remote sensing as it restricts monitoring over regional, continental, and global scales simultaneously with long-term time series analysis (Malenovsky et al., 2019; Verrelst et al., 2019).

A general taxonomy for mapping biophysical and biochemical

[☆] This article is part of a Special issue entitled: 'Plant Functional Traits Study' published in Ecological Informatics.

* Corresponding author.

E-mail address: nabinh@ils.vast.vn (N.A. Binh).

<https://doi.org/10.1016/j.ecoinf.2025.103351>

Received 20 April 2025; Received in revised form 17 July 2025; Accepted 23 July 2025

Available online 24 July 2025

1574-9541/© 2025 The Authors. Published by Elsevier B.V. This is an open access article under the CC BY-NC license (<http://creativecommons.org/licenses/by-nc/4.0/>).

variables separates empirical approaches and physical-based radiative transfer models (RTMs) (Verrelst et al., 2015a; Verrelst et al., 2019). Empirical methods, relying on statistical relationships, can provide robust retrievals for well-defined areas with ample training data. Conversely, RTMs, grounded in physical principles, require less ground-measured data and offer greater transferability, particularly in diverse environments, while also enhancing scientific understanding of plant-light interactions. Hybrid retrieval strategies combine the strengths of both approaches by using RTMs to generate training data as input to empirical models. These frameworks are further improved by integrating machine learning regression algorithms (MLRAs) for superior estimation (Verrelst et al., 2015a; Verrelst et al., 2019). One common hybrid retrieval method is the PROSAIL toolbox within the Sentinel Application Platform (SNAP), utilizes Artificial Neural Network (ANN) to derive key biophysical and biochemical variables across a wide range of vegetation covers (Weiss et al., 2020a). However, the SNAP biophysical processor is trained on synergistic datasets encompassing diverse vegetation types, from forests to agricultural crops. In contrast, custom RTM models offer greater flexibility by allowing users to configure input variables, thus optimizing the representation of specific vegetation characteristics.

In the selection of MLRAs for hybrid model training, Artificial Neural Network (ANN) is among the most popular choices (Verrelst et al., 2015a; Verrelst et al., 2019). The well-established algorithm applies a multi-neural layers network to capture nonlinear relationships between spectroscopic data and vegetation traits. This capability has been demonstrated through applications of not only data-driven non-parametric regression but also RTM-based methods (Verrelst et al., 2015a; Verrelst et al., 2019). However, limitations in ANN include multicollinearity problems and time-consuming computation when complex structures (Verrelst et al., 2019). Unlike ANNs, nonlinear nonparametric Gaussian processes regression (GPR) can effectively tackle these challenges and frequently surpasses their performance in retrieving biophysical and biochemical variables (Danner et al., 2021; Verrelst et al., 2012b; Verrelst et al., 2015b). Furthermore, the Gaussian models were demonstrated with robust performance incorporating various sensors and vegetation ecosystems (Berger et al., 2020; Caballero et al., 2023; Camps-Valls et al., 2016; Ghosh et al., 2022; Verrelst et al., 2012a; Verrelst et al., 2012b; Verrelst et al., 2013a; Xie et al., 2021). A challenge in standard GPR is its assumption of independent noise variance (homoscedasticity), which limits its practical use, particularly in hybrid retrieval methods applied to remotely sensed data. To address this, the refined version of Variational Heteroscedastic Gaussian Process Regression (VHGPR) was developed within the Bayesian framework. VHGPR retains the advantages of GPR for excellent plant trait estimations while explicitly modeling input-dependent noise scenarios (heteroscedasticity) (Estévez et al., 2020; Estévez et al., 2021; Lázaro-Gredilla et al., 2014).

In light of model transferability for spatiotemporal-explicit estimation, evaluating both goodness-of-fit metrics and quantifying uncertainty is critical for optimal model selection (García-Soria et al., 2024; Kaminski et al., 2017). While field measurements are essential for model validation, exhaustive ground truth collection for every pixel in a remote sensing-based spatial approach is impractical. Uncertainty estimation may address this challenge by providing a pixel-wise quality metric, effectively a mask that identifies potential outliers and errors in the derived products. This is particularly valuable for hybrid retrieval models, especially where trait estimation is expected to be highly challenging across heterogeneous landscapes (Malenovsky et al., 2019; Martínez-Ferrer et al., 2022).

The remaining challenge in hybrid models is the noise resulting from the combination of remote sensing observations, RTM-based simulated data, and field measurements (Caicedo et al., 2014; de Sá et al., 2021). Recent studies have emphasized the varying model performance in the presence of noise. For example, using S2 imagery to estimate various traits (leaf area index, chlorophyll content, equivalent water thickness,

and dry matter content), de Sá et al. (2021) showed that GPR performed best in the absence of noise. However, GPR was significantly sensitive to noise levels of 5 % and above 20 %, becoming less effective than random forests (RF), ANN, and multi-task neural networks (MTN) (de Sá et al., 2021). In nitrogen estimation using a simulated hyperspectral signal, both kernel-based GPR and kernel ridge regression (KRR) demonstrated high accuracy and saturation adaptation with noise up to 10 %, outperforming ANN, least squares linear regression (LSLR), partial least squares regression (PLSR), support vector machine (SVR), and quantile regression forests (QRF) (García-Soria et al., 2024). To date, assessing hybrid model noise effects for mapping biophysical and biochemical variables has received limited attention, due to the difficulties in controlling and mitigating its implications (Caicedo et al., 2014; de Sá et al., 2021; García-Soria et al., 2024).

Focusing on rice, a globally critical food crop, paddy chlorophyll content serves as a direct indicator of the plant's photosynthetic capacity and nitrogen status, reflecting its ability to capture light and produce biomass essential for growth and yield (Liu et al., 2024). Upscaling for robust estimation of paddy chlorophyll using remote sensing is thus crucial for large-scale agricultural resource management, tracking phenological shifts, and enhanced crop growth monitoring over time series, thereby optimizing yield. Regarding near real-time applications, such timely and reliable insights into plant health and photosynthetic capacity enable optimized fertilization strategies, early detection of nutrient stress, and more efficient use of agricultural inputs. Ultimately, monitoring these traits is particularly vital for climate change response, aiding in the assessment of crop resilience to environmental stressors (Berger et al., 2022; Sishodia et al., 2020; Weiss et al., 2020b).

Recognizing the challenges associated with providing pixel-wise uncertainty estimates alongside ready-to-use vegetation trait products, and addressing the impact of noise when linking Earth observation constellations with physical radiative transfer models, we ask the following research questions: (1) how does noise affect the performance of remotely sensed plant traits prediction?, and (2) how do hybrid retrieval models produce accurate agricultural trait products, accompanied by uncertainty indicators for reliable operational applications? To address these questions, we evaluated the performance of GPR and VHGPR for estimating paddy chlorophyll pigment concentrations under varying noise levels (0, 5, and 10 %). The validated models were applied to Sentinel-2 (S2) surface reflectance data for full mapping across heterogeneous agricultural landscapes, representing diverse land cover types. Achieved results allow us to provide feedback on the capacity of the approach and discuss its potential for supporting smart, sustainable, and precise agriculture based on remote sensing.

2. Materials and methods

Fig. 1 visualizes the technical workflow undertaken in this study. Details of each step are further described in the following sections.

2.1. Study area

The study area is situated within the Red River Delta of northern Vietnam, also known as the national second-largest rice production. Despite paddy cultivation dominating the region, it exhibits a fragmented pattern interspersed with other land uses. Farming practices typically involve two distinct rice cropping seasons: the spring crop, extending from early November to late May, and the autumn crop, commencing in late May and concluding in mid-November (Yuen et al., 2021). However, significant temporal variability in crop calendars is evident across different locations, with planting and harvesting times demonstrating considerable variation even within individual paddy parcels.

We focused on a heterogeneous landscape spanning a total of 174.82 km² (latitude 21°13'–21°18'N and longitude 105°18'–105°26'E). The study land covers are categorized into seven specific types: water, built-

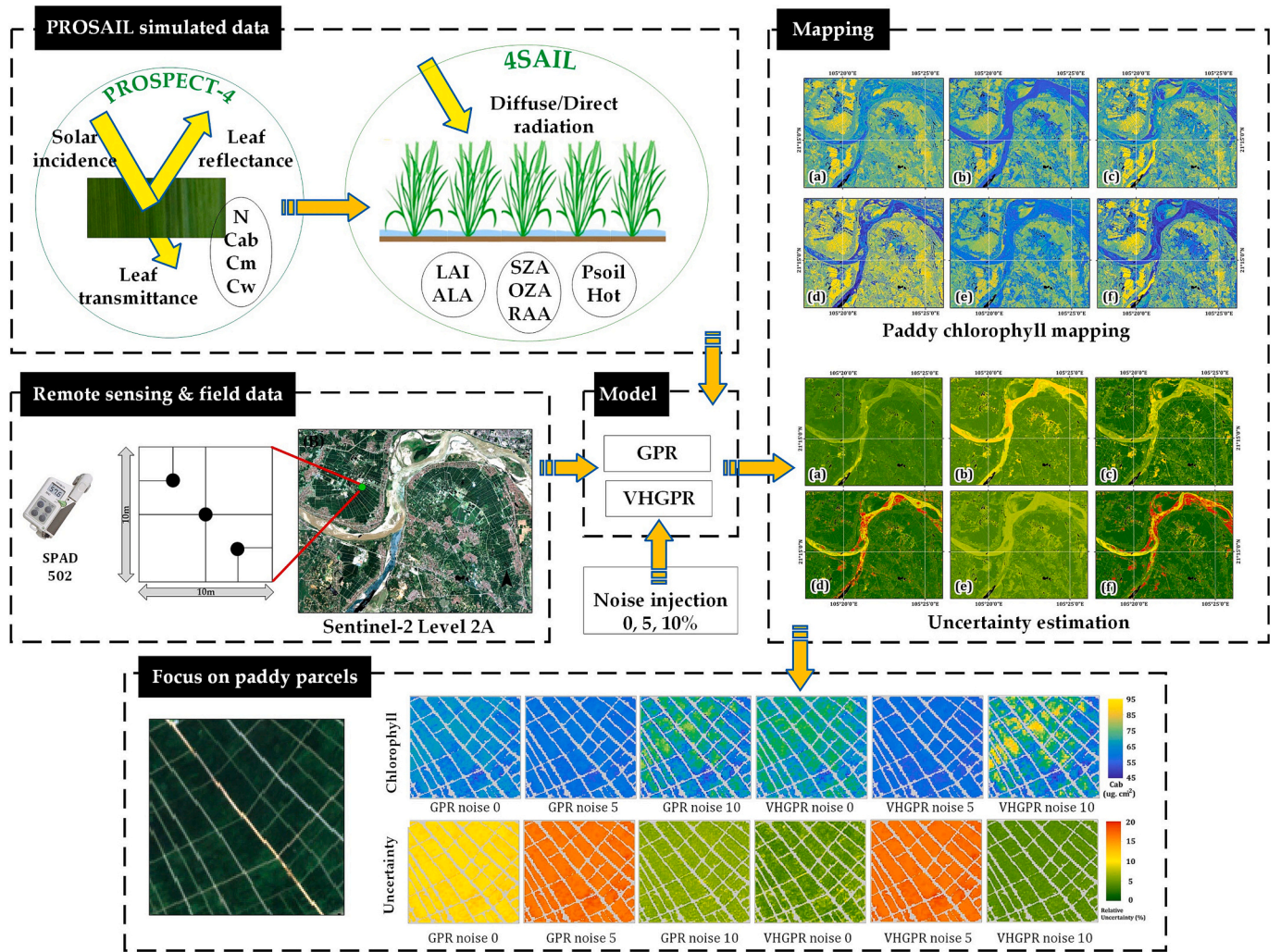


Fig. 1. Flowchart of the study.

up, rice, other crops, woody crops/orchards, barren, and forest. Statistically, rice dominated the landscape, accounting for over 44 % of the total area, followed by water bodies (21 %), built-up areas (19 %), and other crops (12 %). The remaining land cover types (i.e., woody crops/orchards, barren, and forest) collectively comprised a relatively

negligible proportion of the study area (Fig. 2).

2.2. Field campaign and measurement

The field campaign, designed to acquire ground truth data, was

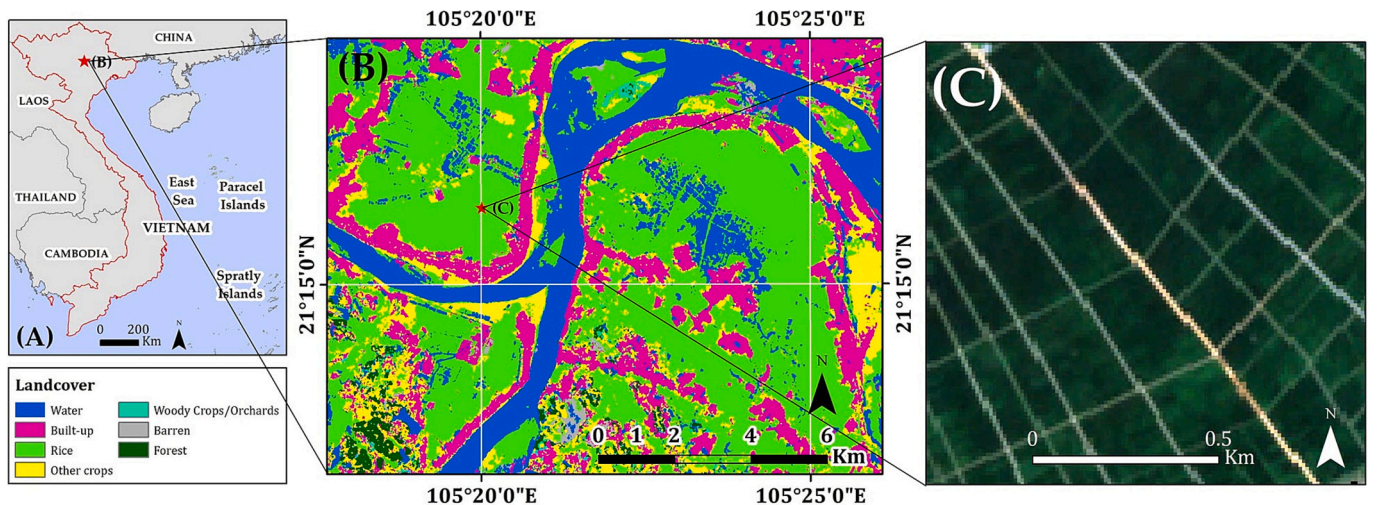


Fig. 2. Location of the study area (A), land cover map over the study site (B), with zoom in paddy parcels using S2 true color composition (C).

carried out between 3 and 6 May 2022. Across the paddy fields, 60 square sampling paddy plots (10 × 10 m) were carefully selected, characterized by a high proportion of foliage cover to minimize the influence of bare soil and water. The size and position of each plot were verified for spatial consistency with the pixel size of S2 imagery. Within each plot, paddy chlorophyll content was recorded using a non-destructive SPAD-502 chlorophyll meter. Three randomly distributed measurements were taken per plot, ensuring for homogeneous canopy, to obtain a representative mean chlorophyll value. The resulting SPAD values were subsequently transformed into chlorophyll content, expressed in $\mu\text{g}\cdot\text{cm}^{-2}$, using the following mathematical relationship (Rossi et al., 2022):

$$\text{LCC in } \mu\text{g}\cdot\text{cm}^{-2} = 0.21752 \times \text{SPAD}^{2.1129} \quad (1)$$

2.3. Sentinel-2 L2A and preprocessing

S2 (satellite 2B) Level 2 A surface reflectance images acquired on May 3, 2022, were downloaded for subsequent analysis, coinciding with the date of field measurements. It is important to acknowledge that cloud contamination is an annual occurrence in this region, influenced by the tropical monsoon climate background. To ensure data quality, rigorous quality control measures were implemented. All defective pixels were removed utilizing both the quality assessment band (QA60) and the S2 cloud probability dataset (s2cloudless) (Skakun et al., 2022; Zupanc, 2017). Ten spectral bands with spatial resolutions ranging from 10 to 20 m were selected. The six bands with an original resolution of 20 m (B5, B6, B7, B8A, B11, B12) were subsequently resampled to 10 m using the Sen2Res plugin available in SNAP (Brodu, 2017), while the remaining bands (B2, B3, B4) were maintained at their native 10 m resolution.

2.4. Physical radiative transfer model

The physical-based leaf-to-canopy radiative transfer model PROSAIL (PROSPECT + 4SAIL) was employed to simulate the relationship between paddy traits and spectral response (Jacquemoud et al., 2009). Integration of PROSPECT with 4SAIL facilitates the transformation of leaf-level to bi-directional reflectance, aligning with satellite observations over the vegetation canopy (Verhoef et al., 2007). A simulation of 1000 samples was conducted using Latin Hypercube Sampling (McKay et al., 1979), ensuring a uniform distribution across the entire range of each input variable. The number of simulations is balanced between optimal Gaussian performance and computational efficiency (Danner et al., 2021; Rasmussen and Williams, 2005), as already successfully demonstrated (Estévez et al., 2020; Estévez et al., 2022). The range of model parameters was configured based on field campaign measurements, complemented by referencing relevant prior works (Adeluyi et al., 2021; Campos-Taberner et al., 2016; Rossi et al., 2022; Xu et al., 2019), as detailed in Table 1. The processing steps were implemented using the ARTMO toolbox developed based on the MATLAB platform

Table 1
Configuration of PROSAIL.

Variables	Acronym	Units	Range	Distribution
PROSPECT				
Leaf structure	N		1.2–2.5	Uniform
Chlorophyll a + b	Cab	$\mu\text{g}\cdot\text{cm}^{-2}$	20–80	Uniform
Water thickness	Cw	$\text{g}\cdot\text{cm}^{-2}$	0.0001–0.05	Uniform
Dry matter	Cm	$\text{g}\cdot\text{cm}^{-2}$	0.0001–0.03	Uniform
4SAIL				
Leaf area index	LAI	$\text{m}^2\cdot\text{m}^{-2}$	0–6	Uniform
Average leaf angle	ALA	degree	0–90	Uniform
Diffuse/Direct radiation			0–100	Uniform
Soil brightness	psoil		0–1	Uniform
Hot spot effect	Hot		0–1	Uniform

(<https://artmotoolbox.com/>).

2.5. Gaussian-based algorithms

Gaussian process-based regression algorithms were employed to model the nonlinear relationships between independent variables (i.e., S2 surface reflectance bands) and the dependent variable (i.e., paddy chlorophyll content). We focused on two variants including the standard GPR and the refined VHGP. A key distinction between these algorithms lies in their treatment of data noise (Camps-Valls et al., 2016). GPR assumes homoscedastic noise, meaning the noise variance is constant across all input values. In contrast, VHGP accommodates heteroscedastic noise, allowing the noise variance to be a function of the input values.

Theoretically, GPR assumes that the function mapping remote sensing data ($\mathbf{x}$) to paddy chlorophyll content $f(\mathbf{x})$ is drawn from a Gaussian Process ($\mathcal{GP}$):

$$f(\mathbf{x}) \sim \mathcal{GP}(\mathbf{m}(\mathbf{x}), \mathbf{k}(\mathbf{x}, \mathbf{x}')) \quad (2)$$

Where $\mathbf{m}(\mathbf{x})$ is the mean function assumed to be zero, $\mathbf{k}(\mathbf{x}, \mathbf{x}')$ is the covariance function between any two input points $\mathbf{x}$ and $\mathbf{x}'$.

For training purpose, assuming that the observed chlorophyll content ($\mathbf{y}$) is a combination of the true underlying chlorophyll content $f(\mathbf{x})$ and noise (ϵ):

$$\mathbf{y} = f(\mathbf{x}) + \epsilon, \text{ and } \epsilon \sim \mathcal{N}(0, \sigma_n^2) \quad (3)$$

Where $\mathcal{N}(0, \sigma_n^2)$ is the normal distribution of zero mean and the variance. For GPR, σ_n^2 represents the constant noise variance (homoscedasticity assumption).

Consider a test point $\mathbf{x}_*$, the joint distribution of observed training outputs $\mathbf{y}$ and the unobserved test output $f_* = f(\mathbf{x}_*)$ is a multivariate Gaussian:

$$\begin{pmatrix} \mathbf{y} \\ f_* \end{pmatrix} \sim \mathcal{N}\left(0, \begin{pmatrix} \mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I} & \mathbf{K}(\mathbf{X}, \mathbf{x}_*) \\ \mathbf{K}(\mathbf{X}, \mathbf{x}_*)^{\text{transpose}} & \mathbf{k}(\mathbf{x}_*, \mathbf{x}_*) \end{pmatrix}\right) \quad (4)$$

Where $\mathbf{K}(\mathbf{X}, \mathbf{X})$ is the kernel matrix evaluated for all training input pairs, $\mathbf{K}(\mathbf{X}, \mathbf{x}_*)$ is the kernel vector between training inputs $\mathbf{X}$ and the test input $\mathbf{x}_*$, $\mathbf{I}$ is the identity matrix, $\mathbf{k}(\mathbf{x}_*, \mathbf{x}_*)$ is the variance of the test point with itself.

For prediction purposes, the GPR is obtained by computing the posterior distribution for f_* :

$$p(f_* | \mathbf{X}, \mathbf{y}, \mathbf{x}_*) = \mathcal{N}(\mu_*, \Sigma_*) \quad (5)$$

$$\mu_* = \mathbf{K}(\mathbf{X}, \mathbf{x}_*)^{\text{transpose}} [\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}]^{-1} \mathbf{y} \quad (6)$$

$$\Sigma_* = \mathbf{k}(\mathbf{x}_*, \mathbf{x}_*) - \mathbf{K}(\mathbf{X}, \mathbf{x}_*)^{\text{transpose}} [\mathbf{K}(\mathbf{X}, \mathbf{X}) + \sigma_n^2 \mathbf{I}]^{-1} \mathbf{K}(\mathbf{X}, \mathbf{x}_*) \quad (7)$$

Where μ_* is predicted chlorophyll content, Σ_* is predictive variance (confidence intervals), using for quantifying uncertainty.

Considering VHGP, the noise variance σ_n^2 becomes $\sigma_n^2(\mathbf{x})$, indicating that the noise variance can vary depending on the input $\mathbf{x}$:

$$\mathbf{Y} = f(\mathbf{x}) + \epsilon, \text{ and } \epsilon \sim \mathcal{N}(0, \mathbf{e}^{g(\mathbf{x})}) \quad (8)$$

Here, $f(\mathbf{x})$ is the $\mathcal{GP}$ for the mean function, and $g(\mathbf{x})$ is another latent $\mathcal{GP}$ that models the logarithm of the noise variance:

$$f(\mathbf{x}) \sim \mathcal{GP}(\mathbf{m}_f(\mathbf{x}), \mathbf{k}_f(\mathbf{x}, \mathbf{x}')) \quad (9)$$

$$g(\mathbf{x}) \sim \mathcal{GP}(\mathbf{m}_g(\mathbf{x}), \mathbf{k}_g(\mathbf{x}, \mathbf{x}')) \quad (10)$$

The log-transformation ensures that the noise variance is always positive. To perform inference in VHGP, the Marginalized Variational approximation approach is employed. An analytical expression for the Kullback–Leibler divergence between a proposal distribution and the true posterior distribution of $f(\mathbf{x})$ and $g(\mathbf{x})$ was provided. Minimizing this

divergence yields an accurate approximation of the true posterior distribution. This optimization simultaneously facilitates model selection.

Regarding band importance analysis in Gaussian-based algorithms, Automatic Relevance Determination (ARD) is a technique used in Bayesian models to identify and prioritize the most relevant input features for a given task. With input vector $\mathbf{x}$ has D dimensions (i.e., spectral bands), the ARD Squared Exponential kernel is defined as:

$$k(\mathbf{x}_i, \mathbf{x}_j) = v^2 \exp\left(-\frac{1}{2} \sum_{d=1}^D \frac{(\mathbf{x}_{id} - \mathbf{x}_{jd})^2}{l_d^2}\right) \quad (11)$$

Where v is a scaling factor, $\mathbf{x}_{id}$ and $\mathbf{x}_{jd}$ is the value of the d -th spectral band for input point $\mathbf{x}_i$ and $\mathbf{x}_j$, l_d is the length-scale per input bands.

During the GPR training process, the hyperparameters (v and l_d) are optimized by maximizing the marginal likelihood of the observed data. The obtained l_d represents the relevance of each spectral band. A lower band l_d indicates higher feature importance for the predictions and vice versa.

According to the mentioned theoretical basis, there are two properties used for uncertainty quantification and band relevance analysis in this study:

(1) Both probabilistic Gaussian models allow for delivering uncertainty level (or confidence interval) per pixel accompanied by trait predictions (Camps-Valls et al., 2016; Verrelst et al., 2013b). The pixel-wise absolute uncertainty, calculated by predictive variance (Σ_*), which can be expressed as the standard deviation (SD) around the mean trait value (μ). To compare retrieval models at a consistent scale, we applied the coefficient of variation (CV) as a measure of relative uncertainty, represented by percentage:

$$CV = (SD/\mu) \times 100 \quad (12)$$

(2) The length-scale per input bands (l_d) derived from the Gaussian models to rank relevant S2 bands (Verrelst et al., 2016b). As aforementioned, a lower l_d indicates higher importance. Consequently, these values were inverted to achieve a positive representation, with higher values denoting greater importance. A logarithmic transformation was then applied to normalize the scale, effectively handling the substantial differences in absolute values across bands. Finally, the relative importance was computed to facilitate clearer interpretation and comparative analysis of spectral relevance among the six models (Binh et al., 2024):

$$\text{Relative importance} = 1 - \log_{10}\left(\frac{l_d^2}{(\max l_d^2)}\right) \quad (13)$$

We introduce multiplicative Gaussian noise to improve the practical utility of the retrieval model and prevent overfitting. Prior to training, we applied noise at increasing levels of 5 % and 10 % to assess its impact compared to the noise-free scenario (i.e., 0 %). Importantly, this allows us to investigate how noise affects model uncertainty and the contribution of individual spectral bands to the retrieval process.

Due to the ill-posedness of RTM inversion (i.e., multiple configured parameters may produce similar spectral responses) and the presence of noise, significant challenges arise when linking PROSAIL data with S2 observation. Hence, we integrated random sampling techniques with Gaussian algorithms to improve model accuracy (Berger et al., 2021; Verrelst et al., 2016a). Finally, given that PROSAIL is exclusively designed for vegetation simulation, spectral reflectances of non-vegetated covers (i.e., built-up areas, water, and barren land), were extracted directly from S2 imagery and incorporated into the training dataset with zero chlorophyll content. This strategic inclusion allows for training a model that performs effectively on actual satellite imagery, which inherently features coexisting vegetated and non-vegetated surfaces, thereby enabling robust mapping across the entire S2 scene. This methodology is also consistent with established precedents in prior studies that aim to build robust models for practical applications (Binh et al., 2022a; Estévez et al., 2020; Estévez et al., 2022; García-Soria

et al., 2024).

2.6. Statistical metrics

The performance of Gaussian-based models under varying noise levels was evaluated against in situ data using linear regression analysis. Three statistical criteria were employed for model assessment: the coefficient of determination (R^2), the relative root mean square error (RRMSE), and the index of agreement (IOA). RRMSE is expressed as a percentage of the mean, providing a normalized measure of model error compared to absolute RMSE, while IOA quantifies the overall agreement between model predictions and observed data, ranging from 0 (no agreement) to 1 (perfect agreement), facilitating a better assessment of the predicted models (Willmott, 1981).

$$R^2 = 1 - \frac{\sum_{i=1}^n (\mathbf{x}_i - \mathbf{y}_i)^2}{\sum_{i=1}^n (\mathbf{x}_i - \bar{\mathbf{x}})^2} \quad (14)$$

$$\text{RRMSE} = 100 \times \frac{\sqrt{\frac{1}{n} \sum_{i=1}^n (\mathbf{y}_i - \mathbf{x}_i)^2}}{\bar{\mathbf{x}}} \quad (15)$$

$$\text{IOA} = 1 - \frac{\sum_{i=1}^n (\mathbf{y}_i - \mathbf{x}_i)^2}{\sum_{i=1}^n (|\mathbf{y}_i - \bar{\mathbf{y}}| + |\mathbf{x}_i - \bar{\mathbf{x}}|)^2} \quad (16)$$

Where, n is the number of samples ($n = 60$), $\mathbf{y}_i$ and $\mathbf{x}_i$ are the estimated and measured paddy chlorophyll values, $\bar{\mathbf{x}}$ are the mean of measured paddy chlorophyll values.

3. Results

3.1. Field validation

Fig. 3 exemplifies linear regression plots and goodness-of-fit metrics for chlorophyll retrieval compared between various Gaussian-based algorithms with varying levels of noise. Overall, all the models achieved robust performance with RRMSE below 5 %. Notably, the standard GPR with noise-free demonstrated the best performance with all out-performed metrics ($R^2 = 0.725$, RRMSE = 3.824 %, and IOA = 0.919). While all models maintained reasonable performance with noise injection, a significant change in relative uncertainty was observed. For instance, relative uncertainty rose to approximately 25 % with 5 % noise in both traditional GPR and refined VHGP models. Evidence from all error metrics also confirmed VHGP with 10 % noise performed slightly better compared to the same model with 5 % noise, yet the difference is negligible. Comparison between standard GPR and VHGP indicates that standard GPR generally provides higher accuracy, while both are sensitive to noise.

The magnitude of noise levels demonstrably impacted the relative contribution of spectral bands in each specific Gaussian retrieval algorithm (Fig. 4). Specifically, with the exception of the 5 % noise condition, the blue (B2–490 nm) and green (B3 – 560 nm) bands emerged as more influential than the red band (B4–665nm) within the Sentinel-2 visible spectral range, as confirmed by prior studies (Verrelst et al., 2012b; Verrelst et al., 2016b). In the red-edge region, the bands centered at 705 nm (B5) and 740 nm (B6) displayed the highest sensitivity, consistent with previous investigations (Clevers and Gitelson, 2013; Delegido et al., 2011). In the near-infrared (NIR) region, the contribution of band B8 (wide spectral width, 10 m resolution) and 8 A (narrow spectral width, 20 m resolution) to the model varied depending on the specific scenario. Furthermore, the inclusion of shortwave infrared (SWIR) bands at 1610 nm (B11) and 2190 nm (B12) significantly enhanced model performance, attributed to their diminished atmospheric interference, as well as heightened sensitivity to soil and vegetation cover characteristics rather than the visible bands (Delloye et al., 2018; Verrelst et al., 2012b).

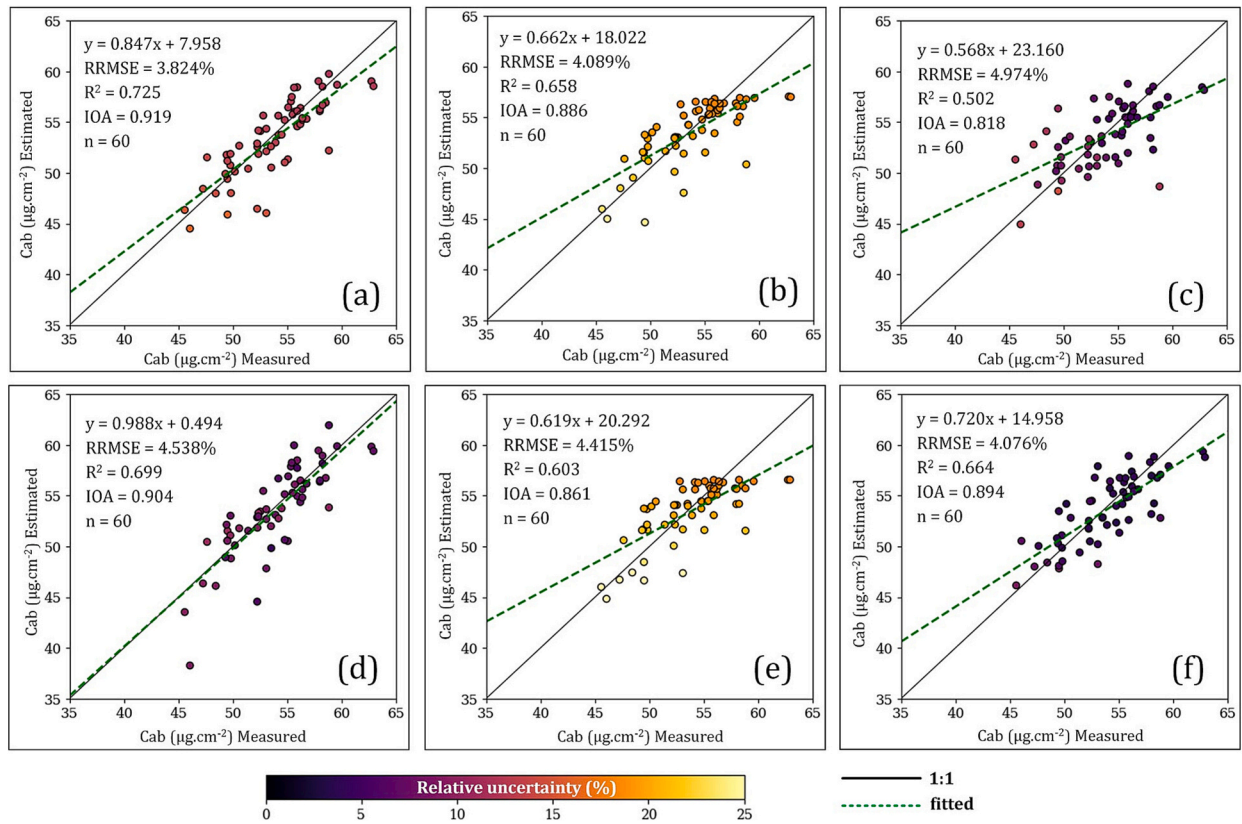


Fig. 3. Performance of Gaussian-based algorithms under varying noise conditions: (a-c) GPR with 0.5, 10, 10 % noise, (d-f) VHGR with 0.5, 10, 10 % noise.

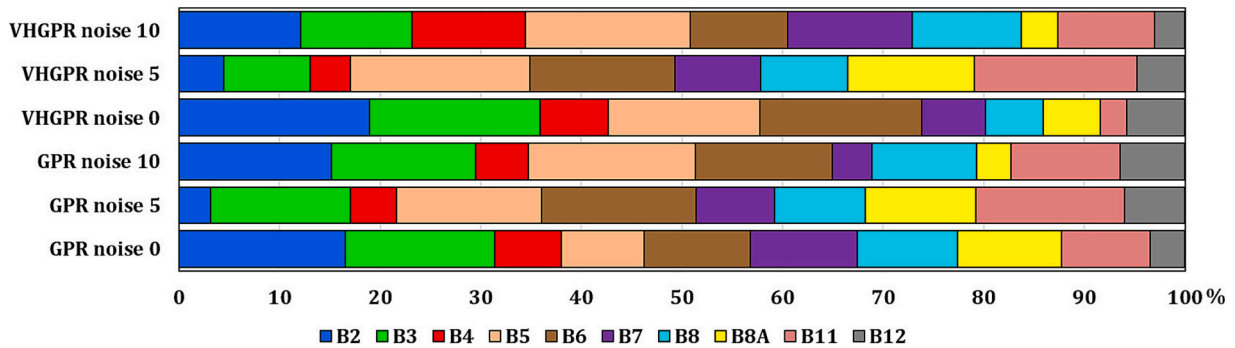


Fig. 4. Hierarchical partitioning represents the percentage of relative importance explained in different retrieval models.

3.2. Mapping performance

Chlorophyll mapping applied to the entire S2 scene accurately represented paddy characteristics compared to remaining land cover characteristics, as well as the differences following specific noise-included Gaussian algorithms (Fig. 5). Model training data may be biased towards paddy crops, hampering accurate trait retrieval on disturbed private land cover properties. Consistency patterns of rice areas were observed in all models, with the range of chlorophyll content from approximately 50 to nearly $95 \mu\text{g.cm}^{-2}$, while the non-vegetation pixel (i.e., built-up, water bodies) likely leads to lower trait levels ($<10 \mu\text{g.cm}^{-2}$).

Along with the chlorophyll mapping procedure, the models allow for generating pixel-wise relative uncertainty, facilitating for comparison between different landcover attributes over complex agricultural environments (Fig. 6). The spatial distribution of relative uncertainties revealed that the highest values belong to water bodies, followed by low

vegetation cover such as built-up areas. Major paddy fields and other vegetation cover performed lower relative uncertainties, implying trained models achieved higher retrieval fidelity over the vegetation pixel compared to non-vegetation pixels. The pronounced differences between vegetation and non-vegetation types highlight the impacts of heterogeneous land cover using spatial and spectral specifications of S2. To the noise-injected Gaussian algorithms, the most fidelity performance was achieved in GPR with 0 % noise (on average, relative uncertainty <20 %), while the highest uncertainties were detected clearly in VHGR in the presence of both noise-free and 10 % included noise (relative uncertainty up to 100 %).

We sampled all exhibited paddy pixels over the study site and performed the correlation matrix (Fig. 7). This enables the relationship analysis between different specific noise-included Gaussian algorithms. For trait maps, results indicated a moderate-to-high correlation, with the lowest correlation found between noise-free VHGR and VHGR with 5 % added noise ($R = 0.58$), and the highest through the comparison

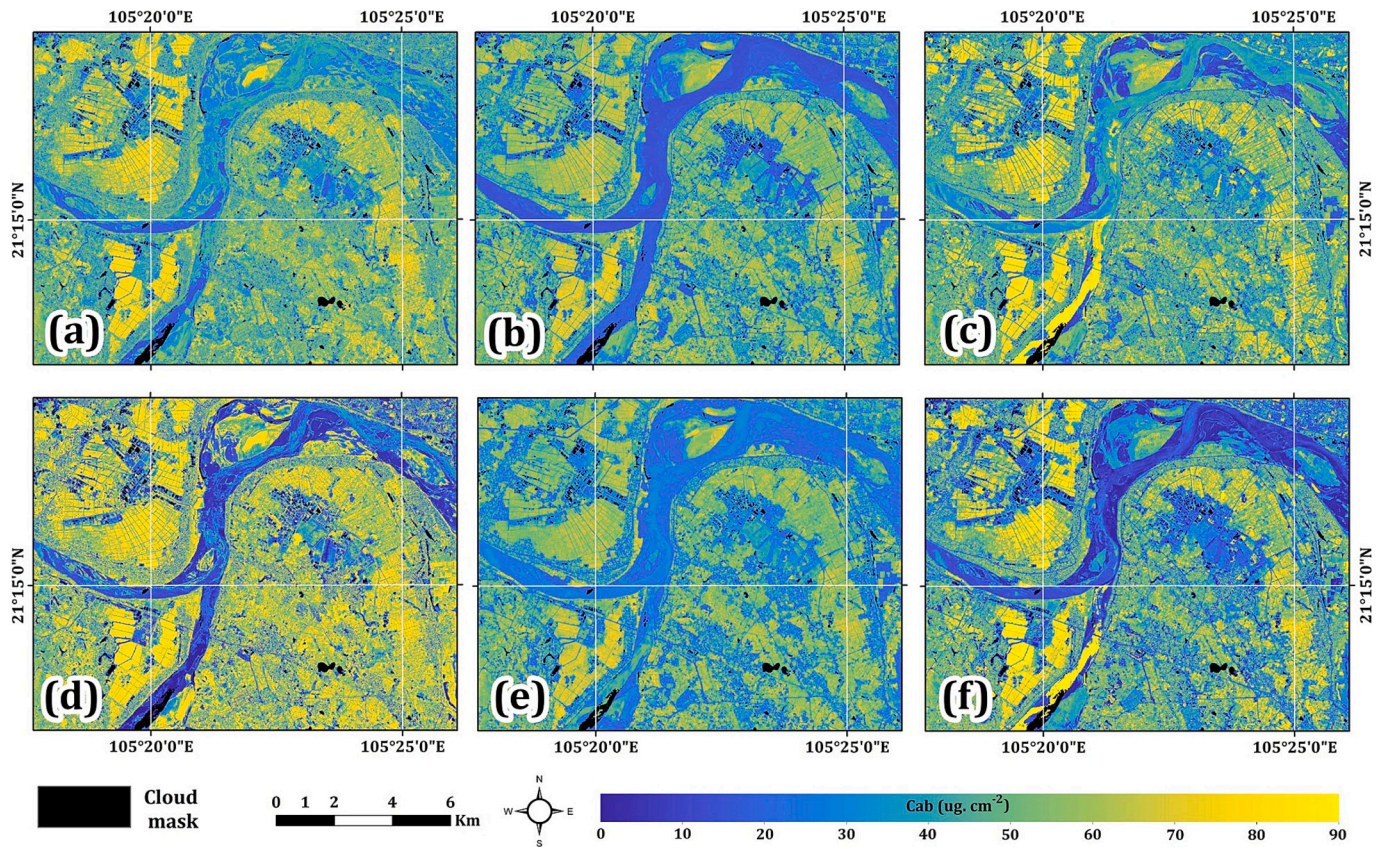


Fig. 5. Visual comparison of chlorophyll predicted by noise-embedded Gaussian model: (a-c) GPR with 0.5, 10 % noise, (d-f) VHGR with 0.5, 10 % noise.

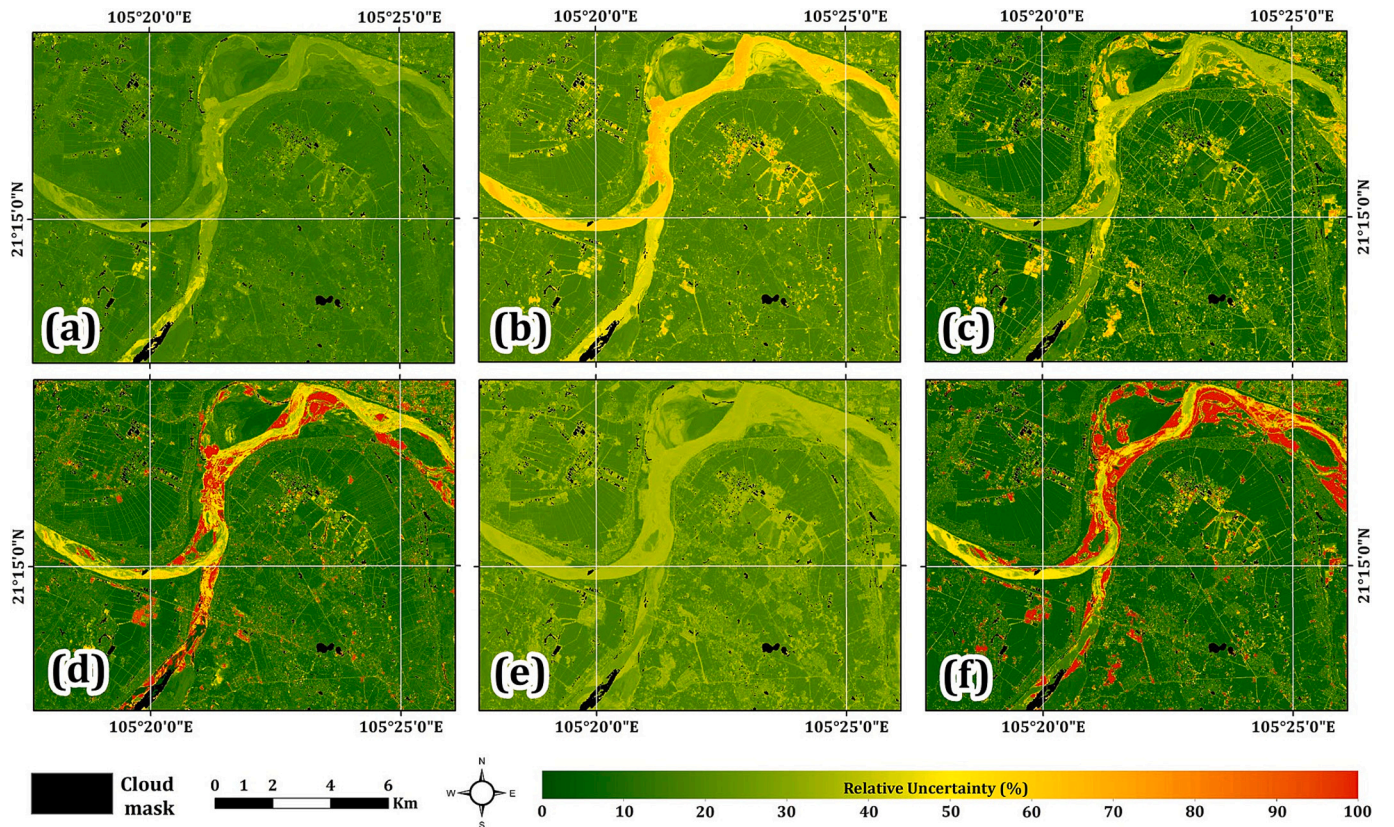


Fig. 6. Visual comparison of relative uncertainty predicted by noise-embedded Gaussian model: (a-c) GPR with 0.5, 10 % noise, (d-f) VHGR with 0.5, 10 % noise.

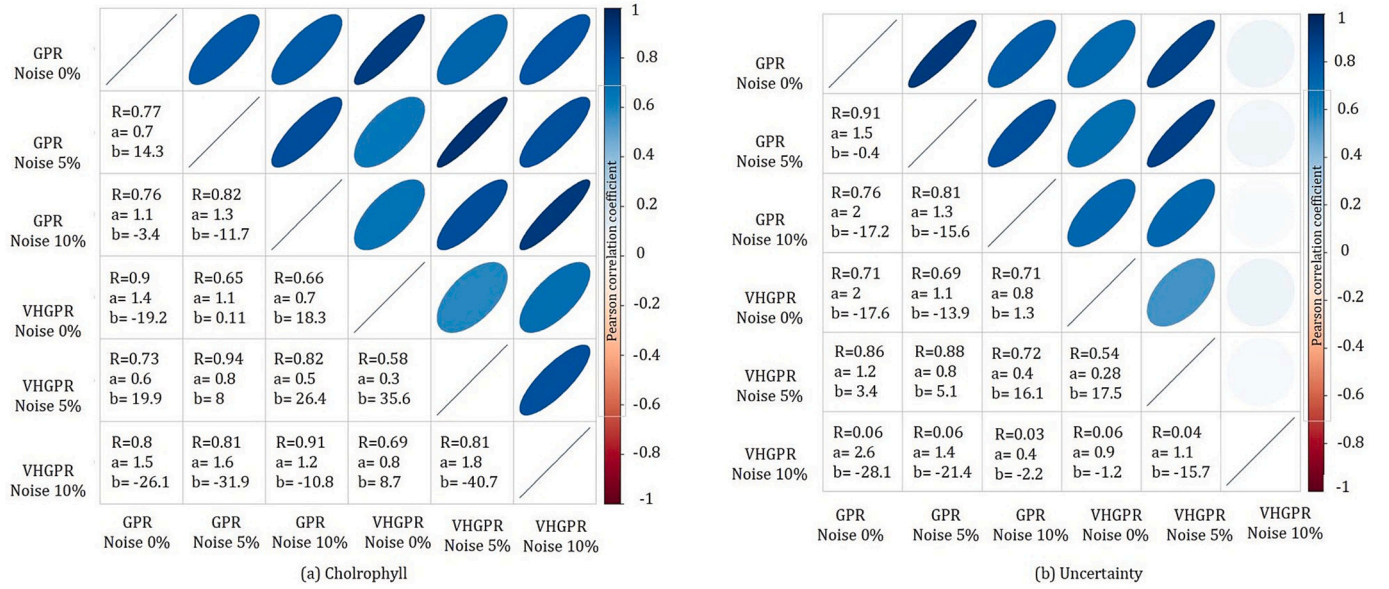


Fig. 7. Correlation matrix between chlorophyll and uncertainty prediction of paddy pixels over the study area.

between GPR and VHGP with the same 5 % noise ($R = 0.94$). Concerning relative uncertainty, the range correlation coefficient varies from the lowest 0.03 (noise 10 %) to the highest 0.91 (GPR noise 0 % and 5 %). Notably, VHGP with 10 % noise included is the only model shows no significant correlation with other models.

It should be noted that for the heterogeneous landscape, the rice crops are mixed with non/other vegetation cover. Therefore, a better illustration of mean uncertainty compared between different land cover attributes and in situ data is depicted in Fig. 8. All vegetation cover (i.e., rice, forest, and orchards) achieved the lowest relative uncertainty with negligible differences, as well as the same performance against in situ data. Whereas the highest mean relative uncertainty occurred in barren areas (exceeding 50 % in VHGP noise 10 %), the land cover characteristic also exhibited the largest variation in different noise levels and Gaussian-based estimations compared to the remaining classes. The variation of relative uncertainty in VHGP is generally higher than GPR, while GPR noise 0 % achieved the lowest mean relative uncertainty across all land cover types. In summary, most of the uncertainty estimated for not only paddy but also other vegetation (i.e., orchards and forests) is equivalent to the in-situ uncertainty, indicating a high fidelity obtained by all plant chlorophyll models.

To precisely focus on active rice cultivation, we removed boundary pixels from paddy parcels based on relative uncertainty thresholds. We then analysed the pixel density distribution leveraging the cumulative distribution function to statistically compare the range of paddy chlo-

rophyll with relative uncertainty (Fig. 9). The model VHGP noise 10 % displayed the largest dispersion between minimum and maximum chlorophyll content, spanning from the lowest $48 \mu\text{g}\cdot\text{cm}^{-2}$ to the highest $95 \mu\text{g}\cdot\text{cm}^{-2}$. Noise 5 % included in both GPR and VHGP showed minimum dispersion with pixel densities (80 %) falling within the range of $55\text{--}65 \mu\text{g}\cdot\text{cm}^{-2}$. Accompanied by trait values, the relative uncertainty range obtained in GPR noise 10 %, VHGP noise 0 % and 10 % (2–7 %) was somewhat lower than the case of 5 % noise injected to both Gaussian algorithms (16–18 %). Notwithstanding the lowest relative uncertainty achieved in VHGP noise 10 % (on average: 5 %), there are some outliers in chlorophyll estimation, implying the anomaly introduced by noise influences. All in all, confirmation for the trustful retrieval models within rice crops accompanied by major low relative uncertainty was achieved in all cases (i.e., the highest around 20 %).

4. Discussion

Based on the achieved outcomes, we discuss notable aspects encompassing (1) quantifying uncertainty in satellite-derived paddy chlorophyll over heterogeneous agricultural landscapes, (2) the influence of noise injection on Gaussian algorithms within hybrid retrieval models, and (3) scope and challenges in linking physics-based knowledge of plant-light interaction to remote sensing observations for spatiotemporal vegetation trait monitoring.

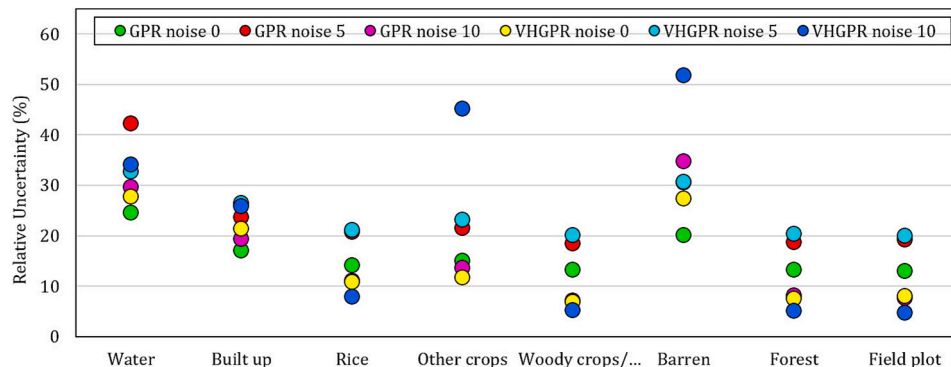


Fig. 8. Visual plot shows the mean relative uncertainty of land cover attributes across all noise-injected Gaussian algorithms.

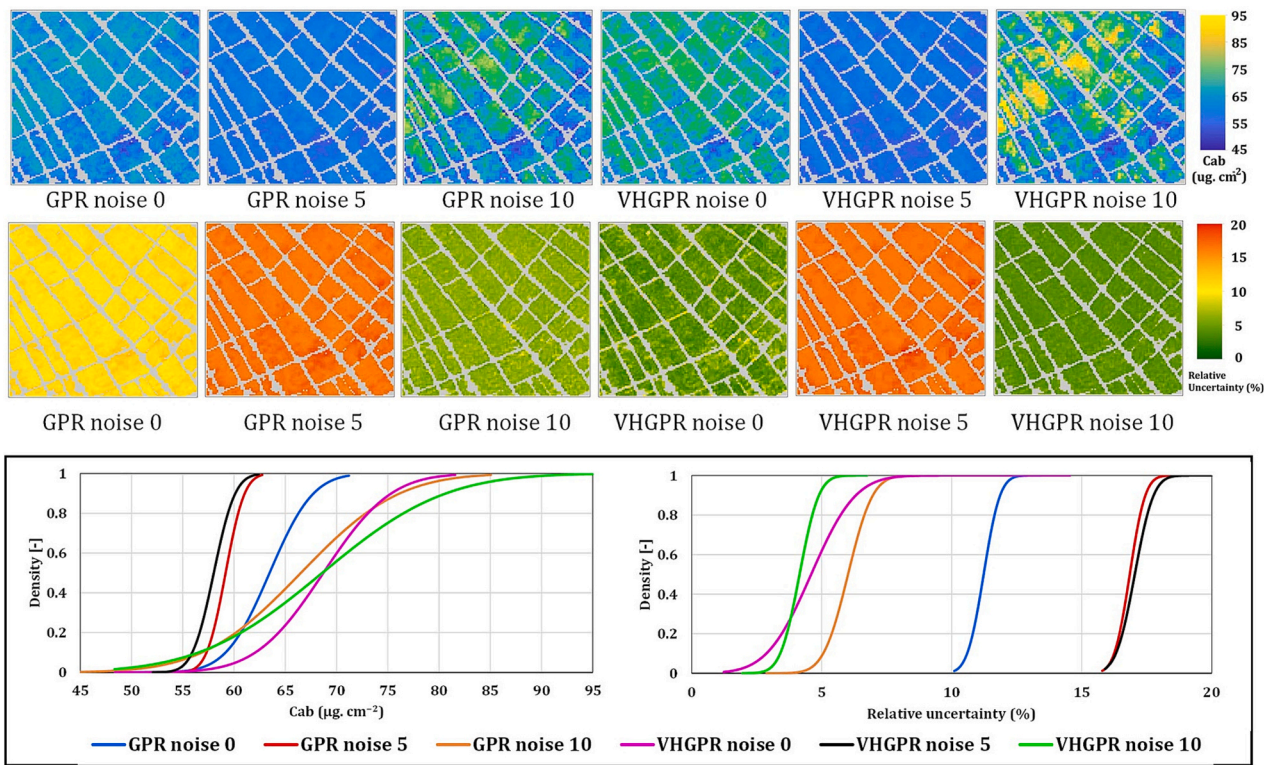


Fig. 9. Spatial distribution of chlorophyll content (top), relative uncertainty (middle), and cumulative distribution function (bottom) over paddy parcels.

4.1. Quantifying uncertainty

We first emphasize the role of uncertainty mapping coupled with satellite-derived plant trait products. Recent studies have demonstrated the benefit of incorporating uncertainty to enhance retrieval model transferability. Estévez et al. (2022) applied this approach to homogeneous agricultural foliage across different geographical areas, while Binh et al. (2024) tested it on structurally complex mangrove forests for time series prediction. These highlight the role of uncertainty in evaluating model transferability across space and time dimensions without further field campaigns. In this work, building upon the insights from uncertainty at different geographical and temporal scales, we explore a different facet of uncertainty to assess the capacity of the retrieval model within a small local area characterized by complex mixed land cover. Specifically, the uncertainty estimates provided by GPR/VHGP can signal areas where the model is less confident due to spectral variability, which is especially valuable in heterogeneous landscapes. This enhancement allows for more nuanced interpretations of the output chlorophyll maps, even when transferring the model to areas that might differ from the primary training region or for time series predictions.

Our aforementioned results indicated the remarkable influence in mapping the whole complex agricultural environment (Fig. 5, Fig. 6, and Fig. 8). Here, pixel-level uncertainty represents an indication of model trustworthiness across various land cover attributes, from non-vegetated areas (built-up areas, water bodies, and barren land) to diverse vegetation types (paddy rice, other crops, forests, and orchards). Additionally, an uncertainty threshold was implemented to eliminate edge pixels from separated paddy parcels, thereby ensuring our analysis concentrated on pure paddy pixels (Fig. 9). Consequently, the uncertainty maps can be used as quality masks corresponding with ready-to-use trait products. This combination allows for reliable analysis of the target vegetation and provides end-users guidelines for appropriate use (Kaminski et al., 2017; Malenovsky et al., 2019).

Our study focused on model uncertainty (Malenovsky et al., 2019), reflecting the combined sensitivity of the radiative transfer model and

the Gaussian algorithms. GPR and VHGP models, trained using data simulated by PROSAIL (Table 1), were developed specifically for paddy rice, representing the high uncertainty for non-vegetation land covers. At present, our model confidence analysis is confined to the Bayesian properties inherent in GPR and VHGP (Rasmussen and Williams, 2005). However, recent techniques like Quantile Regression Forest, Drop-out, and Bootstrapping offer quantified uncertainty applicable to various MLRAs (García-Soria et al., 2024). This opens avenues for further investigation into the applicability of these uncertainty measures when deploying diverse MLRAs for retrieving biophysical and biochemical variables. Moreover, future research should encompass both epistemic uncertainty (i.e., model uncertainty) and aleatoric uncertainty (i.e., data uncertainty), which currently represents a significant limitation for evaluating the reliability of most state-of-the-art remote sensing trait products (Martínez-Ferrer et al., 2022).

4.2. Noise variability

Beyond the inherent model uncertainty, it is crucial to acknowledge additional sources of errors, including image pre-processing and field data collection. To mitigate potential issues in image pre-processing, we utilized orthorectified atmospherically corrected S2 surface reflectance data (level 2 A) alongside pixel-based quality masking (see Section 2.3), which diminishes the influences of clouds, aerosols, and atmospheric gases in the optical wavelengths (Gascon et al., 2017). Furthermore, the field sampling protocol may introduce accumulated errors due to intra-leaf variability, canopy structure, soil/water background interference, and uncertainties in the conversion formula for non-destructive SPAD measurements. While we ensured homogeneous sampling and minimized background effects (see Section 2.2), future work should more explicitly account for these error sources. This includes conducting a comprehensive uncertainty analysis of field measurements and rigorously calibrating and validating the conversion formula for specific experimental setups. Importantly, recognizing the source of aforementioned errors may impact model performance, we therefore introduced

noise into the training data and assessed the natural variability in model performance, pixel uncertainty, and bands contribution.

Noise demonstrated notable impacts on model operation in rice chlorophyll prediction. While evaluation against in situ data is almost acceptable in all cases (Fig. 3), noise affected the spectral relevance bands contributed to models (Fig. 4). In addition, the presence of noise in mapping performance taking into account all land cover types is further amplified in both trait (Fig. 5) and uncertainty thereof (Fig. 6), as well as density scatterplots of correlation (Fig. 7). This highlights the specific noise-handling characteristics of specific algorithm architecture (de Sá et al., 2021; García-Soria et al., 2024).

The unusual variation observed at the 5 % noise level can be partly explained by the intrinsic mathematical properties of inversion methods (i.e., Gaussian algorithms) and parameterization of the rice simulation model (i.e., PROSAIL). In this study, we employed Gaussian algorithms as a benchmark regression model due to their demonstrated capability to handle noise variance (Estévez et al., 2021; García-Soria et al., 2024; Lázaro-Gredilla et al., 2014). For PROSAIL configuration, simulated data were primarily generated based on field-derived rice chlorophyll values, which may not fully reflect complex agricultural environments. For instance, the 5 % noise level when calculated across the entire scene appears to be uniform. This is explained by the low variability of chlorophyll value across all land cover types (Fig. 5b, e), resulting in low SD and reduced relative uncertainty (Fig. 6b, e). However, when focusing specifically on paddy parcels, the 5 % noise level exhibits the highest uncertainty compared to the remaining models (Figs. 7, 9). This indicates the rice models at this noise level are the least reliable, consistent with the accuracy assessment (Fig. 3).

4.3. Scope and challenges

Our retrieval approach leverages the strengths of MLRAs to link a physics-based radiative transfer foundation and remotely sensed observation, forming effective hybrid models. Notwithstanding the inherent robustness of empirical methodologies, hybrid strategies offer promising potential (Verrelst et al., 2019). Firstly, hybrid models bridge gaps in connecting fundamental physical principles with empirical observations, thereby enhancing both scientific knowledge and practical applications. Secondly, hybrid methods address the inherent limitations of empirical models, which are often constrained to specific times and locations. Finally, the integration of uncertainty estimation further enhances the model transferability, enabling seamless retrieval across temporal and spatial dimensions. This facilitates robust long-term tracking and large-scale environmental monitoring without repeated field sampling.

Applying retrieval models to freely accessible multispectral S2 images unlocks the spatiotemporal monitoring of key paddy rice traits, a crucial agricultural practice in Vietnam. This is likely the optimal solution currently given the scarcity of imaging spectroscopy data over the examined region (e.g., hyperspectral missions). The frequent revisits of the S2 constellation enable phenological analysis of seasonal paddy crops, identified as dual or triple seasons depending on the currently complex farming mechanism. Noteworthy, the recent deployment of satellite Sentinel 2C enhances a substantial increase in observation frequency, shortening the revisit period from approximately 5 days to roundly 2–3 days, facilitating tracking rapidly changing phenomena over paddy crops (Jia et al., 2024). Beyond Sentinel-2, exploring multi-source data strategies is also a key direction for future research to achieve holistic and precise paddy trait monitoring. This includes hyperspectral data with finer spectral detail (Moharana and Dutta, 2016; Tagliabue et al., 2022), high spatial resolution imagery acquired by unmanned aerial vehicle (UAV) platforms (Ban et al., 2022; Xu et al., 2019), combining it with Light Detection and Ranging (LiDAR) for three-dimensional plant chlorophyll mapping (Xu et al., 2024), and data fusion techniques for extending multi-decadal data series beyond individual satellite acquisition limits (D.Kovács et al., 2024). The pervasive

cloud cover characteristic of tropical monsoon climates in Vietnam presents a further challenge to optical observation applications (Binh et al., 2022b; Binh et al., 2024; Hoa et al., 2025). This necessitates post-processing techniques or multi-sensor approaches to obtain gap-free multispectral-derived products (Caballero et al., 2023; Kandasamy et al., 2013; Kovács et al., 2023).

In the present study, chlorophyll pigment, one of the Essential Biodiversity Variables representing vegetation photosynthesis (Kissling et al., 2018), was selected as an exemplified function trait. We targeted rice as one of the most worldwide food plants and presented a valuable case study in Vietnam, currently ranked third in global rice production. Nonetheless, the presented methodology offers a generalized framework for vegetation trait retrieval and uncertainty estimation to a wide range of functional plant traits and multispectral Earth observation missions. The model's inherent adaptability suggests its potential for mapping diverse agronomical variables in Vietnam, which is presently characterized by patchwork agricultural landscapes with numerous small-scale farming patterns. Future work will involve acquiring comprehensive training datasets for various vegetation types to enable accurate and robust mapping in these heterogeneous environments.

Unlike purely “black-box” models, Gaussian-based algorithms can be considered a “grey-box” or partially interpretable model (Martinez-Ferrer et al., 2021; Mateo-Sanchis et al., 2021). Here, their interpretability is primarily demonstrated through two components: uncertainty quantification and variable importance analysis. Regarding uncertainty, its explicit quantification offers critical insight into model confidence for each prediction (i.e., per different land cover attributes), thereby enhancing interpretability and proving vital for guiding targeted supplementary field investigations to improve overall model accuracy. Concerning the assessment of variable importance, quantifying the contribution of different input variables (e.g., specific spectral bands or vegetation indices) provides significant insights for remote sensing-based chlorophyll estimation in rice. This not only facilitates a deeper understanding of the input-output relationships within the model, but also reinforces and validates domain knowledge, ultimately supporting the optimization of future data collection strategies (Gevaert, 2022; Höhl et al., 2024).

5. Conclusion

We presented a comprehensive workflow for mapping paddy chlorophyll content using S2 data and hybrid models, explicitly addressing the influence of noise and pixel-level uncertainty within heterogeneous agricultural landscapes. The models were trained using physically-based PROSAIL simulations, incorporating six combinations derived from two Gaussian algorithms (GPR and VHGR) and three noise levels (0 %, 5 %, and 10 %). While all models achieved comparable performance (on average, IOA > 0.8, RRMSE < 5 %), their specific algorithm architectures showed different responses to the imposed noise levels in terms of accuracy, band relevance, and associated uncertainties. Mapping performance over paddy pixels demonstrated high confidence in chlorophyll prediction, with mean relative uncertainties below 25 %. However, transferring these models for full trait mapping across diverse land cover attributes revealed mean relative uncertainties reached more than 50 % for non-vegetated areas. The generated relative uncertainty layers associated with the trait maps provide a valuable masking tool for operational applications, such as the effective removal of non-vegetation features within complex environmental landscapes. Ultimately, we emphasized that quantifying the impact of noise on plant functional trait maps, including uncertainty information is a prerequisite for meeting spatiotemporal-explicit standards in agricultural monitoring. This is fundamental for remote sensing applications in sustainable, smart, and precision agriculture.

CRediT authorship contribution statement

Giang Thi Phuong Thao: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Nguyen An Binh:** Writing – review & editing, Writing – original draft, Visualization, Validation, Supervision, Software, Resources, Project administration, Methodology, Funding acquisition, Formal analysis, Data curation, Conceptualization. **Tong Si Son:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Resources, Investigation, Formal analysis, Data curation, Conceptualization. **Le Thi Thu Ha:** Writing – review & editing, Writing – original draft, Validation, Supervision, Software, Resources, Methodology, Data curation, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

This study was supported by project VAST05.02/24-25 from the Vietnam Academy of Science and Technology (VAST), for which we are very thankful. We thank the Vietnam Academy of Science and Technology (VAST) for funding the research project grant number VAST05.02/24-25.

Data availability

All data in this study are available in <https://doi.org/10.5281/zenodo.15811808>

References

- Adeluyi, O., Harris, A., Verrelst, J., Foster, T., Clay, G.D., 2021. Estimating the phenological dynamics of irrigated rice leaf area index using the combination of PROSAIL and Gaussian process regression. *Int. J. Appl. Earth Obs. Geoinf.* 102. <https://doi.org/10.1016/j.jag.2021.102454>.
- Ban, S., Liu, W., Tian, M., Wang, Q., Yuan, T., Chang, Q., Li, L., 2022. Rice leaf chlorophyll content estimation using UAV-based spectral images in different regions. *Agronomy* 12. <https://doi.org/10.3390/agronomy12112832>.
- Berger, K., Verrelst, J., Féret, J.B., Hank, T., Wocher, M., Mauser, W., Camps-Valls, G., 2020. Retrieval of aboveground crop nitrogen content with a hybrid machine learning method. *Int. J. Appl. Earth Obs. Geoinf.* 92. <https://doi.org/10.1016/j.jag.2020.102174>.
- Berger, K., Pablo, J., Caicedo, R., Martino, L., Wocher, M., Hank, T., 2021. A survey of active learning for quantifying vegetation traits from terrestrial earth observation data. *Remote Sens.* 13, 1–23.
- Berger, K., Machwitz, M., Kycko, M., Kefauver, S.C., Van Wittenberghe, S., Gerhards, M., Verrelst, J., Atzberger, C., van der Tol, C., Damm, A., Rascher, U., Herrmann, I., Paz, V.S., Fahrner, S., Pieruschka, R., Prikaziuk, E., Buchailot, M.L., Halabuk, A., Celesti, M., Koren, G., Gormus, E.T., Rossini, M., Foerster, M., Siegmann, B., Abdelbaki, A., Tagliabue, G., Hank, T., Darvishzadeh, R., Aasen, H., Garcia, M., Pôças, I., Bandopadhyay, S., Sulis, M., Tomelleri, E., Rozenstein, O., Filchev, L., Stancile, G., Schlerf, M., 2022. Multi-sensor spectral synergies for crop stress detection and monitoring in the optical domain: a review. *Remote Sens. Environ.* <https://doi.org/10.1016/j.rse.2022.113198>.
- Binh, N.A., Hauser, L.T., Viet Hoa, P., Thi Phuong Thao, G., An, N.N., Nhut, H.S., Phuong, T.A., Verrelst, J., 2022a. Quantifying mangrove leaf area index from Sentinel-2 imagery using hybrid models and active learning. *Int. J. Remote Sens.* 1–22. <https://doi.org/10.1080/01431161.2021.2024912>.
- Binh, N.A., Hoa, P.V., Thao, G.T.P., Duan, H.D., Thu, P.M., 2022b. Evaluation of chlorophyll-a estimation using sentinel 3 based on various algorithms in southern coastal Vietnam. *Int. J. Appl. Earth Obs. Geoinf.* 112, 102951. <https://doi.org/10.1016/j.jag.2022.102951>.
- Binh, N., Hauser, L.T., Salinero-Delgado, M., Viet Hoa, P., Thi Phuong Thao, G., Verrelst, J., 2024. Monitoring mangrove traits through optical earth observation: towards spatio-temporal scalability using cloud-based Sentinel-2 continuous time series. *ISPRS J. Photogramm. Remote Sens.* 214, 135–152. <https://doi.org/10.1016/J.ISPRSJPRES.2024.06.007>.
- Brodu, N., 2017. Super-resolving multiresolution images with band-independent geometry of multispectral pixels. *IEEE Trans. Geosci. Remote Sens.* 55. <https://doi.org/10.1109/TGRS.2017.2694881>.
- Brown, L.A., Camacho, F., García-Santos, V., Origo, N., Fuster, B., Morris, H., Pastor-Guzman, J., Sánchez-Zapero, J., Morrone, R., Ryder, J., Nightingale, J., Boccia, V., Dash, J., 2021. Fiducial reference measurements for vegetation biogeophysical variables: an end-to-end uncertainty evaluation framework. *Remote Sens.* 13. <https://doi.org/10.3390/rs13163194>.
- Caballero, G., Pezzola, A., Winschel, C., Sanchez Angonova, P., Casella, A., Orden, L., Salinero-Delgado, M., Reyes-Muñoz, P., Berger, K., Deleigido, J., Verrelst, J., 2023. Synergy of Sentinel-1 and Sentinel-2 time series for cloud-free vegetation water content mapping with multi-output Gaussian processes. *Remote Sens.* 15. <https://doi.org/10.3390/rs15071822>.
- Caicedo, J.P.R., Verrelst, J., Munoz-Mari, J., Moreno, J., Camps-Valls, G., 2014. Toward a semiautomatic machine learning retrieval of biophysical parameters. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 7. <https://doi.org/10.1109/JSTARS.2014.2298752>.
- Camps-Taberner, M., García-Haro, F.J., Camps-Valls, G., Grau-Muedra, G., Nutini, F., Crema, A., Boschetti, M., 2016. Multitemporal and multiresolution leaf area index retrieval for operational local rice crop monitoring. *Remote Sens. Environ.* 187. <https://doi.org/10.1016/j.rse.2016.10.009>.
- Camps-Valls, G., Verrelst, J., Munoz-Mari, J., Laparra, V., Mateo-Jimenez, F., Gomez-Dans, J., 2016. A survey on Gaussian processes for earth-observation data analysis: a comprehensive investigation. *IEEE Geosci. Remote Sens. Mag.* <https://doi.org/10.1109/MGRS.2015.2510084>.
- Clevers, J.G.P.W., Gitelson, A.A., 2013. Remote estimation of crop and grass chlorophyll and nitrogen content using red-edge bands on sentinel-2 and-3. *Int. J. Appl. Earth Obs. Geoinf.* 23. <https://doi.org/10.1016/j.jag.2012.10.008>.
- Danner, M., Berger, K., Wocher, M., Mauser, W., Hank, T., 2021. Efficient RTM-based training of machine learning regression algorithms to quantify biophysical & biochemical traits of agricultural crops. *ISPRS J. Photogramm. Remote Sens.* 173, 278–296. <https://doi.org/10.1016/j.isprsjprs.2021.01.017>.
- de Sá, N.C., Baratchi, M., Hauser, L.T., van Bodegom, P., 2021. Exploring the impact of noise on hybrid inversion of proSAIL RTM on sentinel-2 data. *Remote Sens.* 13. <https://doi.org/10.3390/rs13040648>.
- Deleigido, J., Verrelst, J., Alonso, L., Moreno, J., 2011. Evaluation of sentinel-2 red-edge bands for empirical estimation of green LAI and chlorophyll content. *Sensors* 11, 7063–7081. <https://doi.org/10.3390/s110707063>.
- Delloye, C., Weiss, M., Defourny, P., 2018. Retrieval of the canopy chlorophyll content from Sentinel-2 spectral bands to estimate nitrogen uptake in intensive winter wheat cropping systems. *Remote Sens. Environ.* 216. <https://doi.org/10.1016/j.rse.2018.06.037>.
- Estévez, J., Vicent, J., Rivera-Caicedo, J.P., Morcillo-Pallarés, P., Vuolo, F., Sabater, N., Camps-Valls, G., Moreno, J., Verrelst, J., 2020. Gaussian processes retrieval of LAI from Sentinel-2 top-of-atmosphere radiance data. *ISPRS J. Photogramm. Remote Sens.* 167. <https://doi.org/10.1016/j.isprsjprs.2020.07.004>.
- Estévez, J., Berger, K., Vicent, J., Rivera-Caicedo, J.P., Wocher, M., Verrelst, J., 2021. Top-of-atmosphere retrieval of multiple crop traits using variational heteroscedastic gaussian processes within a hybrid workflow. *Remote Sens.* 13. <https://doi.org/10.3390/rs13081589>.
- Estévez, J., Salinero-Delgado, M., Berger, K., Pipia, L., Rivera-Caicedo, J.P., Wocher, M., Reyes-Muñoz, P., Tagliabue, G., Boschetti, M., Verrelst, J., 2022. Gaussian processes retrieval of crop traits in Google earth engine based on Sentinel-2 top-of-atmosphere data. *Remote Sens. Environ.* 273. <https://doi.org/10.1016/j.rse.2022.112958>.
- García-Soria, J.L., Morata, M., Berger, K., Pascual-Ventoe, A.B., Rivera-Caicedo, J.P., Verrelst, J., 2024. Evaluating epistemic uncertainty estimation strategies in vegetation trait retrieval using hybrid models and imaging spectroscopy data. *Remote Sens. Environ.* 310, 114228. <https://doi.org/10.1016/J.RSE.2024.114228>.
- Gascon, F., Bouzinac, C., Thépaut, O., Jung, M., Francesconi, B., Louis, J., Lonjou, V., Lafrance, B., Massera, S., Gaudel-Vacaresse, A., Languielle, F., Alhamoud, B., Viallefont, F., Pflug, B., Bieniarz, J., Clerc, S., Pessiot, L., Trémas, T., Cadau, E., De Bonis, R., Isola, C., Martimort, P., Fernandez, V., 2017. Copernicus sentinel-2A calibration and products validation status. *Remote Sens.* 9. <https://doi.org/10.3390/rs9060584>.
- Gevaert, C.M., 2022. Explainable AI for earth observation: a review including societal and regulatory perspectives. *Int. J. Appl. Earth Obs. Geoinf.* <https://doi.org/10.1016/j.jag.2022.102869>.
- Ghosh, S.S., Dey, S., Bhogapurapu, N., Homayouni, S., Bhattacharya, A., McNairn, H., 2022. Gaussian process regression model for crop biophysical parameter retrieval from multi-polarized C-band SAR data. *Remote Sens.* 14. <https://doi.org/10.3390/rs14040934>.
- Hoa, P.V., Binh, N.A., Thao, G.T.P., An, N.N., Trinh, P.T., Tuan, N.Q., Hanh, N.C., 2025. Long-term trend and seasonal cycles of gap-free downscaled diurnal/nocturnal LST and the interaction to functional plant trait under tropical monsoon climate. *Earth and Space Science* 12, e2024EA003888. <https://doi.org/10.1029/2024EA003888>.
- Höhl, A., Obadić, I., Fernández-Torres, M.-Á., Najjar, H., Oliveira, D.A.B., Akata, Z., Dengel, A., Zhu, X.X., 2024. Opening the black box: a systematic review on explainable artificial intelligence in remote sensing. *IEEE Geosci. Remote Sens. Mag.* 12, 261–304. <https://doi.org/10.1109/MGRS.2024.3467001>.
- Jacquemoud, S., Verhoef, W., Baret, F., Bacour, C., Zarco-Tejada, P.J., Asner, G.P., François, C., Ustin, S.L., 2009. PROSPECT + SAIL models: a review of use for vegetation characterization. *Remote Sens. Environ.* 113, S56–S66. <https://doi.org/10.1016/j.rse.2008.01.026>.
- Jia, K., Hasan, U., Jiang, H., Qin, B., Chen, S., Li, D., Wang, C., Deng, Y., Shen, J., 2024. How frequent the Landsat 8/9-sentinel 2A/B virtual constellation observed the earth for continuous time series monitoring. *Int. J. Appl. Earth Obs. Geoinf.* 130, 103899. <https://doi.org/10.1016/J.JAG.2024.103899>.
- Kaminski, T., Pinty, B., Voßbeck, M., Lopatka, M., Gobron, N., Robustelli, M., 2017. Consistent retrieval of land surface radiation products from EO, including traceable

- uncertainty estimates. *Biogeosciences* 14. <https://doi.org/10.5194/bg-14-2527-2017>.
- Kandasamy, S., Baret, F., Verger, A., Neveux, P., Weiss, M., 2013. A comparison of methods for smoothing and gap filling time series of remote sensing observations - application to MODIS LAI products. *Biogeosciences* 10. <https://doi.org/10.5194/bg-10-4055-2013>.
- Kissling, W.D., Walls, R., Bowser, A., Jones, M.O., Kattge, J., Agosti, D., Amengual, J., Basset, A., van Bodegom, P.M., Cornelissen, J.H.C., Denny, E.G., Deudero, S., Egloff, W., Elmendorf, S.C., Alonso García, E., Jones, K.D., Jones, O.R., Lavorel, S., Lear, D., Navarro, L.M., Pawar, S., Pirzl, R., Rüger, N., Sal, S., Salguero-Gómez, R., Schigel, D., Schulz, K.S., Skidmore, A., Guralnick, R.P., 2018. Towards global data products of essential biodiversity variables on species traits. *Nat. Ecol. Evol.* <https://doi.org/10.1038/s41559-018-0667-3>.
- Kovács, D.D., Reyes-Muñoz, P., Salinero-Delgado, M., Mészáros, V.I., Berger, K., Verrelst, J., 2023. Cloud-free global maps of essential vegetation traits processed from the TOA Sentinel-2 catalogue in Google earth engine. *Remote Sens.* 15. <https://doi.org/10.3390/rs15133404>.
- Kovács, D., Reyes-Muñoz, P., Berger, K., Mészáros, V.I., Caballero, G., Verrelst, J., 2024. Multi-decadal temporal reconstruction of Sentinel-3 OLCI-based vegetation products with multi-output Gaussian process regression. *Ecol. Inform.* 83, 102816. <https://doi.org/10.1016/j.ecoinf.2024.102816>.
- Lázaro-Gredilla, M., Titsias, M.K., Verrelst, J., Camps-Valls, G., 2014. Retrieval of biophysical parameters with heteroscedastic Gaussian processes. *IEEE Geosci. Remote Sens. Lett.* 11. <https://doi.org/10.1109/LGRS.2013.2279695>.
- Liu, X., Xie, Y., Zhu, B., Song, K., 2024. Rice leaf chlorophyll content estimation with different crop coverages based on Sentinel-2. *Ecol. Inform.* 81, 102622. <https://doi.org/10.1016/j.ecoinf.2024.102622>.
- Malenovsky, Z., Homolová, L., Lukeš, P., Buddenbaum, H., Verrelst, J., Alonso, L., Schaepman, M.E., Lauret, N., Gastellu-Etchegorry, J.P., 2019. Variability and uncertainty challenges in scaling imaging spectroscopy retrievals and validations from leaves up to vegetation canopies. *Surv. Geophys.* <https://doi.org/10.1007/s10712-019-09534-y>.
- Martínez-Ferrer, L., Piles, M., Camps-Valls, G., 2021. Crop yield estimation and interpretability with Gaussian processes. *IEEE Geosci. Remote Sens. Lett.* 18. <https://doi.org/10.1109/LGRS.2020.3016140>.
- Martínez-Ferrer, L., Moreno-Martínez, Á., Campos-Taberner, M., García-Haro, F.J., Muñoz-Mari, J., Running, S.W., Kimball, J., Clinton, N., Camps-Valls, G., 2022. Quantifying uncertainty in high resolution biophysical variable retrieval with machine learning. *Remote Sens. Environ.* 280. <https://doi.org/10.1016/j.rse.2022.113199>.
- Mateo-Sanchis, A., Piles, M., Amorós-López, J., Muñoz-Marí, J., Adsua, J.E., Moreno-Martínez, Á., Camps-Valls, G., 2021. Learning main drivers of crop progress and failure in Europe with interpretable machine learning. *Int. J. Appl. Earth Obs. Geoinf.* 104. <https://doi.org/10.1016/j.jag.2021.102574>.
- McKay, M.D., Beckman, R.J., Conover, W.J., 1979. Comparison of three methods for selecting values of input variables in the analysis of output from a computer code. *Technometrics* 21, 239–245. <https://doi.org/10.1080/00401706.1979.10489755>.
- Moharana, S., Dutta, S., 2016. Spatial variability of chlorophyll and nitrogen content of rice from hyperspectral imagery. *ISPRS J. Photogramm. Remote Sens.* 122. <https://doi.org/10.1016/j.isprsjprs.2016.09.002>.
- Povey, A.C., Grainger, R.G., 2015. Known and unknown unknowns: uncertainty estimation in satellite remote sensing. *Atmos. Meas. Tech.* 8. <https://doi.org/10.5194/amt-8-4699-2015>.
- Rasmussen, C.E., Williams, C.K.I., 2005. *Gaussian Processes for Machine Learning*. The MIT Press. <https://doi.org/10.7551/mitpress/3206.001.0001>.
- Rossi, M., Candiani, G., Nutini, F., Gianinetto, M., Boschetti, M., 2022. Sentinel-2 estimation of CNC and LAI in rice cropping system through hybrid approach modelling. *Eur. J. Remote Sens.* <https://doi.org/10.1080/22797254.2022.2117651>.
- Sishodia, R.P., Ray, R.L., Singh, S.K., 2020. Applications of remote sensing in precision agriculture: a review. *Remote Sens.* 12. <https://doi.org/10.3390/rs12193136>.
- Skakun, S., Wevers, J., Brockmann, C., Doxani, G., Aleksandrov, M., Batić, M., Frantz, D., Gascon, F., Gómez-Chova, L., Hagolle, O., López-Puigdollers, D., Louis, J., Lubej, M., Mateo-García, G., Osman, J., Peressutti, D., Pflug, B., Puc, J., Richter, R., Roger, J.C., Scaramuzza, P., Vermote, E., Vesel, N., Zupanc, A., Žust, L., 2022. Cloud mask Intercomparison eXercise (CMIX): an evaluation of cloud masking algorithms for Landsat 8 and Sentinel-2. *Remote Sens. Environ.* 274. <https://doi.org/10.1016/j.rse.2022.112990>.
- Tagliabue, G., Boschetti, M., Bramati, G., Candiani, G., Colombo, R., Nutini, F., Pompilio, L., Rivera-Caicedo, J.P., Rossi, M., Rossini, M., Verrelst, J., Panigada, C., 2022. Hybrid retrieval of crop traits from multi-temporal PRISMA hyperspectral imagery. *ISPRS J. Photogramm. Remote Sens.* 187. <https://doi.org/10.1016/j.isprsjprs.2022.03.014>.
- Verhoef, W., Jia, L., Xiao, Q., Su, Z., 2007. Unified optical-thermal four-stream radiative transfer theory for homogeneous vegetation canopies. *IEEE Trans. Geosci. Remote Sens.* 45. <https://doi.org/10.1109/TGRS.2007.895844>.
- Verrelst, J., Alonso, L., Camps-Valls, G., Delegido, J., Moreno, J., 2012a. Retrieval of vegetation biophysical parameters using Gaussian process techniques. *IEEE Trans. Geosci. Remote Sens.* 50. <https://doi.org/10.1109/TGRS.2011.2168962>.
- Verrelst, J., Muñoz, J., Alonso, L., Delegido, J., Rivera, J.P., Camps-Valls, G., Moreno, J., 2012b. Machine learning regression algorithms for biophysical parameter retrieval: opportunities for Sentinel-2 and -3. *Remote Sens. Environ.* 118. <https://doi.org/10.1016/j.rse.2011.11.002>.
- Verrelst, J., Alonso, L., Rivera Caicedo, J.P., Moreno, J., Camps-Valls, G., 2013a. Gaussian process retrieval of chlorophyll content from imaging spectroscopy data. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* 6. <https://doi.org/10.1109/JSTARS.2012.2222356>.
- Verrelst, J., Rivera, J.P., Moreno, J., Camps-Valls, G., 2013b. Gaussian processes uncertainty estimates in experimental Sentinel-2 LAI and leaf chlorophyll content retrieval. *ISPRS J. Photogramm. Remote Sens.* 86. <https://doi.org/10.1016/j.isprsjprs.2013.09.012>.
- Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J.P., Veroustraete, F., Clevers, J.G., P.W.P.W., Moreno, J., 2015a. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties - a review. *ISPRS J. Photogramm. Remote Sens.* 108, 273–290. <https://doi.org/10.1016/j.isprsjprs.2015.05.005>.
- Verrelst, J., Rivera, J.P., Veroustraete, F., Muñoz-Marí, J., Clevers, J.G.P.W., Camps-Valls, G., Moreno, J., 2015b. Experimental Sentinel-2 LAI estimation using parametric, non-parametric and physical retrieval methods - a comparison. *ISPRS J. Photogramm. Remote Sens.* 108. <https://doi.org/10.1016/j.isprsjprs.2015.04.013>.
- Verrelst, J., Dethier, S., Rivera, J.P., Muñoz-Mari, J., Camps-Valls, G., Moreno, J., 2016a. Active learning methods for efficient hybrid biophysical variable retrieval. *IEEE Geosci. Remote Sens. Lett.* 13, 1012–1016. <https://doi.org/10.1109/LGRS.2016.2560799>.
- Verrelst, J., Rivera, J.P., Gitelson, A., Delegido, J., Moreno, J., Camps-Valls, G., 2016b. Spectral band selection for vegetation properties retrieval using Gaussian processes regression. *Int. J. Appl. Earth Obs. Geoinf.* 52. <https://doi.org/10.1016/j.jag.2016.07.016>.
- Verrelst, J., Malenovsky, Z., Van der Tol, C., Camps-Valls, G., Gastellu-Etchegorry, J.P., Lewis, P., North, P., Moreno, J., 2019. Quantifying vegetation biophysical variables from imaging spectroscopy data: a review on retrieval methods. *Surv. Geophys.* 40, 589–629. <https://doi.org/10.1007/s10712-018-9478-y>.
- Weiss, Baret, F., Jay, S., 2020a. *S2ToolBox Level 2 Products: LAI, FAPAR, FCOVER (Version 2.0)*. Institut National de la Recherche Agronomique (INRA), Avignon.
- Weiss, Jacob, F., Duveiller, G., 2020b. Remote sensing for agricultural applications: a meta-review. *Remote Sens. Environ.* 236. <https://doi.org/10.1016/j.rse.2019.111402>.
- Willmott, C.J., 1981. On the validation of models. *Phys. Geogr.* 2, 184–194. <https://doi.org/10.1080/02723646.1981.10642213>.
- Xie, R., Darvishzadeh, R., Skidmore, A.K., Heurich, M., Holzwarth, S., Gara, T.W., Reusen, I., 2021. Mapping leaf area index in a mixed temperate forest using Fenix airborne hyperspectral data and Gaussian processes regression. *Int. J. Appl. Earth Obs. Geoinf.* 95. <https://doi.org/10.1016/j.jag.2020.102242>.
- Xu, X.Q., Lu, J.S., Zhang, N., Yang, T.C., He, J.Y., Yao, X., Cheng, T., Zhu, Y., Cao, W.X., Tian, Y.C., 2019. Inversion of rice canopy chlorophyll content and leaf area index based on coupling of radiative transfer and Bayesian network models. *ISPRS J. Photogramm. Remote Sens.* 150. <https://doi.org/10.1016/j.isprsjprs.2019.02.013>.
- Xu, L., Shi, S., Gong, W., Chen, B., Sun, J., Xu, Q., Bi, S., 2024. Mapping 3D plant chlorophyll distribution from hyperspectral LiDAR by a leaf-canopyradiative transfer model. *Int. J. Appl. Earth Obs. Geoinf.* 127. <https://doi.org/10.1016/j.jag.2024.103649>.
- Yuen, K.W., Hanh, T.T., Quynh, V.D., Switzer, A.D., Teng, P., Lee, J.S.H., 2021. Interacting effects of land-use change and natural hazards on rice agriculture in the Mekong and Red River deltas in Vietnam. *Nat. Hazards Earth Syst. Sci.* 21. <https://doi.org/10.5194/nhess-21-1473-2021>.
- Zupanc, A., 2017. Improving Cloud Detection With Machine Learning [WWW Document]. <https://medium.com/sentinel-hub/improving-cloud-detection-with-machine-learning-c09dc5d7cf13> accessed 6.16.23.