

Retrieval of crop traits using PROSAIL-based hybrid radiative transfer model and EnMAP hyperspectral data[☆]

Prachi Singh^{a,b}, Prashant K. Srivastava^{a,*}, Prakash Kumar Jha^b, Jochem Verrelst^{c,d},
Pashupati Nath Singh^a, Rajendra Prasad^d

^a Remote Sensing Laboratory, Institute of Environment and Sustainable Development, Banaras Hindu University, Varanasi, India

^b Department of Plant and Soil Sciences, Mississippi State University, Mississippi State, MS -39762, USA

^c Image Processing Laboratory (IPL), University of Valencia, Valencia, Spain

^d Department of Physics, Indian Institute of Technology (BHU), Varanasi, India

ARTICLE INFO

Keywords:

EnMAP
Radiative transfer model
Machine learning
Hybrid algorithm

ABSTRACT

Implementing high spectral resolution imaging from the Environmental Mapping and Analysis Program (EnMAP) paved the way for detailed retrieval of agricultural traits for accurate crop monitoring and management. The proposed methodology involves the integration and detailed analysis of Radiative Transfer Modelling (RTM) with an integrated approach of machine learning (ML) and Active Learning (AL) algorithms for the retrieval of the Leaf Chlorophyll Content (LCC), Carotenoids (Car) and Leaf Area index (LAI) of wheat cropland from the continuous three years of the dataset. Reflectance values of leaf were collected using Analytical Spectral Device (ASD) – Spectroradiometer data ranging from 350–2500 nm and EnMAP satellite hyperspectral data extends spectral data range varies between 420 nm to 1000 nm in the visible and near-infrared (VNIR) of EMR region, and from 900 nm to 2450 nm in the shortwave infrared (SWIR) region for crop parameters mapping for a larger spatial area of Varanasi district, Uttar Pradesh, India. The PROSPECT + SAIL (PROSAIL) RTM was employed to simulate spectral (reflectance) data, and fourteen ML algorithms were assessed for implementation into a hybrid model. Kernel Ridge regression (KRR) was combined with Euclidean-based Diversity (EBD) algorithms to retrieve crop characteristics due to their exceptional accuracy and reduced uncertainty. Spectral profiles were further used to train hybrid models using PCA (Principal Component Analysis) feature selection, and combined techniques (ML + AL) were applied to retrieve LCC, Car, and LAI. Afterwards, biophysical and biochemical spatial large-scale estimation were provided through atmospherically corrected, and noise-removed EnMAP hyperspectral data with the help of a trained and tested hybrid (ML + AL) model and validated with the ground-measured datasets. The performance indicators showed significantly very high values of correlation during calibration (LCC = 0.99, Car = 0.74, and LAI = 0.99) and validation (LCC = 0.66, Car = 0.57, and LAI = 0.88). The work showed that the optimized hybrid (KRR + AL) models customized for EnMAP hyperspectral data can efficiently estimate the wheat biophysical and biochemical parameters in near-real time therefore, expanding this workflow to agricultural fields may enable more effective monitoring and management of wheat crops.

1. Introduction

Around 60 % of the Earth's surface remains capable of supporting biological life. When the proportion of forest lands is discounted, about 50 million km² of land is used for agricultural purposes (Zabel et al. 2014; Adamu et al. 2015; Asner 1998; Marandi et al. 2022; Parida and Kumari 2021). Retrieval of primary vegetation parameters at the leaf and canopy scales, like crop pigments LCC, Car, and LAI, affects

vegetation growth and development, such as photosynthesis and crop yield production (Gitelson 2018; Gitelson et al. 2005). LCC and Car are crucial for photosynthesis due to their capacity to transform solar energy into the metabolic energy required for carbon fixation (Hosikian et al. 2010; Fernández-Marín et al. 2015; Gonzalez-Dugo et al. 2015). It is crucial to monitor the health and stress of crops using remotely sensed imagery for effective crop management (Shah et al. 2019; Baret et al. 1994; Richards and Stokes, 2004; Shah et al. 2019).

[☆] This article is part of a special issue entitled: 'Hyperspectral Rem Sens' published in International Journal of Applied Earth Observation and Geoinformation.

* Corresponding author.

E-mail address: prashant.iesd@bhu.ac.in (P.K. Srivastava).

<https://doi.org/10.1016/j.jag.2025.104769>

Received 22 March 2025; Received in revised form 29 July 2025; Accepted 29 July 2025

Available online 10 August 2025

1569-8432/© 2025 The Author(s). Published by Elsevier B.V. This is an open access article under the CC BY-NC-ND license (<http://creativecommons.org/licenses/by-nc-nd/4.0/>).

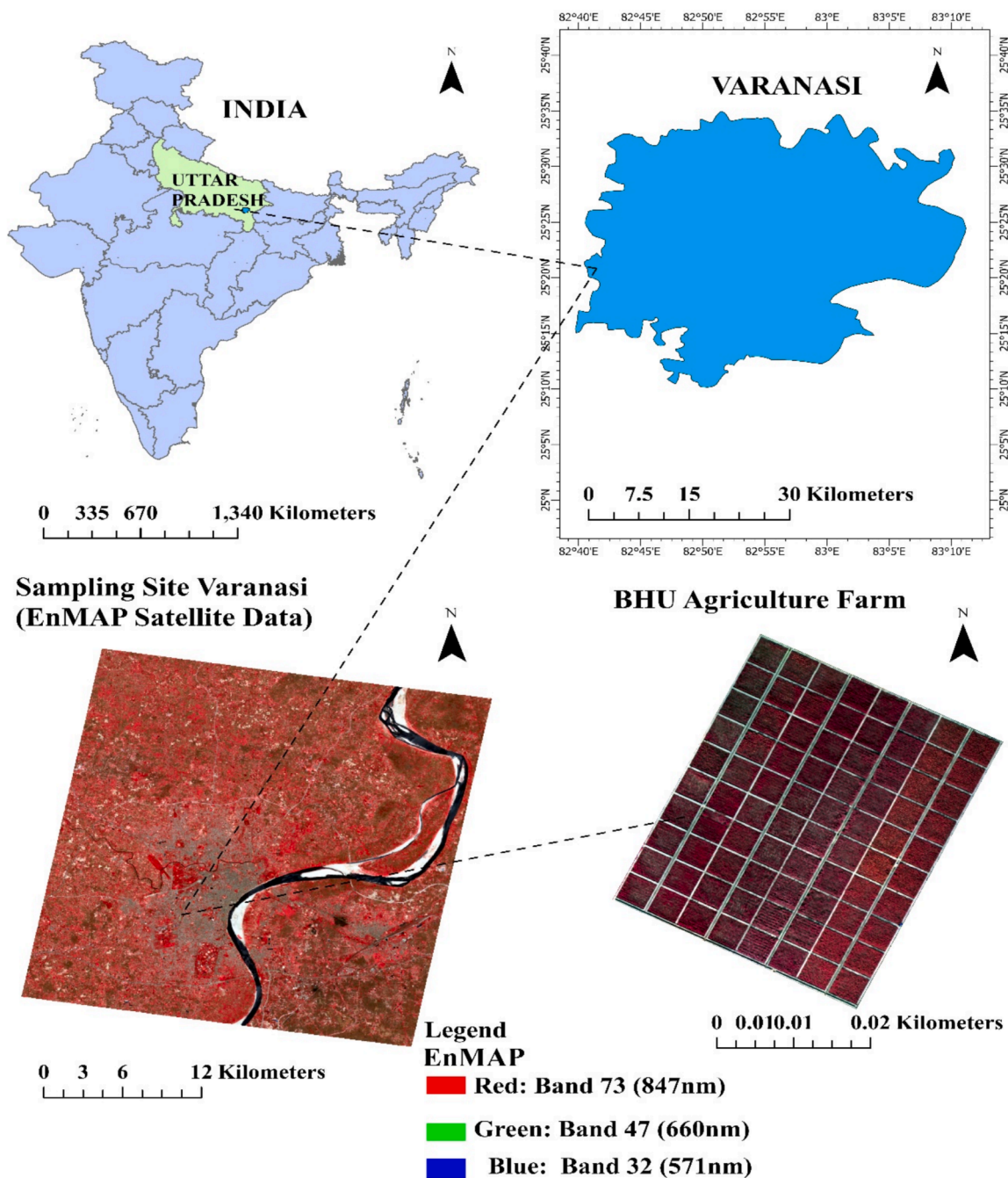


Fig. 1. Geographical location of the study area with EnMAP bands 73, 47 and 32 combinations.

The assessment of LCC and Car is vital for monitoring plant health, regulating fertilizer, and other inputs, as it directly affects crop production levels (Gupta et al. 2022(a); Gupta et al. 2021; Singh et al. 2020). The conventional destructive techniques used to determine LCC and Car require intricate solvent extraction operations and complex in vitro spectrophotometric analysis, which are resource-consuming (Verma et al. 2022). Foliar pigment contents—such as chlorophyll $a+b$ (LCC) and carotenoids (Car), along with leaf area index (LAI), are key traits linked to ecosystem structure and its connection to biodiversity (Miraglio et al., 2019). Collecting data at larger spatiotemporal scales is also not feasible since it requires time-consuming sampling and processing methods (Zhu et al. 2020; Gupta et al. 2021). To overcome these problems, optical satellite data has emerged as a compelling alternative for crop management. It is crucial to monitor the health and stress of

crops using satellite imagery for effective crop modelling and management (Asner 1998; Baret et al. 1994). Satellite data are promising for monitoring spatiotemporal dynamics of vegetation attributes (Zeng et al., 2020). The progress in high-resolution, newly launched satellite imaging data, such as PRISMA (PRecursore IperSpettrale della Missione Applicativa) and EnMAP, offers more accurate information about the biochemical composition of crops (Belanger et al. 1995; Hill et al. 2019; Galvão et al. 2018; Mariotto et al., 2013; Sahoo et al. 2015). The limited number of bands in multispectral sensors is constraining their capacity to offer comprehensive detail concerning ground parameters. In contrast, hyperspectral-enabled sensors deliver data with contiguous and adjacent bands, and provide more accurate spectral analysis of field features compared to broadband data of multispectral images (Mananze et al. 2018; Thenkabail et al. 2018).

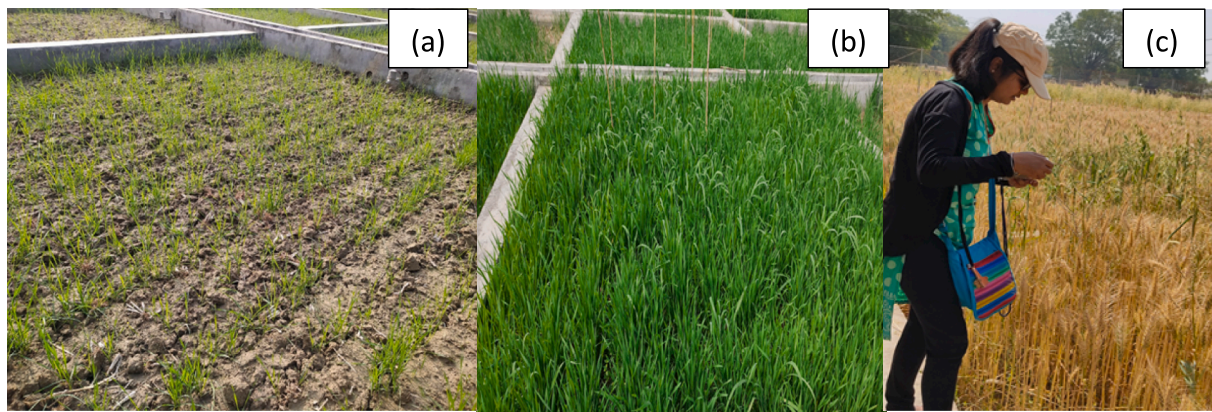


Fig. 2. Data collection photos in the study area during the different stages (a) tillering, (b) heading, and (c) maturity.

Table 1

Solvent equations for the measurement of CHL-a, CHL-b, and C x + c in ($\mu\text{g/ml}$) (Sumanta et al. (2014)).

Solvents	Equations/Formula
80 % Acetone (Solvent)	$\text{CHL-a}_{(\text{CHL-a})} = 12.25A_{663.2} - 279A_{646.8}$ $\text{CHL-b}_{(\text{CHL-b})} = 21.5A_{646.8} - 5.1A_{663.2}$ $\text{LCC} = \text{CHL-a} + \text{CHL-b}$ $\text{Cx} + \text{c}(\text{total Car}) = (1000A_{470} - 1.82\text{CHL-a} - 85.02\text{CHL-b})/198$ Where, A_b is the absorbance bands

The ideal retrieval method is to create direct statistical correlations between canopy reflectance data from airborne or spaceborne sensors and ground-measured biochemical contents. However, these correlations are site and species-specific, making it challenging to apply findings directly to different locations or species (Gupta et al. 2021). Retrieval approaches based on physical modeling adhere to radiative transfer (RT) principles across leaf, canopy, and atmospheric levels, as shown in RTMs. These methods can be broadly applicable and are useful in mapping applications. Mapping RTM input variables involves simulating the RTM with spectral observations until a reduced state is achieved for variable retrieval. However, both statistical and physically-based techniques have their respective advantages and disadvantages (Verrelst et al. 2015; Demarez and Gastellu-Etchegorry 2000). Physical modelling-based inversion methods are scientifically valid as they compare observed spectral data with that produced by RTM. At the leaf level, models like PROSPECT and SAIL, collectively known as PROSAIL, are used to predict leaf parameters (Gupta et al. 2021; Jacquemoud et al. 1995). Similarly, studies using related leaf and canopy radiative transfer-based models investigate how regulating factors affect the reflectance qualities of leaves at the canopy level. These studies are also valuable for estimating the biochemical content in leaves (Feret et al. 2008; Jacquemoud et al. 1995; Zhang et al. 2008; Jay et al. 2016). These techniques are computationally intensive, as it require multiple model runs per observation to reach a minimum. In contrast, hybrid model-based inversion is faster due to the pre-computed LUT, which stores simulated reflectance and input parameters. LUT-based spectral simulations and MLRA methods have been widely used to retrieve key vegetation traits across growth stages. (Darvishzadeh et al. 2011; Zhang et al. 2008). To enhance the result accuracy, more advanced AL has also been introduced. Therefore, this work focuses on the development, optimization, and validation of a hybrid combination of machine learning with AL models for the retrieval of the wheat vital parameters such as LCC, Car, and LAI using a new advanced high-resolution EnMAP hyperspectral dataset. Therefore, the main objectives of this study are (i) to examine the effectiveness of the combined approach of MLRA + AL to predict wheat crop parameters using radiometer and EnMAP

hyperspectral data. (ii) to parameterize the MLRA model to retrieve the best AL sampling methods for retrieving important crop parameters such as LCC, Car, and LAI. (iii) to generate large-scale mapping of plant biochemical and biophysical parameters by using newly launched EnMAP satellite data.

2. Study area

In this study, agricultural areas of Varanasi in Northern India were considered for model development. Varanasi lies between latitudes $25^{\circ}10'30''$ and $25^{\circ}35'15''$ N and longitudes $82^{\circ}40'50''$ and $83^{\circ}12'18''$ E and is part of the Indo-Gangetic alluvial plain. The Ganga River constitutes its natural boundary in the east and southeast, while the Gomati River defines its northern boundary. Varanasi district has a total area of 1,535 km^2 , of which 1,371.22 km^2 is rural and 163.78 km^2 is urban (Adinehvand and Rai, 2016). The average annual rainfall in the research area is roughly 1,100 mm, the temperature series 7°C to 46°C , and the average yearly temperature is 25°C . Rabi and Kharif crops are the most commonly grown crops in the research area (Nistor et al. 2020). Groundwater media mainly consist of sedimentary deposits. Due to its favourable living conditions, including fertile soil, a flat landscape, and ample water resources, it can sustain a high population density (Younger, 2009). The study area map is shown in Fig. 1.

3. Field and satellite dataset

3.1. Sampling data

Extensive field data collection for LCC, Plant height, and LAI measurement was carried out from 2020 to 2023 in the BHU agriculture field of Varanasi. The GPS instrument was used to gather ground coordinates. At the stages of tillering, heading, and maturity, 60 samples were collected every year from 2020 to 2023. The wheat crop leaf samples were also collected to measure fresh and dry weight, which was used to calculate the water thickness. Leaf was also collected to measure the LCC and Car using destructive measurement. Additionally, the crop LAI was obtained using an LI-COR LAI-2000 instrument, specifically designed for analysing plant canopies. Pictures of the field data collection are shown in Fig. 2.

i) Destructive measurement of chlorophyll and carotenoids

In this study, destructive measurements of leaf chlorophyll content and carotenoids were carried out in the laboratory. The most useful stages for crop tillering, heading, and maturity were chosen. Throughout all sampling times, leaf samples were obtained, and measurements were taken for both fresh weight and dried weight. The leaf samples were initially washed extensively with distilled water in the laboratory, subsequently dried at room temperature (18°C), and analysed to determine the levels of chlorophylls (CHL-a and CHL-b) and Car.

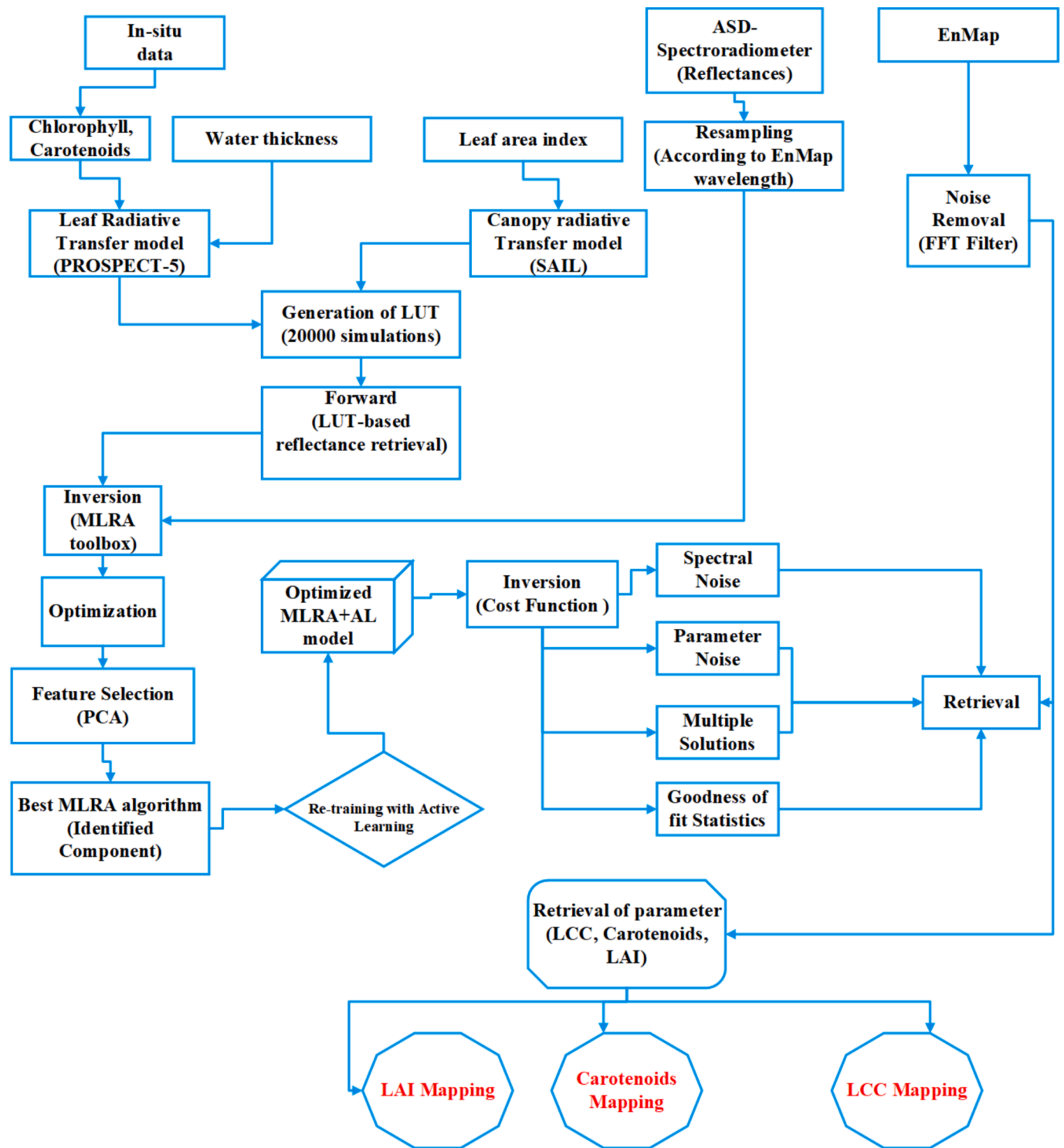


Fig. 3. Graphical presentation of the proposed methodology.

ii) Analytical procedure

A precise weight of 0.10 g of a freshly harvested plant leaf sample was measured. Subsequently, the sample was blended in the tissue homogenizer with 2 ml of a diverse extractant solvent. The sample mixture, which had been made uniform in composition, was subjected to centrifugation at a rate of 10,000/m rpm for 15 min at a set temperature of 40 °C. After the supernatant was separated, 0.51 ml of it was mixed with 0.9 ml of the suitable solvent. The measurement of values at specific absorbance bands of CHL-a, CHL-b, and Car in the solution mixture

was made using a spectrophotometer (Parkin). The mathematical formula used to measure the amounts of CHL-a, CHL-b, and Car using 80 % Acetone solvent is shown in Table 1.

3.2. Satellite data (EnMAP)

The EnMAP satellite was launched on 1st April 2022, at 6:24p.m. (CEST) by SpaceX using a Falcon 9 rocket from Cape Canaveral, making it the largest payload launched by the U.S. space business. EnMAP is a

Table 2

Comprehensive list of parameters used by PROSAIL model in ARTMO with their ranges.

Leaf: PROSPECT 5			Canopy: 4SAIL		
Input	Unit	Minimum/Maximum	Input	Unit	Minimum/Maximum
Structure of leaf* (N)	—	1–5	Total leaf area (LAI)	m ² /m ²	0–5
Total Chl (LCC)	µg/g	0–76	Distribution of angle of leaf*	Degree	0–20
Water thickness (C _w)	g/cm ²	0.001–0.05	Hotspot parameter *	—	0–1
Carotenoids (Car)	µg/g	0–3	Direct Energy *	—	0–50
			Angle*	Degree	0°

* The range was based on the information provided in the literature (Beal et al. 1982).

hyperspectral satellite developed by German institutions, led by the German Aerospace Centre (DLR). It features a spectral sample interval of 6.5 nm for VNIR and 10 nm for SWIR, with up to 99 VNIR bands and 163 SWIR bands. It also has high signal-to-noise ratios, over 500 in VNIR and 150 in SWIR. Processing options include L0, L1B, L1C, and L2A, with a 30 km swath and a ground sampling distance of 30 × 30 m (Storch et al. 2020). The EnMAP data is available to the public through the EnMAP Ground Section Instrument Planning Subsystem (IPS) Portal, which can be accessed via the EOWEB Geoportal (EGP) at no cost (Brell et al. 2021; Pascual-Venteo et al. 2024; Stuffer et al. 2007). In the current study, orthorectified reflectance data (L2A) were downloaded from the website (https://www.enmap.org/data_access/), and noise removal was done using a Fast Fourier transform filter. The EnMAP toolbox in QGIS 3.40.2 was further used to convert data into GEOTIFF format. These GEOTIFF tiles were then used in the ARTMO toolbox for the retrieval of crop parameters.

4. Methodology

The leaf-based PROSPECT model associated with the canopy-based model SAIL into ARTMO (Rivera et al., 2013) was employed to derive crop LCC, Car, and LAI. A simulated database was generated using a combined model named PROSAIL after 20,000 simulations to generate an LUT and to perform forward modelling. These simulations involved various combinations of input variables, including LCC, Water thickness (C_w), leaf structure, and LAI, along with their corresponding predicted reflectance values. Further, ASD-Spectroradiometer reflectance data with 2101 bands were used and resampled according to EnMAP satellite data with 242 bands to match the spectral resolution of satellite data. Simulated spectra and resampled radiometer data were then further used as input data for ARTMO's MLRA-based inversion toolbox. The measured data were divided into two subsets: calibration (or training) and validation. In our study, the data was split into two parts 70 % and 30 % for calibration and validation, respectively. The MLRA toolbox was further used for all the optimization and to retrieve parameters. This calibration involved spectral redundancy in the MLRA using Principal Component Analysis (PCA). The RT model was calibrated using the obtained leaf-based collected reflectance spectra, as described by Verrelst et al., 2013. To enhance the model performance further, the best MLRA model was combined with AL to invert crop parameters. Then, finally, the remaining 30 % data was used for independent testing of the final model performance. Subsequently, goodness-of-fit statistics and various outcomes are computed to assess the precision of the obtained LCC, Car, and LAI. Ultimately, a comprehensive crop parameters mapping were conducted using the recently launched EnMAP satellite data, which offers high-resolution imagery at a large scale. The detailed methodology and inversion workflow are shown in Fig. 3.

4.1. PROSAIL LUT simulation

The inversion analysis was exclusively conducted using the ARTMO version 3.30 (Verrelst et al., 2013). ARTMO is a graphical user interface (GUI) toolbox developed in MATLAB. It integrates and harmonizes the record of commonly used leaf and canopy radiative transfer models

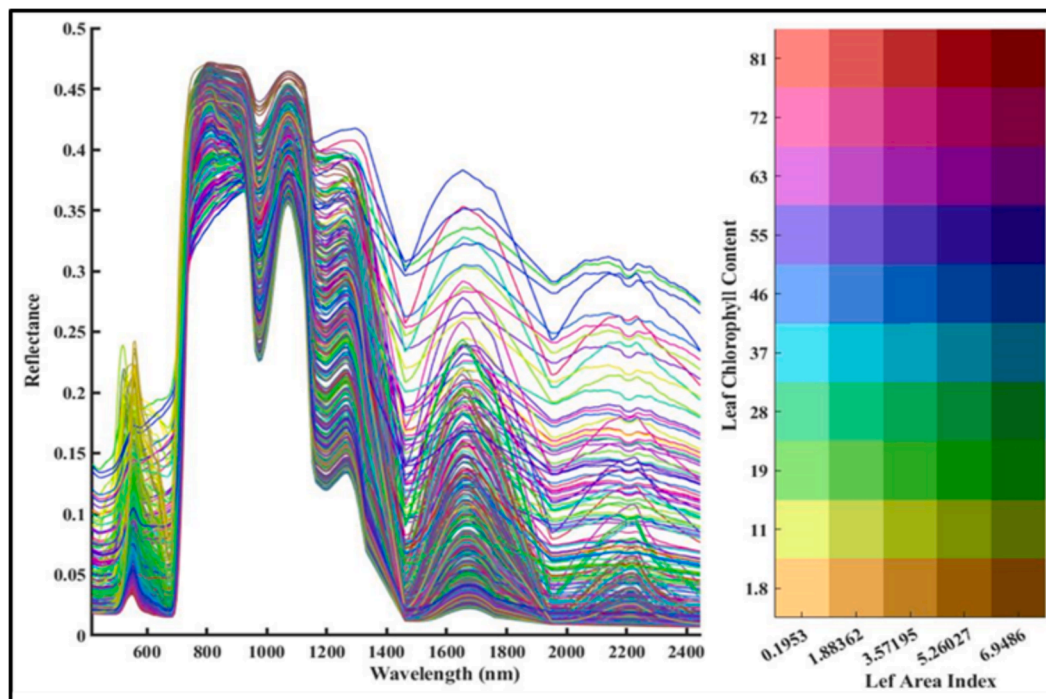


Fig. 4. The simulated spectra using PROSAIL-LUT based on the LCC and LAI parameters.

Table 3

MLRA Performances for Crop Parameters (a- LCC, b- Car, c- LAI) retrieval.

MLRA	RMSE ($\mu\text{g/g}$)	NRMSE (%)	Correlation coefficient (R)
Weighted k-nearest neighbor regression	0.89	18.94	0.79
K-nearest neighbor regression	0.51	10.77	0.96
Adaptive Regression Splines	0.88	18.72	0.80
Kernel ridge Regression	0.51	10.77	0.97
Random Forest (TreeBagger)	0.90	19.04	0.79
Regularized least-squares regression	1.01	21.50	0.87
Bagging trees	1.22	25.93	0.75
Gaussian Processes Regression	1.48	31.57	0.66
Regression tree	1.49	31.64	0.66
Elastic Net regression	1.49	31.68	0.66
Boosting trees	1.57	33.34	0.71
Least squares linear regression	1.65	35.02	0.73
Principal components regression	1.66	35.26	0.72
Partial least squares regression	1.86	39.55	0.52
(a) LCC			
MLRA	RMSE ($\mu\text{g/g}$)	NRMSE (%)	Correlation coefficient (R)
Weighted k-nearest neighbor regression	15.28	20.34	0.76
K-nearest neighbor regression	15.32	20.39	0.75
Adaptive Regression Splines	15.38	20.47	0.76
Kernel ridge Regression	14.41	19.84	0.85
Random Forest (TreeBagger)	16.41	21.84	0.70
Regularized least-squares regression	16.98	22.60	0.75
Bagging trees	18.05	24.03	0.70
Gaussian Processes Regression	19.54	26.01	0.56
Regression tree	19.80	26.35	0.53
Elastic Net regression	20.30	27.02	0.49
Boosting trees	20.37	27.12	0.60
Least squares linear regression	21.84	29.07	0.47
Principal components regression	21.84	29.07	0.46
Partial least squares regression	23.60	31.418	0.40
(b) Car			
MLRA	RMSE (m^2/m^2)	NRMSE (%)	Correlation coefficient (R)
Weighted k-nearest neighbor regression	4.07	31.87	0.45
K-nearest neighbor regression	3.97	31.14	0.40
Adaptive Regression Splines	4.02	31.51	0.43
Kernel ridge Regression	3.92	30.77	0.54
Random Forest (TreeBagger)	4.22	33.08	0.61
Regularized least-squares regression	4.58	35.92	0.63
Bagging trees	4.61	36.16	0.53
Gaussian Processes Regression	5.56	43.63	0.03
Regression tree	5.62	44.12	0.35
Elastic Net regression	5.62	42.12	0.13
Boosting trees	5.87	46.08	0.49
Least squares linear regression	3.95	26.70	0.37
Principal components regression	5.95	46.70	0.47
Partial least squares regression	11.40	89.42	0.42
(c) LAI			

(RTMs) (Rivera et al., 2013). This toolbox includes the necessary elements for implementing inverse canopy reflectance models. It offers features such as the ability to introduce the parameter and spectrum noise to the simulated spectral reflectance and the option to use the average or middle value of multiple best-performing models for the inversion process, among other functionalities (Verrelst et al., 2013). The accuracy of retrieved variables using the spectral simulation database (Weiss et al. 2000). Usually, running the inversion algorithm on large simulated spectral data with more than 10,000 entries enhances the precision of parameter estimation. However, this improvement comes at the cost of a longer execution time. A spectral simulation was produced by conducting 20,000 simulations of canopy reflectance, each with corresponding effort variable standards. PROSAIL parameterization was conducted using either published values or variable ranges obtained from field measurements during field samplings described in Table 2. In a published study of (Martínez-Huitle et al. 2023), the value was set for the leaf structure index N ranges from 1 to 4. The LCC values were obtained by wet lab analysis and ranged from 0 to 75. These values were treated as external data since they were gathered in the laboratory

employing the Spectrophotometer instrument. Water thickness (C_w) was measured using the difference between the fresh weight and dry weight of a leaf, then dividing that difference by the leaf's area. In this study, the hotspot parameter was defined within a range of 0 to 1, consistent with standard practices in radiative transfer modelling (e.g., Verhoef & Bach, 2007). The sun zenith angle (SZA) values were predefined based on the specific acquisition conditions of the EnMAP satellite overpass. EnMAP captures imagery in a near-nadir viewing configuration, the observation zenith angle (OZA) was set to 0° , representing a directly overhead sensor geometry (Asrar and Myneni, 1991). The creation of LUTs involved the execution of PROSAIL in forward mode. The variable distributions were either Gaussian or uniform. To generate the LUT, we utilized the most commonly used LHS technique named Latin Hypercube-based sampling (LHS), which is a bedded Monte Carlo-based data collection technique. This technique was chosen as an alternative to gridded sampling and involved selecting a subgroup of 20,000 entries. The implementation of LHS was done in ARTMO. The LHS approach requires the lowest and highest boundary values of the input variables (Verrelst et al. 2016). The MySQL database underlying ARTMO

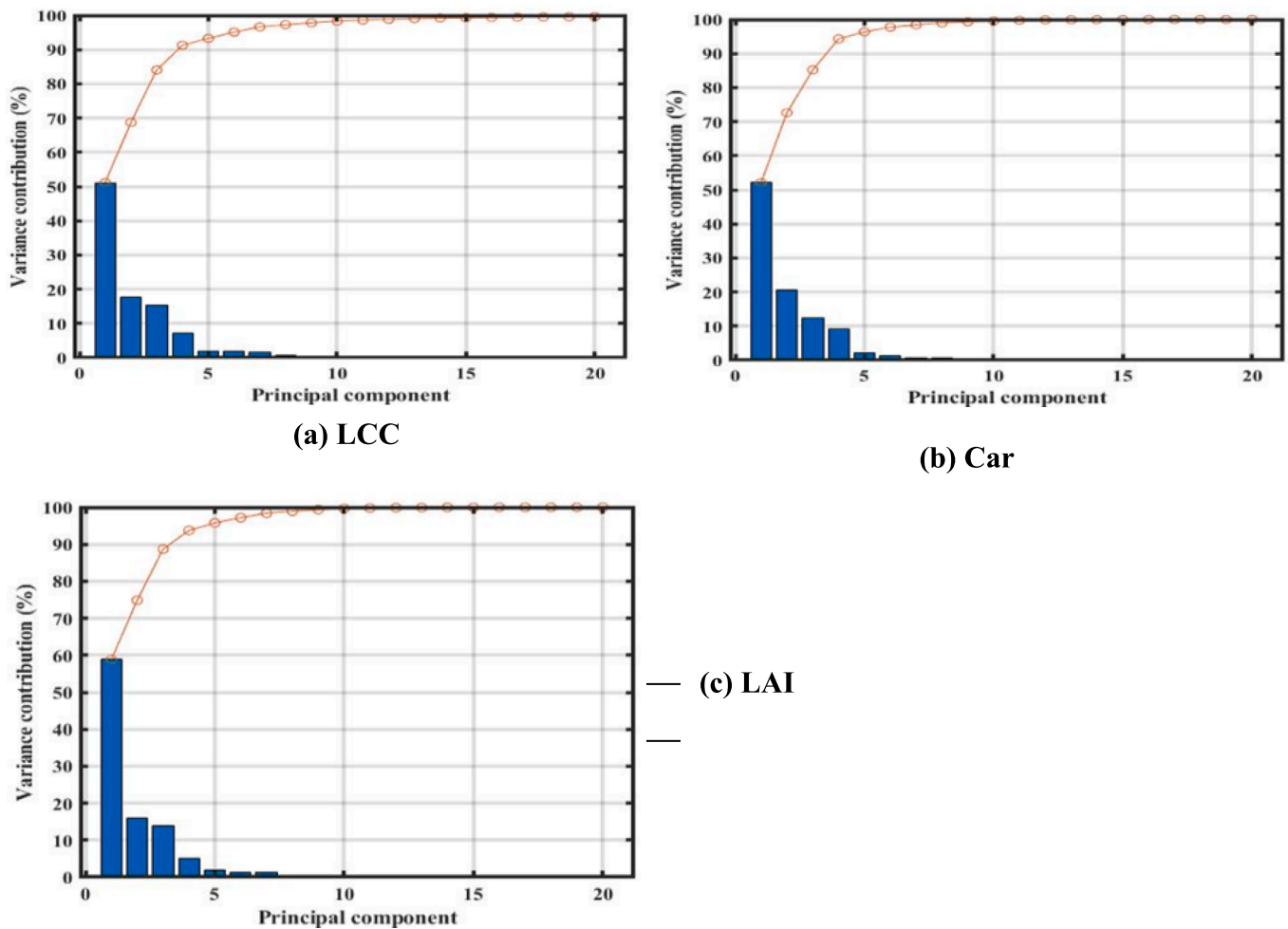


Fig. 5. Variable contribution plot using PCA for LCC (a), Car (b), and LAI (c).

continuously stores every simulation and any inversion results.

Simulated canopy reflectance data were further used with the validation data of the resampled radiometer reflectance data for model inversion (MLRA-retrieval). The best MLRA model with a combination of AL was used to predict crop parameters at a larger scale. The LUT group took 200 s, and the overturn process took 1 h and 30 min for image generation (size- 1.8 GB) in MATLAB.

4.2. Kernel ridge regression (KRR)

In this research, 14 types of MLRA algorithms such as Weighted k-nearest neighbor regression (W-NNR), K-nearest neighbour regression (K-NNR), Adaptive Regression Splines (ARS), Kernel ridge Regression (KRR), Random Forest (TreeBagger), Regularized least-squares regression (RLS), Bagging trees (BT), Gaussian Processes Regression (GPR), Regression tree (RT), Elastic Net regression (ENR), Boosting trees, Principal Component Regression (PCR), and Partial least squares regression (PLS) were optimized to find the best MLRA algorithm to retrieve crop parameters for the selected area. The KNN algorithm was further selected for model optimization. KRR is a regression method that minimises the squared residuals, making it a kernel-based version of regularised least squares regression (Ji et al., 2022). The KRR evaluation for this study was performed with the ARTMO (Aghababaei et al., 2022). To enable KRR training, it was necessary to modify two parameters: the σ parameter for the Gaussian kernel function and the C parameter for regularisation. The parameters were optimised using conventional cross-validation.

4.3. PCA for dimensionality reduction

The ASD-Spectroradiometer hyperspectral data comprises 2101 bands, which may result in slow performance in machine learning models due to spectral band redundancy, also known as the Hughes phenomenon. It is essential to employ an appropriate dimensionality reduction technique to decrease duplication and processing time while simultaneously enhancing accuracy. The ARTMO MLRA toolbox incorporates eleven techniques for reducing dimensionality, which are used to find the most significant statistical variables. Based on a comprehensive analysis of several scenarios, it has been determined that employing PCA with twenty elements is the optimal approach for extracting multiple crop features from hyperspectral data using KRR, as stated by Rivera-Caicedo et al., 2017.

4.4. Active Learning

To decrease the sampling extent without compromising the performance of the model's outputs, it is possible to reject noisy and uncooperative samples by employing a reliable sample selection criterion (Rivera-Caicedo et al., 2017). When predicting agricultural traits using large datasets, two commonly used active learning methods are uncertainty-based and diversity-based approaches. Pool, entropy query by bagging, and residual active learning (PAL, EQB, RSAL) are all methods that fall under the uncertainty approach. The diversity approach consists of three components: Euclidean-based, Angle-based, and Clustering-based diversity (EBD, ABD, CBD). The ARTMO MLRA toolbox offers six types of active learning (AL) approaches, along with

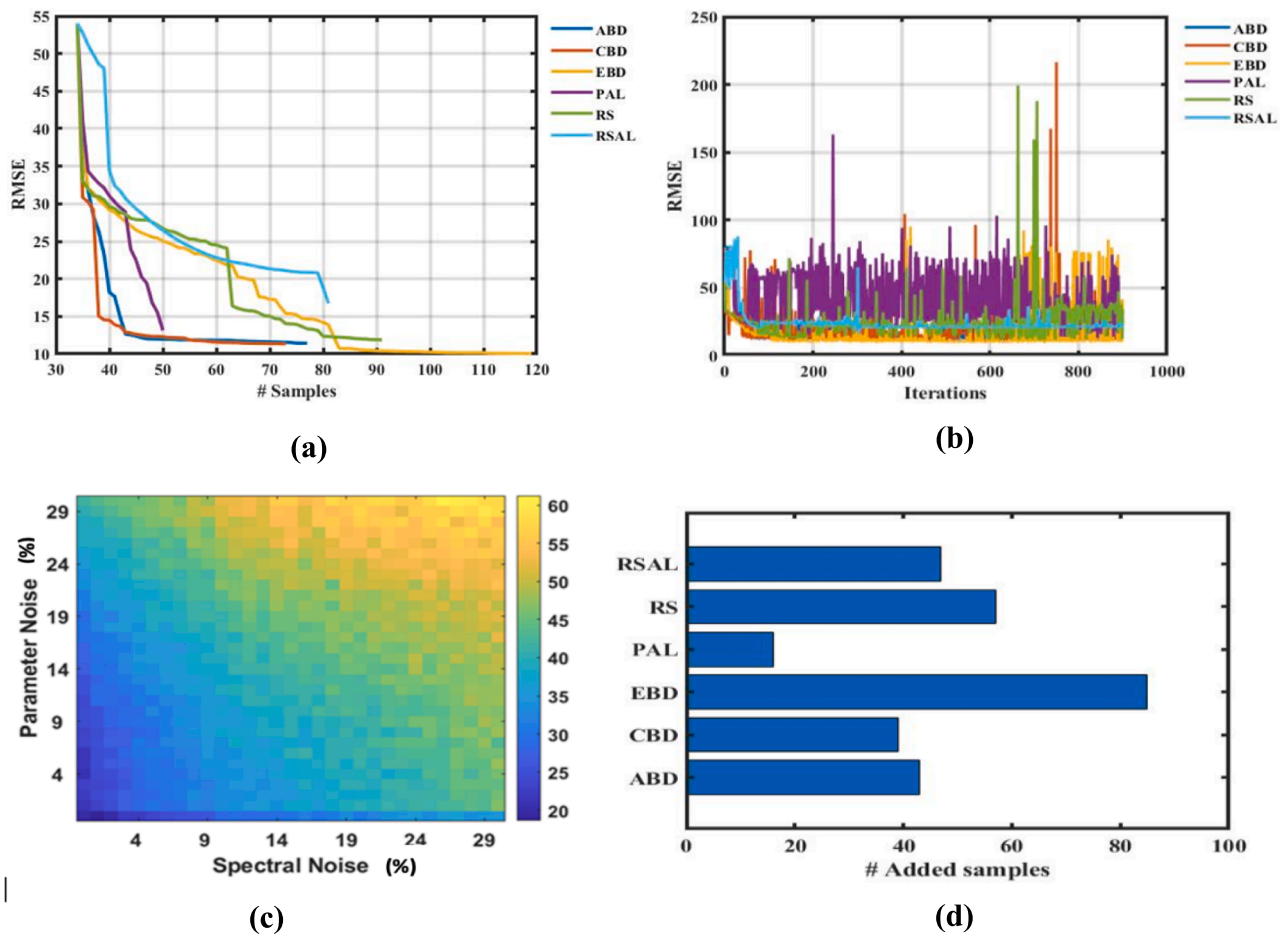


Fig. 6. AL samples performance (a), Iteration (b), Correlation Matrix plot between Spectral and Parameter Noise (c) and Added samples (d) for LCC.

random sampling strategies, for analysing PROSAIL-simulated spectra using a regression algorithm. In the case of the unpredictability strategy, the samples were evaluated according to their level of uncertainty, and the samples with the lowest level of certainty were selected as the result (ElRafey and Wojtusiak 2018; Douak et al. 2011; Douak et al. 2013). These AL algorithms use the most valuable training data to develop and predict models with optimal accuracy. According to multiple studies, the EBD technique, when combined with the MLRA model, is effective for retrieving crop parameters. The active learning technique chooses a sample from the pooled data to train the model. This process is performed iteratively to ascertain whether the data chosen from the pool enhances the model's performance or not. Therefore, samples exhibiting minimal progress are excluded from the last dataset, and this procedure is iterated until all samples in the collective dataset have been assessed against the validated dataset. For validation, we used a dataset that was acquired in the field. The accuracy was assessed by employing commonly used statistics, including Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and Normalised RMSE (NRMSE). In summary, a comprehensive analysis was conducted on five different active learning (AL) approaches: ABD, CBD, EBD, PAL, and RSAL. The findings of this analysis were to find the optimal strategy for retrieving parameters.

5. Results & discussion

5.1. Results

5.1.1. Spectral simulation

The PROSAIL model was employed to develop LUT simulations after

setting up with input data. During these simulations, the process involves combining leaf simulations with canopy simulations, resulting in a rapid and significant increase in the number of runs, as indicated by the supplied range. More than 20,000 simulations were conducted using the refreshed ASD-spectroradiometer data. Fig. 4 displays the generated LUT spectra, where the colour varies according to the factors LCC and LAI.

5.1.2. Model evaluation and selection of KRR

Simulated resampled spectra were further used as input in the MLRA inversion toolbox. In this research, 14 types of MLRA algorithms such as W-NNR, K-NNR, ARS, KRR, RF, RLS, BT, GPR, RT, ENR, BT, PCR, and PLS were optimized to find the best MLRA algorithm to retrieve the crop parameters for selected area. RTM-generated spectra and Resampled ASD-Spectroradiometer data were used for the model training and validation. The calibration involved spectral redundancy in the MLRA training using Principal Component Analysis (PCA). Model calibration was done for all parameters, using MLRA such as LCC, Car, and LAI. All MLRA algorithm performance can be seen in Table 3. Among all MLRA models, KNN outperformed and statistical metrics such as RMSE = 0.51 $\mu\text{g/g}$, NRMSE = 10.77 %, and Correlation coefficient (R) = 0.97 were found for LCC. Separately, the similar calibration was also done for the carotenoids, and performance metrics were calculated as RMSE = 14.41 $\mu\text{g/g}$, NRMSE = 19.84 %, and Correlation coefficient (R) = 0.85. Last, for LAI, RMSE = 3.92 m^2/m^2 , NRMSE = 30.77 %, and Correlation coefficient (R) = 0.54. During the model calibration, dimensionality reduction was also provided through PCA for the selection of 20 components to reduce the model processing time (Fig. 5). Based on a comprehensive analysis of several scenarios, it has been shown that

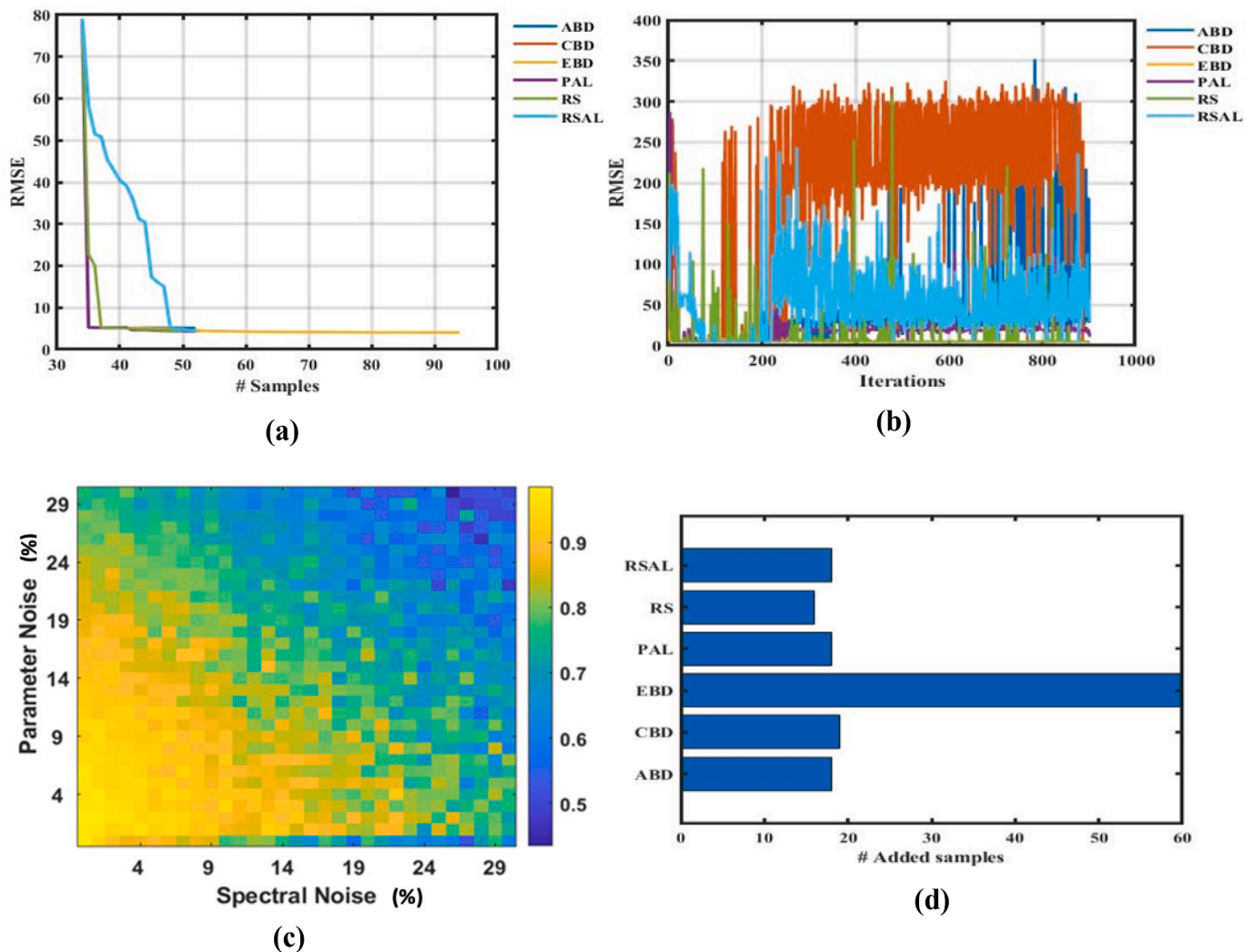


Fig. 7. AL samples performance (a), Iteration (b), Correlation Matrix plot between Spectral and Parameter Noise (c) and Added samples (d) for Car.

employing PCA with 20 components is the optimal approach for extracting many crop features from hyperspectral observations using Kernel Ridge Regression (KRR). This method provides a balance between accuracy and runtime.

5.1.3. Evaluation of MLRA + AL (Calibration)

For enhancement of the performance of the models, it is recommended to retrain the model using advanced AL approaches, namely, ABD, CBD, EBD, PA, RS, and RSAL. The active learning (AL) approach selects items from the collective dataset and uses them to train the model. This method is carried out to determine whether the samples taken from the pool improve the performance of the model or not. This investigation used 70 % dataset for calibration purposes using the MLRA + AL technique to enhance the model performance. The performance of the combined approach of MLRA + AL for parameter retrieval was found to be better in comparison with the inclusion of AL techniques. During the optimization, the spectral and parameter noise were also added to cope with ground and atmospheric noises. 85 samples and 900 iterations were done during the calibration for LCC. Sample Performance, Iteration, added samples, and inclusion of spectral and parameter noise information can be seen in Fig. 6 (a-d) for LCC. For car 60 samples and 900 iterations were done during the optimization shown in Fig. 7(a-d). Finally for the LAI, 58 samples and 886 iterations were done during the optimization, and shown in Fig. 8(a-d).

5.1.4. Validation of MLRA + AL model

70 % of the data was utilized for calibrating the MLRA + AL model remaining 30 % was allocated for validation purposes using MLRA + AL approach. Among all AL techniques, EBD gives the most effective inversion method for all parameters with KRR, as determined by various goodness-of-fit statistics, including RMSE, NRMSE, and Correlation coefficient (R). The recovered leaf reflectance yielded the accurate results for LCC during the calibration ($R = 0.99$, $RMSE = 2.64 \mu\text{g/g}$, and $NRMSE = 2.97 \%$) (Fig. 9 (a)) and validation ($RMSE = 16.39$, $Correlation\ coefficient\ (R) = 0.66$, $NRMSE = 17.05 \%$), as shown in (Fig. 9 (b)). The Car was estimated using leaf reflectance data, with a satisfactory level of accuracy ($Correlation\ coefficient\ (R) = 0.74$, $RMSE = 4.73 \mu\text{g/g}$ and $NRMSE = 18.05 \%$) during the calibration (Fig. 10(a)) and ($Correlation\ coefficient\ (R) = 0.57$, $RMSE = 4.38 \mu\text{g/g}$ and $NRMSE = 5.69 \%$) during the validation (Fig. 10(b)). Similarly, the LAI was estimated with high accuracy ($Correlation\ coefficient\ (R) = 0.99$, $RMSE = 0.21 \text{ m}^2/\text{m}^2$ and $NRMSE = 3.56 \%$) during calibration (Fig. 11(a)) and ($Correlation\ coefficient\ (R) = 0.88$, $RMSE = 0.91 \text{ m}^2/\text{m}^2$ and $NRMSE = 1.01 \%$) during the validation (Fig. 11(b)).

5.1.5. Crop parameters estimation at a larger scale

The most effective inversion approach was then used to analyse the EnMAP satellite data for LCC, Car, and LAI mapping. The ARTMO LUT was resampled to accomplish this based on the EnMAP wavelengths. The implemented inversion techniques and cost function for retrieving LCC,

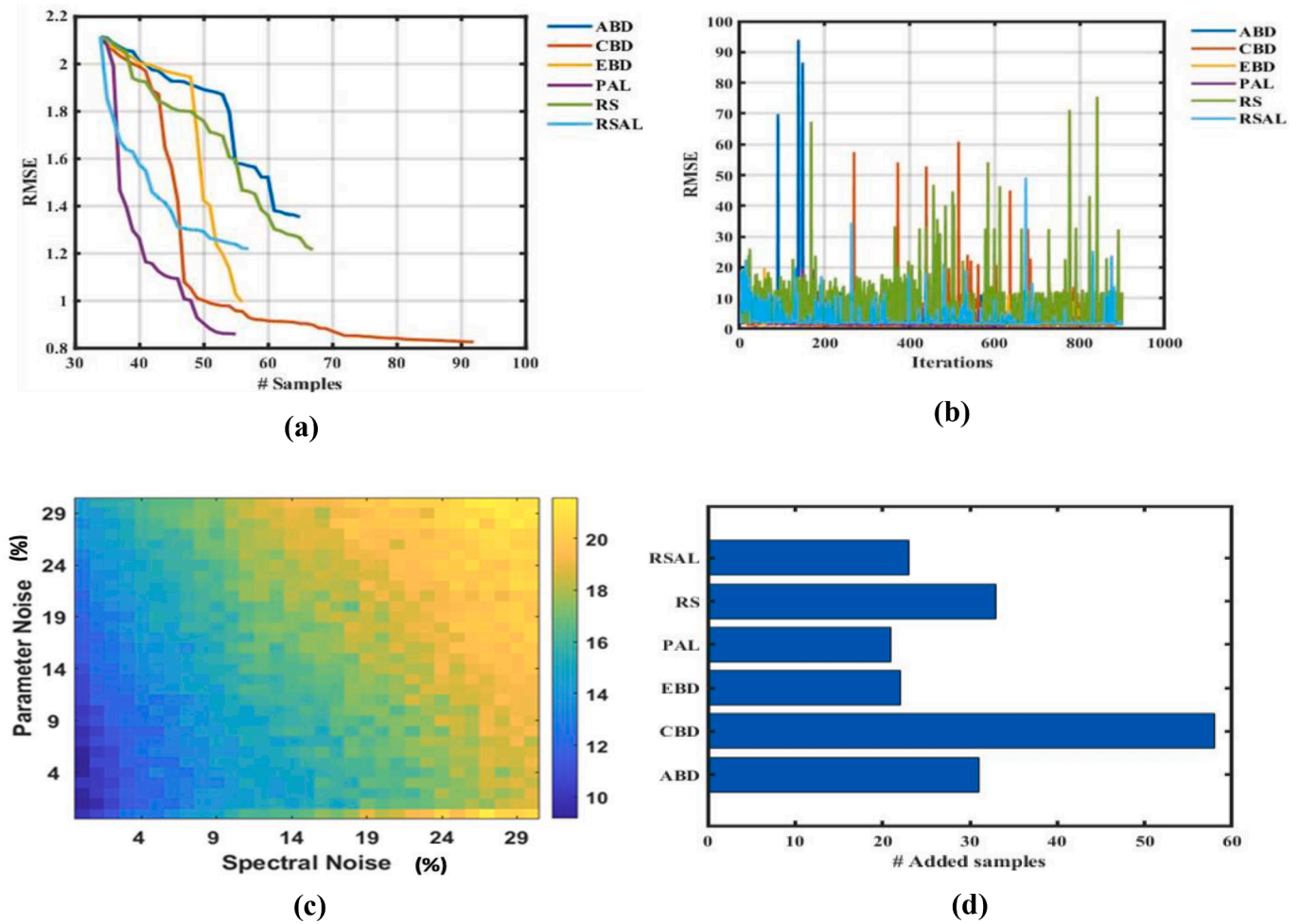


Fig. 8. AL samples performance (a), Iteration (b), Correlation Matrix plot between Spectral and Parameter Noise (c) and Added samples (d) for LAI.

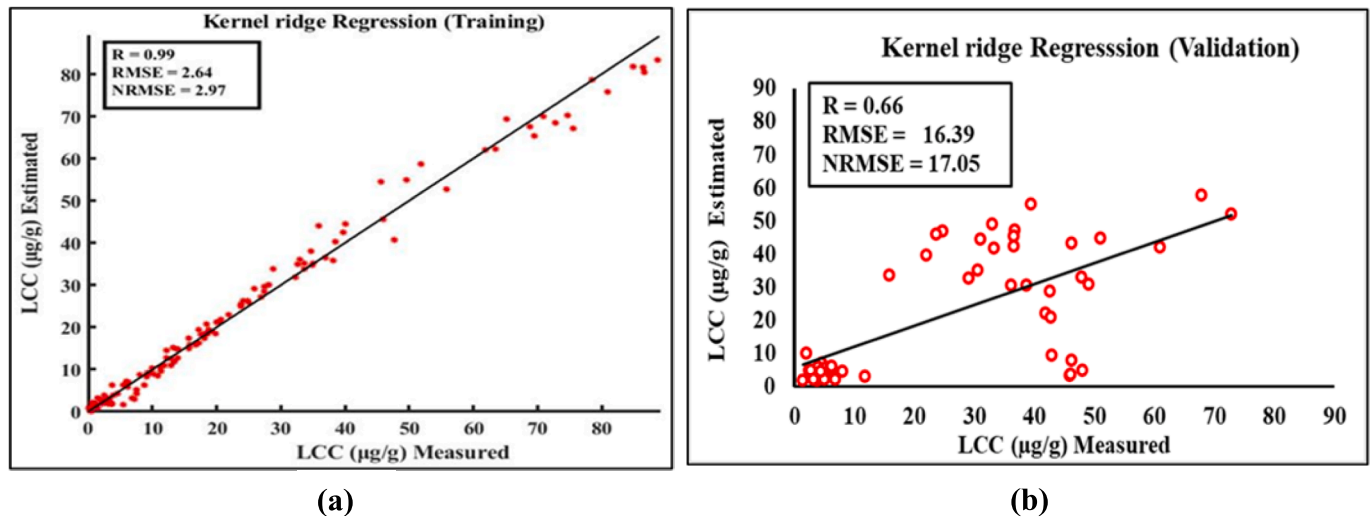


Fig. 9. LCC performance with KRR + EBD (a) training and (b) validation.

Car, and LAI produced some excellent results during validation with the simulated LUT-based data. Subsequently, the most effective MLR + AL inversion approach was utilized to process the EnMAP data and produce accurate maps of LCC, Car, and LAI for the chosen study area. The most highly correlated hybrid model was then used to retrieve the LCC over the area using EnMAP and is shown in Fig. 12. The estimated LCC ranged

from 0 to 70 $\mu\text{g/g}$. The Car range varied from 0 to 1.2 $\mu\text{g/g}$, as depicted in Fig. 13. The most recent LAI mapping was also conducted using EnMAP satellite data, with values varying from 0 to 7 m^2/m^2 (Fig. 14). The LCC, Car, and LAI values obtained by the MLRA-based retrieval technique are strongly linked with the collected datasets in the field. Based on the results obtained and comprehensive research of the

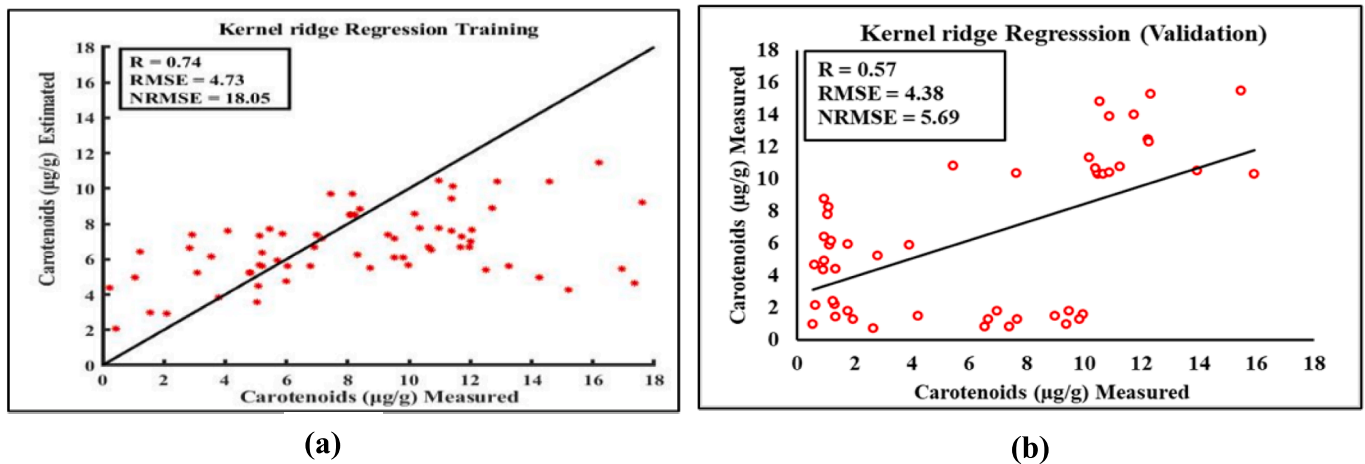


Fig. 10. Estimated Car performance with KRR + EBD model (a) training and (b) validation.

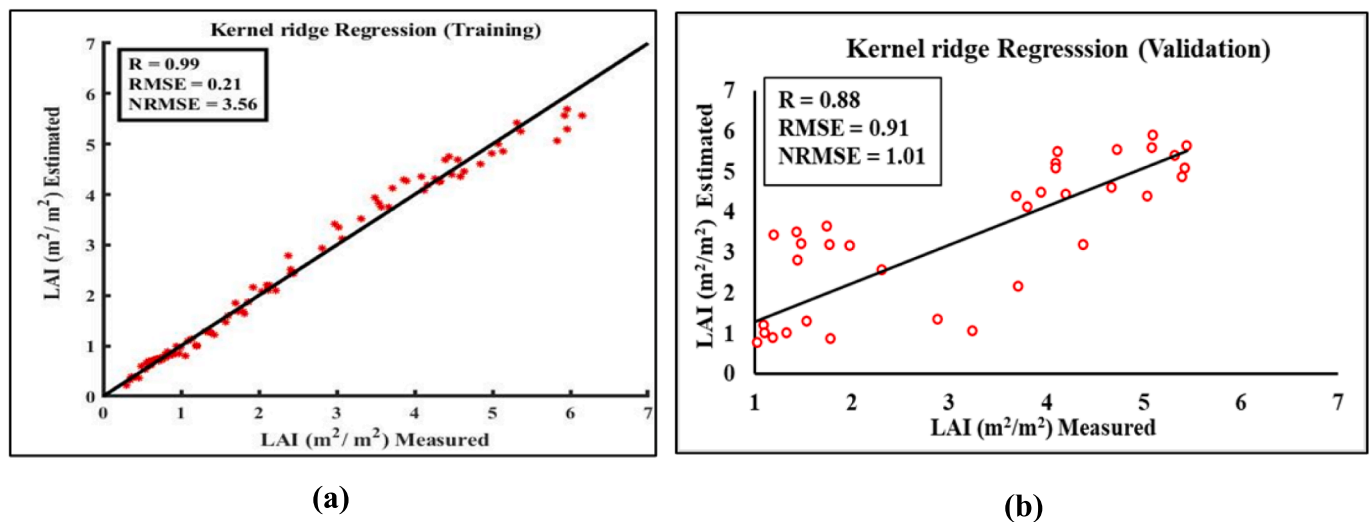


Fig. 11. LAI performance with KRR + EBD model (a) training and (b) validation.

different inversion performances, it is evident that retrieving crop parameters is a highly useful mapping approach that relies purely on crop reflectance information. In summary, the map that was created showed that how the MLRA-based retrieval model can be used to generate high spatial resolution maps of LCC, carotenoids, and LAI at a larger scale.

5.2. Discussion

This study highlights the potential of combining machine learning (ML) and active learning (AL) strategies with radiative transfer model (RTM) simulations using spaceborne hyperspectral EnMAP data to accurately retrieve key crop biophysical parameters, specifically LCC, Car and LAI. Among the models assessed, kernel ridge regression (KRR) achieved the highest accuracy for all parameters, particularly when integrated with an AL strategy based on expected model change with disagreement (EBD). The performance indicators showed significantly high correlation values during calibration (LCC = 0.99, Car = 0.74, and LAI = 0.99) and validation (LCC = 0.66, Car = 0.57, and LAI = 0.88). The EBD-based AL notably improved retrieval performance for LCC, Car, and LAI, demonstrating the effectiveness of AL in reducing model training requirements and maximizing accuracy with fewer labeled samples.

Our findings align with a growing body of literature supporting hybrid approaches that combine RTMs with ML or AL to enhance

retrieval accuracy and efficiency. For example, [Sahoo et al. \(2023\)](#) used simulated hyperspectral data and Gaussian Process Regression (GPR) to estimate LAI and canopy chlorophyll content (CCC), achieving high validation scores ($R^2 = 0.889$ for LAI and 0.656 for CCC). Similarly, [Tsele and Ramoelo \(2024\)](#) and [Ma et al. \(2025\)](#) utilized Sentinel-2 and UAV hyperspectral data alongside RTM-based active learning approaches, showcasing their effectiveness in selecting the most informative training samples from large, simulated datasets. These studies collectively underscore the value of hybrid techniques in addressing the limitations of sparse field data and improving retrieval robustness. In another study, [Jin et al. \(2024\)](#) explored a deep learning–RTM hybrid approach to estimate crop chlorophyll content from GF-5 hyperspectral satellite data, reporting strong accuracy and illustrating the adaptability of hybrid frameworks to various sensors. [Zhang et al. \(2008\)](#) combined PROSAIL simulations with support vector regression (SVR) to estimate rice canopy nitrogen and LAI using PRISMA hyperspectral imagery. Their results contributed to the growing consensus on the effectiveness of RTM–ML integration for trait retrieval at regional scales. However, none of these studies reported a hybrid approach for the simultaneous retrieval of LCC, Car, and LAI using spaceborne EnMAP data. Our study addresses this gap by demonstrating that integrating EBD-based active learning with KRR enables efficient, accurate multi-trait estimation even in data-limited scenarios, offering a novel and operationally scalable solution. This study supports next-generation missions like EnMAP,

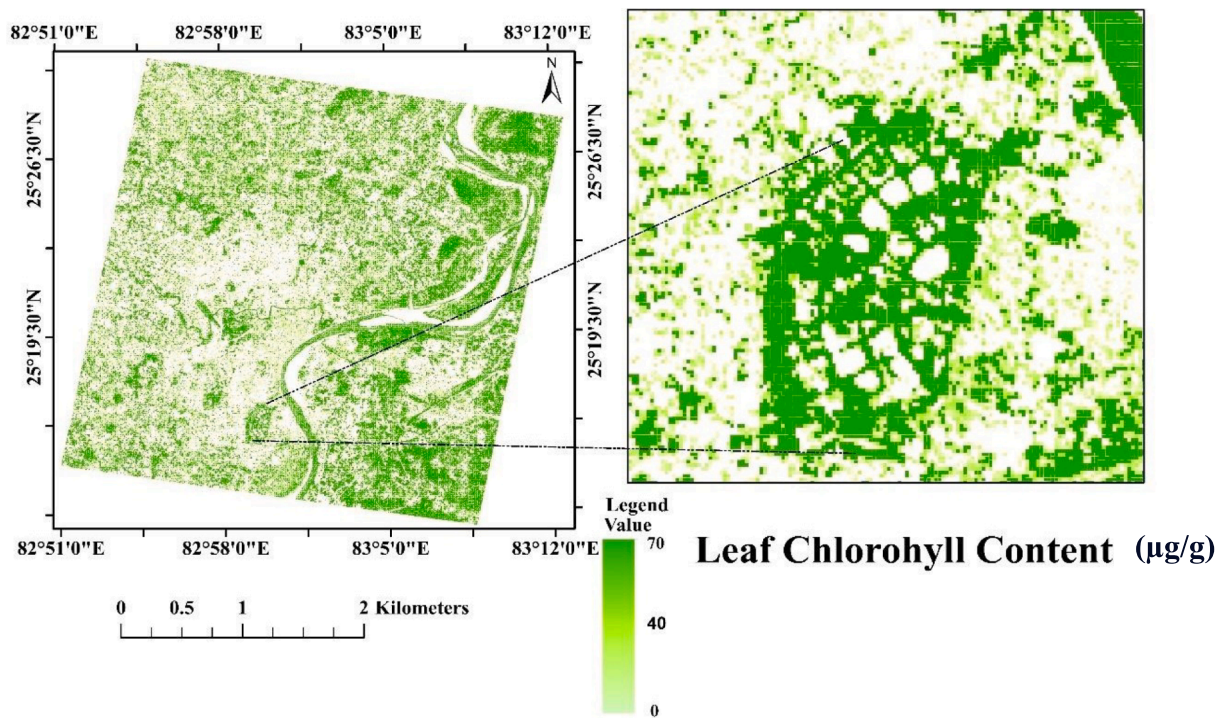


Fig. 12. LCC mapping for whole area left (panel) and subset right (panel) over the test site.

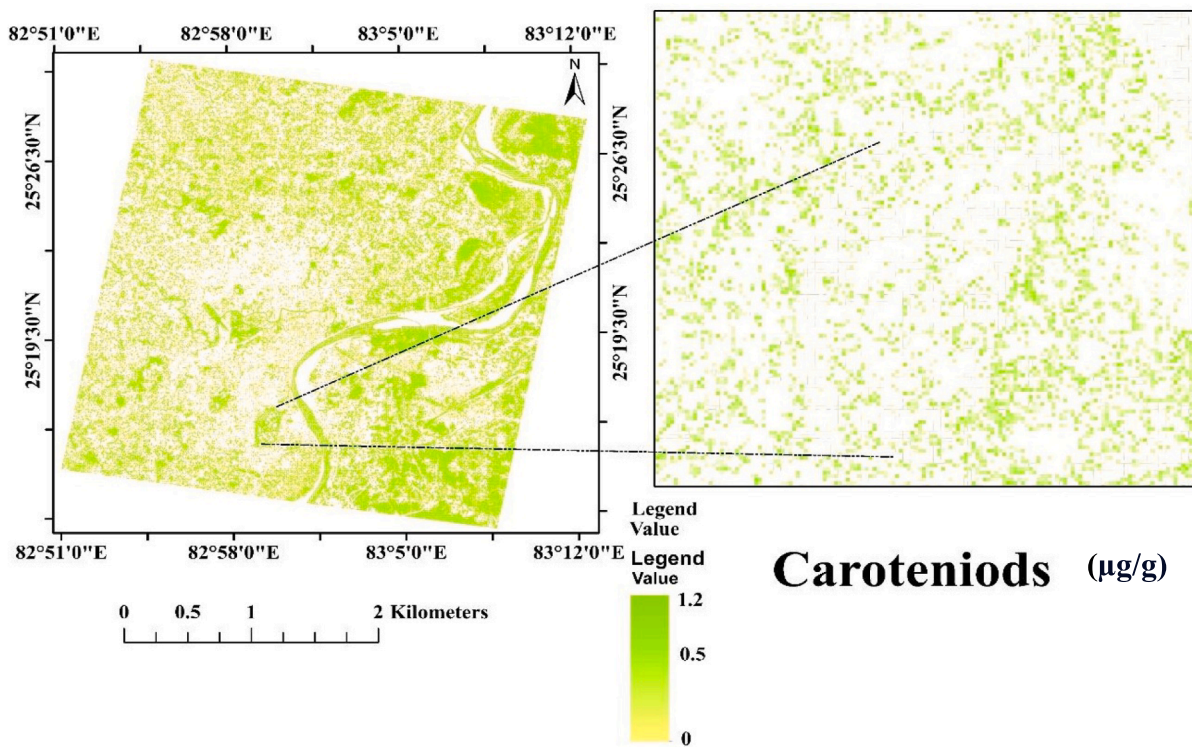


Fig. 13. Car mapping for whole area left (panel) and subset right (panel) over the test site.

PRISMA, and NASA's SBG program, which aim to enhance terrestrial ecosystem monitoring with high-resolution, spectral data. These missions facilitate the quantitative estimation of vegetation biochemical traits on regional to global scales. The developed framework shows how to effectively utilize these datasets using AL-RTM hybrid methods, especially in data-scarce areas.

Hybrid modelling approaches encounter challenges, such as the computational intensity involved in generating synthetic reflectance libraries covering diverse biophysical parameters. The high dimensionality and nonlinear nature of hyperspectral data also complicate sample identification. As noted by Verrelst et al. (2016), the effectiveness of RTM-based simulations relies on the calibration and representativeness

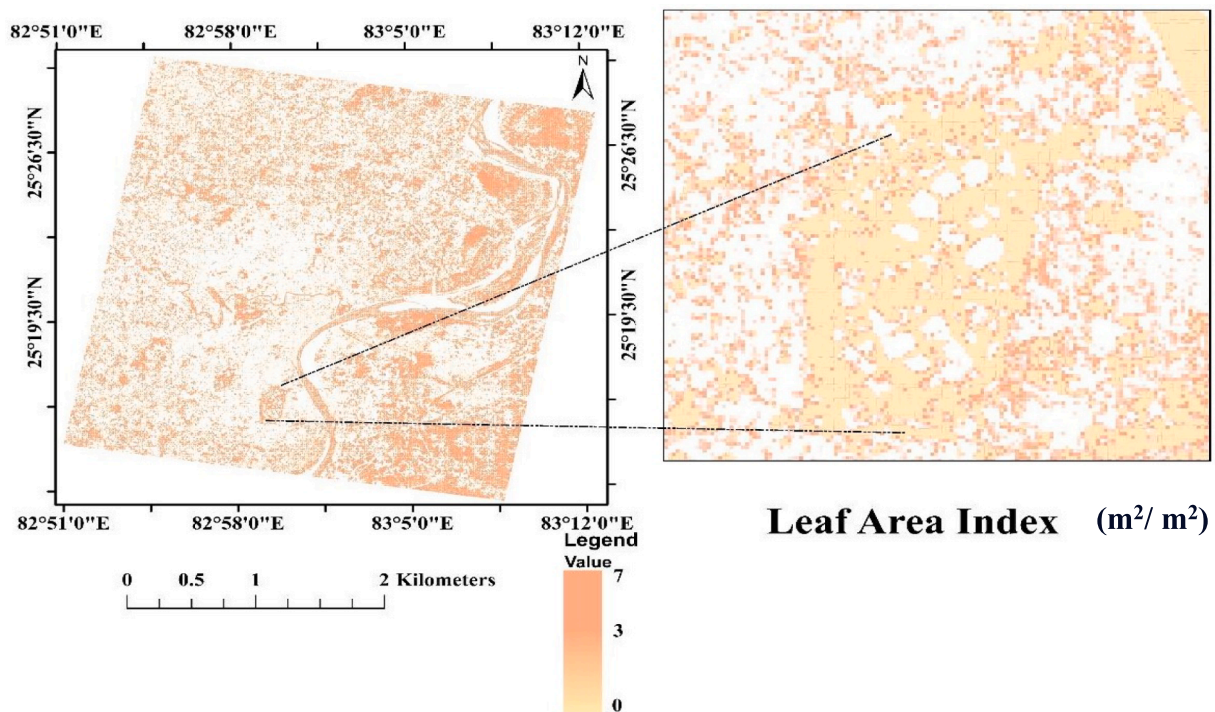


Fig. 14. LAI mapping for whole area left (panel) and subset right (panel) over the test site.

of input parameters. This study tackles these issues by integrating EBD-based AL, which optimally selects impactful training samples, reducing the need for extensive ground datasets. The proposed hybrid framework is scalable and adaptable for estimating additional traits like nitrogen content, anthocyanins, or water stress indicators. As future hyperspectral missions improve in resolution, such frameworks will prove increasingly valuable for agricultural monitoring and environmental modelling.

In conclusion, this study contributes to the evolving field of hyperspectral remote sensing by demonstrating an effective, data-efficient approach for biophysical parameter retrieval. By aligning with the capabilities of current and future hyperspectral missions, this work advances scalable solutions for large-scale vegetation monitoring and precision agriculture.

6. Conclusion

The research shows the effectiveness of PROSAIL, RTM-based MLRA inversion, in mapping crop metrics over the agricultural region. To accomplish this, MLRA-based inversion strategies using 14 types of MLRA algorithms such as Weighted k-nearest neighbor regression (W-NNR), K-nearest neighbour regression (K-NNR), Adaptive Regression Splines (ARS), Kernel ridge Regression (KRR), Random Forest (TreeBagger), Regularized least-squares regression (RLS), Bagging trees (BT), Gaussian Processes Regression (GPR), Regression tree (RT), Elastic Net regression (ENR), Boosting trees, Principal Component Regression (PCR), and Partial least squares regression (PLS) were thoroughly assessed. These strategies were enhanced with advanced active learning, including spectral and parameter noise, and dimensionality reduction using PCA. The evaluation was conducted using ASD-Spectroradiometer and EnMAP satellite reflectance data. Out of many models tested, the most effective approach with AL was used for the inversion and implemented with the EnMAP satellite data to produce large-scale maps of the LCC, Car, and LAI for the chosen study area. Out of all the MLRA assessed, the “KRR” yielded the highest level of accuracy in estimating LCC, Car, and LAI. To enhance the model’s accuracy, KRR was integrated with AL using the EBD. The approach provided a significantly improved retrieval

of the wheat biophysical and biochemical parameters. A comprehensive crop parameters map was then generated with the EnMAP satellite datasets with good precision, which indicates that the hybrid-based methods are generally useful and can be extended for the optically active phyto-pigments.

CRediT authorship contribution statement

Prachi Singh: Writing – review & editing, Writing – original draft, Visualization, Validation, Software, Methodology, Investigation, Formal analysis, Conceptualization. **Prashant K. Srivastava:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Prakash Kumar Jha:** Writing – review & editing. **Jochem Verrelst:** Writing – review & editing, Resources, Data curation. **Pashupati Nath Singh:** Writing – review & editing. **Rajendra Prasad:** Visualization, Validation.

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

We are thankful to the Institute of Eminence, Banaras Hindu University, for funding this research.

Data availability

Data will be made available on request.

References

- Adamu, B., Tansey, K., Ogutu, B., 2015. Using vegetation spectral indices to detect oil pollution in the Niger Delta. *Remote Sens. Lett.* 6 (2), 145–154.

- Adinehvand, M., Rai, P.K., 2016. Assessment of Nutritional Status and Deficiency Disease through Geographical Survey: a Case Study of Varanasi District in India. In: *Forum Geografic*, Vol. 15, No. 2.
- Aghababaei, M., Ebrahimi, A., Naghipour, A.A., Asadi, E., Pérez-Suay, A., Morata, M., Garcia, J.L., Rivera Caicedo, J.P., Verrelst, J., 2022. Introducing ARTMO's Machine-Learning Classification Algorithms Toolbox: Application to plant-type detection in a semi-steppe Iranian landscape. *Remote Sens. (Basel)* 14 (18), 4452.
- Asner, G.P., 1998. Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sens. Environ.* 64 (3), 234–253.
- Asrar, G., & Myneni, R. B. (1991). Applications of radiative transfer models for remote sensing of vegetation conditions and states. In *Photon-Vegetation Interactions: Applications in Optical Remote Sensing and Plant Ecology* (pp. 537–558).
- Baret, F., Vanderbilt, V.C., Steven, M.D., Jacquemoud, S., 1994. Use of spectral analogy to evaluate canopy reflectance sensitivity to leaf optical properties. *Remote Sens. Environ.* 48 (2), 253–260.
- Bealde, C.L., Talbot, H., Jarvis, P.G., 1982. Canopy structure and leaf area index in a mature Scots pine forest. *Forestry: an International Journal of Forest Research* 55 (2), 105–123.
- Belanger, M.J., Miller, J.R., Boyer, M.G., 1995. Comparative relationships between some red edge parameters and seasonal leaf chlorophyll concentrations. *Can. J. Remote. Sens.* 21 (1), 16–21.
- Brell, M., Guanter, L., Segl, K., Scheffler, D., Bohn, N., Bracher, A., Soppa, M.A., Foerster, S., Storch, T., Bachmann, M., Chabrillat, S., 2021. The EnMAP satellite–Data product validation activities. In: 2021 11th workshop on hyperspectral imaging and signal processing: Evolution in remote sensing (WHISPERS). IEEE, pp. 1–5.
- Darvishzadeh, R., Atzberger, C., Skidmore, A., Schlerf, M., 2011. Mapping grassland leaf area index with airborne hyperspectral imagery: a comparison study of statistical approaches and inversion of radiative transfer models. *ISPRS J. Photogramm. Remote Sens.* 66 (6), 894–906.
- Demarez, V., Gastellu-Etchegorry, J.P., 2000. A modeling approach for studying forest chlorophyll content. *Remote Sens. Environ.* 71 (2), 226–238.
- Douak, F., Benoudjit, N., Melgani, F., 2011. A two-stage regression approach for spectroscopic quantitative analysis. *Chemom. Intel. Lab. Syst.* 109 (1), 34–41.
- Douak, F., Melgani, F., Benoudjit, N., 2013. Kernel ridge regression with active learning for wind speed prediction. *Appl. Energy* 103, 328–340.
- ElRafey, A., Wojtusiak, J., 2018. A hybrid active learning and progressive sampling algorithm. *Int. J. Mach. Learn. Comput* 8 (5).
- Feret, J.B., François, C., Asner, G.P., Gitelson, A.A., Martin, R.E., Bidol, L.P., Ustin, S.L., Le Maire, G., Jacquemoud, S., 2008. PROSPECT-4 and 5: Advances in the leaf optical properties model separating photosynthetic pigments. *Remote Sens. Environ.* 112 (6), 3030–3043.
- Fernández-Marín, B., Artetxe, U., Barrutia, O., Esteban, R., Hernández, A., García-Plazaola, J.I., 2015. Opening Pandora's box: cause and impact of errors on plant pigment studies. *Front. Plant Sci.* 6, 148.
- Galvão, L.S., Epifanio, J.C.N., Breunig, F.M., Formaggio, A.R., 2018. Crop type discrimination using hyperspectral data: advances and perspectives. *Biophysical and Biochemical Characterization and Plant Species Studies* 183–210.
- Gitelson, A.A., 2018. Recent developments in remote estimation of crop biophysical and biochemical properties at various scales. *Biophysical and Biochemical Characterization and Plant Species Studies* 3–24.
- Gitelson, A.A., Viña, A., Ciganda, V., Rundquist, D.C., Arkebauer, T.J., 2005. Remote estimation of canopy chlorophyll content in crops. *Geophys. Res. Lett.* 32 (8).
- Gonzalez-Dugo, V., Hernandez, P., Solis, I., Zarco-Tejada, P.J., 2015. Using high-resolution hyperspectral and thermal airborne imagery to assess physiological condition in the context of wheat phenotyping. *Remote Sens. (Basel)* 7 (10), 13586–13605.
- Gupta, A., Singh, P., Srivastava, P.K., Pandey, M.K., Anand, A., Chandra Sekar, K., Shanker, K., 2022. Development of hyperspectral indices for anti-cancerous Taxol content estimation in the Himalayan region. *Geocarto Int.* 37 (25), 7699–7715.
- Gupta, A., Srivastava, P.K., Petropoulos, G.P., Singh, P., 2021. Statistical Unfolding Approach to Understand Influencing Factors for Taxol Content Variation in High Altitude Himalayan Region. *Forests* 12 (12), 1726.
- Hill, J., Buddenbaum, H., Townsend, P.A., 2019. Imaging spectroscopy of forest ecosystems: perspectives for the use of space-borne hyperspectral earth observation systems. *Surv. Geophys.* 40 (3), 553–588.
- Hosikian, A., Lim, S., Halim, R., Danquah, M.K., 2010. Chlorophyll extraction from microalgae: a review on the process engineering aspects. *Int. J. Chem. Eng.* 2010 (1), 391632.
- Jacquemoud, S., Baret, F., Andrieu, B., Danson, F.M., Jaggard, K., 1995. Extraction of vegetation biophysical parameters by inversion of the PROSPECT+ SAIL models on sugar beet canopy reflectance data. Application to TM and AVIRIS sensors. *Remote Sens. Environ.* 52 (3), 163–172.
- Jay, S., Bendoula, R., Hadoux, X., Féret, J.B., Gorretta, N., 2016. A physically-based model for retrieving foliar biochemistry and leaf orientation using close-range imaging spectroscopy. *Remote Sens. Environ.* 177, 220–236.
- Ji, S., Gu, C., Xi, X., Zhang, Z., Hong, Q., Huo, Z., Zhao, H., Zhang, R., Li, B., Tan, C., 2022. Quantitative monitoring of leaf area index in rice based on hyperspectral feature bands and ridge regression algorithm. *Remote Sens. (Basel)* 14 (12), 2777.
- Jin, Z., Liu, H., Cao, H., Li, S., Yu, F., Xu, T., 2024. Hyperspectral Remote Sensing Estimation of Rice Canopy LAI and LCC by UAV coupled RTM and Machine Learning. *Agriculture* 15 (1), 11.
- Ma, Y., Song, X., Zhang, J., Pan, D., Feng, H., Yang, G., Qiu, C., Sun, H., Zheng, C., Li, P., 2025. Inversion of biophysical parameters of potato based on an active learning pool-based sampling strategy. *Int. J. Remote Sens.* 46 (1), 320–339.
- Mananze, S., Pôças, I., Cunha, M., 2018. Retrieval of maize leaf area index using hyperspectral and multispectral data. *Remote Sens. (Basel)* 10 (12), 1942.
- Marandi, M., Parida, B.R., Ghosh, S., 2022. Retrieving vegetation biophysical parameters and GPP using satellite-driven LUE model in a National Park. *Environ. Dev. Sustain.* 24 (7), 9118–9138.
- Mariotto, I., Thenkabail, P.S., Huete, A., Slonecker, E.T., Platonov, A., 2013. Hyperspectral versus multispectral crop-productivity modeling and type discrimination for the HypSIRI mission. *Remote Sens. Environ.* 139, 291–305.
- Martínez-Huitle, C.A., Rodrigo, M.A., Sirés, I., Scialdone, O., 2023. A critical review on latest innovations and future challenges of electrochemical technology for the abatement of organics in water. *Appl. Catal. B* 328, 122430.
- Miraglio, T., Adeline, K., Huesca, M., Ustin, S., Briottet, X., 2019. Monitoring LAI, chlorophylls, and carotenoids content of a woodland savanna using hyperspectral imagery and 3D radiative transfer modeling. *Remote Sens. (Basel)* 12 (1), 28.
- Nistor, M.M., Rai, P.K., Dugesar, V., Mishra, V.N., Singh, P., Arora, A., Kumra, V.K., Carebia, I.A., 2020. Climate change effect on water resources in Varanasi district, India. *Meteorological Applications* 27 (1), e1863.
- Parida, B.R., Kumari, A., 2021. Mapping and modeling mangrove biophysical and biochemical parameters using Sentinel-2A satellite data in Bhitarkanika National Park, Odisha. *Model. Earth Syst. Environ.* 7, 2463–2474.
- Pascual-Venteo, A.B., García, J.L., Berger, K., Estévez, J., Vicent, J., Pérez-Suay, A., Van Wittenberghe, S., Verrelst, J., 2024. Gaussian process regression hybrid models for the top-of-atmosphere retrieval of vegetation traits applied to PRISMA and EnMAP imagery. *Remote Sens. (Basel)* 16 (7), 1211.
- Richards, K.R., Stokes, C., 2004. A review of forest carbon sequestration cost studies: a dozen years of research. *Clim. Change* 63 (1), 1–48.
- Rivera, J.P., Verrelst, J., Leonenko, G., Moreno, J., 2013. Multiple cost functions and regularization options for improved retrieval of leaf chlorophyll content and LAI through inversion of the PROSAIL model. *Remote Sens. (Basel)* 5 (7), 3280–3304.
- Rivera-Cañedo, J.P., Verrelst, J., Muñoz-Marí, J., Camps-Valls, G., Moreno, J., 2017. Hyperspectral dimensionality reduction for biophysical variable statistical retrieval. *ISPRS J. Photogramm. Remote Sens.* 132, 88–101.
- Sahoo, R.N., Gakhar, S., Rejith, R.G., Verrelst, J., Ranjan, R., Kondraju, T., Meena, M.C., Mukherjee, J., Daas, A., Kumar, S., Kumar, M., 2023. Optimizing the retrieval of wheat crop traits from UAV-borne hyperspectral image with radiative transfer modelling using gaussian process regression. *Remote Sens. (Basel)* 15 (23), 5496.
- Sahoo, R.N., Ray, S.S., Manjunath, K.R., 2015. Hyperspectral remote sensing of agriculture. *Curr. Sci.* 848–859.
- Shah, S.H., Angel, Y., Houborg, R., Ali, S., McCabe, M.F., 2019. A random forest machine learning approach for the retrieval of leaf chlorophyll content in wheat. *Remote Sens. (Basel)* 11 (8), 920.
- Singh, P., Srivastava, P.K., Malhi, R.K.M., Chaudhary, S.K., Verrelst, J., Bhattacharya, B. K., Raghunathi, A.S., 2020. Denoising AVIRIS-NG data for generation of new chlorophyll indices. *IEEE Sens. J.* 21 (5), 6982–6989.
- Storch, T., Honold, H.P., Alonso, K., Pato, M., Mücke, M., Basili, P., Chabrillat, S., Fischer, S., 2020. Status of the imaging spectroscopy mission EnMAP with radiometric calibration and correction. *ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci.* 1, 41–47.
- Stuffer, T., Kaufmann, C., Hofer, S., Förster, K.P., Schreier, G., Mueller, A., Eckardt, A., Bach, H., Penné, B., Benz, U., Haydn, R., 2007. The EnMAP hyperspectral imager—An advanced optical payload for future applications in Earth observation programmes. *Acta Astronaut.* 61 (1–6), 115–120.
- Sumanta, N., Haque, C.I., Nishika, J., Suprakash, R., 2014. Spectrophotometric analysis of chlorophylls and carotenoids from commonly grown fern species by using various extracting solvents. *Res J Chem Sci* 2231, 606X.
- Thenkabail, P.S., Lyon, J.G., Huete, A., 2018. Advances in hyperspectral remote sensing of vegetation and agricultural crops. In: *Fundamentals, Sensor Systems, Spectral Libraries, and Data Mining for Vegetation*. CRC Press, pp. 3–37.
- Tsele, P., Ramoelo, A., 2024. Hybrid retrieval of grass biophysical variables based-on radiative transfer, active learning and regression methods using Sentinel-2 data in Marakele National Park. *Geocarto Int.* 39 (1), 2387087.
- Verhoef, W., Bach, H., 2007. Coupled soil–leaf–canopy and atmosphere radiative transfer modeling to simulate hyperspectral multi-angular surface reflectance and TOA radiance data. *Remote Sens. Environ.* 109 (2), 166–182.
- Verma, B., Prasad, R., Srivastava, P.K., Singh, P., Badola, A., Sharma, J., 2022. Evaluation of simulated AVIRIS-NG imagery using a spectral reconstruction method for the retrieval of leaf chlorophyll content. *Remote Sens. (Basel)* 14 (15), 3560.
- Verrelst, J., Camps-Valls, G., Muñoz-Marí, J., Rivera, J.P., Veroustraete, F., Clevers, J.G., Moreno, J., 2015. Optical remote sensing and the retrieval of terrestrial vegetation bio-geophysical properties—a review. *ISPRS J. Photogramm. Remote Sens.* 108, 273–290.
- Verrelst, J., Rivera, J.P., Leonenko, G., Alonso, L., Moreno, J., 2013. Optimizing LUT-based RTM inversion for semiautomatic mapping of crop biophysical parameters from Sentinel-2 and -3 data: Role of cost functions. *IEEE Trans. Geosci. Remote Sens.* 52 (1), 257–269.
- Verrelst, J., Sabater, N., Rivera, J.P., Muñoz-Marí, J., Vicent, J., Camps-Valls, G., Moreno, J., 2016. Emulation of leaf, canopy and atmosphere radiative transfer models for fast global sensitivity analysis. *Remote Sens. (Basel)* 8 (8), 673.
- Weiss, M., Baret, F., Myneni, R., Pragnère, A., Knyazikhin, Y., 2000. Investigation of a model inversion technique to estimate canopy biophysical variables from spectral and directional reflectance data. *Agronomie* 20 (1), 3–22.
- Younger, P.L., 2009. Groundwater in the environment: an introduction. John Wiley & Sons.
- Zabel, F., Putzenlechner, B., Mauser, W., 2014. Global agricultural land resources—a high resolution suitability evaluation and its perspectives until 2100 under climate change conditions. *PLoS One* 9 (9), e107522.

- Zeng, L., Wardlow, B.D., Xiang, D., Hu, S., Li, D., 2020. A review of vegetation phenological metrics extraction using time-series, multispectral satellite data. *Remote Sens. Environ.* 237, 111511.
- Zhang, Y., Chen, J.M., Miller, J.R., Noland, T.L., 2008. Leaf chlorophyll content retrieval from airborne hyperspectral remote sensing imagery. *Remote Sens. Environ.* 112 (7), 3234–3247.
- Zhu, Y., Jeon, S., Sung, H., Kim, Y., Park, C., Cha, S., Jo, H.W., Lee, W.K., 2020. Developing UAV-based forest spatial information and evaluation technology for efficient forest management. *Sustainability* 12 (23), 10150.