

# Estimation of Rice Crop Height From X- and C-Band PolSAR by Metamodel-Based Optimization

Onur Yuzugullu, Esra Erten, *Member, IEEE*, and Irena Hajnsek, *Fellow, IEEE*

**Abstract**—Rice crops are important in global food economy and are monitored by precise agricultural methods, in which crop morphology in high spatial resolution becomes the point of interest. Synthetic aperture radar (SAR) technology is being used for such agricultural purposes. Using polarimetric SAR (PolSAR) data, plant morphology dependent electromagnetic scattering models can be used to approximate the backscattering behaviors of the crops. However, the inversion of such models for the morphology estimation is complex, ill-posed, and computationally expensive. Here, a metamodel-based probabilistic inversion algorithm is proposed to invert the morphology-based scattering model for the crop biophysical parameter mainly focusing on the crop height estimation. The accuracy of the proposed approach is tested with ground measured biophysical parameters on rice fields in two different bands (X and C) and several channel combinations. Results show that in C-band the combination of the HH and VV channels has the highest overall accuracy through the crop growth cycle. Finally, the proposed metamodel-based probabilistic biophysical parameter retrieval algorithm allows estimation of rice crop height using PolSAR data with high accuracy and low computation cost. This research provides a new perspective on the use of PolSAR data in modern precise agriculture studies.

**Index Terms**—Agriculture, optimization methods, polarimetry, polynomial approximation, synthetic aperture radar (SAR).

## I. INTRODUCTION

**R**ICE is an important crop, whose cultivation is an important building blocks of several economies in the world. In the history of the agriculture, rice goes back to 8000 BC and even today, it keeps its economical importance for the rural communities. Therefore, like many other crops, rice needs to be monitored frequently to optimize its competitiveness in the global food market. Because of these reasons, there is a wealth of knowledge about rice in the literature, including its morphology, phenology, and the impacts of environmental

factors on growth. Efficient monitoring techniques such as remote sensing exist for precise agricultural monitoring.

Remote sensing systems provide solutions over large areas. As an advantage, synthetic aperture radar (SAR) systems provide day-and-night and all-weather operation capability and unlike other remote sensing systems, they are sensitive on the physical properties of the target in the three-dimensional (3-D) space. Thus, SAR becomes the perfect candidate for precise agriculture studies in terms of [1], [2] biophysical parameter estimation.

Considering the precise agriculture concept, the estimation of crop parameters by means of SAR includes various techniques such as polarimetry (PolSAR) [3], [4], interferometry (InSAR) [5], [6], differential interferometry (DInSAR) [7], [8], polarimetric interferometry (PolInSAR) [9], [10], and tomography (TOMO-SAR) [11]. In principle these methods, except PolSAR, allow the estimation of crop parameters such as canopy height. Also, they do not require any *a priori* information. However, it is not always so trivial to acquire such SAR data. At this point by combining the knowledge of the agricultural and environmental principles with the fundamentals of the SAR polarimetry, it is possible to increase the accuracy of low-cost analysis. In this case, it is possible to combine *a priori* information such as biological growth rules and boundaries of a specific crop, with PolSAR methods which focus on figuring out the interaction of the electromagnetic waves with the canopy. In this paper, polarimetric channel and frequency-based comparisons are provided for a scattering model inversion algorithm, which estimates the observables from the measurements, to exploit the efficiency of polarimetric SAR data in estimation of the crop morphology by focusing on the stalk height.

The growth stage of agricultural crops can be coarsely determined using methods that employ single [12]–[14] or multiple [15]–[17] acquisitions. However, none of these methods are capable of explaining the crop morphology. To estimate the morphology, a deeper relation is required. In the literature, this explanation is covered by two different approaches: backward and forward. The backward approach [12], [18] shows the relation of the biophysical parameters and the backscattering coefficients under different frequencies and incidence angles. On the other hand, the forward approach uses the scattering models [19]–[22] to approximate the backscattering coefficients from a given complex plant morphology. Due to the underdetermined character of such morphology-based scattering models, analytical inversion is extremely challenging. Towards solving this issue, Zhang *et al.* [4] have applied a genetic algorithm (GA) to estimate the height and density of paddy rice from RADARSAT-2 data. Apart from the methods present in the literature, this study

Manuscript received December 18, 2015; revised March 15, 2016; accepted May 30, 2016. This work was supported by the Scientific and Technological Research Council of Turkey under Project 113Y446, by the German Aerospace Center under Project XTILAND1476, by the Canadian Space Agency under Project SOAR-EI-5151, and by the Scientific Research Projects Coordination of Istanbul Technical University under Project 37496.

O. Yuzugullu is with the Institute of Environmental Engineering, ETH Zurich, Zurich CH-8093, Switzerland (e-mail: yuzugullu@ifu.baug.ethz.ch).

E. Erten is with the Istanbul Technical University, Geomatics Engineering, Istanbul 34469, Turkey (e-mail: eserten@gmail.com).

I. Hajnsek is with the Institute of Environmental Engineering, ETH Zurich, Zurich CH-8093, Switzerland, and also with the Microwaves and Radar Institute, German Aerospace Center, Wessling 82234, Germany (e-mail: irena.hajnsek@dlr.de).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/JSTARS.2016.2575362

focuses on the multi-dimensional distribution of the possible morphologies by following a probabilistic approach using the well-known particle swarm optimization (PSO) [23].

Morphology-based scattering models consider the real physical structure of the plants and, therefore, have complex algorithms. In this complexity, the importance of the parameters should be known, which is achieved by means of global sensitivity analysis (GSA). This allows understanding of the physics behind the scattering behavior. The model proposed by [19] is a well-known example of such scattering models. For the chosen model, the GSA analysis has been done using Sobel' indices using a high-dimensional model representation (HDMR), in terms of the polynomial chaos expansion (PCE) metamodel. It showed that canopy height and the vertical structural density are the most important biophysical parameters for the model output [24]. As a forward step, in this paper, we focus on the height estimation accuracy of the complex scattering model under different frequencies and channel combinations by taking coarse growth phase information as *a priori*.

The paper is structured in four main sections. The proposed methodology as shown in Fig. 1 is explained including the details of the scattering model and the metamodel in Section II. Section III provides a brief explanation of the ground and the SAR data, and the results of the frequency and the polarimetric channel comparisons are presented in Section IV. Finally, the outcomes are summarized in Section V.

## II. METHODOLOGY

### A. Theoretical Model and the PCE Metamodel

Temporal changes in the SAR backscattering signature of a rice canopy is closely related to variations in the biophysical properties (e.g., stalk, leaf, and panicle structure) and the water content of the canopy in time. This section focuses on the HDMR of a morphology-based incoherent theoretical scattering model with the PCE metamodel.

The model used in this study to simulate the rice canopy is one of the more complex ones in the literature, which considers complete plant structure. To start with, it considers the complete plant morphology by allowing for the inclusion of the location of the scatterers and their quantitative densities. Moreover, it also takes into account the backscattering enhancement and the clustering effects of the scatterers [25]. The structural randomness of the canopy is considered by means of Monte Carlo (MC) simulations. Thus, the final backscattering coefficients for different polarimetric channels are calculated by MC simulations over a group of dielectric cylinders and elliptical disks over a dielectric half space [26].

The model simulates a unit illuminated area,  $A$ . Inside  $A$ , the plants are randomly placed and made sure that none of them are overlapping. In area  $A$ , there are  $n_s$  plants with  $n_t$  stalks with a height of  $h_s$  and a diameter of  $d_s$ . Each stalk has  $n_l$  leaves with a length of  $l_l$  and a width of  $w_l$ . Last but not the least in the list of plant structures, there are  $n_p$  panicles with a length of  $h_p$  and a diameter of  $d_p$ . In addition, complex dielectric constants ( $\epsilon_{s,l}$ ) include the effect of the moisture content. Finally, flooded ground is characterized by the complex dielectric constant  $\epsilon_g$ .

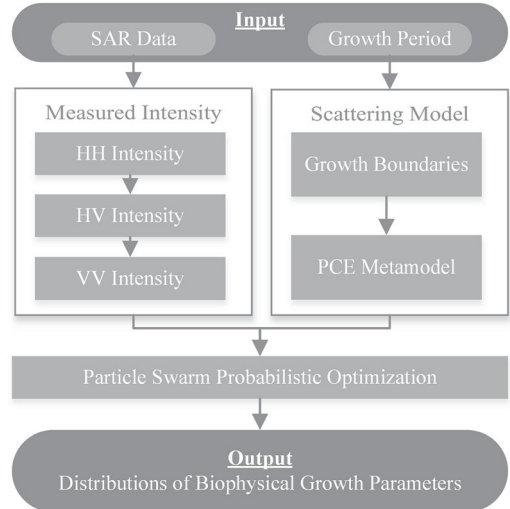


Fig. 1. Block diagram of the proposed approach.

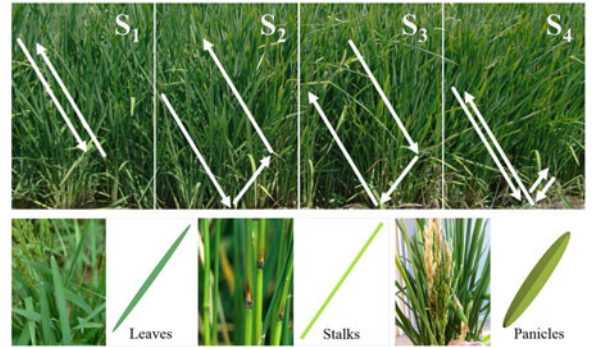


Fig. 2. Scattering mechanisms of an electromagnetic wave and the corresponding plant elements involved in the chosen backscattering model.

The corresponding theoretical backscattering algorithm considers four independent scattering mechanisms ( $S_n$ ), visualized in Fig. 2.

$S_1$ : Direct scattering from the scatterers.

$S_2$ : Scattering from the canopy followed by reflection from the ground.

$S_3$ : Reflection from the ground followed by scattering from the canopy.

$S_4$ : Reflection from the ground followed by scattering from the canopy and followed by reflection from the ground.

The sum of these four scattering mechanisms relate the incident ( $i$ ) wave  $\vec{E}^i$  to the scattered ( $s$ ) wave  $\vec{E}^s$  as follows:

$$E_q^s(\vec{r}) = \frac{e^{ikr}}{r} (S_1 + S_2 + S_3 + S_4) \vec{E}_p^i. \quad (1)$$

In (1),  $q$  and  $p$  correspond to transmitted and received polarization channels, respectively. Lastly, the backscattering coefficients for the different channels,  $qp$ , are approximated from the ratio between  $E_q^s$  and  $E_p^i$  while the wave travels through the

distance between the sensor and the target,  $r$ , as follows:

$$\sigma_{qp}^o = \frac{4\pi r^2 \langle |E_q^s|^2 \rangle}{A |E_p^i|^2}. \quad (2)$$

The chosen complex morphology-based scattering model is computationally expensive due to the included MC simulations, requiring long computation times. In this paper, this issue is handled by introducing HDMR methods. Such high-dimensional representations substitute any mathematical model to an input–output metamodel by applying a polynomial expansion. Such that, for a function  $f$  with input vector  $x$  the expansion is written as

$$f(x) = f_0 + \sum_{i=1}^n f_i(x_i) + \sum_{i \neq j} f_{ij}(x_i, x_j) + \dots + f_{12\dots n}(x_1, x_2, \dots, x_n). \quad (3)$$

The metamodels are the response surface of a surrogate model with unknown coefficients. Since they are polynomial expansions, their coefficients can be estimated by means of optimization algorithms. Fortunately, when they are compared to the original model, they provide significantly similar behavior. With their polynomial-based algorithm, they are memory efficient and computationally inexpensive. It should be mentioned that using such models as a substitute to the complex models boosts the efficiency up to  $10^4$  times. In this study, PCE, developed by [27], was implemented in MATLAB using the Uncertainty Quantification Laboratory (UQLab) toolbox [28]. In this way, it reduces the computational cost of the approach by handling the multidimensionality of the retrieval problem.

### B. Probabilistic PSO

Agricultural crops are biophysically complex structures through their growth cycle. Inverting the complex mathematical relation between the morphology and an electromagnetic wave travelling inside the canopy creates an underdetermined problem. In such problems, the number of unknowns (i.e., morphological parameters) are higher than the number of equations (i.e., polarimetric backscattering intensities). Selecting optimization algorithms is a good tool to overcome this shortcoming. Using optimization, the extrema (e.g., minima and maxima) points of the model can be determined empirically. However, use of data with high variation may reduce their accuracy by leading to a local extrema. To reduce this uncertainty, a probabilistic approach is preferred, which considers a constrained region in the parameter space and provides a set of solutions that includes the function extrema as well. In the probabilistic methods, multiple initiations have been done aiming the same output to have an idea of the distribution of the solution set.

In the literature, there are several optimization methods to handle multivariate problems. Methods like GA, artificial neural networks, and evolution strategy are some examples for flexible and intelligent algorithms. However, due to the complex logic behind, they require numerous adjustments. In this study, to keep the inversion approach simple and effective, PSO is considered. The PSO algorithm optimizes a problem by iteratively

improving the solution using a population of candidates based on the initial velocity and position of each particle. During an iteration, each particle is affected by the particle that has the best position in the search space. Thus, the particle swarm continuously moves to find the local/global extremes inside given boundaries [23].

In this study, the fitness function for the PSO algorithm is developed using the PCE metamodel of the morphology-based backscattering model. As the PCE metamodel mimics the scattering model, morphology vector  $X$ , given in (4), is taken as an input and the backscattering intensities,  $\sigma_{qp}$  is produced as output

$$X = [h_s, d_s, n_s, l_l, w_l, n_l, l_p, d_p, n_p]. \quad (4)$$

The PSO algorithm is used to minimize the output of the fitness function by varying the  $X$  vector in every  $i^{\text{th}}$  initiation. The full-polarization fitness function of the corresponding optimization problem for the  $i^{\text{th}}$  iteration is given in (5). The structure of the designed fitness function is flexible and it can be modified for different channel combinations (e.g., single-, dual-, or quad-polarization). In other words, based on the preferred polarimetric channels, unused components of the fitness function can be eliminated from the linear equation. The idea behind using multiple polarimetric channels in the same fitness function is to consider the same morphological structure and, therefore, to fix the parameter space. In (5),  $\sigma_{qp}$  and  $\bar{\sigma}_{qp}$  correspond to measured and PCE estimated backscattering intensities in several  $qp$  (HH, HV, VV) channels, respectively

$$\min \mathcal{C}_i = (\sigma_{\text{HH}} - \bar{\sigma}_{\text{HH}_i})^2 + (\sigma_{\text{HV}} - \bar{\sigma}_{\text{HV}_i})^2 + (\sigma_{\text{VV}} - \bar{\sigma}_{\text{VV}_i})^2. \quad (5)$$

Almost all of the optimization algorithms require boundary conditions, also called constraints, to reduce the complexity of the problem and to avoid the presence of misleading results. In this study, constraints are chosen based on the biophysical principles of the agricultural crops. The PSO algorithm, used in this paper, considers the following three given constraints:

- 1) *Positivity Constraint*: In the physical world with the biological growth rules, a plant dimension cannot be negative. Therefore, this constraint forces the optimizer to keep the sizes of the all morphological parameters positive under all conditions.
- 2) *Min–Max Constraint*: Based on the coarse growth phase of the crop, the biophysical parameters have specific minimum and maximum values. The ranges are given in Fig. 5. For this constraint, the growth boundaries are extended by 10% to take the extreme conditions into account.
- 3) *Morphological Constraint*: Within a crop type, each biophysical parameter evolves in accordance with the others following biological growth rules. Because of this, some crop morphologies are biophysically impossible to observe under healthy conditions. Such as, it is not possible to observe a rice plant with 2-m-height and 1 leaf with 10 cm in length. Thus, they are eliminated using the convex hull obtained from the ground-based morphology dataset. By this, biophysically possible morphologies are obtained.



Fig. 3. Location of the test area (TA) (Ipsala, Turkey). The test fields for the ground measurement are also marked in the figure. Photos at the left represent rice fields in given dates for the field marked with red color.

Such constraints bound the optimizer, thus, reduce the computational cost. Additionally, the use of morphological growth boundaries increases the accuracy of the analysis. The optimization procedure is stopped after the percent change in the elements of the morphological vector is less than 1%. In this study, the PSO algorithm has been initialized 200 times to generate representative parameter distributions. Lastly, the mode of the resulting distribution is used as the extrema point of the fitness function.

### III. DATASETS

#### A. Ipsala Test Site and Ground Data

The investigated area is near the Lake Gala National Park in the Thrace region, north west of Istanbul, Turkey. It is centered at  $37^{\circ}7'53''N$  and  $6^{\circ}19'32''E$  and has a flat topography. It is one of the major rice cultivation areas in Turkey with an area of  $20\text{ km} \times 30\text{ km}$ . In the region, single season rice cultivation is done from late April to early October. The Meriç (Evros or Maritsa) river flows close to the rice fields and provides the majority of the fresh water requirement of the irrigation districts. Fig. 3 visualizes the location of the test site and the fields with example photos from a test field.

The ground surveys were conducted synchronous to the SAR data acquisitions to measure the rice crop morphology through the full growth cycle. A total of five test fields were selected from the area of interest, which were georeferenced by a GPS before the campaign. At each test site a total of seven biophysical parameters were monitored: stalk height above water surface, stalk diameter, leaf length and width, number of plants per  $\text{m}^2$ , number of tiller per plant, and number of leaves per tiller. Additionally, the growth stages are reported in terms of the BBCH, which increases from 0 to 99 during the phenological cycle. Fig. 4 visualizes the temporal trend of the BBCH stage of the test sites.

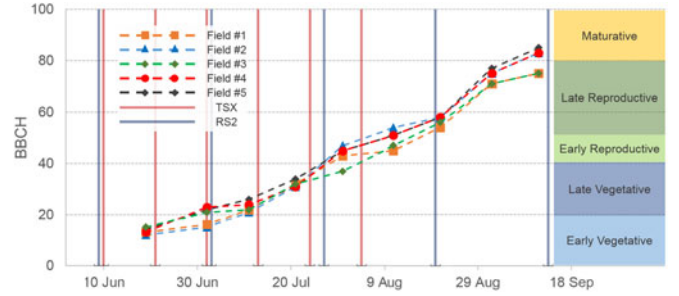


Fig. 4. Acquisition dates for SAR data in 2014, given with the temporal variation of the BBCH stages for test fields.

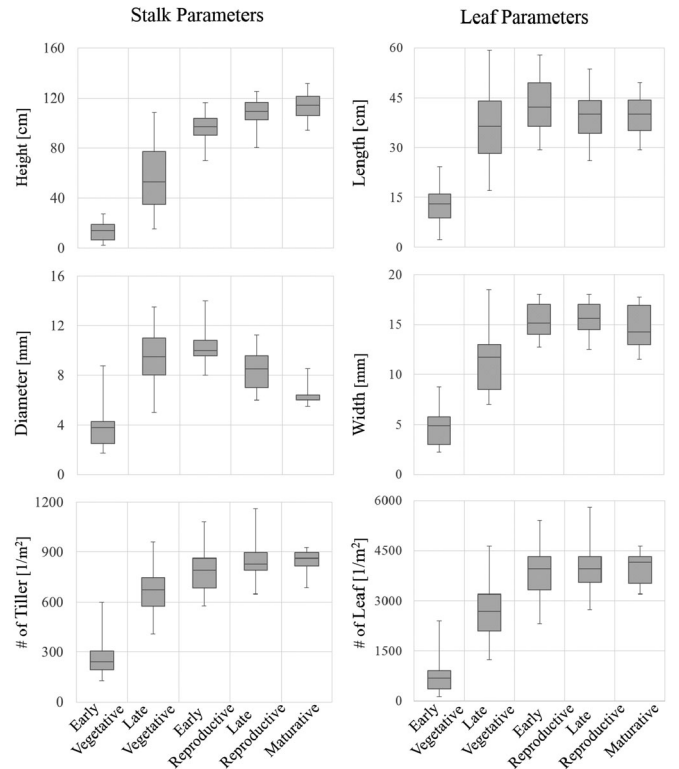


Fig. 5. Temporal variations of the rice crop biophysical parameters from 2014 Ipsala ground campaign with respect to the growth stage as a Box-and-Whisker plot. Box presents the information for the quartiles while the whiskers present minimum and maximum values.

Fig. 5 shows the temporal changes in biophysical parameters as a function of the BBCH. For each parameter, a quasi-linear increase is observed until midlate reproductive stage. Later, while most of the morphological parameters tend to stabilize, the stalk diameter starts to decrease due to reduced water content. The differences within descriptors at a specific growth stage can also be related to the differences in the surrounding environment.

#### B. SAR Data

The Ipsala test site was observed with the TerraSAR-X (TSX) and RADARSAT-2 (RS2) in 2014. The acquisition dates of the TSX and RS2 data are shown in Fig. 4. The TSX data were acquired by the German Aerospace Center (DLR) and the RS2

TABLE I  
INPUT PARAMETERS THAT ARE KEPT CONSTANT FOR THE SIMULATIONS GIVEN FOR BOTH TESTED FREQUENCIES

Parameter	X-band value	C-band value
Central frequency [GHz]	9.65	5.35
Dielectric constant, plant	25.7 + 8j [25]	
Dielectric constant, ground	74 + 21j [25]	
Average incidence angle [deg]	31	29
Distance to target [km]	514	798
Illuminated area x-size [km]	2.6	7.6
Illuminated area y-size [km]	1.8	5.2
Number of realizations	250	

by the Canadian Space Agency. The data is delivered in single look complex (SLC) format and coregistered by using bilinear interpolation.

TSX and RS2 systems have two major differences. First, TSX operates in the central frequency of 9.65 GHz, while RS2 operates in the 5.35 GHz. Due to this difference, TSX is more sensitive to small-scaled biophysical changes. Also, the penetration depth inside the canopy increases with decreasing frequency. Therefore, RS2 interacts more with the lower parts of the canopy. Second, TSX and RS2 have temporal resolution of 11 and 24 days, respectively. Such that, TSX allows more frequent monitoring than the RS2 system.

#### IV. RESULTS AND DISCUSSIONS

This paper presents two important comparisons for the estimation of biophysical parameters using a proposed probabilistic inversion approach. The first comparison considers the effect of frequency (see Section IV-A) between X- and C-band SAR data. The second comparison takes the possible polarimetric channel combinations into account (see Section IV-B). For the second comparison, quad-pol RS2 data are used. The accuracy of comparisons are not only evaluated over the correctness of the estimation, but also the convergence rate of the optimization-based algorithm.

This rate of convergence can be quantified using three basic statistical parameters, which tests the similarities between resulting set of solution and its relationship to the delta function. The first parameter, range, is the difference between the maximum and the minimum value of the distribution. Higher range means that distribution is widely spread. In terms of convergence, it means that the algorithm has a comparable lower rate of convergence. The second parameter is absolute value of the one minus the ratio between the mean and the median of the resulting distribution. As this value gets closer to zero, skewness of the distribution reduces and convergence increases. The third and the last parameter is the standard deviation of the resulting distribution, which is sign of a higher convergence rate with its lower value.

A PCE metamodel representing a theoretical electromagnetic backscattering model is used for the forward approach which simulates from the biophysical parameters to obtain backscattering coefficients. The parameters given in Table I were assumed constant during the evaluations, either due to low degree of sen-

TABLE II  
MEASURED SAR VERSUS THE ESTIMATED THEORETICAL BACKSCATTERING INTENSITIES GIVEN BY  $R^2$  AND RMSE VALUES FOR DIFFERENT CHANNELS

		$R^2$	RMSE [dB]
TSX	HH	0.871	2.02
X-band	VV	0.846	1.91
	HH	0.813	2.48
RS2	HV	0.782	2.85
C-band	VV	0.831	2.23

sitivity of the backscattering model to the parameter or because they are system parameters [29].

Based on the constant parameters shown in Table I, simulations for the available ground data are completed. The accuracy analysis of the simulations are provided in Table II. Interpreting the results, simulations for the individual band showed that copolar X-band channels has a slightly higher accuracy than C-band. Additionally, X-band HH channel and C-band VV channel provided higher  $R^2$  (coefficient of determination) and lower root mean square error (RMSE) values. The lowest accuracy was observed in C-band cross-polarization channel, HV.

The chosen backscattering metamodel is strongly affected by the changes that take place in biophysical parameters. GSA based on the PCE variance decomposition showed that the majority of the uncertainty in the metamodel output is due to the canopy height and the vertical density of the scatterers. Additionally, starting from the late reproductive stage, the density of the panicles significantly contributes to the metamodel output [24].

The scattering metamodel and the proposed probabilistic inversion algorithm considers complete morphological structure. Therefore, each height value is supported with a full physical structure including stalk density, leaf length and width, panicle length and diameter and density descriptors as number of plants per  $m^2$  and number of tiller, leaves, and panicles per plant. Consequently, the results need to be analyzed by taking this information into account.

The following sections present the estimation accuracy of the height parameter, which has the highest uncertainty among the model input parameters, in the scattering metamodel. The results were obtained using the proposed probabilistic inversion algorithm over the metamodel using the PSO given for different frequencies (see Section IV-A) and different polarization combinations (see Section IV-B).

##### A. Estimation Accuracy: X-Band and C-Band

As shown in Fig. 6, the probabilistic inversion of the scattering algorithm for height estimation are calculated for three distinct times that have different coarse growth phases. The first group of results belong to the early vegetative and the following two groups correspond to late vegetative and early reproductive stages, respectively. In order to have a viable study, similar resolutions in TSX and RS2 data and HH–VV channel combination fitness function are used for the optimization analysis. For this,  $13 \times 13$  and  $11 \times 11$  window sizes are used for TSX and

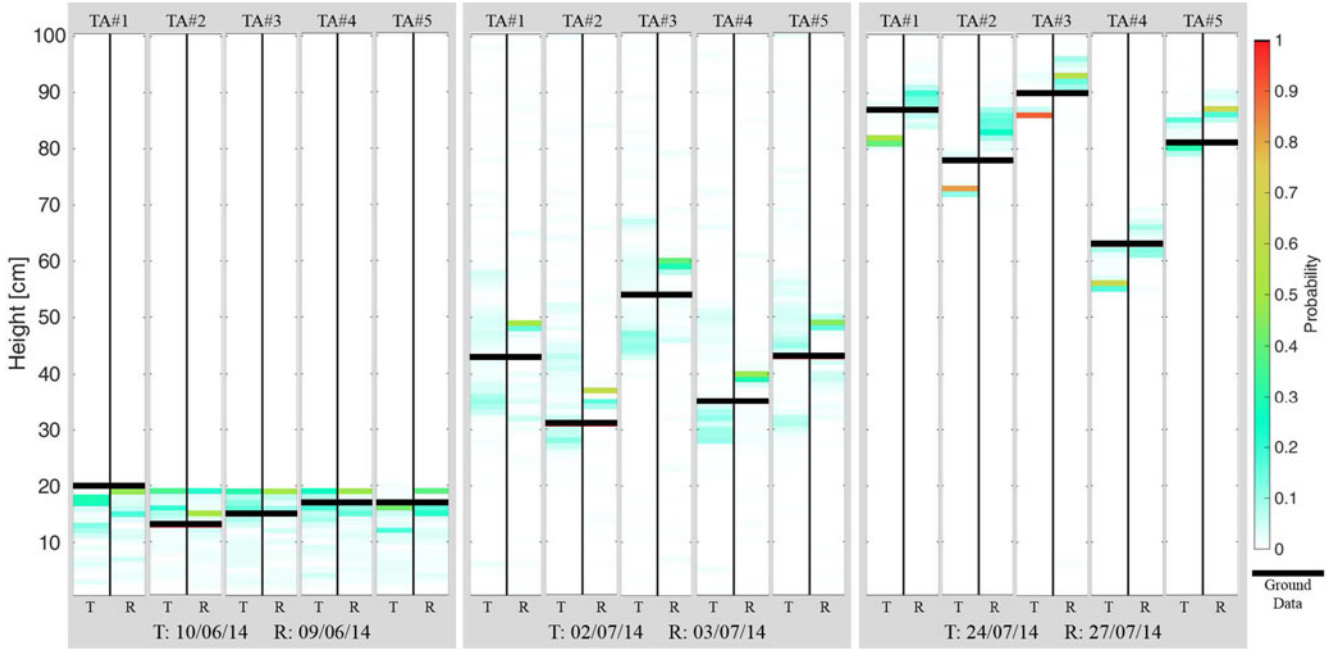


Fig. 6. Height estimation accuracy distribution plots for three different dates and for five spatially independent Test Areas (TA). T and R correspond to TerraSAR-X and RADARSAT-2, respectively. Number of iterations = 200.

TABLE III  
FIELD AVERAGE CONVERGENCE RATE VALUES X- AND C-BAND DATA IN THE EARLY VEGETATIVE PHASE

	X-band	C-band
Range	$0.19 \pm 0.02$	$0.15 \pm 0.03$
abs (1-(Mean/Median))	$0.15 \pm 0.01$	$0.05 \pm 0.03$
Standard deviation	$0.071 \pm 0.015$	$0.042 \pm 0.007$

RS2, respectively. The results are presented over five independent test fields. The color scale in Fig. 6 represents the detection probability of the canopy height with the black line representing the value of the field-averaged ground measurement.

During the early vegetative stage, the probabilistic inversion algorithm for each frequency has consistent results. In this stage, stalk height varies between 5 and 30 cm. There is no obvious over- or under-estimation trend for the full test fields. The majority of the deviation of the estimation from the ground measurement varies in the range of  $\pm 10$  cm. Besides, as shown in Table III, the proposed algorithm has better rate of convergence in C-band compared to X-band. This may originate from the sensitivity of X-band to smaller scaled morphological changes, in which the growth of plant structures affect X-band backscattering behavior stronger than C-band.

After the early vegetative stage, the plant goes into the late vegetative phase. During this stage, the plant height evolves from 15 to 110 cm, while the backscattering intensities increase 10 dB in average for each individual channel (e.g. HH, HV, and VV). Moreover, as shown in Fig. 5, this phase has the largest available morphological parameter space. Therefore, as shown in Table IV, the convergence rate for different morphological

TABLE IV  
FIELD AVERAGE CONVERGENCE RATE VALUES X- AND C-BAND DATA IN THE LATE VEGETATIVE PHASE

	X-band	C-band
Range	$0.80 \pm 0.06$	$0.83 \pm 0.07$
abs (1-(Mean/Median))	$0.06 \pm 0.03$	$0.05 \pm 0.03$
Standard deviation	$0.151 \pm 0.012$	$0.134 \pm 0.017$

TABLE V  
FIELD AVERAGE CONVERGENCE RATE VALUES X- AND C-BAND DATA IN THE EARLY REPRODUCTIVE PHASE

	X-band	C-band
Range	$0.31 \pm 0.07$	$0.29 \pm 0.06$
abs (1-(Mean/Median))	$0.03 \pm 0.02$	$0.03 \pm 0.02$
Standard deviation	$0.032 \pm 0.007$	$0.033 \pm 0.008$

parameters are reduced in both frequencies, which causes a higher deviation in the inversion results. Additionally, in this stage the scattering algorithm tends to underestimate for X-band and over estimate for C-band. This condition has both model-based and physical reasons. From the model point of view, variation in the model parameters changes the value of the optimum solution. Besides, the difference in the data calibration methods for the TSX and RS2 data may effect this situation. The other important issue is the frequency. Since the frequency is important for the interaction with the target, it may cause under- or over-estimated results. Thus, based on these two reasons, small degree of deviation ( $\leq 10$  cm) is accepted.

Finally, the last case of the frequency comparison analysis belongs to the early reproductive stage. During this stage, the

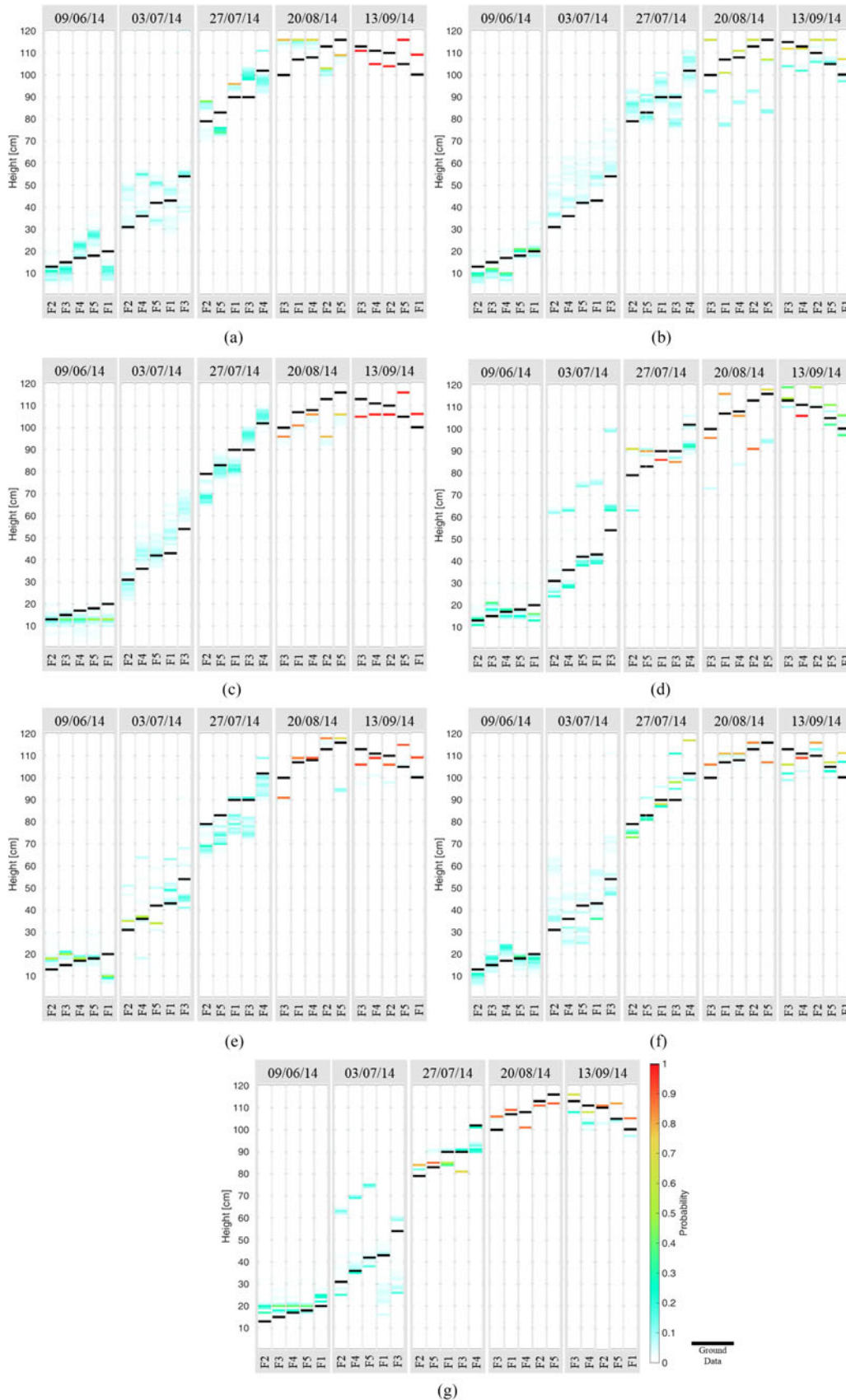


Fig. 7. Height estimation accuracy distribution plots for seven different polarization channel combinations, five different dates for five spatially independent TAs. Number of runs for the PSO = 200. The x-axis is sorted with respect to the increasing height. (a) Single Pol: HH, (b) Single Pol: HV, (c) Single Pol: VV, (d) Dual Pol: HH&HV, (e) Dual Pol: HV&VV, (f) Dual Pol: HH&VV, and (g) Quad Pol: HH&HV&VV.

canopy height increases from 75 to 120 cm. Moreover, by initiation of the ears and the panicles slows down the decreasing trend of the backscattering intensities and turns them to an increasing trend. In terms of the convergence rate of the metamodel, given in Table V, an antipodean situation is observed due to given morphological reasons with respect to the late vegetative stage.

In overall, under the same conditions for the incidence angle and the local resolution, the convergence rate of C-band is observed to be higher than X-band in the vegetative stage of the rice fields. However, in the early reproductive stage, the degree of convergence increases significantly for X-band, while it decreases for C-band due to small-scaled changes in the plant morphology. In addition, since the model is not capable of perfectly explaining the backscattering behavior, the amount of deviation in the mode of the distribution resulting from the ground measurement values can be considered acceptable for any industrial monitoring applications.

### B. Estimation Accuracy: Single-, Dual-, and Quad-Pol

The estimation accuracy analyses for the different channels have been carried out using seven cases corresponding to different single polarimetric (HH, HV, and VV) or multiple polarimetric channels (HH&HV, HV&VV, HH&VV and HH&HV&VV). In this section, the probabilistic inversion algorithm for the biophysical parameter retrieval is applied over the fully polarimetric RS2 data, which was acquired through the full growth cycle of rice crops in total of five acquisitions.

The results of the probabilistic inversion are presented in Fig. 7 for each fitness function and test field. The color scale used in Fig. 7 is based on the observance probability of the results, with the black line emphasizing the ground measured value. The outcomes are thoroughly analyzed in Fig. 8. The analysis covers the standard deviation and the bias between the mean of the distribution and the ground measured value. A summary presenting the distance with the mean of the results and the optimum solution, point (0,0), is given in Table VI. The channel comparison over the optimization procedure is discussed for each growth phase (e.g., early vegetative, late vegetative, early reproductive, late reproductive, and maturative) separately to provide a clear view about the effect of the distinct morphological differences.

1) *Early Vegetative (St.1)*: The highest height retrieval accuracy is observed with the single channel fitness function of the HH polarization, while the lowest is observed with the VV channel. From the optimization with dual-polarization configurations, all options have provided similar distributions. However, the quad-polarization fitness function has resulted in a lower accuracy compared to dual-polarization fitness functions. Therefore, the probabilistic retrieval of the height parameter in the first phase of the growth can be achieved solely using the HH intensity. This condition is supported by the plant morphology. The reason is that the stalks are mostly underwater with all the leaves above the water surface. This results in weaker interaction of electromagnetic waves with the vertical structures of the canopy. Consequently, the canopy height is underestimated more for the VV polarimetric channel.

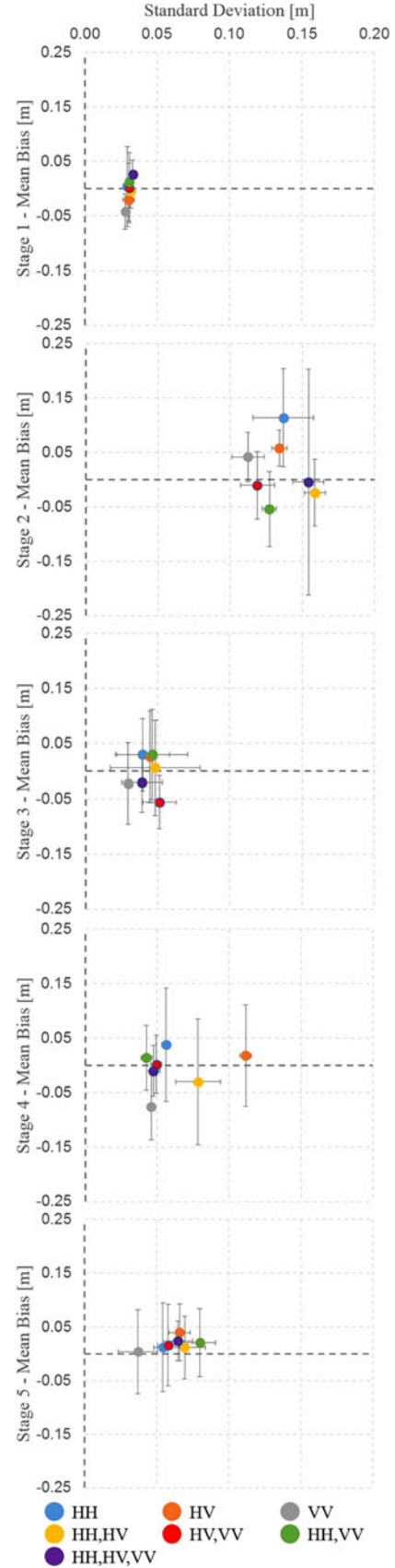


Fig. 8. Stage average of standard deviations versus scene average of mean bias values are given with error bars representing the corresponding standard deviations.



2) *Late Vegetative (St.2)*: In order to decide upon the estimation quality of the stalk height retrieval algorithm in the second stage, the growth boundaries should be remembered. Due to the wide range of possible solutions, the results tend to have a higher variance around the optimum solution as observed in ground measurements. Vertical structures of the canopy are formed during the late vegetative phase of the growth. Thus, the results with the VV including fitness functions have higher accuracy. The highest accuracies are observed in the VV, HV&VV based fitness functions. However, in contrast to the early vegetative phase of the growth, the lowest accuracy was observed in the single-polarization HH function due to strong attenuation from the leaves.

3) *Early Reproductive (St.3)*: The early reproductive stage can be described as a growth-wise saturation stage for the plant height. The rate of increase for the canopy height slows down significantly in this phase. However, the plant wet biomass continues to increase. Since the physical structure of the canopy is continuing to be mainly vertical, VV channel fitness function continues to have the highest accuracy. Besides, the height retrieval from the fitness function of the HV channel in combination with VV has the lowest accuracy.

4) *Late Reproductive (St.4)*: In the late reproductive stage of the growth cycle, the plant reaches its maximum wet biomass and structural density. Therefore, in this dense structural environment none of the single polarization-based fitness functions have high accuracy. Besides, the HV channel is significantly different from the ground value compared to the rest of the cases. In this growth phase, the fitness functions of the HV&VV, HH&VV, and HH&HV&VV combinations showed the highest accuracies.

5) *Maturative (St.5)*: As the rice field reaches the end of its growth cycle, the water content decreases and the vertical structure becomes easily observable again. This condition is also observed in the probabilistic retrieval of the stalk height parameter. The fitness function that takes the VV channel into account has the highest accuracy compared to the others. On the other hand, the function for the HH&VV channel combination has the lowest accuracy.

The extensive analysis of the proposed probabilistic inversion of a morphology-based electromagnetic scattering model approach shows that there is no single fitness function present for the retrieval of the stalk height in rice fields. From the set of single channel fitness functions, the function of the VV channel shows superiority in most of the growth phases. However, a solution obtained through a single channel fitness function (e.g., HH, HV, and VV) may result in a stronger deviation from the optimum solution with respect to the function of the polarimetric channel combinations. Therefore, to provide the morphological consistency in the 3-D space, the channel combinations with at least two channels should be preferred. So, until the maturation stage, the fitness function for the HH&VV combination has lower average standard deviation and lower bias between the solution and the ground measurement, based on Table VI.

The observed deviation in the results mainly occurs from three reasons. The first reason is in the accuracy of the scattering model in explaining the scattering mechanisms. Since the

TABLE VI  
EUCLIDEAN DISTANCE [CM] BETWEEN THE MEAN OF THE ANALYSIS RESULT AND THE GROUND MEASUREMENT

Polarization	St.1	St.2	St.3	St.4	St.5
HH	2.97	17.81	4.95	6.80	5.57
HV	3.64	14.60	5.16	11.32	7.72
VV	5.05	12.02	3.67	8.88	3.72
HH,HV	3.21	16.04	4.86	8.41	7.06
HV,VV	3.09	11.94	7.59	4.96	6.02
HH,VV	3.31	13.82	5.49	4.48	8.27
HH,HV,VV	4.22	15.43	4.38	4.84	6.91

Lower distance indicates higher accuracy.

model cannot consider the environmental impacts and has simplifying assumptions about the plant morphology, it would be mathematically impossible to expect perfect inversion results. The second reason is the presence of the speckle noise. Even with spatial smoothing, the presence of the noise in the SAR data affects the backscattering intensities. The third and the least significant reason is the accuracy of the PCE metamodel in fitting to the backscattering model. Considering all the reasons, the majority of the deviation originates from the first reason, the forward approach accuracy of the backscattering model. It can be solved by using an alternative scattering model, which needs to be able to provide a better explanation to the scattering behavior of the rice canopy.

The drawbacks of the proposed approach lie in two important points. The first point is the requirement of a preclassification for the decision of the coarse growth phase and the second one is the requirement of the biological growth boundaries for the rice crop. It needs to be mentioned that first issue has already been studied in the literature and several statistical methods have already been proposed such as in [16], [17], and [30]. For the second case, the morphological information about all major crops, including rice, are available in the literature to avoid impossible structures such as plants with 2 m stalk height and 2 cm stalk diameter.

## V. CONCLUSION AND FUTURE WORK

In this paper, a probabilistic inversion method for a scattering model has been proposed, applied and tested for biophysical parameter retrieval such as stalk height of rice crops from polarimetric SAR data. The proposed probabilistic inversion scheme is designed to be computationally efficient using HDMR methods, specifically a PCE metamodel. The approach requires a preclassification step for the coarse growth phase and the biological growth boundaries of the rice crops. In this method, unlike the previous ones, the biophysical parameters of the targets are estimated in terms of their probability distributions for a specific polarimetric backscattering intensity. In other words, the retrieval algorithm is developed to estimate the multivariate distribution of the possible plant structures that correspond to a measured backscattering intensity under different conditions like different frequencies and polarimetric channels.

The accuracy of the proposed methodology was tested in rice fields for two different frequencies (X- and C-band) and seven

different cases of polarimetric channel combinations. In terms of the frequency comparison, inversion outcome in X-band provided higher accuracy than C-band with higher sensitivity to morphological variations. This higher sensitivity in X-band can be explained by the small-scaled developments in the plant morphology compared to the wavelength. Additionally, the channel combination comparison in C-band pointed out that, due to vertical structure of the canopies, use of VV channel or the combination of the HH&VV channels has higher overall height estimation accuracy through the growth cycle.

Since the scattering model considers the physical changes in 3-D space including all components of the plant morphology, the applicability of the proposed approach is expected to be valid for different locations. However, the performance of the probabilistic inversion algorithm mainly depends on the accuracy of the forward scattering model in explaining the scattering behavior. Lower accuracy of the forward electromagnetic scattering model may lead to stronger deviations from the real value of the biophysical parameter. It is also possible to apply the given approach for the determination of full plant morphology such as canopy density and corresponding dimensionality of the leaves or panicles. However, following the outcomes of the GSA the research is focused on the stalk height, which is the most significant parameter for the polarimetric backscattering intensity in rice crops.

Finally, the novel approach presented in this study emphasizes the use of PolSAR for biophysical parameter estimation using a metamodel based and, therefore, computationally inexpensive morphology-based electromagnetic scattering model. Using such metamodels allows use of normally computationally expensive algorithms in industrial applications for much lower costs.

In the future, it is planned to broaden the use of the proposed algorithm on different crops under the concept of the precision agriculture. Moreover, the improvement of the probabilistic retrieval algorithm and HDMM methods to eliminate the preclassification step will be a subject of future research.

#### ACKNOWLEDGMENT

The authors would like to thank the support of Dr. S. Marelli and Prof. Dr. B. Sudret from Chair of Risk, Safety & Uncertainty Quantification from ETH Zurich with the Polynomial Chaos Expansion metamodel and to the Directorate of Trakya Agricultural Research Institute for proving the ground measurements from Turkey.

#### REFERENCES

- [1] D. Lamb and R. B. Brown, "PaPrecision agriculture: Remote-sensing and mapping of weeds in crops," *J. Agricultural Eng. Res.*, vol. 78, no. 2, pp. 117–125, 2001.
- [2] D. G. Blumberg, "High resolution X-band SAR imagery for precise agriculture and crop monitoring," in *Proc. 3rd Int. Workshop Sci. Appl. SAR Polarimetry Polarimetric Interferometry*, 2007.
- [3] L.-F. Wang, "Monte Carlo simulation model for electromagnetic scattering from vegetation and inversion of vegetation parameters," Ph.D. dissertation, Dept. Elect. Eng. Comput. Sci., MIT, Cambridge, MA, USA, 2007.
- [4] Y. Zhang, X. Liu, S. Su, and C. Wang, "Retrieving canopy height and density of paddy rice from Radarsat-2 images with a canopy scattering model," *Int. J. Appl. Earth Observ. Geoinform.*, vol. 28, no. 0, pp. 170–180, 2014.
- [5] U. Wegmüller and C. Werner, "Retrieval of vegetation parameters with SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 1, pp. 18–24, Jan. 1997.
- [6] M. E. Engdahl, M. Borgeaud, and M. Rast, "The use of ERS-1/2 tandem interferometric coherence in the estimation of agricultural crop heights," *IEEE Trans. Geosci. Remote Sens.*, vol. 39, no. 8, pp. 1799–1806, Aug. 2001.
- [7] E. Erten, C. Rossi, and O. Yuzugullu, "Polarization impact in TanDEM-X data over vertical-oriented vegetation: The paddy-rice case study," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 7, pp. 1501–1505, Jul. 2015.
- [8] C. Rossi and E. Erten, "Paddy rice monitoring using TanDEM-X," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 900–910, Feb. 2015.
- [9] J. Ballester-Berman, J. Lopez-Sanchez, and J. Fortuny-Guasch, "Retrieval of biophysical parameters of agricultural crops using polarimetric SAR interferometry," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 4, pp. 683–694, Apr. 2005.
- [10] M. Pichierri, I. Hajnsek, and K. P. Papathanassiou, "A Multibaseline Pol-InSAR Inversion Scheme for Crop Parameter Estimation at Different Frequencies," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4952–4970, 2016.
- [11] J. Hannah, M. Pardini, I. Hajnsek, and K. P. Papathanassiou, "First multi-frequency investigation of SAR tomography for vertical structure of agricultural crops," in *Proc. 10th Eur. Conf. Synthetic Aperture Radar*, 2014, pp. 1–4.
- [12] Y. Inoue *et al.*, "Season-long daily measurements of multifrequency (Ka, Ku, X, C, and L) and full-polarization backscatter signatures over paddy rice field and their relationship with biological variables," *Remote Sens. Environ.*, vol. 81, no. 2, pp. 194–204, 2002.
- [13] J. Lopez-Sanchez, S. Cloude, and J. Ballester-Berman, "Rice phenology monitoring by means of SAR polarimetry at X-Band," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 7, pp. 2695–2709, Jul. 2012.
- [14] J. Lopez-Sanchez, F. Vicente-Guijalba, J. Ballester-Berman, and S. Cloude, "Polarimetric response of rice fields at C-Band: Analysis and phenology retrieval," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2977–2993, May 2014.
- [15] F. Vicente-Guijalba, T. Martinez-Marin, and J. Lopez-Sanchez, "Crop phenology estimation using a multitemporal model and a Kalman filtering strategy," *IEEE Geosci. Remote Sens. Lett.*, vol. 11, no. 6, pp. 1081–1085, Jun. 2014.
- [16] C. Kucuk, G. Taskin, and E. Erten, "Paddy-rice phenology classification based on machine-learning methods using multitemporal co-polar X-Band SAR images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. PP, no. 99, pp. 1–11, 2016.
- [17] O. Yuzugullu, S. Marelli, E. Erten, B. Sudret, and I. Hajnsek, "A metamodel-based inversion for determining rice growth stage from SAR data," *Remote Sens. Environ.*, in review.
- [18] Y. Inoue and E. Sakaiya, "Relationship between X-Band and backscattering coefficients from high-resolution satellite SAR and biophysical variables in paddy rice," *Remote Sens. Lett.*, vol. 4, no. 3, pp. 288–295, 2013.
- [19] M. Karam and A. Fung, "Electromagnetic scattering from a layer of finite length, randomly oriented, dielectric, circular cylinders over a rough interface with application to vegetation," *Int. J. Remote Sens.*, vol. 9, no. 6, pp. 1109–1134, 1988.
- [20] J. M. Stiles and K. Sarabandi, "Electromagnetic scattering from grassland. I. A fully phase-coherent scattering model," *IEEE Trans. Geosci. Remote Sens.*, vol. 38, no. 1, pp. 339–348, Jan. 2000.
- [21] J. Koay *et al.*, "Paddy fields as electrically dense media: Theoretical modeling and measurement comparisons," *IEEE Trans. Geosci. Remote Sens.*, vol. 45, no. 9, pp. 2837–2849, Sep. 2007.
- [22] Y. Oh, S. Hong, Y. Kim, J. Hong, and Y. Kim, "Polarimetric backscattering coefficients of flooded rice fields at L- and C-bands: measurements, modeling, and data analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 8, pp. 2714–2721, Aug. 2009.
- [23] R. Eberhart and J. Kennedy, "A new optimizer using particle swarm theory," in *Proc. 6th Int. Symp. Micro Mach. Hum. Sci.*, New York, NY, USA, 1995, vol. 1, pp. 39–43.
- [24] O. Yuzugullu, S. Marelli, E. Erten, B. Sudret, and I. Hajnsek, "Global sensitivity analysis of a morphology based electromagnetic scattering

- model,” in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2015, pp. 1017–1020.
- [25] T. Le Toan *et al.*, “Rice crop mapping and monitoring using ERS-1 data based on experiment and modeling results,” *IEEE Trans. Geosci. Remote Sens.*, vol. 35, no. 1, pp. 41–56, Jan. 1997.
- [26] L. Tsang, K. Ding, G. Zhang, C. Hsu, and J. Kong, “Backscattering enhancement and clustering effects of randomly distributed dielectric cylinders overlying a dielectric half space based on Monte Carlo simulations,” *IEEE Trans. Antennas Propag.*, vol. 43, no. 5, pp. 488–499, May 1995.
- [27] G. Blatman and B. Sudret, “Efficient computation of global sensitivity indices using sparse polynomial chaos expansions,” *Rel. Eng. Syst. Safety*, vol. 95, no. 11, pp. 1216–1229, 2010.
- [28] S. Marelli and B. Sudret, *UQLab: A Framework for Uncertainty Quantification in MATLAB*, 2014, ch. 257, pp. 2554–2563.
- [29] L. Wang, J. Kong, K. Ding, T. Le Toan, F. Ribbes, and N. Floury, “Electromagnetic scattering model for rice canopy based on Monte Carlo simulation,” *Progress Electromagn. Res.*, vol. 52, pp. 153–171, 2005.
- [30] O. Yuzugullu, E. Erten, and I. Hajnsek, “Rice growth monitoring by means of X-Band co-polar SAR: Feature clustering and BBCH scale,” *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 6, pp. 1218–1222, Jun. 2015.



**Onur Yuzugullu** received the B.Sc. and M.Sc. degrees in environmental engineering from the Middle East Technical University, Ankara, Turkey, in 2008 and 2011, respectively. He is currently working toward the Ph.D. degree at the Earth Observation and Remote Sensing Group, Institute of Environmental Engineering, Swiss Federal Institute of Technology, ETH Zurich, Zurich, Switzerland.

His Master thesis focused on the determination of the chlorophyll-a concentrations in a eutrophic lake using space-borne multispectral data. His research

interest includes the growth stage determination of rice crops considering the relation between crop morphology and the polarimetric signatures from a stochastic point of view.



**Esra Erten** (S’06–M’10) received the Ph.D. degree in computer vision and remote sensing from the Department of Computer Engineering and Microelectronics, Berlin University of Technology, Berlin, Germany, in 2010.

She was with the High-Frequency Institute, German Aerospace Center, Oberpfaffenhofen, Germany, from 2008 to 2010, where she worked on information theory for multichannel SAR images. From 2010 to 2012, she was the Chair of Earth Observation and Remote Sensing, Institute of Environmental Engineering, ETH Zurich, Zurich, Switzerland, where she worked on applied radar remote sensing for environmental parameter estimation. She is currently an Associated Professor at the Department of Geomatics Engineering, Faculty of Civil Engineering, Istanbul Technical University, Istanbul, Turkey. Her research interests include information extraction and image understanding from SAR and optical images; in particular information theory, multivariate statistics, polarimetry, and interferometry.



**Irena Hajnsek** (A’01–M’06–SM’09–F’14) received the Diploma degree (Hons.) from the Free University of Berlin, Berlin, Germany, in 1996, and the Dr. degree (Hons.) from the Friedrich Schiller University of Jena, Jena, Germany, in 2001.

Since November 2009, she has been a Professor of Earth Observation at the Institute of Environmental Engineering, Swiss Federal Institute of Technology, ETH Zurich, Zurich, Switzerland, and also the Head of the Polarimetric SAR Interferometry Research Group at the Microwaves and Radar Institute, German Aerospace Center, Wessling, Germany. Her main research interests include in electromagnetic propagation and scattering theory, radar polarimetry, SAR and interferometric SAR data processing techniques, environmental parameter modeling and estimation. Since 2010, she is the Science Coordinator of the German Satellite Mission TanDEM-X. She was the Technical Program Cochair of the IEEE International Geoscience and Remote Sensing Symposium 2012, Munich, Germany. Since 2013 she has been a Member of the IEEE Geoscience and Remote Sensing Society AdCom.