



Estimation of Crop Yield From Combined Optical and SAR Imagery Using Gaussian Kernel Regression

Yeshanbele Alebele, *Graduate Student Member, IEEE*, Wenhui Wang, Weiguo Yu, Xue Zhang, Xia Yao, *Senior Member, IEEE*, Yongchao Tian , Yan Zhu, Weixing Cao, and Tao Cheng , *Senior Member, IEEE*

Abstract—The synthetic aperture radar (SAR) interferometric coherence can complement optical data for the estimation of crop growth parameters, but it has not been yet investigated for predicting crop yield. Many studies have used machine-learning methods, such as neural networks, random forest, and Gaussian process regression, to estimate crop yield from remotely sensed data. However, their performance depends on the amount of available ground truth data. This study proposed Gaussian kernel regression for rice yield estimation from optical and SAR imagery using a limited amount of ground truth data. The main objective was to investigate the synergetic use of Sentinel-2 vegetation indices and Sentinel-1 interferometric coherence data through Gaussian kernel regression for estimating rice grain yield. The prediction accuracy was assessed using *in situ* measured yield data collected in 2019 and 2020 over Xinghua county in Jiangsu Province, China. In all cases, Gaussian kernel regression outperformed the probabilistic Gaussian regression and Bayesian linear inference. With the independently used optical and SAR data, a better prediction accuracy was achieved with the optical red edge difference vegetation index (RDVI1) ($r^2 = 0.65$, RMSE = 0.61 t/ha) than with the interferometric coherence ($r^2 = 0.52$ and RMSE = 0.79 t/ha). The highest prediction accuracy can be achieved by combining RDVI1 with interferometric coherence at the heading stage ($r^2 = 0.81$ and RMSE = 0.55 t/ha). The results suggest the value of synergy between satellite interferometric coherence and optical indices for crop yield mapping with Gaussian kernel regression.

Index Terms—Gaussian regression, kernels, optical vegetation indices (VIs), SAR interferometric coherence, Sentinel-1, Sentinel-2, yield.

I. INTRODUCTION

RICE is one of the most staple food crops feeding about half of the world population [3]. Rising global demand

Manuscript received March 1, 2021; revised August 11, 2021; accepted October 6, 2021. Date of publication October 8, 2021; date of current version October 27, 2021. This work was supported in part by the National Key R&D Program of China under Grant 2019YFE0126900, in part by the National Natural Science Foundation of China under Grants 41871259 and 32021004, and in part by the Jiangsu Collaborative Innovation Center for Modern Crop Production jointly sponsored by Province and Ministry and Jiangsu Agricultural Science and Technology Innovation Fund under Grant CX(20) 3072. (*Corresponding author: Tao Cheng.*)

The authors are with the National Engineering and Technology Center for Information Agriculture (NETCIA), MOE Engineering Research Center of Smart Agriculture, MARA Key Laboratory of Crop System Analysis and Decision Making, Jiangsu Key Laboratory for Information Agriculture, Institute of Smart Agriculture, Nanjing Agricultural University, Nanjing 210095, China (e-mail: 2017201102@njau.edu.cn; 2017201078@njau.edu.cn; 2019201091@njau.edu.cn; 2018101167@njau.edu.cn; yaoxia@njau.edu.cn; yctian@njau.edu.cn; yanzhu@njau.edu.cn; caow@njau.edu.cn; tcheng@njau.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2021.3118707

for agricultural production places growing pressure on agro-ecosystems and the food chain as a whole, generating a new scenario for agricultural policy and scientific study. Rice has been extensively cultivated, especially in China, to provide food for the growing population. China, being the largest rice producer globally, contributed 23% of the total of 19.26 billion tons of rice production worldwide from 1994 to 2016 [6]. Rice crop yield prediction that is prompt and accurate before harvesting plays a role in rice decision-making, development, forecasting grain market prices, and ensuring food security [8], [9]. The accurate and detailed rice yield mapping can also help locate areas associated with low grain yield, guiding farmers in their decision-making when selecting proper land management practices. Experienced agronomists can achieve a rough yield estimate based on agrometeorological models by compiling survey information during the growing season. Although agrometeorological model based estimates are difficult for the reconciliations with field observations and consume time and energy. Estimates based on field experience, however, are arbitrary and not relevant to large-scale issues. The application of remote-sensing (RS) data to agriculture and crop production has been popular, especially based on predictive empirical models. It is possible to efficiently and quantitatively estimate crop yield by such data [12], [13]. In recent times, the availability of RS imagery enables accurate crop monitoring, improved understanding of the effects of farming practices, and early warnings of low yields [14], [15]. A broad range of RS data and related products are available. The most suitable type will rely on the classification of crops, modeling of vegetation, and understanding water dynamics [16]. In this respect, it is a unique means of providing daily visits to large areas with information on crop status, spatially explicit, and temporally resolved yield maps to be extracted. Most research works on the use of RS data for crop yield estimation rely on visible and infrared sensors, such as advanced high-resolution radiometers, Landsat, or spectroradiometers with moderate resolution imaging (MODIS) [17]. Unfortunately, because of the dispute between temporal resolution and spatial resolution, satellite data methods using these data types have limited potential for high-resolution yield estimation. More recently, high temporal and spatial resolution data are freely available from the constellations of Sentinel-1 and Sentinel-2 of the European Space Agency (ESA) under the Copernicus program [18]. These data types open a new opportunity for high-resolution crop monitoring activity, although their capacity has not been fully established.

Both optical and synthetic aperture radar (SAR) vegetation indices (VIs) have been extensively used for crop yield estimation [19], [20]. Among the used indices, the normalized difference vegetation index (NDVI) values get saturated at dense vegetation resulting in loss of sensitivity to vegetation parameters. In comparison, the enhanced vegetation index (EVI) improves the sensitivity over dense vegetation overcoming the problem of saturation in NDVI. It alleviates the problem of atmospheric and soil background noise [21]. Rice grain showed a positive relationship with backscatter coefficients at the end of the tillering stage and a negative relationship at the end of the grain filling stage. The correlations at late growth stages are not strong enough to produce a highly accurate yield estimate [22]. Most recent studies also focused on the combined exploitation of optical and SAR imagery data for improved yield estimation. Improved yield estimations were reported from the combined usage of optical (MODIS time series) and microwave (vegetation optical depth) data from soil moisture active passive [17].

SAR data have unique advantages in the remote prediction of crop yield because SAR imagery is accessible independent of weather conditions and can complement optical data under adverse weather conditions. For the highest efficiency, SAR data with rich structural information can be combined effectively with optical data with high spectral information about the given target [23]. However, backscatter coefficients derived from SAR imagery are prone to saturation problems at dense vegetation as optical indices. Most recent studies showed that the interferometric coherence derived from SAR could be effectively used for soil moisture and vegetation parameter estimation [24]. The interferometric coherence, which represents a cross-correlation product calculated from two coregistered complex SAR images over a small window of pixels, can complement the optical indices. Interferometric coherence has an opposite trend with optical indices and backscatter coefficient for vegetation over the growth stages [11], [24]. In this study, we propose a novel fusion approach for exploiting interferometric coherence from Sentinel-1A and optical indices from Sentinel-2A data collections for the problem of yield predictions. Direct exploitation of these RS data types for yield prediction in itself is challenging. In this regard, different machine-learning techniques have been developed and improved yield estimation accuracy [3], [19], [25], [26]. The estimation methods are selected based on the tradeoff between the performance in terms of given target parameters, interpretability of results, and computational time. The most extensively used crop yield prediction methods include multiple linear regression, random forest, and neural network [27], [28]. Previous studies have shown that Gaussian process regression (GPR) outperforms other machine-learning approaches for yield prediction.

Gaussian-process-based regression predictions are based on Gaussian distribution assumptions, resulting in a solution that can be encoded as a mean vector and covariance matrix. However, the predictor variable is usually nonlinear observations with high uncertainty. In this case, the standard Gaussian-process-based regression with covariance functions is a poor choice. Due to the uncertainty in observations, the limited number and

sensitivity of observations to some part of the state vectors impact the predictions of the Gaussian-process-based regression model [29]. In these cases, approaches based on sampling are preferred. Samples from the posterior probabilistic density functions (pdf) are drawn as a solution, allowing the solution to be nonnormal multimodal. The probabilistic Gaussian regression model is derived by Bayesian formulation from the standard Gaussian regression model. For the probabilistic Gaussian, we used Markov chain Monte Carlo (MCMC) as a sampling method to generate random samples from the posterior pdf. The sampling priors in MCMC act as an extra constraint on the inference which enables to limit the solution space [30].

In this study, we introduced two Gaussian-based regression approaches, Gaussian kernel regression based on kernel values as weighing functions, and PGPR based on MCMC sampling. The two Gaussian-based approaches were compared with their linear extension, Bayesian linear inference regression to relate SAR and optical derived metrics with field-measured crop yield. Thus, the objectives of this study were as follows:

- 1) to identify the optimal growth stage for which the selected optical indices and interferometric coherence can be used for the prediction of yield;
- 2) to determine if the combination of interferometric coherence and optical indices will improve the estimation of crop yield using selected regression models; and
- 3) to validate the estimation performance of Gaussian kernel regression models with limited ground truth samples and to compare with other Bayesian regression methods.

II. DATA AND PREPROCESSING

A. Study Area and In Situ Datasets

The study was conducted in Xinghua County, Jiangsu province (114° 38' 00" E-122° 00' 00" E, 30° 00' 00" -35° 30' 00" N), China, as shown in Fig. 1. The study area is characterized by hot and rainy summers, cold and rainy winters, and dry and windy springs. Rice and wheat are the major cereal crops grown in the study area. Rice cultivation begins around the end of May to the end of October. The sampling was done over 60 points in 2019 and 2020 crop seasons. Over two transects were selected for vegetation sampling at each sampling field, including plant height, leaf area index, dry biomass, and grain yield through destructive and nondestructive methods.

Rice growing in Xinghua was during the yellow maturing stage at the end of September. Yield sampling measurements were performed for the 2019 and 2020 rice seasons in parallel with Sentinel-1/2 acquisitions over the reference plots. About 120 sample points were established over Xinghua, and the coordinates of each plot were traced using a global positioning system (GPS) Trimble Geoxh receiver.

B. RS Data

The Sentinel-1A SAR data from the European radar C-band imaging system is collected from June to September of each field trip year. The data used were collected by a single lookup complex in interferometry width mode with a 250-km swath

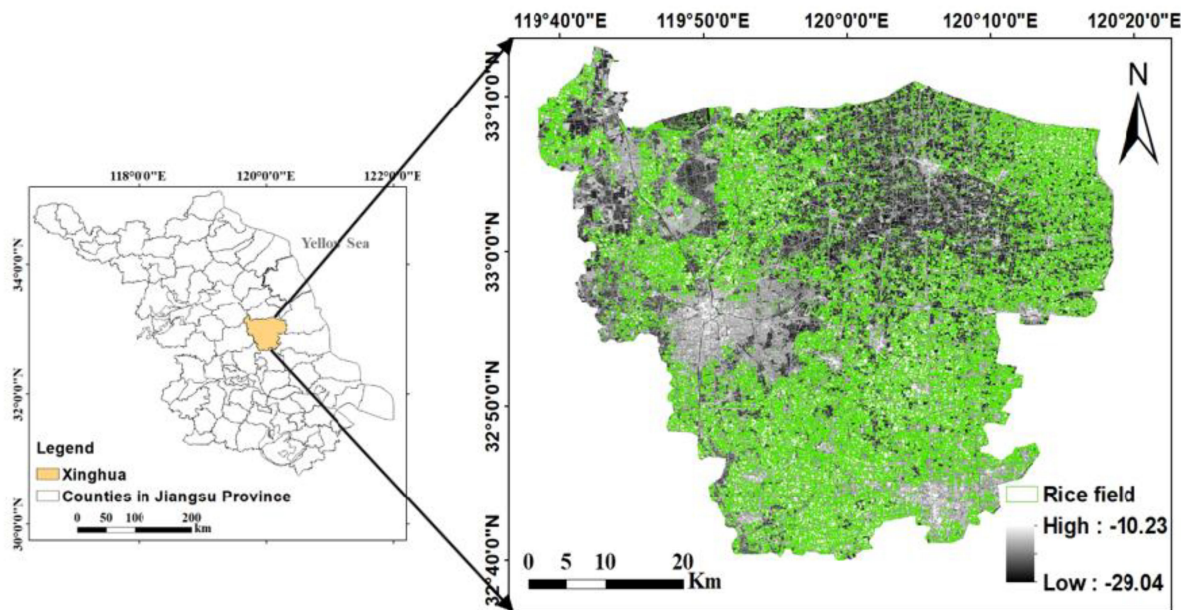


Fig. 1. Location map of the study area in Jiangsu province, China. Also shown is the map for rice fields extracted from Sentinel-2 imagery with the OCSVC method (OCSVM).

width. After range-Doppler terrain correction, a 10-m square pixel size was achieved. Sentinel-1A imagery has VV and VH polarizations and a 12-day repeat cycle. Sentinel-2 sensor carried a multispectral instrument (MSI) with 13 available bands and spectral range of 0.4–2.4 μm . The Sentinel-2 MSI imagery has four bands in 10 m, six bands in 20 m and three bands in 60 m.

All Sentinel images were acquired on similar dates and downloaded from Copernicus Open Access Hub.¹ We applied a range of preprocessing procedures to derive the interferometric coherence from Sentinel-1A SAR imagery using each of the downloaded raw SAR images. The main preprocessing steps included S1-tops splitting, orbit file implementation, S1-back geocoding, coherence estimation, S1-tops deburst, multilooking, speckle filtering, and range to Doppler correction. Each step was implemented using the ESAs SNAP software.² Sentinel-2 images were acquired in the form of fixed cartographic geometry at level-1C with a coordinate system of WGS84 in the UTM projection. The main preprocessing procedures done on Sentinel-2 include atmospheric correction, radiometric calibration, and thermal noise removal. After preprocessing, an RGB composite was created to increase separation among the land cover types for interpretation of the data. From the stack of both Sentinel-1A and Sentinel-2A imagery, we created the RGB composite with images at different dates and polarizations. The RGB composite with field-measured GPS points was used to extract rice fields (see Fig. 2). The extraction of rice fields for the whole study area was done using our field visit in combination with the RGB composite from Sentinel-1 and Sentinel-2 imagery and one-class support vector classification (OCSVC) [31].

C. Used VIs and Features Selection

As a proxy for observed photosynthetically active radiation relative to total biomass, optically derived indices have been widely used for yield estimation problems [32], [33]. Above-ground biomass (AGB) at each specific growth stage can be an indicator of the expected yield. Different VIs have different strengths for the estimation of plant growth parameters and yield. For example, soil and atmospheric noise can be alleviated using EVI, which can reduce the background signal noise by subtracting the blue reflectance band [21], [32]. The early saturation problem common in NDVI can be reduced by using RDVI, as the red edge vegetation band is highly sensitive to AGB than the infrared index in NDVI [23], [34]. The optical saturation, soil, and atmospheric noise can also be alleviated by fusing optical indices with SAR-derived parameters. In this study, different optical indices and SAR interferometric coherence were selected and analyzed for the estimation of yield at different growth stages (see Table I). The suitability of both optical and SAR indices varies at different growth stages and combination scenarios. Studies have shown that VIs are more sensitive to crop growth parameters and yield than the raw spectral bands [35], [36]. However, VIs become saturated from medium to dense vegetation. We proposed joint exploitation of SAR interferometric coherence and optical VIs to solve the problem of spectral saturation at dense vegetation. The interferometric coherence has an opposite trend with VIs complementing the VIs to be used at dense vegetation [11]. Accordingly, we used a regression tree for piecewise linear estimation of variable importance and associations for different SAR interferometric coherence and optical indices at different growth stages. The calculated variable importance (VIMP) was used to select the optimal indices and growth stages for the estimation of yield. VIMP for x_v variable

¹[Online]. Available: <https://scihub.copernicus.eu/dhus/#/home>

²[Online]. Available: <http://step.esa.int/main/download/>

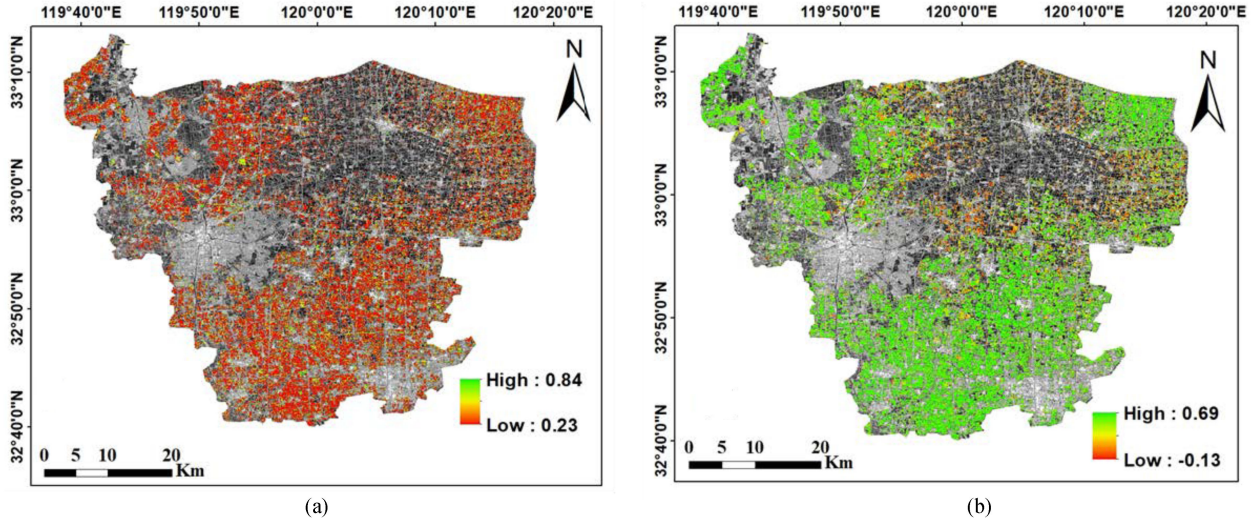


Fig. 2. (a) Interferometric coherence VH from Sentinel-1A, August 2, 2019 and (b) RDVI1 from Sentinel-2A MSI data, September 5, 2020 extracted over rice fields in Xinghua county.

TABLE I
REFERENCES AND EQUATIONS OF VIS USED IN THIS STUDY

Vegetation indices	Formulas	References
Normalized difference vegetation index (NDVI)	$\frac{b7 - b4}{b7 + b4}$	[1]
Enhanced vegetation index (EVI)	$2.5 \left(\frac{b7 - b2}{2.4 * b7 + b2 + 1} \right)$	[2]
Red edge difference vegetation index 1 (RDVI1)	$\frac{b6 - b5}{b6 + b5}$	[4, 5]
Red edge difference vegetation index 2 (RDVI2)	$\frac{b7 - b5}{b7 + b5}$	[4, 7]
Normalized difference water index (NDWI)	$\frac{b8 - b11}{b8 + b11}$	[10]
Interferometric coherence VH & VV		[11]

is the difference between the error of the prediction when x_v is noised up and the error of the prediction otherwise. VIMP helps to explain the difference between prediction error under noise and without noise. What can be called a random assignment of the left-right daughter node corresponds to a variable noise. To determine the paired importance and association between two variables, we consider the joint VIMP of a pair of variables (x_v, x_w) in a binary tree (t). Let $t = (v, w)$ and write x_t to indicate (x_v, x_w). The paired VIMP for x_t is the difference between prediction error when x_t is noised up versus the prediction error otherwise [37]. The fit resemble MATLAB function was used for the calculation of variable importance, and predAssociation MATLAB function was also used for paired importance and association calculation. Paired importance and association are the predictive measures that yield the predictive measure of associations across selected indices at each specific growth stage. The inferred strength of the relationship between paired predictors (indices) was used to combine SAR interferometric and optical indices at the optimal growth stage. The combination of VIs and SAR interferometric coherence was used as input for

the regression models. First, we computed the unbiased predictor importance score for all selected VIs and coherence_VH & VV. Second, VIs with the highest predictor importance score were chosen to be used as input for the selected regression models. Finally, the best performing indices from each sensor type were combined and used as input for the proposed regression models.

III. PROPOSED REGRESSION METHODS

A. Bayesian Linear Regression

For a formal linear regression of a given data, the goal is to find the slope and intercept of for linear equation. From a Bayesian perspective, our goal is not to find the point estimate for the slope and intercept. Instead, we should calculate the posterior distribution for the slope and the intercept, given the data. For linear regression model given as

$$y = \beta_0 + \beta_1 x. \quad (1)$$

Here, β_0 is the y -intercept and β_1 is the slope. Our goal is to calculate the posterior distribution for β_0 and β_1 given the data. First, we obtained the likelihood of these parameters by using the Gaussian sampling distribution assumption, which is the built-in function in python SciPy.Stats. The likelihood is the probability distribution of observed y_i conditioned on the current model parameters (β_0, β_1 , and σ), and feature x_i . $P(y_i | x_i, \beta_0, \beta_1)$ gives the likelihood of the parameters by the Gaussian distribution assumption of y_i and x_i . The prior distribution is the current distribution of the parameters (β_0, β_1) with the observed x_i and y_i , without any Gaussian or binomial distribution assumption. Having the likelihood and priors of β_0 and β_1 , we learned the posterior distribution of these parameters for newly observed x_{i+1} and y_{i+1} pairs, $p(\beta_0, \beta_1) | \text{prior}(\beta_0, \beta_1, x_i, y_i)$. Baye's rule (parameter estimation) tells us how to calculate the posterior distribution with our parameters and given data:

$$p(\beta_0, \beta_1 | y_1, y_2, \dots, y_N) \propto p(\beta_0, \beta_1) p(y_1, y_2, \dots, y_N | \beta_0, \beta_1). \quad (2)$$

There are N data points, and so there are N y -values in the dataset, called (y_1, y_2, \dots, y_N) . This describes our beliefs about the connection between the data and the parameters, without which it would be impossible to learn anything from the data. If we knew the true values of β_0 and β_1 , then we would predict the y -values to be scattered around the straight line. Expressly, we will assume that each point departs from the straight line by an amount ϵ_i , which has a $(N(0, \sigma^2))$ probability distribution. For now, we will assume σ , the standard deviation of the scatter is known. In approximate notation, this can be written as

$$y_i \sim N(\beta_0 + \beta_1 x_i, \sigma^2). \quad (3)$$

It is implied that all the data values are independent (given the parameters). Therefore, the likelihood can be written as a product of normal densities, one for each data point

$$p((y_1, y_2, \dots, y_n) | \beta_0, \beta_1) = \prod_{i=1}^N \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2\sigma^2}(y_i - (\beta_0 + \beta_1 x_i))^2\right]. \quad (4)$$

The above analytical results made the unrealistic assumption that the scatter's standard deviation was known. In practice, σ usually needs to be estimated from the data as well. Therefore, we should include it as an extra unknown parameter in the Bayesian framework. Our parameters are now β_0 , β_1 , and σ . Using the same data and likelihood as before. For Bayesian regression with parameter and data x , we can predict the new data x' by calculating the posterior predictive distribution, which is just the probabilistic distribution for x' given x

$$p(x'|x) = \int p(\theta|x) p(x'|x, \theta) d\theta. \quad (5)$$

The first term inside the integral is posterior, and therefore, the whole integral is an expectation value with respect to the posterior distribution.

B. Bayesian Formulation From GPR

GPR models are nonparametric kernel-based probabilistic models. The mean vector of this joint distribution is generally assumed to be a zero vector, and the covariance matrix is obtained using the covariance function defined over a pair of input values. Using the training set $((x_i, y_i); i = 1, 2, \dots, n)$, where $x \in R^d$ and $y_i \in R$, drawn from an unknown distribution. A GPR model addresses the question of predicting the values of a response variable y_{new} , given the new input vector x_{new} , and the training data. The GPR model establishes a relation between the input $x \in R^d$ and the output variable $y \in R$ of the form $\hat{y} = f(x) = \sum_{i=1}^N \alpha_i k(x_i, x) + \alpha_0$ where $(x_i)_{i=1}^N$ is the derivative ratio used in the training phase, $\alpha_0 \in R$ is the weight assigned to each one of them, α_0 is the basis in the regression function, and k is a function evaluating the similarity between the test data and all N training samples.

GPR is an infinite collection of random variables, any subset of which has a Gaussian distribution. A realization of the Gaussian process is a modeling of a function, $f \sim Gp(m(x), k(x, x'))$.

Depending on the type of data, we can fit different covariance functions with the Gaussian process. For GPR, the kernel function k is used for checking the similarity between the test and training samples. In this study, the automatic relevance determination vector kernel was used

$$k(x, x') = \eta \exp\left[-\frac{(x - x')^2}{2\rho^2}\right] \quad (6)$$

where η and ρ are the unknowns for the hyperparameters of the covariance functions. Equation (1) generates covariance matrices from the prior with no data, or Gp prior is sampled here.

To make predictions with probabilistic GPR (PGPR), we stand from the Bayes formula, which is stated as follows:

$$\Pr(\theta|y) \propto \prod_{i=1}^N \Pr(y_i|\theta) \Pr(\theta) \quad (7)$$

where θ is unknown to be determined from Gp prior, $\Pr(y_i|\theta)$ is the likelihood, and $\Pr(\theta|y)$ is the posterior. The posterior predictive distribution is given by

$$\Pr(y^{\text{new}}|y) = \int \Pr(y^{\text{new}}|\theta) \Pr(\theta|y) d\theta \quad (8)$$

where $\Pr(\theta|y)$ is the posterior and $\Pr(y^{\text{new}}|\theta)$ is the likelihood. y^{new} , prediction from, can be derived as follows:

$$\Pr(y^{\text{new}}|x^{\text{new}}, y, x) = N(\mu^{\text{new}}, \Sigma^{\text{new}}) \quad (9)$$

where μ^{new} is the mean vector and Σ^{new} is the covariance matrix and both of which are determined from the covariance functions. Having the posterior and likelihood, the MCMC method is used to generate the posterior distributions from random samples and to generate the hyperparameter values. The implementation of Gaussian-based regression algorithms was done in Python programming language.

C. Gaussian Kernel Regression

GPR attempts to approximate the target output $f(x)$ where $X \in R^d$ by interpreting it as a probability distribution function. It addresses predicting the values of response variables using the training set drawn from the unknown distribution and given the new input vector. A detailed description of GPR and its application for RS data can be found in most prominent studies [38]–[40], and here, we provide a Gaussian-kernel-based regression on which kernel values are used to derive weights and, finally, to use them to predict outputs from the given inputs [41]. A kernel function must be symmetrical and mathematically this property can be expressed as

$$k(-u) = k(+u). \quad (10)$$

The symmetric property of the kernel function enables its maximum value $\max(k(u))$ to lie in the middle of the curve. The area under the curve of the function must be equal to one. Mathematically, this property is expressed as

$$\int_{-\infty}^{+\infty} k(u) du = 1. \quad (11)$$

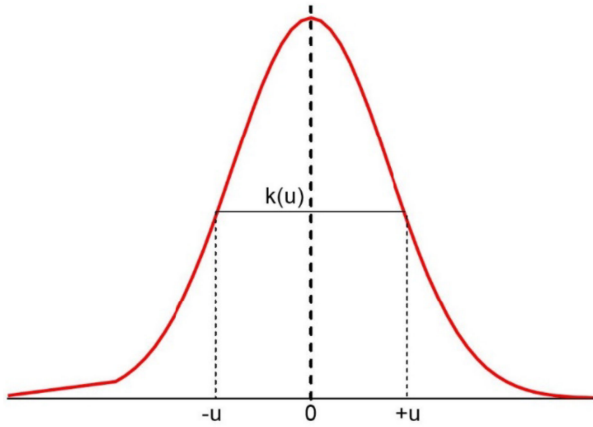


Fig. 3. Illustration of a Gaussian kernel function.

Values of the kernel function cannot be negative, i.e., $k(u) \geq 0$ for all $-\infty < u < +\infty$. The graph of the Gaussian function has a symmetric property of “bell curve” shape, which coincides with the symmetric property of a kernel function (see Fig. 3). The fundamental calculation behind kernel regression is to estimate the weighted sum of all observed y values for a given predictor value x_i . Kernel regression estimates each point using a weighted local sample that is centered around the point of interest. The local sample is weighted using a kernel function that possesses several useful properties. A kernel function defines each observation’s weight within a (typically) symmetric predetermined bandwidth [42]. Unlike other regressions that make no distinction of where the data are located when estimating the conditional expectation, kernel regression will estimate the point of interest using a bandwidth. Before introducing the kernel estimators, we first derive a kernel function. Considering x our point of interest, we can write an indicator function such that data fall within a range h (our bandwidth) around x

$$n_x = \sum_{i=1}^n 1 \left\{ x - \frac{h}{2} \leq x_i \leq x + \frac{h}{2} \right\}. \quad (12)$$

The corresponding probability of falling in this box (centered on x) is thus n_x/n . This indicator function can be rewritten as

$$n_x = \sum_{i=1}^n \left(\frac{1}{2} \right) 1 \left\{ \left| \frac{x_i - x}{h} \right| \leq 1 \right\}. \quad (13)$$

This function is better known as a uniform kernel (Table II) and is more commonly written as

$$k(\varphi) = \begin{cases} 1/2, & \text{if } |\varphi| \leq 1 \\ 0, & \text{otherwise} \end{cases} \quad (14)$$

where $k(\varphi)$ is defined as $\frac{(x_i - x)}{h}$, and represents how local the observation x_i is relative to x .

In this study, the Gaussian kernel is chosen to calculate kernels for the given data points. The expanded form of the Gaussian

 TABLE II
COMMONLY USED SECOND-ORDER KERNEL FUNCTIONS

Kernel	$k(\varphi)$	$k_2 k$
Uniform (S=0)	$\frac{1}{2} 1\{ \varphi \leq 1\}$	1/3
Epanechnikov (S=1)	$\frac{3}{4} (1 - \varphi^2) 1\{ \varphi \leq 1\}$	1/5
Biweight (S=2)	$\frac{15}{16} (1 - \varphi^2)^2 1\{ \varphi \leq 1\}$	1/7
Triweight (S=3)	$\frac{35}{32} (1 - \varphi^2)^3 1\{ \varphi \leq 1\}$	1/9
Gaussian (S= ∞)	$\frac{1}{h\sqrt{2\pi}} e^{-0.5(\varphi)^2}$	1

kernel equation is

$$k(x) = \frac{1}{h\sqrt{2\pi}} e^{-0.5\left(\frac{x-x_i}{h}\right)^2} \quad (15)$$

where x_i is the observed data point, x is the value where kernel function is computed, and h is called the bandwidth. The bandwidth in the kernel regression is used as a smoothing parameter since it controls variance and bias in the output. Kernels are developed at each value of x_i and used to estimate the weighted sum of all observed y values for a given single predictor x_i . Three inputs are required to construct a kernel curve around the given data points. Observation data point (x_i), the value of bandwidth (h), and linearly spaced series of data points that include observed data points. The kernel values are scaled between 0 and 1 and used as weights. The following equation is used to scale the kernel values between 0 and 1

$$w_i = \frac{k_i}{\sum_{i=1}^n k_i} \quad (16)$$

where w_i is the weight for input, and i and n are the number of data points.

A bandwidth h of the kernel may alter the density estimate, and it can accordingly affect the goodness of fit of the density function to the unknown underlying target [43]. When the distance between the data point and the query point is 0, the function produces its highest value; however, as the distance between the data point and the query point increases, the weight value decreases exponentially. When performing kernel regression, we will compute the weighted average across all training points using (16). However, only data points close to the query will have a major impact on the output. In this study, we consider applying a kernel of locally adaptive bandwidth obtained by iteratively computing the optimal bandwidth using the kernel bandwidth optimization method. Given $[x_i]_{i=1}^N$ be our data. Let the kernel with bandwidth h be $k_h(t)$. Then, we compute a formula

$$C_h = \sum_{i,j} \int k_h(x - x_i) k_h(x - x_i) dx - 2 \sum_{i \neq j} k_h(x_i - x_j). \quad (17)$$

We compute h that minimizes C_h . The method estimates a bandwidth that minimizes the expected loss between the kernel estimate and the unknown underlying function from the observed data.

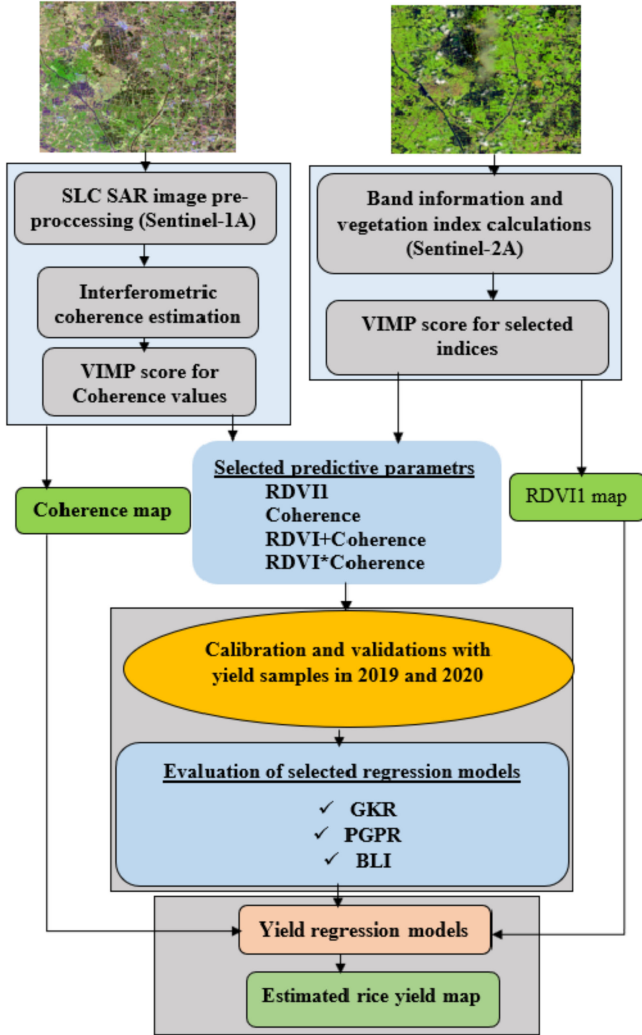


Fig. 4. Summary of the method used in this study for the estimation of rice grain yield at field scale.

D. Statistical Metrics for Evaluating Model Performance

As described in Section II and Fig. 4, the field visits for a sampling of yield parameters were done over Xinghua County in 2019 and 2020. Field measured yield samples collected in 2019 and 2020 were divided into training and validation samples for the chosen regression models. The performance analysis of selected machine-learning methods (Gaussian kernel regression, PGPR, and Bayesian linear regression) for the retrieval of rice grain yield measurements was evaluated through the root-mean-square error (RMSE) and squared correlation coefficient (r^2), which were calculated as

$$r^2 = \frac{\sum_i^n (x_i - \bar{x})^2 (y_i - \bar{y})^2}{\sum_i^n (x_i - \bar{x}) \sum_i^n (y_i - \bar{y})} \quad (18)$$

$$\text{RMSE} = \sqrt{\frac{\sum_i (y_i - \hat{y}_i)^2}{n}} \quad (19)$$

where y_i and \hat{y}_i are the ground truth measurement and estimated yield variables, respectively, for point i , and \bar{y} is the

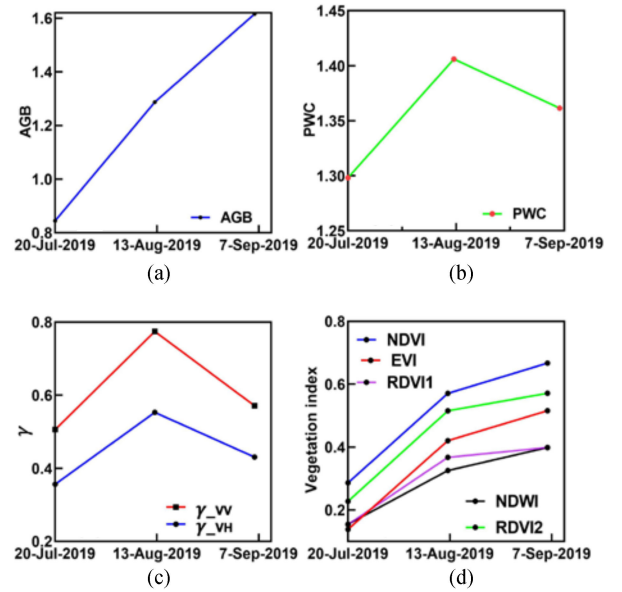


Fig. 5. Temporal variation of (a) AGB, (b) PWC, (c) interferometric coherence VH & VV, and (d) VIs NDVI1, EVI, RDVI1, RDVI2, and NDWI.

TABLE III
CALCULATED VIMP AS A PREDICTIVE MEASURE FOR SELECTED SAR INTERFEROMETRIC COHERENCE AND OPTICAL INDICES

Feature	Jointing	Booting	Heading
	VIMP (%)	VIMP (%)	VIMP (%)
NDVI	-0.12	0.12	-0.23
EVI	0.24	-0.11	0.27
RDVI1	-0.01	0.27	0.45
RDVI2	-0.05	0.13	0.19
NDWI	0.26	-0.13	-0.20
γ_{VH}	-0.21	0.34	0.44
γ_{VV}	0.11	0.01	0.31

arithmetic means of estimated variables, and n is the number of sample points. A good retrieval result contains a low RMSE value and a high correlation coefficient.

IV. RESULTS

A. Trends of VIs and Feature Importance Score

Optical indices for the three main growth stages followed similar trends of AGB (see Fig. 5). The interferometric coherence (γ) showed a consistent relationship with plant water content (PWC) with a continuous increase until August 2019. After August 2019, the interferometric coherence and PWC showed a consistent decline. The interferometric coherence and optical indices had different responses to crop growth parameters at different growth stages. The paired predictive importance of the selected parameters for the prediction of yield at different growth stages was indicated in Table III. The unbiased predictor

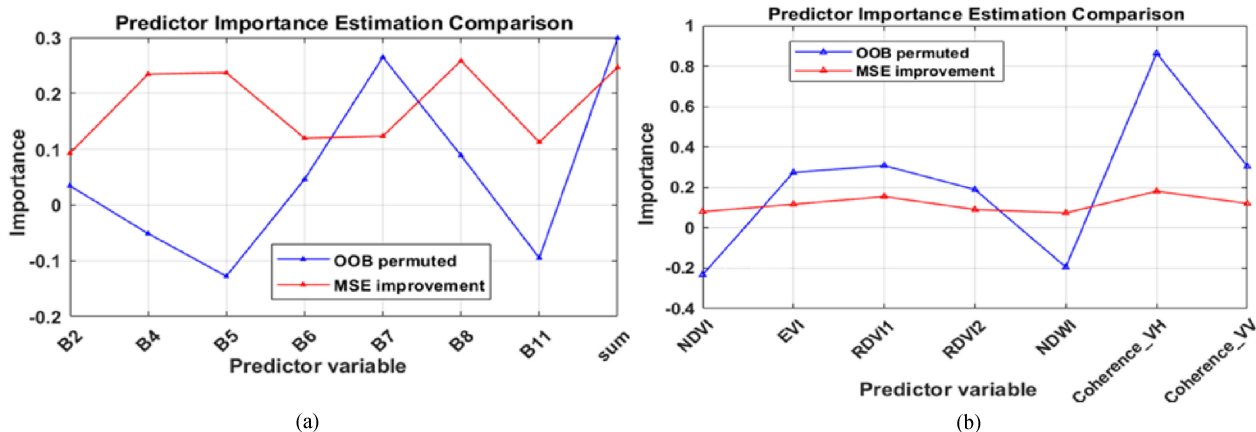


Fig. 6. Predictor importance estimation comparison. (a) Predictor importance comparison among selected spectral bands. (b) Predictor importance comparison for selected optical indices and interferometric coherence VH and VV.

importance estimates showed that NDWI and EVI showed the most important predictive value at the jointing stage. The interferometric coherence and RDVI1 variables resulted in higher performance at the booting stage. Overall, the highest predictive score was observed using interferometric coherence and RDVI1 variables at the heading stage. The higher comparative predictive score infers the predictor variables' strength with selected SAR and optical derived metrics. The predictor importance estimation comparison scores ranked stagewise from the heading stage (highest), then the booting stage and jointing stage (lowest) (see Table III).

B. Relationship of Yield With Selected Indices

SAR and optical vegetation descriptors have shown high sensitivity to AGB and PWC. Among the identified growth stages, the highest feature importance score in percentage was archived at the heading stage. Comparing Vis with spectral bands, we found a higher predictor importance score for VIs than the spectral bands (see Fig. 6). EVI, RDVI1, and RDVI2 resulted in the highest predictive importance score in comparison to spectral bands. We observed a lower MSE improvement value of spectral bands than VIs. The best performing VIs RDVI1, RDVI2, and Coherence_VH resulted in a predictor importance score >0.4, which is higher than that of the spectral bands. Doing correlation between SAR and optical vegetation descriptors with yield parameters resulted in a relatively best correlation with optical RDVI1 ($r^2 = 0.29$ followed by interferometric Coherence $r^2 = 0.23$). The best performing interferometric Coherence and RDVI1 parameters were combined, and give the best predictive feature with a sum of Coherence, and RDVI1 $r^2 = 0.41$, and with the multiplication of Coherence with RDVI1, $r^2 = 0.24$ both at the heading stage (see Fig. 7).

C. Estimation of Yield by Regression Algorithms

The optical index (RDVI1) from Sentinel-2 and interferometric coherence from Sentinel-1A was used as input for Bayesian linear regression to predict yield at different growth stages. From the independently used parameters, the best performance was

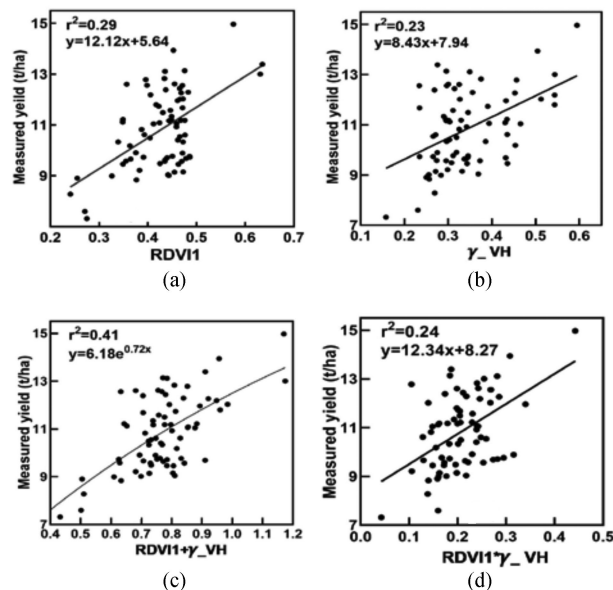


Fig. 7. Relationship between field-measured yield data and various VIs derived from Sentinel-1A SAR and Sentinel-2A MSI data at heading stage.

achieved using RDVI1 metrics with $r^2 = 0.48$ and RMSE = 0.78 t/ha followed by interferometric coherence with VH polarization ($r^2 = 0.47$ and RMSE = 0.89 t/ha) both at the heading stage. From the combined metrics (interferometric coherence and RDVI1), the best statistical results were achieved with a sum of RDVI1 and interferometric coherence ($r^2 = 0.57$ and RMSE = 0.76 t/ha) than the product of interferometric coherence and RDVI1 ($r^2 = 0.54$ and RMSE = 0.83 t/ha) (Table IV).

Selected optical and SAR vegetation descriptors were used as input for the PGPR regression workflow to estimate yield at various growth stages. Table III shows that the interferometric coherence from Sentinel-1A SAR data generally has strong yield predictive capacity at both booting and heading stages (VIMP = 34%, at the booting stage, and VIMP = 44%, at heading stage). RDVI1 outperformed other selected optical and SAR vegetation descriptors, especially at the heading stage with

TABLE IV
YIELD ESTIMATES DERIVED USING RDVII & INTERFEROMETRIC COHERENCE_VH OF DIFFERENT GROWTH STAGES USING GAUSSIAN KERNEL, PGPR, AND BLI MODELS

Vegetation indices	Growth stage	BLI		PGPR		GPKR	
		r^2	RMSE	r^2	RMSE	r^2	RMSE
EVI	Jointing	0.31	0.89	0.35	0.90	0.51	0.74
NDWI	Jointing	0.41	0.64	0.48	0.53	0.47	0.65
RDVII	Booting	0.46	0.61	0.59	0.61	0.65	0.61
γ_{VH}	Booting	0.44	0.89	0.42	0.59	0.41	0.78
RDVI+ γ_{VH}	Booting	0.56	0.67	0.59	0.72	0.64	0.68
RDVII* γ_{VH}	Booting	0.51	0.70	0.53	0.83	0.57	0.72
RDVII	Heading	0.48	0.78	0.52	0.81	0.58	0.67
γ_{VH}	Heading	0.47	0.89	0.51	0.91	0.52	0.79
RDVII+ γ_{VH}	Heading	0.57	0.76	0.77	0.78	0.81	0.55
RDVII* γ_{VH}	Heading	0.54	0.83	0.63	0.77	0.71	0.71

VIMP = 45% (see Table III). The RDVII and interferometric coherence at VH polarization, which was sampled at the heading stage and independently used as input for PGPR, resulted in higher r^2 and lower RMSE. The statistical results derived using the RDVII ($r^2 = 0.52$ and $\text{RMSE} = 0.81$ t/ha) were higher than from interferometric coherence ($r^2 = 0.52$ and $\text{RMSE} = 0.91$ t/ha) at the heading stage. A combination of coherence and RDVII used as input for PGPR improved the estimation performance than those independently used metrics (see Table IV). Among the combination predictive metrics, the highest and the lowest RMSE was achieved using a sum of RDVII and interferometric coherence ($r^2 = 0.77$ and $\text{RMSE} = 0.78$ t/ha) than the multiplication of interferometric coherence and RDVII, with $r^2 = 0.63$ and $\text{RMSE} = 0.77$ t/ha, both at the heading stage.

The highest prediction accuracy was achieved using the Gaussian kernel regression algorithm than other used regression models in this study. EVI, NDWI, RDVII, and interferometric coherence were used as input for Gaussian kernel regression at different growth stages to predict yield. The highest r^2 and lowest RMSE were achieved using RDVII, $r^2 = 0.65$, and $\text{RMSE} = 0.61$ t/ha followed by coherence, with $r^2 = 0.52$ and $\text{RMSE} = 0.79$ t/ha. The combined parameters from coherence and RDVII resulted in the overall best prediction accuracy for all growth stages and regression algorithms. Gaussian kernel regression with a sum of Coherence at VH polarization and RDVII parameters resulted in the highest prediction accuracy with $r^2 = 0.81$ and $\text{RMSE} = 0.55$ t/ha followed by multiplication of Coherence and RDVII parameter with $r^2 = 0.71$ and $\text{RMSE} = 0.71$ t/ha (see Fig. 8). The predictive capacity of the used parameters (VIs) and GPR model were poor at early growth stages (jointing stage) than those at later growth stages (booting and heading stages). Fig. 9 shows the map of rice grain yield estimated with the best performance parameter, which is a sum of interferometric coherence and RDVII with Gaussian kernel regression.

V. DISCUSSION

A. Growth Stage-Specific Predictive Parameters for Grain Yield Estimation

Different growth stage information must be considered for more accurate yield predictions, as crop growth parameters vary at different phenological stages and have different contributions for yield prediction [13], [44]. We have done three field visits at different rice growth stages, and all sampling points are assumed to have homogenous phenological stages as field management activities, and environmental conditions are assumed to be similar. Due to their sensitivity to important crop growth indicators like crop phenology, AGB, and LAI, optical indices have been extensively used for the prediction of yield. Optical indices suffer a saturation problem from medium to high LAI, and biomass conditions and being a big limitation for using these data types [21], [45]. SAR data are considered to complement optical data types, especially for saturation issues, which are prevalent in using optical data types at later growth stages. However, backscatter coefficients derived from SAR imagery also suffer saturation problems from medium to high LAI and biomass conditions [46]. The interferometric coherence, which represents a cross-correlation product calculated from coregistered complex SAR images, has an opposite trend with optical indices [47]. The formulation of yield in plant growth greatly depends on the synchronicity of PWC and accumulated biomass at each specific growth stage [17], [48]. In this regard, the interferometric coherence, which follows the trends of the PWC and selected optical index (RDVII), which follows the trends accumulated biomass over the growth stage can be combined for the prediction of yield. Transformed VIs can reduce part of noise caused by different perturbing factors, such as atmospheric conditions, topographic illuminations, sensor calibration, and soil background. VIs are taken as a means to remove or reduce these perturbing effects from the raw spectral bands. VIs have often been preferred over the usage of spectral bands in that they

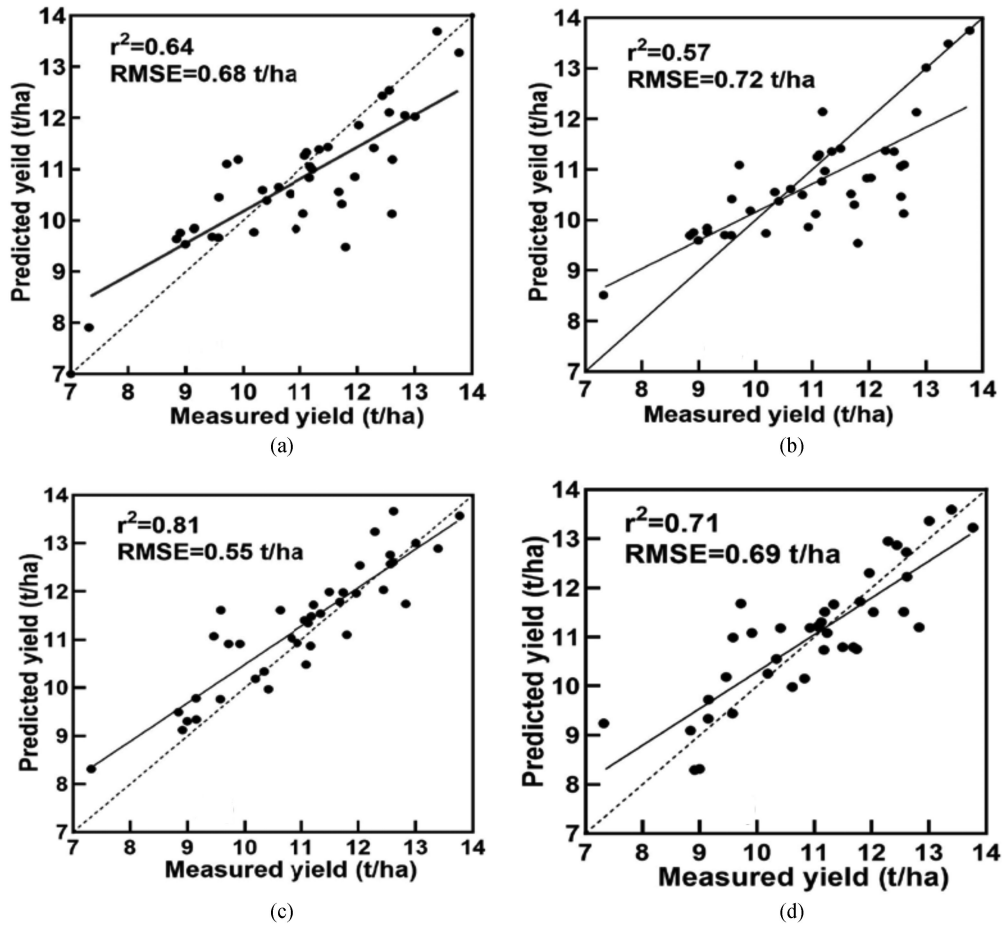


Fig. 8. Comparison of measured and predicted rice crop yield using Gaussian kernel regression (a) sum of interferometric coherence and RDVII, (b) product of interferometric coherence with RDVII, both (a) and (b) at booting stage, (c) sum of interferometric coherence and RDVII, (d) product of interferometric coherence with RDVII, both (c) and (d) at heading stage.

simplify data analysis into a single metric while at the same time they are normalizations that help to reduce data errors due to poor viewing geometry or hazy atmosphere [49]. The normalization also allows for easier comparison across different sensors. This study focused on the joint exploitation of VIs and interferometric coherence for rice yield estimation.

Optical indices and interferometric coherence have different responses to plant growth parameters at different growth stages resulting in different variable importance scores to predict yield. The calculated VIMP from the piecewise linear regression tree reveals that NDWI (26%) and EVI (24%) were the highest VIMP at the jointing stage than other optical indices and the interferometric coherence. The better results of NDWI and EVI at the jointing stage can be explained by the high sensitivity of EVI and NDWI to high biomass conditions. Due to the favorable growth conditions in the study area, most of the observed LAI were more than 3, and due to this, NDVI and other selected optical indices saturated early, resulting in less VIMP score for the prediction of yield. The EVI is an optimized index designed to enhance the vegetation signal in high-biomass regions with enhanced sensitivity by decoupling the canopy background signal and reducing the effect of atmospheric and soil background noise.

On the other hand, NDWI lost its sensitivity at later growth stages for LAI more than 4.5 [45]. At the booting and heading stage, relatively higher VIMP scores were observed using RDVII than other indices derived using visible and NIR bands. The red-edge bands are affected by other leaf and canopy parameters (e.g., leaf chlorophyll content and leaf angle distributions) across growth stages than the visible and NIR bands [45]. Using the interferometric coherence, we observed high sensitivity at the booting and heading stage. The interferometric coherence at VH polarization resulted in a higher VIMP score than the interferometric coherence at VV polarization. However, the correlation of interferometric coherence with crop growth parameters and yield needs further investigation.

B. Comparison of Predictor Variables

Among the identified growth stages highest VIMP score was observed at the heading stage using RDVII and interferometric coherence at VH polarization (see Table III). This result conforms with previous findings, which reveal the high sensitivity of optical indices for yield parameters at later growth stages though saturation still exists at this stage [50]. This study hypothesized that the interferometric coherence related to PWC and RDVII,

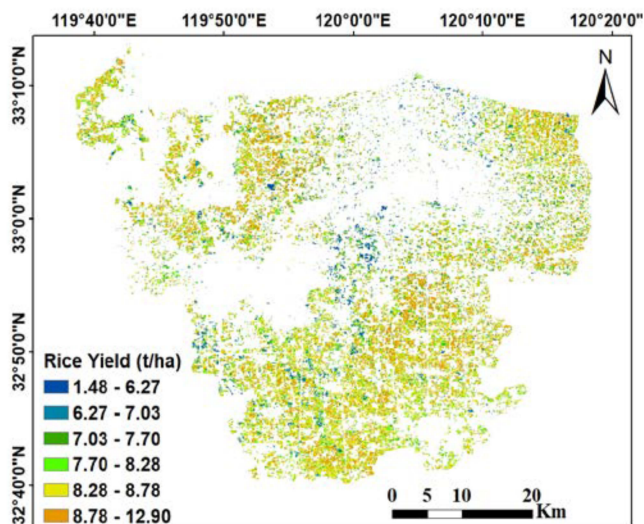


Fig. 9. Map of rice grain yield estimated by the sum of interferometric coherence and RDVI using Gaussian kernel regression.

which can explain the AGB at each specific growth stage, can be related to yield. The interferometric coherence, which is taken as an indicator for the changes to AGB water uptake dynamics, predicted yield with $r^2 = 0.52$ and $RMSE = 0.79$ t/ha using Gaussian kernel regression. RDVI1, which is more sensitive to AGB change at each specific growth stage, predicted yield with $r^2 = 0.65$ and $RMSE = 0.61$ t/ha using PGPR. The good prediction accuracy, achieved from the interferometric coherence, can be explained by the sensitivity of coherence to crop phenology, soil moisture, and irrigation levels in the fields [11].

The optical indices suffer from a saturation problem, especially after the vegetative growth stage [34], [50]. The synergetic metric from the two parameters has complementary information and can be effectively used to predict yield. Using data across optical and microwave spectral ranges can enhance crop yield estimation capability beyond what can be achieved using an individual sensor [17]. Among the used metrics, the proposed sum of RDVI1 and Coherence at VH polarization outperforms with $r^2 = 0.81$ and $RMSE = 0.56$ t/ha and is followed by the product of interferometric coherence at VH polarization and RDVI1 with $r^2 = 0.71$ and $RMSE = 0.71$ t/ha. We contend that this accuracy level is parallel or better than the previous major findings that relied upon deep learning and machine-learning techniques [34], [50]. The best statistical results from combined metrics suggest that interferometric coherence and RDVI1 metrics contain complementary information worth exploiting jointly. The combination parameter sum of RDVI1 and interferometric coherence resulted in higher statistical results than other used parameters for all used regression models (see Table IV). The disparity between Sentinel-2 derived indices and the interferometric coherence, which showed the opposite relationship at the peak growth stage (see Fig. 5), can be explained by the reason for providing complementary information, which is worth exploiting. The interferometric coherence is a measure of cross-correlation between master and slave SAR images 0

(low) to 1 (high) and is sensitive to PWC, crop structure, crop phenology, and soil moisture content [47]. Having information on the AGB (RDVI1), water content change, and soil moisture, interferometric coherence can facilitate obtaining a useful model for crop yield estimation.

C. Advantage of Kernel and Probabilistic Gaussian Regression

In this study, we tested the PGPR model, which is developed from standard GPR based on probabilistic theory and MCMC random sampling techniques [54]. Our proposed PGPR has a conditional property that can answer the question “for the given one value, what do we expect for unobserved values under the multivariate normal.” The prediction in PGPR is improved even with a limited amount of ground truth data as PGPR uses maximum posterior distribution (MAP) for the prediction of a new point on which the posterior $pr(\theta|y)$ form Gaussian distribution and the observed $y_{new}|\theta$ as a likelihood. The PGPR used in this study is based on a Bayesian formulation from standard GPR using MCMC for generating random samples of parameter values from the posterior distributions [55], [56]. Our results suggest that PGPR based on MCMC sampling to quantify the hyperparameters’ uncertainty resulted in satisfactory prediction accuracy greater than or equal to the accuracy of other related studies that used standard GPR based on sampling from covariance functions. Bayesian MCMC is drawn as a solution to quantify the uncertainty of the hyperparameters to be nonnormal multimodal than the GPR based on covariance functions. The uncertainty, limited sensitivity, and a limited number of observations used for training and validation impact the prediction accuracy of nonprobabilistic and nonparametric regression models used so far. In the training phase, GPR can maximize the likelihood with respect to the given target data [57]. The estimated likelihood helps to solve the uncertainty and limited sensitivity of the training inputs, which allows the GPR methods to exploit the priors in the prediction regression (see Fig. 10) [54], [58]. The role of priors in PGPR is to limit the solution space and act as an extra constraint on the inference, limiting the solution space. Yield estimates using different VIs and proposed algorithms agreed with ground validation samples. The used VIs and combination of indices ($RDVI1 + Coherence_{VH}$) provided prediction uncertainty for yield estimation. These uncertainties are useful to draw a conclusion about the quality of retrievals and used inputs. It is worth noting that the used training input is the only factor that influences the prediction uncertainties of the used models. If the used training input is similar or judged to be similar, the resulting prediction uncertainties are reduced because we are working with similar input features. As a result, prediction uncertainty must be viewed as a qualitative variable connected with used training input [59]. The advantage of using MCMC with probabilistic Gaussian regression proposed in this study is to derive the uncertainty of the posterior distributions with random samples so that optimal hyperparameter values can be derived (see Fig. 10).

Most prominent studies for crop yield estimation from RS data using Gaussian-based regression algorithms have reported

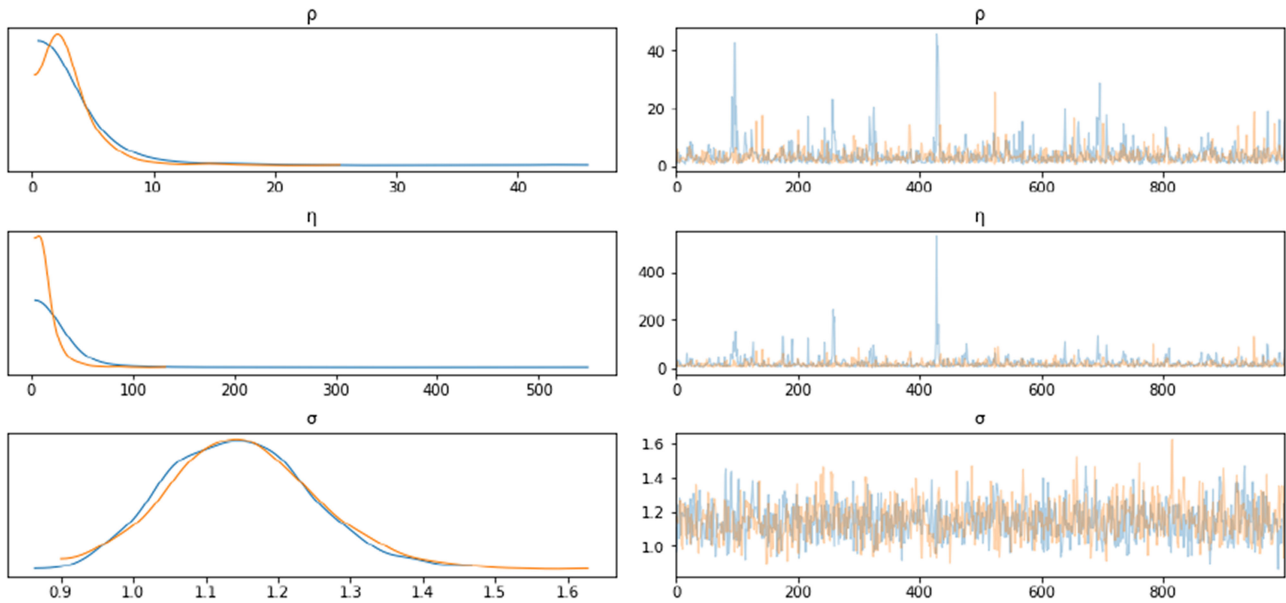


Fig. 10. Posterior distribution of the estimated hyperparameters (left), and posterior uncertainties for the estimated parameters values with MCMC sampling (right) for probabilistic GPR.

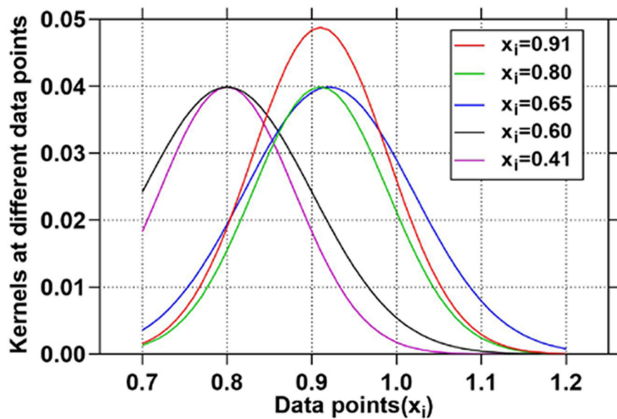


Fig. 11. Estimated kernel values from Gaussian kernel regression at different data points.

satisfactory accuracies [27], [28]. However, most of the existing studies relied on large geographic areas with a high amount of inputs for training and validations. Nonlinear kernel-based regression has been used for crop yield estimation from RS data from yield samples nationwide [17]. In this study, we proposed a method for crop yield estimation based on kernels and their weights with a limited amount of ground truth data. We tested the efficiency of Gaussian kernel regression with ground truth validation data collected in 2019 and 2020. The proposed Gaussian kernel regression has better performance than other kernel-based regression methods used so far. Kernel values are used to derive weights to predict outputs from given input based on (13) and (14) as indicated in Fig. 11. The derived kernels and weights are used for predicting each data point using other observed data values, allowing the Gaussian kernel regression model to have high interpretability for the understanding of

data characteristics [51], [52]. The capability of the Gaussian kernel regression model to capture spatiotemporal patterns from spatiotemporal data through the application of Gaussian kernel function with kernel curves, which are built for each data points (x_i) as a mean and the bandwidth (h) as a standard deviation, which is as in the case of Gaussian distributions [26], [53].

VI. CONCLUSION

Gaussian kernel regression and PGPR derived from Bayesian formulation of standard GPR were used to estimate yield from Sentinel-1A interferometric coherence and Sentinel-2A VIs. This study investigated the capacity of interferometric coherence to complement the information from optical data. From the piecewise linear estimation of variable importance, the best score was observed at the heading stage (RDVII 45% and interferometric coherence_VH 44%) and used as input for the regression models. Our results demonstrated the capability of Gaussian kernel regression for crop yield prediction with a limited amount of ground truth data $r^2 = 0.81$ and RMSE = 0.51 t/ha. We found that MCMC-based random sampling for the estimation of best hyperparameters improved the estimation performance of PGPR. Optical indices and SAR backscatter coefficients had a limitation of saturation after the booting stage. In contrast, the interferometric coherence can be used at the later growth stages. It can be concluded that the Gaussian kernel regression with a sum of RDVII and interferometric coherence metrics can be of particular interest for an operational crop yield prediction.

REFERENCES

- [1] J. Rouse, R. Haas, J. Schell, and D. Deering, "Monitoring vegetation systems in the great plains with ERTS," *NASA Special Publication*, vol. 351, pp. 309–317, 1974.

- [2] H. Q. Liu and A. Huete, "A feedback based modification of the NDVI to minimize canopy background and atmospheric noise," *IEEE Trans. Geosci. Remote Sens.*, vol. 33, no. 2, pp. 457–465, Mar. 1995.
- [3] Y. Zhang, B. Yang, X. Liu, and C. Wang, "Estimation of rice grain yield from dual-polarization Radarsat-2 SAR data by integrating a rice canopy scattering model and a genetic algorithm," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 57, pp. 75–85, 2017.
- [4] A. Gitelson and M. N. Merzlyak, "Spectral reflectance changes associated with autumn senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. leaves. Spectral features and relation to chlorophyll estimation," *J. Plant Physiol.*, vol. 143, no. 3, pp. 286–292, 1994.
- [5] D. A. Sims and J. A. Gamon, "Relationships between leaf pigment content and spectral reflectance across a wide range of species, leaf structures and developmental stages," *Remote Sens. Environ.*, vol. 81, no. 2/3, pp. 337–354, 2002.
- [6] F. FAOSTAT, "Cited 28 February 2018," 2018. [Online]. Available: <http://www.fao.org/faostat/en/#data/QC>
- [7] E. Barnes *et al.*, "Coincident detection of crop water stress, nitrogen status and canopy density using ground based multispectral data," in *Proc. 5th Int. Conf. Precis. Agriculture*, vol. 1619, 2000.
- [8] W. G. Bastiaanssen and S. Ali, "A new crop yield forecasting model based on satellite measurements applied across the Indus Basin, Pakistan," *Agriculture, Ecosystems, Environ.*, vol. 94, no. 3, pp. 321–340, 2003.
- [9] M. Parry, C. Rosenzweig, and M. Livermore, "Climate change, global food supply and risk of hunger," *Philos. Trans. Roy. Soc. B: Biol. Sci.*, vol. 360, no. 1463, pp. 2125–2138, 2005.
- [10] B.-C. Gao, "NDWI—A normalized difference water index for remote sensing of vegetation liquid water from space," *Remote Sens. Environ.*, vol. 58, no. 3, pp. 257–266, 1996.
- [11] Z. Bai *et al.*, "Could vegetation index be derive from synthetic aperture radar? – The linear relationship between interferometric coherence and nDVI," *Sci. Rep.*, vol. 10, no. 1, pp. 1–9, 2020.
- [12] P. C. Doraiswamy, S. Moulin, P. W. Cook, and A. Stern, "Crop yield assessment from remote sensing," *Photogrammetric Eng. Remote Sens.*, vol. 69, no. 6, pp. 665–674, 2003.
- [13] M. Mkhabela, P. Bullock, S. Raj, S. Wang, and Y. Yang, "Crop yield forecasting on the Canadian Prairies using MODIS NDVI data," *Agricultural Forest Meteorol.*, vol. 151, no. 3, pp. 385–393, 2011.
- [14] X. Jin *et al.*, "A review of data assimilation of remote sensing and crop models," *Eur. J. Agronomy*, vol. 92, pp. 141–152, 2018.
- [15] A. Veloso *et al.*, "Understanding the temporal behavior of crops using Sentinel-1 and Sentinel-2-like data for agricultural applications," *Remote Sens. Environ.*, vol. 199, pp. 415–426, 2017.
- [16] S. C. Steele-Dunne, H. McNairn, A. Monsivais-Huerta, J. Judge, P.-W. Liu, and K. Papatthassiou, "Radar remote sensing of agricultural canopies: A review," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 5, pp. 2249–2273, May 2017.
- [17] A. Mateo-Sanchis, M. Piles, J. Muñoz-Marí, J. E. Adsuara, A. Pérez-Suay, and G. Camps-Valls, "Synergistic integration of optical and microwave satellite data for crop yield estimation," *Remote Sens. Environ.*, vol. 234, 2019, Art. no. 111460.
- [18] J. Aschbacher and M. P. Milagro-Pérez, "The European Earth monitoring (GMES) programme: Status and perspectives," *Remote Sens. Environ.*, vol. 120, pp. 3–8, 2012.
- [19] P. Bose, N. K. Kasabov, L. Bruzzone, and R. N. Hartono, "Spiking neural networks for crop yield estimation based on spatiotemporal analysis of image time series," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 11, pp. 6563–6573, May 2016.
- [20] A. Stepanov, K. Dubrovin, A. Sorokin, and T. Aseeva, "Predicting Soybean yield at the regional scale using remote sensing and climatic data," *Remote Sens.*, vol. 12, no. 12, pp. 1936–1959, 2020.
- [21] X. Jin *et al.*, "Combined multi-temporal optical and radar parameters for estimating LAI and biomass in winter wheat using HJ and RADARSAT-2 data," *Remote Sens.*, vol. 7, no. 10, pp. 13251–13272, 2015.
- [22] J. Wang *et al.*, "Field-scale rice yield estimation using Sentinel-1A synthetic aperture radar (SAR) data in coastal saline region of Jiangsu Province, China," *Remote Sens.*, vol. 11, no. 19, pp. 2274–2283, 2019.
- [23] Y. Alebele *et al.*, "Estimation of canopy biomass components in paddy rice from combined optical and SAR data using multi-target Gaussian regressor stacking," *Remote Sens.*, vol. 12, no. 16, pp. 2564–2586, 2020.
- [24] N. Ouadi *et al.*, "Monitoring of wheat crops using the backscattering coefficient and the interferometric coherence derived from Sentinel-1 in semi-arid areas," *Remote Sens. Environ.*, vol. 251, pp. 112050–112070, 2020.
- [25] H. Aghighi, M. Azadbakht, D. Ashourloo, H. S. Shahrabi, and S. Radiom, "Machine learning regression techniques for the silage maize yield prediction using time-series images of landsat 8 OLI," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 12, pp. 4563–4577, Dec. 2018.
- [26] J. E. Adsuara, A. Pérez-Suay, J. Muñoz-Marí, A. Mateo-Sanchis, M. Piles, and G. Camps-Valls, "Nonlinear distribution regression for remote sensing applications," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 12, pp. 10025–10035, Dec. 2019.
- [27] E. Kamir, F. Waldner, and Z. Hochman, "Estimating wheat yields in Australia using climate records, satellite image time series and machine learning methods," *ISPRS J. Photogrammetry Remote Sens.*, vol. 160, pp. 124–135, 2020.
- [28] L. Martínez-Ferrer, M. Piles, and G. Camps-Valls, "Crop yield estimation and interpretability with Gaussian processes," *IEEE Geosci. Remote Sens. Lett.*, pp. 1–5, 2020, doi: [10.1109/LGRS.2020.3016140](https://doi.org/10.1109/LGRS.2020.3016140).
- [29] J. Huang *et al.*, "Assimilation of remote sensing into crop growth models: Current status and perspectives," *Agricultural Forest Meteorol.*, vol. 276, 2019, Art. no. 107609.
- [30] W. R. Gilks, "Markov chain Monte Carlo," in *Encyclopedia of Biostatistics*, vol. 4. Hoboken, NJ, USA: Wiley, 2005.
- [31] X. Xu *et al.*, "Evaluation of one-class support vector classification for mapping the paddy rice planting area in Jiangsu Province of China from Landsat 8 OLI imagery," *Remote Sens.*, vol. 10, no. 4, pp. 546–569, 2018.
- [32] Q. Xie *et al.*, "Retrieval of crop biophysical parameters from sentinel-2 remote sensing imagery," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 80, pp. 187–195, 2019.
- [33] A. Kross, H. McNairn, D. Lapen, M. Sunohara, and C. Champagne, "Assessment of Rapideye vegetation indices for estimation of leaf area index and biomass in corn and soybean crops," *Int. J. Appl. Earth Observ. Geoinformation*, vol. 34, pp. 235–248, 2015.
- [34] Q. Yang, L. Shi, J. Han, Y. Zha, and P. Zhu, "Deep convolutional neural networks for rice grain yield estimation at the ripening stage using UAV-based remotely sensed images," *Field Crops Res.*, vol. 235, pp. 142–153, 2019.
- [35] C. Leprieur, M. M. Verstraete, and B. Pinty, "Evaluation of the performance of various vegetation indices to retrieve vegetation cover from AVHRR data," *Remote Sens. Rev.*, vol. 10, no. 4, pp. 265–284, 1994.
- [36] G. Rondeaux, M. Steven, and F. Bare, "Optimization of soil-adjusted vegetation indices," *Remote Sens. Environ.*, vol. 55, no. 2, pp. 95–107, 1996.
- [37] H. Ishwaran, "Variable importance in binary regression trees and forests," *Electron. J. Statist.*, vol. 1, pp. 519–537, 2007.
- [38] J. Verrelst, J. P. Rivera, J. Moreno, and G. Camps-Valls, "Gaussian processes uncertainty estimates in experimental Sentinel-2 LAI and leaf chlorophyll content retrieval," *ISPRS J. Photogrammetry Remote Sens.*, vol. 86, pp. 157–167, 2013.
- [39] L. Pipia, J. Muñoz-Marí, E. Amin, S. Belda, G. Camps-Valls, and J. Verrelst, "Fusing optical and SAR time series for LAI gap filling with multioutput Gaussian processes," *Remote Sens. Environ.*, vol. 235, 2019, Art. no. 111452.
- [40] G. Camps-Valls *et al.*, "Biophysical parameter estimation with adaptive Gaussian processes," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2009, vol. 4, pp. IV-69–IV-72.
- [41] R. Tibshirani and L. Wasserman, "Sparsity, the Lasso, and friends," in *Statistical Machine Learning (Lecture Notes)*. Pittsburgh, PA, USA: Carnegie Mellon Univ., 2017, pp. 1–25.
- [42] D. J. Henderson and A. C. Souto "An introduction to nonparametric regression for labor economists," *J. Labor Res.*, vol. 39, no. 4, pp. 355–382, 2018.
- [43] H. Shimazaki and S. Shinomoto, "Kernel bandwidth optimization in spike rate estimation," *J. Comput. Neurosci.*, vol. 29, no. 1, pp. 171–182, 2010.
- [44] D. K. Bolton and M. A. Friedl, "Forecasting crop yield using remotely sensed vegetation indices and crop phenology metrics," *Agricultural Forest Meteorol.*, vol. 173, pp. 74–84, 2013.
- [45] T. Dong *et al.*, "Estimating crop biomass using leaf area index derived from Landsat 8 and Sentinel-2 data," *ISPRS J. Photogrammetry Remote Sens.*, vol. 168, pp. 236–250, 2020.
- [46] Y. Inoue, E. Sakaiya, and C. Wang, "Capability of C-band backscattering coefficients from high-resolution satellite SAR sensors to assess biophysical variables in paddy rice," *Remote Sens. Environ.*, vol. 140, pp. 257–266, 2014.
- [47] D. Palmisano *et al.*, "Sensitivity of Sentinel-1 interferometric coherence to crop structure and soil moisture," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2019, pp. 6219–6222.

- [48] F. Tian *et al.*, “Coupling of ecosystem-scale plant water storage and leaf phenology observed by satellite,” *Nature Ecol. Evol.*, vol. 2, no. 9, pp. 1428–1435, 2018.
- [49] D. M. Johnson, “An assessment of pre-and within-season remotely sensed variables for forecasting corn and soybean yields in the United States,” *Remote Sens. Environ.*, vol. 141, pp. 116–128, 2014.
- [50] R. Song, T. Cheng, X. Yao, Y. Tian, Y. Zhu, and W. Cao, “Evaluation of Landsat 8 time series image stacks for predicting yield and yield components of winter wheat,” in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2016, pp. 6300–6303.
- [51] R. Sambasivan, S. Das, and S. K. Sahu, “A Bayesian perspective of statistical machine learning for Big Data,” *Comput. Statist.*, vol. 35, no. 3, pp. 893–930, 2020.
- [52] H. Shimazaki and S. Shinomoto, “Kernel bandwidth optimization in spike rate estimation,” *J. Comput. Neurosci.*, vol. 29, no. 1, pp. 171–182, 2010.
- [53] D. Li, P. C. B. Phillips, and J. Gao, “Uniform consistency of nonstationary kernel-weighted sample covariances for nonparametric regression,” *Econometric Theory*, vol. 32, no. 3, pp. 655–685, 2013.
- [54] F. Lindsten, T. B. Schön, A. Svensson, and N. Wahlström, “Probabilistic modeling – Linear regression & Gaussian processes,” Uppsala Univ., Uppsala, Sweden, vol. 7, 2017.
- [55] B. Dumont, V. Leemans, M. Mansouri, B. Bodson, J.-P. Destain, and M.-F. Destain, “Parameter identification of the STICS crop model, using an accelerated formal MCMC approach,” *Environ. Model. Softw.*, vol. 52, pp. 121–135, 2014.
- [56] D. Makowski, D. Wallach, and M. Tremblay, “Using a Bayesian approach to parameter estimation; comparison of the GLUE and MCMC methods,” *Agronomie*, vol. 22, no. 2, pp. 191–203, 2002.
- [57] D. H. Svendsen, L. Martino, M. Campos-Taberner, F. J. García-Haro, G. Camps-Valls, and R. Sensing, “Joint Gaussian processes for biophysical parameter retrieval,” *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1718–1727, Mar. 2018.
- [58] T. Iizumi, M. Yokozawa, and M. Nishimori, “Parameter estimation and uncertainty analysis of a large-scale crop model for paddy rice: Application of a Bayesian approach,” *Agricultural Forest Meteorol.*, vol. 149, no. 2, pp. 333–348, 2009.
- [59] M. Campos-Taberner *et al.*, “Multitemporal and multiresolution leaf area index retrieval for operational local rice crop monitoring,” *Remote Sens. Environ.*, vol. 187, pp. 102–118, 2016.



Yeshanbele Alebele (Graduate Student Member, IEEE) received the B.S. degree in land administration and surveying from Bahir Dar University, Bahir Dar, Ethiopia, in 2013, and the M.S. degree in geodesy and geomatics engineering from Adama Science and Technology University, Adama, Ethiopia, in 2015. Since September 2017, he has been working toward the Ph.D. degree in agricultural informatics with the National Engineering & Technology Center for Information Agriculture (NETCIA), Nanjing Agricultural University, Nanjing, China.

His research interests include microwave remote sensing, and the development of multisensor techniques for enhanced retrievals of vegetation biophysical parameters.



Wenhui Wang was born in Langfang, China, in 1989. He received the M.S. degree in soil and water conservation and desertification prevention from Fujian Agriculture and Forestry University, Fujian, China, in 2017. Since September 2017, he has been working toward the Ph.D. degree in agricultural informatics with the National Engineering and Technology Center for Information Agriculture (NETCIA), Nanjing Agricultural University, Nanjing, China.

His research interests include crop nitrogen monitoring with UAV and satellite imagery.



Weiguo Yu received the B.S. degree in surveying and mapping engineering and the M.S. degree in geodesy and surveying engineering from Anhui University of Science & Technology, Anhui, China, in 2012 and 2016, respectively. Since July 2019, he has been working toward the Ph.D. degree in agricultural informatics with the National Engineering & Technology Center for Information Agriculture (NETCIA), Nanjing Agricultural University, Nanjing, China.

As a joint graduate student, from 2014 to 2016, he was with the Institute of Eco-environment and Agrometeorology, Chinese Academy of Meteorological Sciences. He is the author and coauthor of 11 papers and one software copyright. His current research interests include remote sensing big data and rice productivity prediction.



Xia Yao (Senior Member, IEEE) received the B.S. degree in agronomy, the M.S. degree in agricultural ecology, and the Ph.D. degree in crop informatics from Nanjing Agricultural University, Nanjing, China, in 2001, 2004, and 2009, respectively.

Since December 2015, she has been a Professor with the National Engineering & Technology Center for Information Agriculture (NETCIA) and College of Agriculture, Nanjing Agricultural University. From December 2016 to December 2017, she was a Visiting Scholar with the Department of Geography, University of Hawaii, Honolulu, HI, USA. She is the author and coauthor of more than 60 papers, three monographs, and 12 patents. She supervised 12 M.S. and Ph.D. students, and investigated more than ten projects as a PI. Her research activities are built on the sophisticated multiscale platforms for acquiring timely remotely sensed data over crop fields. Her research interests include crop hyperspectral remote sensing, crop growth/biotic/abiotic stress/senescence monitoring based on SIF, active remote sensing (LiDAR), unmanned aerial vehicles (UAVs), and high throughput crop phenotyping.

Dr. Yao is a recipient of four national- and province-level awards. She is also a recipient of “the Young and Middle-aged Academic Leader in Blue and Indigo Engineering” in Jiangsu province in 2016. Also, she is the reviewer of *Remote Sensing of Environment*, *Field Crop Research*, *ISPRS Journal of Photogrammetry and Remote Sensing*, *International Journal of Applied Earth Observation and Geoinformation*, etc. She has been serving on the Editorial Board of *Remote Sensing* since 2018. She was appointed as the Dean of Smart Agriculture Department in December 2019.



Yongchao Tian received the B.S. degree in agronomy, the M.S. degree in crop cultivation and farming system, and the Ph.D. degree in crop informatics from Nanjing Agricultural University (NAU), Nanjing, China, in 2000, 2003, and 2008, respectively.

Since December 2014, he has been a Professor with the National Engineering & Technology Center for Information Agriculture (NETCIA) and College of Agriculture, NAU. He is currently the Executive Deputy Director of the Institute of Smart Agriculture, NAU. He is the author of more than 90 papers in journals indexed by the Web of Science, and holds 26 patents. He is the Principal Investigator of several national-level projects funded by the Ministry of Science and Technology and the National Natural Science Foundation of China. He works mainly on crop growth monitoring and precision farming management.

Dr. Tian is a recipient of two National Second-Class Awards for Science and Technology Advancement, and won the “Thousand Talents Program” in Jiangxi province in 2018, “the Young and Middle-aged Academic Leader in Qingnan Project” in Jiangsu province in 2014.



Yan Zhu received the Ph.D. degree in crop cultivation and farming from Nanjing Agricultural University, Nanjing, China, in 2003.

From 2012 to 2013, she was a visiting scientist with the Department of Agricultural and Biological Engineering, University of Florida, Gainesville, FL, USA. She is currently a Professor of Information Agronomy and the Dean of the College of Agriculture with Nanjing Agricultural University. She is also heading the National Engineering & Technology Center for Information Agriculture (NETCIA), MOE

Engineering Research Center of Smart Agriculture, MARA Key Laboratory of Crop System Analysis and Decision Making, and Jiangsu Key Laboratory for Information Agriculture. Her research has been broadly focused on smart farming theory and technology and has made significant achievements in a digital simulation of crop productivity formation process, and spectral monitoring, and quantitative diagnosis of crop growth indicators. She is the author of more than 170 papers on Web of Science indexed journals including *Nature Climate Change* and *Global Change Biology*, and three book chapters. She holds eight national invention patents and 16 computer software copyrights. Her H-index is 29.

Prof. Zhu, as the second recipient, has won the Second-class National Award for Progress in Science and Technology twice, and the First-class Ministerial or Provincial Award for Progress in Science and Technology twice. She has also won the First-class Jiangsu Provincial Agricultural Technology Extension Award as the leading recipient. She is a member of the Steering Council for the Agricultural Model Intercomparison and Improvement Project (AgMIP). She was a Guest Editor of Special Issues in the *European Journal of Agronomy* and *Remote Sensing*. She was a recipient of the Youth Science and Technology Award of China and the Young Women Scientist of China. She was also a recipient of the National Science Fund for Distinguished Young Scholars by the National Natural Science Foundation of China in 2017.



Weixing Cao received the Ph.D. degree in crop science from Oregon State University, Corvallis, OR, USA, in 1989.

From 1989 to 1994, he was a Postdoctoral Research Associate and then an Assistant Scientist with the College of Agricultural and Life Sciences, University of Wisconsin, Madison, WI, USA. Since 1994, he has been a Professor with the College of Agriculture, Nanjing Agricultural University, Nanjing. He is currently the Director of Institute of Smart Agriculture and an Honorary Director of the National Engineering

& Technology Center for Information Agriculture (NETCIA), Nanjing Agricultural University, a Co-convenor of the Discipline Evaluation Group in Crop Science under the Degree Committee of the State Council, and the Head of the National Consortium of the Professional Degree in Agricultural Engineering and Information Technology. His academic outputs include more than 200 publications on international peer-reviewed journals, eight monographs and textbooks, and 41 national invention patents. In the past 25 years, he has advised a total of 156 master's and Ph.D. students. He has accomplished outstanding achievements in the specific areas of crop system simulation and design, growth monitoring and diagnosis, and precision crop management. His main research fields include crop ecology, information agronomy, and smart agriculture.

Prof. Cao, as the leading recipient, has won the Second-class National Award for Progress in Science and Technology three times, and the First-class Ministerial or Provincial Award for Progress in Science and Technology five times. He was a recipient of the National Science Fund for Distinguished Young Scholars by the National Natural Science Foundation of China, and the Achievement Award by the Crop Science Society of China.



Tao Cheng (Senior Member, IEEE) received the B.S. degree in geographic information system from Lanzhou University, Lanzhou, China, in 2003, the M.E. degree in photogrammetry and remote sensing from Peking University, Beijing, China, in 2006, and the Ph.D. degree in earth and atmospheric sciences from the University of Alberta, Edmonton, AB, Canada, in 2010.

From January 2011 to December 2013, he was a Postdoctoral Scholar with the Department of Land, Air and Water Resources, University of California, Davis, CA, USA. Since December 2013, he has been a Professor with the National Engineering & Technology Center for Information Agriculture (NETCIA) and College of Agriculture, Nanjing Agricultural University. He is the author of more than 50 papers on international and national journals and one book chapter, and holds two patents. He is the Principal Investigator of several national-level projects funded by the National Key R&D Program and the National Natural Science Foundation of China. His main research interests include crop mapping and monitoring, crop phenotyping, imaging and nonimaging spectroscopy of vegetation, quantitative remote sensing, machine learning, and image analysis.

Dr. Cheng was a Guest Editor of special issues in *Remote Sensing* and *IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING*. He has been an Associate Editor for *CABI Agriculture and Bioscience* since 2021 and serving on the Editorial Boards of *Precision Agriculture*, *ISPRS International Journal of Geo-Information*, and *Drones*. He was the Chair of IEEE GRSS Nanjing Chapter from 2016 to 2020. He was a recipient of the Youth Science & Technology Award from the Crop Science Society of China in 2019.