

# DETERMINING RICE GROWTH STAGE WITH X-BAND SAR: A METAMODEL BASED INVERSION

O. Yüzüğüllü, S. Marelli, E. Erten, B. Sudret and I. Hajnsek



## Data Sheet

---

**Journal:** Remote Sensing

**Report Ref.:** RSUQ-2017-014

**Arxiv Ref.:** –

**DOI:** <https://doi.org/10.3390/rs9050460>

**Date submitted:** 30 March 2017

**Date accepted:** Accepted: 3 May 2017

---

# Determining Rice Growth Stage with X-Band SAR: A Metamodel Based Inversion

Onur Yüzügüllü, Stefano Marelli, Esra Erten, Bruno Sudret and Irena Hajnsek

June 11, 2018

## Abstract

Rice crops are important in the global food economy, and new techniques are being implemented for their effective management. These techniques rely mainly on the changes in the phenological cycle, which can be investigated by remote sensing systems. High frequency and high spatial resolution Synthetic Aperture Radar (SAR) sensors have great potential in all-weather conditions for detecting temporal phenological changes. This study focuses on a novel approach for growth stage determination of rice fields from SAR data using a parameter space search algorithm. The method employs an inversion scheme for a morphology-based electromagnetic backscattering model. Since such a morphology-based model is complicated and computationally expensive, a surrogate metamodel-based inversion algorithm is proposed for the growth stage estimation. The approach is designed to provide estimates of crop morphology and corresponding growth stage from a continuous growth scale. The accuracy of the proposed method is tested with ground measurements from Turkey and Spain using the images acquired by the TerraSAR-X (TSX) sensor during a full growth cycle of rice crops. The analysis shows good agreement for both datasets. The results of the proposed method emphasize the effectiveness of X-band PolSAR data for morphology-based growth stage determination of rice crops.

## 1 Introduction

Temperate climatic conditions with easy access to water sources provide optimum conditions for rice cultivation. In the history of agricultural practices, rice farming goes back to almost 8000 BC. Currently, rice is a major source of income for the rural communities all around the world. According to the International Rice Research Institute (IRRI), worldwide rice production totaled 969 million tons in 2010 GRI SP (Global Rice Science Partnership) (2011). Frequent and efficient monitoring strategies are necessary for the optimization of economic competitiveness and the estimation of the associated environmental impacts (e.g.,

methane emissions). Concerning these issues, researchers have focused on finding sustainable monitoring methods for agricultural fields.

In the last decade, remote sensing techniques have frequently been used as a viable method to monitor agricultural areas Atzberger (2013); Mulla (2013). To this end, two main data sources are presented: optical and Synthetic Aperture Radar (SAR) systems. Optical systems measure reflected sunlight and provide spectral properties of their targets. Additionally, due to short wavelengths ( $\lambda < 2500$  nm), they are mainly susceptible to atmospheric factors. On the other hand, SAR systems are superior compared to optical systems concerning their temporal coverage with their “all-weather” day and night imaging capability. Furthermore, they play a significant role in environmental monitoring with their sensitivity to the physical alterations in monitored objects. However, to provide a clear explanation of such changes, one has to understand the relation between the object and the backscattered energy in SAR systems. This relation can be a complex interaction between the object and the sensor parameters including frequency, polarization setting and incidence angle.

The selection of appropriate sensor parameters is crucial for agricultural monitoring. For instance, with SAR systems, one should match the size of the structural parts of the crops with the available wavelength (frequency) of the system to identify the effects of morphological changes. Furthermore, the use of different frequencies and incidence angles changes the scattering behavior and the attenuation of the waves inside the canopy Lim et al. (2007). Thus, in crop monitoring, high-frequency ( $f > 5$  GHz), in other words low wavelength ( $\lambda < 6$  cm), systems are expected to be more sensitive to morphological changes than low-frequency systems.

Several studies monitoring paddy rice fields show that plant morphological structures have a high correlation with the backscattering behavior of high frequency electromagnetic (EM) waves Ulaby et al. (1984); Bouman (1991); Inoue and Sakaiya (2013); Inoue et al. (2014); Rossi and Erten (2015); Yuzugullu et al. (2017); Erten et al. (2017). There are currently two mainstream approaches to monitor crops, namely backward and forward. The first class of approaches is based on the statistics of the polarimetric parameters Shao et al. (2001); Sakamoto et al. (2005); Lopez-Sanchez et al. (2012); Inoue et al. (2002); Koppe et al. (2013); Vicente-Guijalba et al. (2014); Lopez-Sanchez et al. (2014); Ribbes (1999); Yonezawa et al. (2012); Inoue et al. (2014); Erten et al. (2015). Among these, Inoue et al. (2002, 2014) have completed the most comprehensive work to date by observing the full phenological cycle with different frequencies, polarizations and incidence angles. Such methods also depend on several factors including temporal variations in the structural density, type of crop or the use of seeds with different genotypes. In the literature, similar temporal trends have been observed for polarimetric descriptors during the phenological cycle, such as intensity, entropy, alpha and phase differences Inoue et al. (2002); Lopez-Sanchez et al. (2012). The statistical classification methods in the literature are cost-effective and easy to implement. However,

the high dynamic range of the polarimetric parameters makes it not trivial to develop widely applicable approaches that cover the full growth period in different locations. Consequently, current thresholding-based statistical methods need to be adjusted for each new monitoring campaign depending on the region of the world in which they take place. Additionally, they usually require a full-time series for each new campaign. Therefore, most of the current statistical methods are not capable of explaining the growth cycle of rice crops in different areas while avoiding high training costs.

Unlike the statistical ones, the second class of approaches is based on the scattering behavior of an EM wave inside the canopy. This behavior can be explained by analytical relations between the crop physical structure (i.e., morphology) and the backscattering coefficients Toan et al. (1997); Stiles and Sarabandi (2000); Bouman et al. (1999); Koay et al. (2007); Wang et al. (2005). Such backscattering models take the plant morphology as input and provide the scattering properties of the EM wave (i.e., backscattering coefficients) as output Lin and Sarabandi (1999). Since they consider the geometrically-simplified crop morphology, they usually have sophisticated mathematical algorithms. With such a level of complexity, knowing the sensitivity of the model outputs on the input parameters is essential to understand the dynamics of the model. Global Sensitivity Analysis (GSA) provides the necessary quantitative tools to assess it properly Sobol (1993); Sudret (2008). However, the implementation of GSA can result in high computational costs due to the large number of model evaluations needed in standard sampling-based approaches Sobol (1993). An effective strategy to significantly reduce this computational burden is by using Polynomial Chaos Expansion (PCE), a surrogate modeling technique that has exceptional convergence properties for the estimation of Sobol indices for complex models Sudret (2008). It reduces the total computational cost to that of its training set, which is typically relatively small. Therefore, PCE makes GSA computationally efficient. Finally, after the assessment of the importance of model parameters, it is possible to develop an inversion scheme for the crop morphology from polarimetric observations. Despite their higher complexity (mathematical and computational), inverse methods have the potential to provide a much deeper insight into the actual growth stage of a crop field, as they take into account the quantitative interaction between the plants and the EM field. Additionally, they can provide a continuous estimate of the growth stage, because they are directly sensitive to the underlying plant morphology.

An incoherent EM backscattering model Toan et al. (1997) is chosen for this study. The model considers a simplified plant morphology with a higher number of unknowns than the number of measurables. For such an ill-posed problem, an analytical inversion approach is infeasible. Furthermore, the presence of speckle noise in the measured intensity data makes pixel-sized inversion algorithms less efficient. In this paper, a parameter search space algorithm combined with a PCE surrogate of the full EM model is considered as a powerful option to handle these issues for the model inversion scheme.

This article proposes a new and effective method to determine the growth stage of flooded and broadcast-sown rice fields in large-scale cultivation areas using Polarimetric SAR (PolSAR) data. Apart from the previously-mentioned growth phase determination methods for the rice fields, the proposed method focuses on the effect of the changes in crop morphology on the polarimetric backscattering intensities during the phenological cycle. It extends the a priori growth phase information with an EM backscattering model in a computationally-efficient framework. The EM model inversion consists of a PCE-based parameter space search algorithm. Finally, the results are combined to determine the growth stage of the field from its morphology.

The paper provides the theory behind the proposed inversion approach in detail along with concise information about PCE metamodels and GSA in Section 2. Section 3 covers the test areas with the ground measurements and TerraSAR-X (TSX) campaigns. Section 4 presents the main results of GSA and growth stage estimation. Section 5 summarizes the work with an overall view on the proposed growth stage estimation approach in the context of precise agriculture.

## 2 Growth Stage Determination

Understanding the growth phases of the rice growth cycle is crucial in explaining their effect on the SAR system responses. The most common rice cultivation practice begins by flooding the fields several weeks before sowing. There are two main planting methods: transplanting and broadcasting. In transplanting, the seedlings are prepared and then transplanted to the fields to provide regular spacing between the plants. On the contrary, with broadcast sowing, the seeds are thrown on the flooded fields, resulting in spatial morphological heterogeneity (GRIISP (Global Rice Science Partnership) (2011)).

Three significant growth periods can be identified in rice cultivation: vegetative, reproductive and maturative. The full cycle takes between 120–150 days depending on agricultural and environmental factors. In the literature, the rice growth cycle is defined by two distinct scales: International Rice Research Institute (IRRI) GRIISP (Global Rice Science Partnership) (2011) and Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie (BBCH) Lancashire et al. (1991). The IRRI scale divides the growth cycle into five phases, while the BBCH scale uses 100 stages between 0 and 99. A general overview of the growth cycle is presented in Figure 1 with sample morphologies. In this research, the IRRI scale is chosen as the a priori growth phase information.

- Vegetative period: The crops increase in height and structural density, depending on several factors such as soil properties, temperature and seeding density. The stalk orientation stays mostly vertical. Since the plants are structurally weak, the duration of this period strongly depends on the environmental conditions and the genotype of

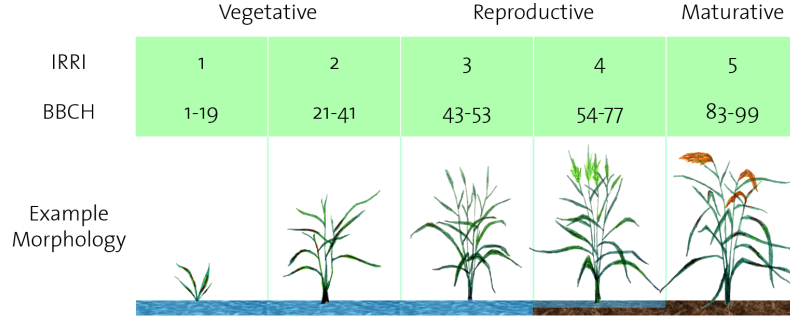


Figure 1: Growth cycle of a rice plant with the corresponding International Rice Research Institute (IRRI), Biologische Bundesanstalt, Bundessortenamt und Chemische Industrie (BBCH) scale and sample structure.

the crops.

- Reproductive period: As the plants become stronger, they become less sensitive to the environmental stresses. Plant height and density continue to increase heterogeneously together with increasing wet biomass, which leads to varying orientation in leaves and stalks due to increasing weight. The flag leaf forms through the end of the period.
- Maturative period: Excess water in the fields is drained, leading to a reduction in plant total biomass due to lower moisture content. Grains become more mature and heavier.

In this study, the growth stages of the fields are determined using the approach shown in Figure 2. This method uses three different inputs: PolSAR data, growth information and the backscattering (EM) model. PolSAR data are used in two steps of the algorithm: feature clustering of polarimetric parameters for BBCH assignment and the parameter space search algorithm. Phenological data with a priori growth phase information are used for determining the growth boundaries and trends and training a first PCE metamodel that predicts the BBCH scale based on the available morphological parameter ( $PCE_{BBCH}$ ). The backscattering model is then surrogated by an additional PCE metamodel for each of the various growth phases ( $PCE_{EM}$ ). Later, the outputs of the feature clustering,  $PCE_{EM}$  and growth trends as natural limitations are integrated into the parameter search space algorithm. Finally, the growth stages are estimated using clustered PolSAR data and the  $PCE_{BBCH}$ .

## 2.1 Backscattering Model

In this study, the canopy is modeled as uniformly-distributed individual plants over a half space that represents the flooded ground used in broadcast seeding. The backscattering coefficients are estimated using the first order solution of the radiative transfer equation. The chosen model provides a structural description of the plants including their simplified crop morphology (e.g., stalks, tillers, leaves and panicles). The model also includes the backscattering enhancements and resulting wave clustering effects from the scatterers Toan

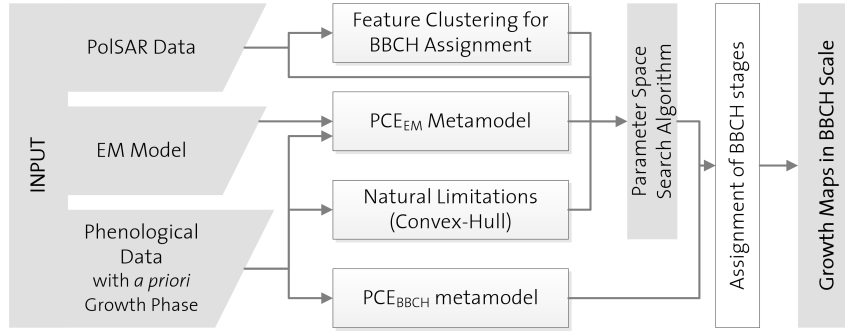


Figure 2: Block diagram of the proposed approach.

et al. (1997). The resulting backscattering intensities are estimated for different polarimetric channels by performing Monte Carlo (MC) simulations using Foldy–Lax multiple scattering equations Tsang et al. (1995).

In the simulations, rice plants with vertically-oriented stalks are placed randomly in a unit area  $A$ , as in broadcast seeding. The locations of the plants are randomized automatically in each iteration of the MC simulation to provide spatial heterogeneity. Inside  $A$ , there are  $n_s$  plants with  $n_t$  tillers with average height  $h_s$  and diameter  $d_s$ . Each tiller has  $n_l$  leaves with length  $l_l$  and width  $w_l$  and  $n_p$  panicles with length  $l_p$  and width  $w_p$ . The complex dielectric constants are  $\epsilon_{s,l}$  for all plant structures (e.g., tillers, leaves and panicles) and  $\epsilon_g$  for the underlying ground.

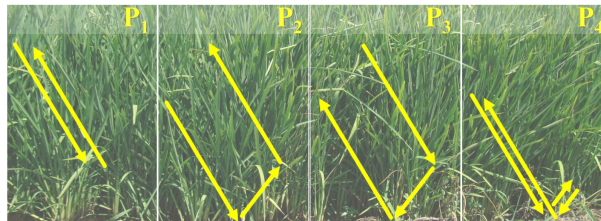


Figure 3: The scattering mechanisms involved in the chosen EM model.

The current first order solution of the electromagnetic scattering problem considers four major scattering mechanisms ( $P_n$ ), visualized in Figure 3:

1. Direct scattering from the scatterers
2. Scattering from the canopy followed by reflection from the ground
3. Reflection from the ground followed by scattering from the canopy
4. Reflection from the ground followed by scattering from the canopy, again followed by reflection from the ground:



$$\begin{aligned}
E_q^s(\mathbf{r}') &= \frac{e^{ikr}}{r} (P_1 + P_2 + P_3 + P_4) E_p^i \\
&= \sum_t \sum_j \left[ f_{qp}^t(\pi - \theta_i, \pi + \phi_i; \theta_i, \phi_i) e^{-M_{qq} \frac{z_j^t}{\cos \theta_i}} e^{-M_{pp} \frac{z_j^t}{\cos \theta_i}} e^{2i(k_x^i x_j^t + k_y^i y_j^t - k_z^i z_j^t)} \right. \\
&\quad + R_q(\theta_i) f_{qp}^t(\theta_i, \pi + \phi_i; \theta_i, \phi_i) e^{+M_{qq} \frac{2h + z_j^t}{\cos \theta_i}} e^{-M_{pp} \frac{z_j^t}{\cos \theta_i}} e^{2i(k_x^i x_j^t + k_y^i y_j^t - k_z^i h)} \\
&\quad + f_{qp}^t(\pi - \theta_i, \pi + \phi_i; \pi - \theta_i, \phi_i) R_p(\theta_i) e^{-M_{qq} \frac{z_j^t}{\cos \theta_i}} e^{+M_{pp} \frac{2h + z_j^t}{\cos \theta_i}} e^{2i(k_x^i x_j^t + k_y^i y_j^t - k_z^i h)} \\
&\quad \left. + R_q(\theta_i) f_{qp}^t(\theta_i, \pi + \phi_i; \pi - \theta_i, \phi_i) R_p(\theta_i) e^{+M_{qq} \frac{2h + z_j^t}{\cos \theta_i}} e^{+M_{pp} \frac{2h + z_j^t}{\cos \theta_i}} e^{2i(k_x^i x_j^t + k_y^i y_j^t - k_z^i (2h + z_j^t))} \right] E_p^i \quad (1)
\end{aligned}$$

The top surface of the canopy is indicated as  $z = 0$  and the underlying surface as  $z = -h$ . It is possible to model the behavior of an incident wave  $\bar{E}^i$  in the direction  $(\theta_i, \phi_i)$ , of the incidence and look angle, using (1). The model follows the finite cylinder approximation Karam and Fung (1988); Karam et al. (1988) for stems, tillers and panicles and the physical optics approximation Karam and Fung (1989) for leaves. Additionally, the model variables are listed as: type of the morphological structure,  $t$ ; scattering matrix element where  $q$  and  $p$  are polarization channels ( $q, p$  for H,V) for scattered and incident waves,  $f_{qp}^t$ ; propagation vector of the incident and scattered wave,  $\bar{k}_p^{i,s}$ ; Fresnel reflection coefficients,  $R_p(\theta)$  and  $R_q(\theta)$ . Lastly, the effect of the attenuation due to mixed structural scatterers inside the canopy is considered by the  $M_{qp}$  term:

$$M_{qp} = \frac{i2\pi n_s n_t}{k_0 A h} (\langle f_{qp}^{tiller} \rangle + n_l \langle f_{qp}^{leaf} \rangle + n_p \langle f_{qp}^{panicle} \rangle) \quad (2)$$

where the angular brackets represent configurational average,  $h$  is the height of the canopy and  $k_0$  is the free space wave number. Structural location vectors are given as:  $k_x^i = k_0 \sin \theta_i \cos \phi_i$ ,  $k_y^i = k_0 \sin \theta_i \sin \phi_i$ ,  $k_z^i = k_0 \cos \theta_i$ . The backscattering coefficients for the polarimetric channel,  $qp$ , are estimated from the ratio between the amplitudes of the scattered and incident electrical waves (3).

$$\sigma_{qp}^o = \frac{4\pi r^2 \langle |E_q^s|^2 \rangle}{A_i |E_p^i|^2} \quad (3)$$

where  $A_i$  is the illuminated area and  $r$  is the distance between the sensor and the target. For the MC simulation, the backscattering coefficients are averaged over 200 realizations.

## 2.2 Polynomial Chaos Expansion and Global Sensitivity Analysis

Sparse polynomial chaos expansions (PCE) are a well-known technique in the uncertainty quantification literature, and they are well suited to inversion problems. Compared to other surrogate modeling techniques such as Gaussian process modeling (also known as kriging, Thomas J. Santner (2013)) or support vector regression Vapnik (1995), they are particularly well suited for the solution of inverse problems. Indeed, their global approximation character combined with the strict relation they share with Sobol variance decomposition Sudret (2008);

Marelli and Sudret (2015), as well as and their built-in error estimators can be directly applied to assess the identifiability of parameters in inverse problems.

### 2.2.1 Polynomial Chaos Expansion

To reduce the high costs associated with the MC simulation of morphology-based scattering, the computational model can be substituted with a metamodel, a computationally inexpensive analytical approximation of the full computational model. Due to its versatility and relatively low training costs, sparse PCE Blatman and Sudret (2010) is an ideal candidate. PCE is a spectral decomposition technique that allows one to represent a finite-variance scalar-output function  $Y = \mathcal{M}(\boldsymbol{\xi})$  as:

$$Y = \mathcal{M}(\boldsymbol{\xi}) = \sum_{j=0}^{\infty} a_j \Psi_j(\boldsymbol{\xi}) \quad (4)$$

where  $\boldsymbol{\xi} \in \mathbb{R}^M$  is the random vector of morphological parameters,  $a_j \in \mathbb{R}$  is a set of scalar coefficients and the  $\Psi_j(\boldsymbol{\xi}) \in \mathbb{R}$  form a polynomial orthonormal basis with respect to the functional scalar product (expectation value):

$$\langle g(\boldsymbol{\xi})h(\boldsymbol{\xi}) \rangle = \int_{\mathcal{D}_{\boldsymbol{\xi}}} g(\boldsymbol{\xi})h(\boldsymbol{\xi})f_{\boldsymbol{\xi}}(\boldsymbol{\xi})d\boldsymbol{\xi} \quad (5)$$

where  $\mathcal{D}_{\boldsymbol{\xi}}$  is the support of  $\boldsymbol{\xi}$  and  $f_{\boldsymbol{\xi}}(\boldsymbol{\xi})$  the Probability Density Function (PDF) of the input random vector  $\boldsymbol{\xi}$ . Due to the linearity of Equation (4), the  $a_j$  coefficients can be non-intrusively and efficiently calculated using compressive-sensing-based least-square minimization techniques (e.g., least angle regression-based selection Blatman and Sudret (2010)) from a training set of full model evaluations of  $\mathcal{M}(\boldsymbol{\xi})$ . The size of the training set determines the maximal complexity and the accuracy of the resulting metamodel. PCE was implemented in MATLAB<sup>®</sup> within the UQLab framework Marelli and Sudret (2014, 2015).

### 2.2.2 Global Sensitivity Analysis: Sobol Indices

GSA allows one to quantify the effect of the variability of each of the input parameters in  $\boldsymbol{\xi}$  on the variability of the model response  $\mathcal{M}(\boldsymbol{\xi})$ . A widely-accepted global sensitivity measure in the uncertainty quantification literature is given by the variance-decomposition-based Sobol indices Sobol (1993).

The basic form of variance decomposition consists in representing a computational model  $\mathcal{M}(\boldsymbol{\xi})$  as a sum of functions depending only on increasingly larger subsets of the input vector  $\boldsymbol{\xi}$  as follows:

$$\mathcal{M}(\boldsymbol{\xi}) = \mathcal{M}_0 + \sum_{i=1}^M \mathcal{M}_i(\xi_i) + \sum_{i \neq j} + \dots + \mathcal{M}_{12\dots M}(\xi_1, \xi_2, \dots, \xi_M) \quad (6)$$

where the  $\mathcal{M}_{i_j\dots_s}$  are scalar functions depending on the subset of input variables  $\{\xi_i, \xi_j, \dots, \xi_s\}$ . In Sobol Sobol (1993), it is demonstrated that such a decomposition exists for every finite-variance functional and that it is orthonormal, hence yielding unique coefficients. Sobol indices are defined as the ratio of the variance of each term  $D_{i_j\dots_s}$  in Equation (6) to the total variance  $D$ :

$$S_{ij\dots s} = D_{ij\dots s}/D. \quad (7)$$

It is demonstrated that a close relation exists between variance decomposition and PCE coefficients, which allows for the calculation of Sobol indices directly from the PCE coefficients  $a_j$  without the need for additional sampling Sudret (2008). Therefore, the total costs of the GSA reduce to the calculation of the PCE training set.

### 2.3 Feature Clustering for BBCH Assignment

Agricultural fields are known to have spatial morphological heterogeneity due to growth competition. This condition is observed mostly in fields with broadcast seeding practices. Therefore, this structural heterogeneity is also expected for all growth stages at any time  $t$  in a field. The definition of the BBCH scale takes this heterogeneity into account and states that BBCH growth stage assignment must be done on the dominant morphology within the field Lancashire et al. (1991). In other words, the assigned BBCH value of a field has to represent at least 50% of all crops. Due to the same growth stage and similar morphology, crops are expected to have similar polarimetric scattering behaviors. Thus, to provide this requirement for the BBCH value assignment, the PolSAR data are clustered to obtain the smallest group with 50% of the samples using the well-known K-means algorithm in the space of statistically independent PolSAR parameters (i.e.,  $\sigma_{\text{HH}}^o$ ,  $\sigma_{\text{VV}}^o$  and  $\rho$ ). Details about the clustering methodology are given in Yuzugullu et al. (2015).

### 2.4 Parameter Space Search Algorithm

The proposed solution is designed as a constrained optimization problem by considering the ill-posed condition of the scattering model. In the literature, there are two different ways to handle similar optimization problems: deterministic and distribution-based approaches. The former approach converges to a single optimum value Smith (1995), whereas the latter converges to a distribution of values Dupacova (2013). In this case, because SAR data have high variance, mainly due to speckle noise, a method that converges to a single intensity value would be ineffective. Especially for deterministic methods, the presence of variance reduces the rate of convergence and increases the degree of classification error. On the other hand, distribution-based approaches are capable of handling problems with different levels of variance. The proposed parameter search space algorithm links the a priori IRRI growth phase (i.e., phase  $S$ ) and the backscattering model Toan et al. (1997). The proposed parameter space search approach follows the flow scheme given in Figure 4.

The method starts with the simulation of the parameter space using the growth-phase-specific  $\text{PCE}_{\text{EM}}$ . At this step, the growth phase,  $S$ , also determines the parametric range (min-max) of the morphological descriptors. The corresponding parameter space,  $\mathbb{P}^S$ , can be visualized as a hyper-grid (8). For any  $S$ , the coordinate of a single point in the grid has the information to define a rice canopy with morphology and structural density. However,

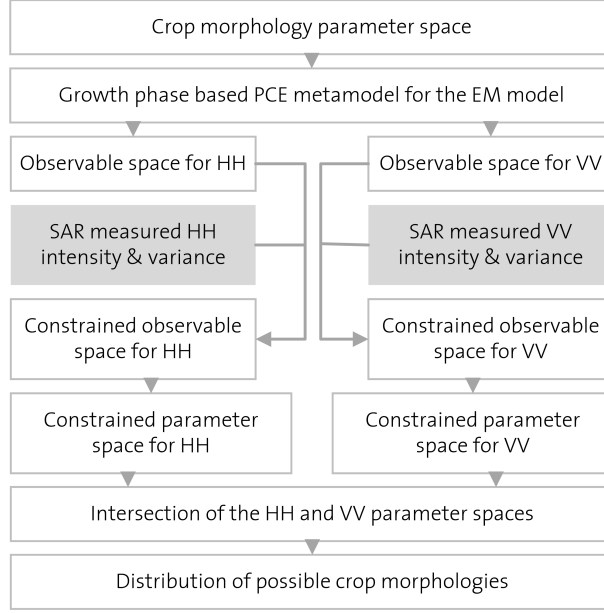


Figure 4: Block diagram of the proposed parameter space search algorithm.

the biologically-impossible structures present in the data cloud of the ground measurement database need to be eliminated. For this purpose, the  $\mathbb{P}^S$  is constrained using the convex-hull method based on the morphological growth information, which was collected from the literature and ground campaigns.

$$\mathbb{P}^S = (\vec{h}_s^S, \vec{d}_s^S, \vec{n}_t^S, \vec{l}_l^S, \vec{w}_l^S, \vec{n}_l^S, \vec{l}_p^S, \vec{w}_p^S, \vec{n}_p^S) \quad (8)$$

After preparing the  $\mathbb{P}^S$ , the  $\text{PCE}_{\text{EM}}$  is used to simulate  $\mathbb{P}^S$  and obtain the corresponding observable spaces for each polarimetric channel  $\mathbb{O}_{\text{qp}}^S$ , including the backscattering intensities. The fitness function of the distribution-based optimization problem in a constrained parameter space is given by:

$$\min \mathbb{Z}_\sigma = \frac{[E(\sigma_{\text{qp}} - \mathbb{O}_{\text{qp}}^S)]^2}{\text{Number of } qp \text{ combinations}} \quad (9)$$

In (9), the  $\sigma_{\text{qp}}$  and  $E(\bullet)$  represent the measured backscattering intensity and the expected value of the difference between measured and observable SAR intensities, respectively. In the proposed method, there are two constraints on the  $\mathbb{P}^S$ , which explain the relation between  $\mathbb{P}^S$  and  $\mathbb{O}_{\text{qp}}^S$ .

1. Backscattering intensity: Each  $\mathbb{O}_{\text{qp}}^S$  covers a wide range of intensities based on the corresponding morphologies in the  $\mathbb{P}^S$ . However, the intensity values obtained from the SAR data only cover a small range of  $\mathbb{O}_{\text{qp}}^S$ . In order to consider the spatial heterogeneity of the field, the mean ( $\mu_{\sigma_{\text{qp}}}$ ) and standard deviation ( $\nu_{\sigma_{\text{qp}}}$ ) of the measured backscattering intensities are calculated for each data cluster of polarimetric channels. Each  $\mathbb{O}_{\text{qp}}^S$  is

then bounded according to the intensity constraints given by (10).

$$\mathbb{C}_{qp} = \mathbb{O}_{qp}^S[\mu_{\sigma_{qp}} - 2\nu_{\sigma_{qp}}, \mu_{\sigma_{qp}} + 2\nu_{\sigma_{qp}}] \quad (10)$$

The confined observable space  $\mathbb{C}_{qp}$  has the same dimensionality as  $\mathbb{O}_{qp}^S$ , but fewer samples. The link between the  $\mathbb{O}_{qp}^S$  and  $\mathbb{P}^S$  is used to select the corresponding morphologies from the  $\mathbb{P}^S$  for each polarimetric channel to obtain the constrained parameter spaces,  $\mathbb{B}_{qp}$ . Thus, the morphological structures that are included in each  $\mathbb{B}_{qp}$  have a similar  $\sigma_{qp}$  with respect to the measured SAR intensity values.

2. Morphological consistency: In PolSAR data, a particular intensity value may correspond to different physical structures. This constraint resolves the ambiguity by taking the intersection of all  $\mathbb{B}_{qp}$  sets of each  $N$  polarimetric channels as seen in (11).

$$\mathbb{I} = \bigcap_{i=1}^N \mathbb{B}_{qp}^i = \mathbb{B}_{qp}^1 \cap \mathbb{B}_{qp}^2 \cap \mathbb{B}_{qp}^3 \cap \dots \quad (11)$$

The resulting set  $\mathbb{I}$  includes the multidimensional parameter distributions for the nine inputs in (8). The morphology vectors are kept intact to preserve the plant morphologies for the last step of the analysis.

## 2.5 Assignment of Growth Stages by $\text{PCE}_{\text{BBCH}}$

The last step of the proposed approach considers the sample distribution of resulting morphologies that are included in set  $\mathbb{I}$ . Since the BBCH stage is not physically measurable and strongly subjective, there is a need for a relation between morphological parameters and the BBCH stage. This link has a complex and non-linear behavior. Moreover, subjective decision criteria of the BBCH scale lead to a high degree of variation due to the variation in biophysical parameters. Therefore,  $\text{PCE}_{\text{BBCH}}$  is trained by taking samples from the morphological measurements as input,  $\mathbf{X}$ , on the corresponding BBCH stages as output,  $\mathbf{Y}$ . This metamodel is then used to estimate the BBCH stages ( $\text{BBCH}_{est}$ ) for the set of  $\mathbb{I}$ . Finally, the growth stage of the field is determined by calculating the mode of the distribution of  $\text{BBCH}_{est}$ .

## 3 Ground Campaign and SAR Data

### 3.1 Test Area and Ground Measurements

This study was carried out in two independent rice cultivation sites located in Spain and Turkey. Figure 5 shows the location of the fields. Both sites are sowed by the broadcast technique over the flooded ground. Figure 6 summarizes the timeline of the ground measurements and SAR acquisitions for both datasets with IRR phases.



Figure 5: Study areas with white framed test fields: (a) Isla Major, Spain, and (b) Ipsala, Turkey.

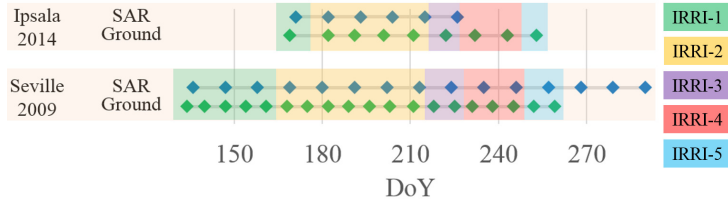


Figure 6: Accordance plot of SAR acquisition and ground measurement dates (day of year) given with color-coded IRRI stages for both test sites.

**Isla Major, Spain** The site is located in the Isla Major region, South of Seville, centered at  $N37^{\circ}7'53''$  and  $E6^{\circ}19'32''$ . The region has a flat topography and covers an area with an average radius of 20 km. The ground campaign was conducted in 2009 to measure the phenological stage, canopy height, plant and tiller density for the whole cultivation period (May–October). Figure 5 (a) presents the location of the test area and the fields that were chosen for the ground measurements.

**Ipsala, Turkey** The site is located in the Thrace region, North West of Istanbul, centered at  $N40^{\circ}47'59''$  and  $E26^{\circ}13'14''$ . The region has a flat topography and covers an area with an average radius of 15 km. The ground campaign was conducted in 2014 to measure the phenological stage, morphological parameters (stalk diameter and length, leaf width and length), plant, tiller and leaf density for the whole cultivation period (May–September). Figure 5 (b) presents the location of the test area and the fields that were chosen for ground measurements.

### 3.2 SAR Dataset

In this study, data from the TerraSAR-X (TSX) mission are used. It operates at a central frequency of 9.65 GHz with a wavelength of 31 mm. As an advantage to the other systems,

TSX allows frequent monitoring of environmental changes with a temporal resolution of 11 days. Therefore, it is one of the best options on the market for agriculture monitoring purposes.

All data were acquired in descending strip map mode and processed by the German Aerospace Center to the standard product Level 1b, i.e., single look complex (SLC) data (16-bit) with a  $\sim 2$ -m pixel-size. Later, the data were co-registered by using bi-linear interpolation with an average root mean squared (RMS) accuracy of 0.1 pixels. Before the analysis, multi-looking was applied on the data with a boxcar of  $11 \times 11$  pixels to reduce the speckle noise. Figure 6 presents the acquisition plan of the dual polarization (HH and VV) SAR data with a central incidence angle of  $31^\circ$  for both test sites.

## 4 Results and Discussion

In this paper, a stack of HH/VV dual-polarization descending TSX images over rice fields located in Spain and Turkey was employed to check the effectiveness of the proposed methodology. This section presents the GSA analysis of the EM backscattering model and the accuracy assessments of the growth stage determination algorithm. Since the algorithm has a stepwise scheme, the accuracy analysis is provided separately for each step. The discussions about the analysis outcomes are given in their specific sub-sections.

Figure 7 presents the boundaries of six crop morphology parameters from the Ipsala 2014 campaign with their quantiles and max-min values for each growth phase,  $S$ . For the  $\mathbb{P}^S$ , boundaries are further extended by 5% as a safety factor to consider morphological anomalies, such as over- or under-growth conditions.

### 4.1 Accuracy Assessment: Backscattering Model

In this study, the theoretical backscattering model estimates the HH and VV backscattering intensities of the rice canopies, at a central frequency of 9.65 GHz and at an average incidence angle of  $31^\circ$  to be consistent with the TSX beam. The effect of the variation in the incidence and look angle during the acquisitions is assumed as constant along the scene.

Before applying the proposed inversion-based classification, GSA was used to identify the parameters that most affect the backscattering coefficients; this step is known as model-reduction. During this step, the sensitivity of the backscattering coefficients to environmental parameters such as plant and ground dielectric values has been evaluated. The results are reported in Table 1. The variability in both the real and imaginary parts of the dielectric constants within the given boundaries ( $[22.0 + 6.0i \sim 30.0 + 10i]$  for the canopy and  $[60.0 + 15i \sim 80.0 + 25.0i]$  for the ground) has a significantly low impact on the model response. Indeed, the corresponding Sobol indices are much lower than, e.g., those of the stalk height parameter. Therefore, the dielectric constants were kept constant during the analysis by

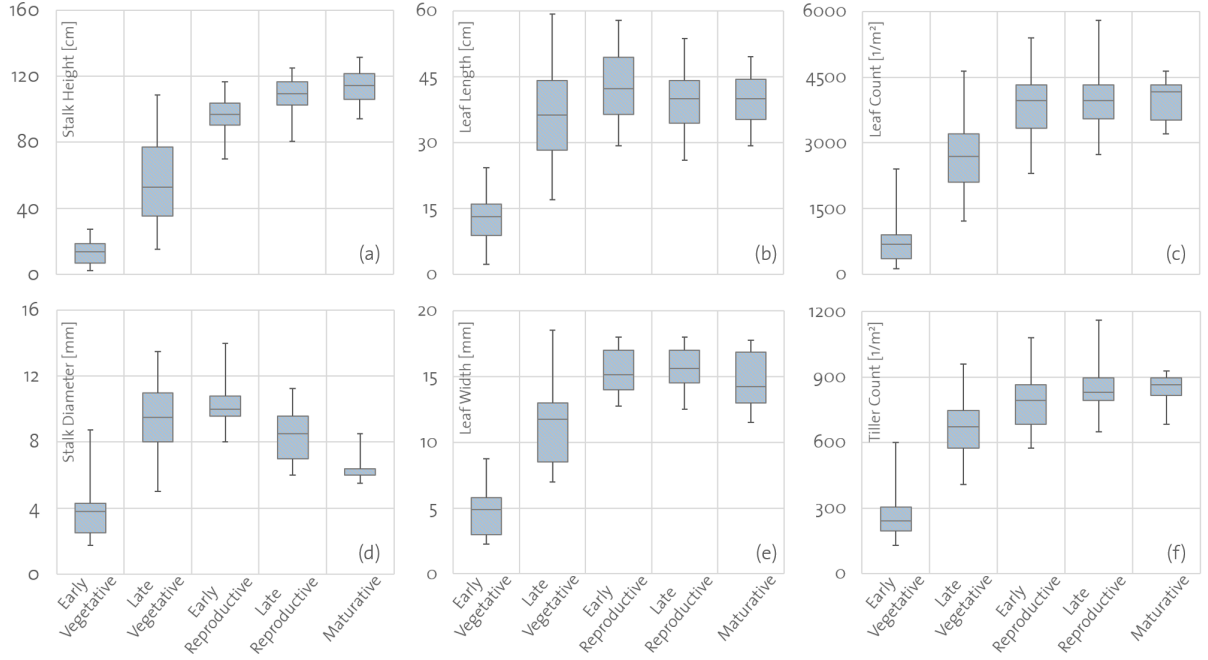


Figure 7: Variation of biophysical parameters in different growth phases. Data are obtained from the 2014 Ipsala ground campaigns. (a) Stalk height [cm], (b) Leaf length [cm], (c) Leaf count [1/m<sup>2</sup>], (d) Stalk diameter [mm], (e) Leaf width [mm], (f) Tiller count [1/m<sup>2</sup>]

setting them to values based on Wang et al. (2005).

Table 2 summarizes the values of the parameters assumed to be constant during the simulations of the ground measured data for the estimation of backscattering intensities. The reported values of the parameters are determined either from the SAR settings or the existing literature Toan et al. (1997). Constant parameters can represent the average properties of a rice canopy compared to real environmental conditions.

Figure 8 shows the correlation between the results of the theoretical morphology-based backscattering model from the simulation of the Ipsala ground measurements from 2014 and the acquired TSX data in dB. The results as the mean and standard deviation of the estimated values are grouped into three available growth phases in two polarimetric channels, i.e., HH and VV. In the corresponding figure, each growth phase is represented by a different color and symbol. Good agreement is obtained between the SAR measurements and the model simulations for both polarimetric channels.

There is clearly a strong correlation between measured and estimated backscattering intensity values for both polarimetric channels. For the full dataset, the 2D coefficient of determination ( $R^2$ ) and the root mean square error (RMSE) are calculated to be 87.1% and 2.02 dB for the HH channel and 84.6% and 1.91 dB for the VV channel. Additionally, when the growth phases are considered separately, for the HH channel, the RMSE values of the first three stages are calculated as 2.69, 1.83 and 1.96 dB, respectively. For the VV channel, they are calculated as 2.21, 1.91 and 1.58. The implementation of the backscattering model



Table 1: Calculated Sobol indices for the real and imaginary parts of the di-electric constant for plants and the underlying ground.

	E.Veg.		L.Veg.		E.Rep.		L.Rep.		Mat.	
	HH	VV	HH	VV	HH	VV	HH	VV	HH	VV
$h_{\text{stalk}}^r$	0.753	0.934	0.361	0.398	0.473	0.446	0.411	0.576	0.223	0.595
$\epsilon_{\text{stalk}}^r$	0.009	0.008	0.007	0.006	0.006	0.007	0.001	0.002	0.003	0.005
$\epsilon_{\text{stalk}}^i$	0.006	0.009	0.003	0.002	0.008	0.010	0.002	0.002	0.007	0.011
$\epsilon_{\text{leaf}}^r$	0.007	0.004	0.005	0.003	0.101	0.008	0.002	0.004	0.003	0.004
$\epsilon_{\text{leaf}}^i$	0.003	0.008	0.002	0.003	0.008	0.003	0.006	0.005	0.006	0.005
$\epsilon_{\text{panicle}}^r$	-	-	-	-	-	-	0.011	0.009	0.012	0.008
$\epsilon_{\text{panicle}}^i$	-	-	-	-	-	-	0.014	0.012	0.009	0.010
$\epsilon_{\text{ground}}^r$	0.112	0.069	0.074	0.053	0.009	0.007	0.008	0.005	0.038	0.013
$\epsilon_{\text{ground}}^i$	0.124	0.085	0.092	0.069	0.013	0.014	0.011	0.009	0.024	0.007

does not include the underlying surface nor the morphological 3D orientation information. The underlying surface is assumed to be water all through the cultivation cycle with high dielectric constant. The 3D orientation of the morphological components is neglected because the model has been shown to provide reasonable accuracy even without their inclusion (see Figure 8). Including the rotational parameters for each plant would result in additional scatter of the EM response, comparable to the effect of the stochastic placement of the plants in the area described in Section 2.1.

## 4.2 PCE<sub>EM</sub> and Global Sensitivity Analysis

A PCE<sub>EM</sub> is generated by preparing a sample of size 2000 of the full model for each of the five growth stages identified previously, hence at a total training cost of 10,000 model evaluations. Note that this is a one-time cost: after the PCE<sub>EM</sub> is trained, no new full model evaluations are necessary to evaluate the PCE<sub>EM</sub> on new sets of morphological parameters. The PCE<sub>EM</sub> surrogates the mean and the variance of the forward backscattering model of the MC simulations in all polarimetric channels.

Regarding computational costs, the evaluation of the PCE<sub>EM</sub> is comparatively inexpensive. In other words, while the original implementation of the backscattering model required approximately 22 hours to calculate the response to 2000 simulations on a computer with 24 GB RAM and 8 cores, the PCE<sub>EM</sub> needed only 0.04 seconds with a single core on the same hardware. Therefore, such an improvement allows increasing the size of the parameter

Table 2: Input parameters that are kept constant for the backscattering model evaluations.

Parameter	Value
Central frequency	9.65 GHz
Dielectric constant ( $\epsilon_{s,l}$ )	$25 + 8j$
Dielectric constant ( $\epsilon_g$ )	$70 + 20j$
Average incidence angle ( $\theta$ )	$31^\circ$
Look angle	$90^\circ$
Distance to target	514 km
Illuminated area x-size	2.58 m
Illuminated area y-size	1.79 m
Number of MC iterations	200

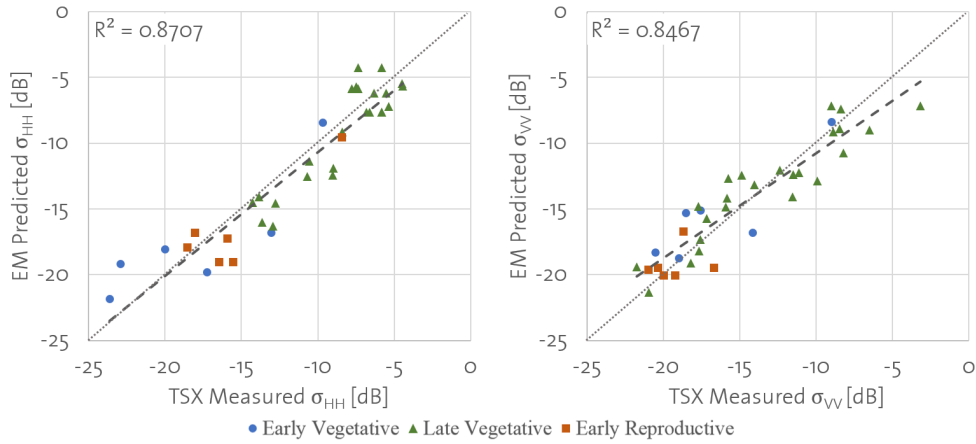


Figure 8: The TSX measured versus the theoretical backscattering model (1) predicted HH and VV channel backscattering intensities from the Ipsala 2014 campaign.

and the observable spaces significantly, which in turn enhances the variation in the crop morphology input vectors.

Figure 9 shows the results of the accuracy and GSA of the  $PCE_{EM}$ . The figure is structured as an array with the first two rows visualizing the accuracy analysis and the last row visualizing the GSA results. Each column corresponds to a growth phase with all chosen crop morphological parameters. Accuracy analysis of the  $PCE_{EM}$  is given with a corresponding polynomial degree (P.Deg.),  $R^2$ , RMSE and Leave-One-Out (LOO) error Blatman and Sudret (2010). The results are discussed below in detail for each growth phase.

Early vegetative: For both polarimetric channels, the stage-specific  $PCE_{EM}$  can approximate the backscattering coefficients perfectly. For HH and VV channels, the  $R^2$  values are calculated to be 99.4% and 98.3%, respectively. The estimated RMSE values are 0.25 dB

and 0.34 dB for HH and VV channels, respectively. The GSA of the theoretical model shows that stalk height is the primary source of the variance in the model output. Besides, the sensitivity to the variation in stalk diameter is observed to be stronger in the HH channel.

Late vegetative: During this stage, the significant growth in the plants increases the dynamic range of the intensity values in both polarimetric channels. This variance is also detected in the stage-specific  $PCE_{EM}$  outputs. The results of the accuracy analysis show that  $R^2$  and RMSE values are calculated to be 89.2% and 1.78 dB for HH and 83.5% and 1.93 dB for the VV channel. GSA shows that the major source of the variance in the model output originates from the stalk height, stalk diameters and the number of tillers. In addition, the HH channel is slightly more sensitive to stalk density compared to the VV channel.

Early reproductive: As the plant enters this phase, head leaves and panicles are observed. The accuracy assessment of the  $PCE_{EM}$  reports the  $R^2$  and RMSE for the HH channel as 89.1% and 0.94 dB and the VV channel as 80.0% and 0.97 dB. Concerning GSA, the model is observed to be sensitive to stalk height in both polarimetric channels. Additionally, the HH channel is sensitive to the changes in the number of tillers. On the other hand, the VV channel is found to be sensitive to the variation in panicle width and number of panicles.

Late reproductive: For each polarimetric channel, the accuracy assessment of the stage-specific  $PCE_{EM}$  provides  $R^2$  and RMSE values as 81.9% and 0.95 dB for the HH channel and 89.9% and 0.56 dB for the VV channel. For the GSA, the source of the model variability is related to the changes in stalk height for both polarimetric channels. Furthermore, the number of tillers and the number of panicles are other sources of variability for the HH and VV channels, respectively.

Maturative: During the last stage of the growth cycle, the accuracy of the stage-specific  $PCE_{EM}$  is estimated for  $R^2$  and RMSE values as 84.4% and 0.96 dB for HH and 89.8% and 0.58 dB for the VV channel. Moreover, the sources of the variation in the model outputs are found to be stalk height for HH and VV and the number of tillers for HH.

To sum up, growth-phase-specific  $PCE_{EM}$  can estimate the outputs of the theoretical backscattering model with high accuracy. Minimum  $R^2$  and maximum RMSE values for the full cycle are calculated to be 80.0% and 1.98 dB. Therefore, the replacement of the backscattering model with the surrogate  $PCE_{EM}$  is acceptable. While the highest accuracy is observed in the early vegetative stage, the lowest accuracy is in the late vegetative stage due to ranges of corresponding parameter spaces as shown in Figure 7. GSA shows that throughout the growth cycle, model outputs in both polarimetric channels (HH and VV) are most sensitive to the stalk height. However, for the HH channel, the number of tillers

and stalks become important starting from the late vegetative phase due to their effect in increasing the attenuation inside the canopy.

### 4.3 Structures of the Parameter and Observable Spaces

The proposed approach follows a search algorithm that depends on two multi-dimensional spaces, parameter  $\mathbb{P}$  and observable  $\mathbb{O}_{qp}$ . The first space is built up based on the morphological parameters as a regular grid. In the current case, the increments of the grid were chosen as 1 cm for stalk height and leaf length, 1 mm for stalk diameter and leaf width, 1 unit for tiller and leaf number and, finally, 10 units for plant number. For each growth phase  $\mathbb{P}$ , a different number of possible samples is obtained. Table 3 summarizes the remaining sample sizes throughout the analysis. The reason behind the different number of samples is due to the varying ranges (maximum and minimum values) of parameters as seen in Figure 7. Later, when the biologically unrealistic morphologies are eliminated in accordance to the crop morphology database, the sample sizes reduce to the values shown in Row 3 of Table 3. Here, it is observed that several morphologies in the parameter space are not biologically favored. In the next step, intensity and matching morphology constraints are implemented. Table 3 provides the average values from this study for the minimum and the maximum number of samples remaining after each constraint is applied. These values can change based on the variance of the data, which is either due to the structural heterogeneity of the region or due to the size of the smoothing window. Finally, the remaining samples are used as an input to the  $\text{PCE}_{\text{BBCH}}$ .

Table 3: An average sample size of parameter space during each step of the search algorithm procedure.

Process Step		Growth Phase				
		1	2	3	4	5
Sample Size	$\mathbb{P}$ Space	48300	1492920	204160	338328	181350
	Pos. Morp.	7010	343800	47330	120780	26330
	Cons. 1: $\mathbb{B}$	2000-2500	25000-40000	8000-11000	17000-23000	4000-7000
	Cons. 2: $\mathbb{I}$	1500-2200	12000-30000	4000-10000	9000-20000	2000-6000

### 4.4 Accuracy Assessment: $\text{PCE}_{\text{BBCH}}$

A random growth stage can correspond to several different plant morphologies. Besides, the variance of the crop morphology can affect the BBCH results. Therefore, this relation is achieved using a PCE metamodel that relates ground measurements to their corresponding

BBCH stages by  $PCE_{BBCH}$ .

Figure 10 shows the results of the accuracy analysis of the  $PCE_{BBCH}$ . The plot is given for the BBCH stages measured in the field versus the ones estimated by the  $PCE_{BBCH}$ . For the training, 200 randomly chosen ground measurements are taken into consideration from the 2014 Ipsala ground campaign. The coefficient of determination is calculated to be 94.0%. The overall RMSE value is found to be 5.80 stages, with the lowest variance in the early vegetative stage and the highest in the maturative stage. This can be explained by the degree of the morphological complexity. In other words, as the structure gets more complicated, a lower accuracy for the  $PCE_{BBCH}$  is observed. Lastly, the proposed scheme can provide a continuous growth trend in the BBCH scale.

#### 4.5 Accuracy Assessment: BBCH Assignment

The growth stage determination method proposed in this study makes the BBCH scale directly available for broadcast seeded rice monitoring using PolSAR. In other words, the phenological stage of a rice field of interest can be estimated by observing its polarimetric response and that of the surrounding area. Therefore, to prove the consistency of the approach, the PCE metamodel-based inversion method is independently applied to each test field present in each TSX acquisition from Isla Major and Ipsala.

The accuracy analysis through correlation plots of the proposed algorithm applied to all test fields of the available TSX data is shown in Figure 11a,b for the Isla Major and the Ipsala sites, respectively. The value of  $R^2$  between ground measured and estimated BBCH is 94.1% for the Isla Major and 84.1% for the Ipsala test sites. Additionally, the RMSE deviation from the measured value is found to be 7.66 BBCH stages for Isla Major and 5.24 BBCH stages for Ipsala data. Unfortunately, the Ipsala data are not available for the full cycle. The analysis shows that, while the proposed algorithm tends to overestimate the earlier stages, it underestimates the later stages. This can be explained by crop morphology and subjective assignment of the BBCH stages. The inclusion of the  $PCE_{BBCH}$ -based growth stage assignment improves the overall accuracy. Field- and full-scale growth maps are visualized in Figures 12 and 13, respectively.

## 5 Conclusions

This paper has demonstrated that X-band HH/VV dual-polarization SAR data are suitable for the estimation of flooded and broadcast-sowed rice field growth stages on a continuous scale, in terms of BBCH. This is due to the sensitivity of the X-band polarimetric descriptors to small-scale morphological changes. The validation of the proposed approach carried out at the field level provided an error of less than 10 BBCH stages. Additionally, the  $R^2$  between the ground measurements and the algorithm estimation is found to be consistently higher

than 80.0%.

Since the proposed methodology gives promising results, it may encourage agriculturists and local authorities to use products based on SAR data for their monitoring purposes. The main strengths, limitations and opportunities of the proposed methodology are:

## 5.1 Strengths

- The algorithm depends on the rice crop morphology. The proposed approach extends the usage of existing classification algorithms. The results of the current classification algorithms can be used instead of the a priori growth phase information as a coarse classifier. The proposed method introduces the  $PCE_{EM}$ -based parameter search space approach, resulting in an estimate of the BBCH based on crop morphology.
- Several genotype variations are available for rice crops. The range of admissible morphological parameters (e.g., crop height vs. leave size) may, therefore, need to be extended should data on new/additional crop morphologies become available. The proposed method can easily be updated automatically with each new crop morphology dataset by appropriately extending the allowed morphological parameter space. Therefore, each new dataset will contribute to the preservation of the plant morphological growth principles for different genotypes. The possibility to extend the base morphological datasets allows the proposed approach to be extended to include new morphologies.
- The proposed method can make detection of in-field heterogeneities possible for observing growth abnormalities. The included feature clustering approach handles polarimetrically similar regions of the field separately, and therefore, spatially-localized problems (e.g., sickness or overgrowth) can be handled, unless they have statistically a representative number of samples.

## 5.2 Limitations

- Even though the results are promising, some aspects were omitted in the chosen backscattering model such as the 3D orientation of the scatterers, the curvature of the leaves and panicles and the agronomical exceptions as extreme water loss from the plants. Besides, according to the Directorate of Trakya Agricultural Research Institute, the rice fields located in Turkey are kept flooded until 10–15 days before harvesting. Therefore, the current implementation of the model only considers the flooded conditions and misses the non-flooded periods.
- The performance of the morphology estimation strongly depends on the performance of the backscattering model and the environmental conditions. The slight bias in the EM model predictions that can be observed in Figure 8 may be related to the

slight bias in the reconstruction response w.r.t. the ground truth in Figure 11a,b. A quantitative study of the effects of model bias on the inversion results would require additional high-quality ground truth measurements. Nevertheless, it is expected that improvements in the model predictivity, especially when effective at reducing model bias, could similarly improve the accuracy of the inversion results.

- Since the proposed approach was developed for fields with flooded or strongly moist underlying surfaces, further studies are needed to assess its applicability for fields with dry or slightly moist soil.

### 5.3 Opportunities

- The chosen theoretical backscattering model can be replaced by any other morphology-based EM backscattering model. The alternative models may lead to higher accuracies with a higher number of parameters. However, the uncertainties of the inputs should also be taken into account. Therefore, it is possible to state that, for an improvement in the inversion accuracy, the alternative models should have lower variance in their outputs, which can be achieved by inclusion of the cross-polarimetric channels (HV and VH). Additionally, the proposed approach is also applicable to the monitoring of different crop types by simulating their morphology and the underlying ground information with the theoretical EM backscattering model.
- With the inclusion of the metamodels, the computational cost of the inversion algorithms decreases significantly. This may lead to the development and integration of new backscattering models with realistic crop morphology.

Future work will focus on the evaluation of the proposed methodology using different frequencies, crop types, as well as incidence angles. Ongoing and future missions such as Tandem-L and Sentinel-1 will be excellent opportunities for these evaluations.

## Acknowledgments

This work has been supported by the Scientific and Technological Research Council of Turkey (TUBITAK) under Project 113Y446 and by the German Aerospace Center (DLR) under project XTILAND1476. Additionally, the authors would like to thank the Federacion de Arroceros de Sevilla and the Directorate of Trakya Agricultural Research Institute for providing the ground measurements from Spain and Turkey, respectively.

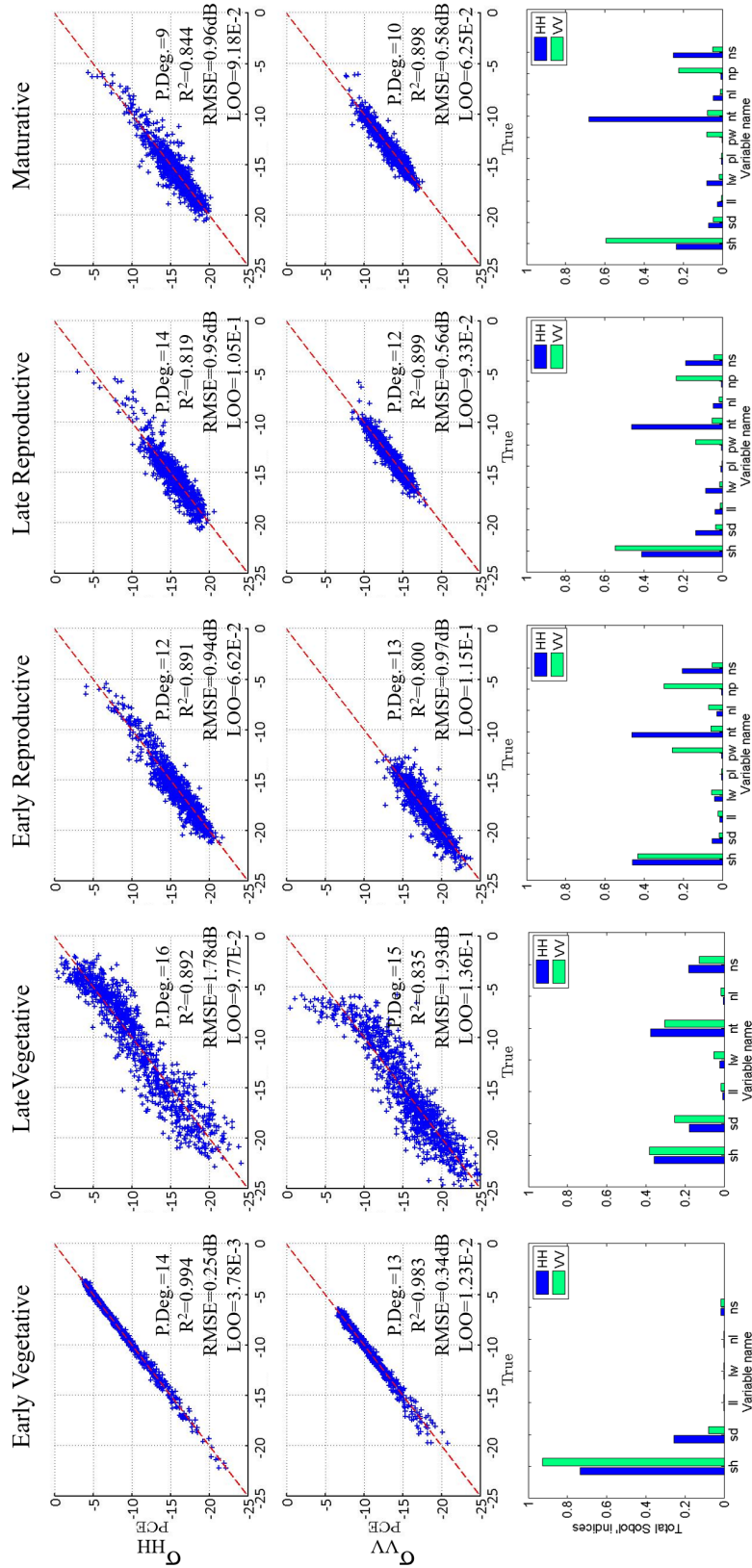


Figure 9: [Row 1&2] Relationship between backscattering model and PCE<sub>EM</sub> simulated  $\sigma^o$  (in dB) values given for five growth stages with corresponding polynomial degree (P.Deg.),  $R^2$ , RMSE and LOO values. True value indicates the scattering model simulated values. [Row 3] GSA results given for five growth stages with corresponding Total Sobol' indices for each input parameter.



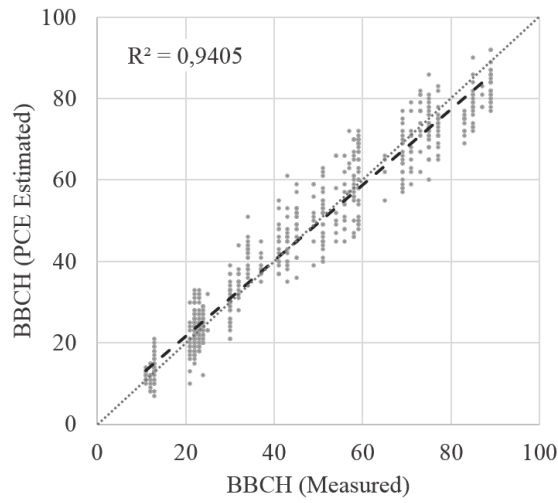


Figure 10: Relationship between ground measured and  $PCE_{BBCH}$  simulated BBCH values with corresponding  $R^2$  for the training data.

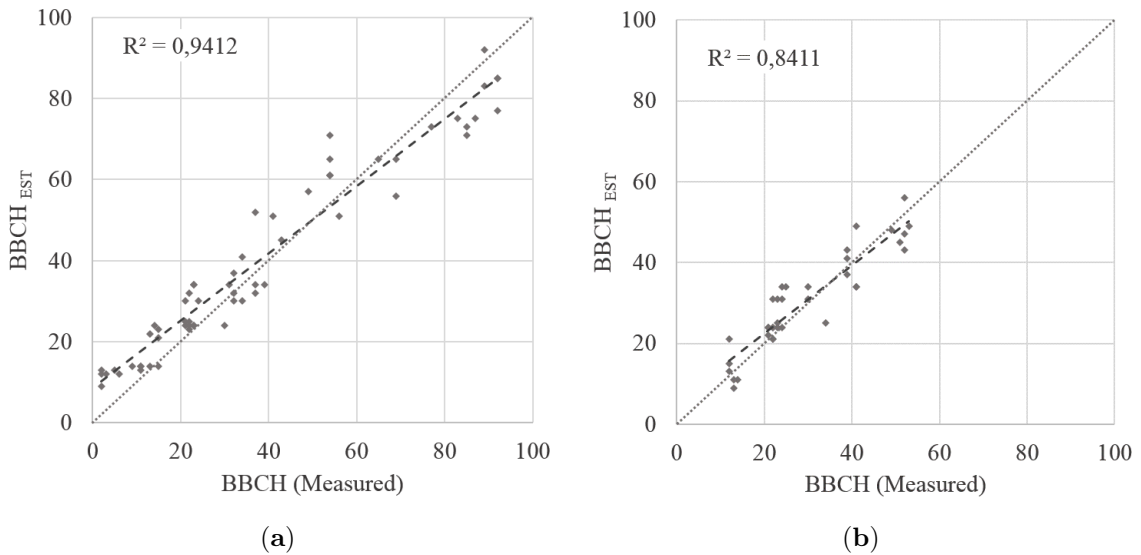


Figure 11: Relationship between ground measured and algorithm estimated BBCH stages with  $R^2$  values. (a) Isla Major; (b) Ipsala.

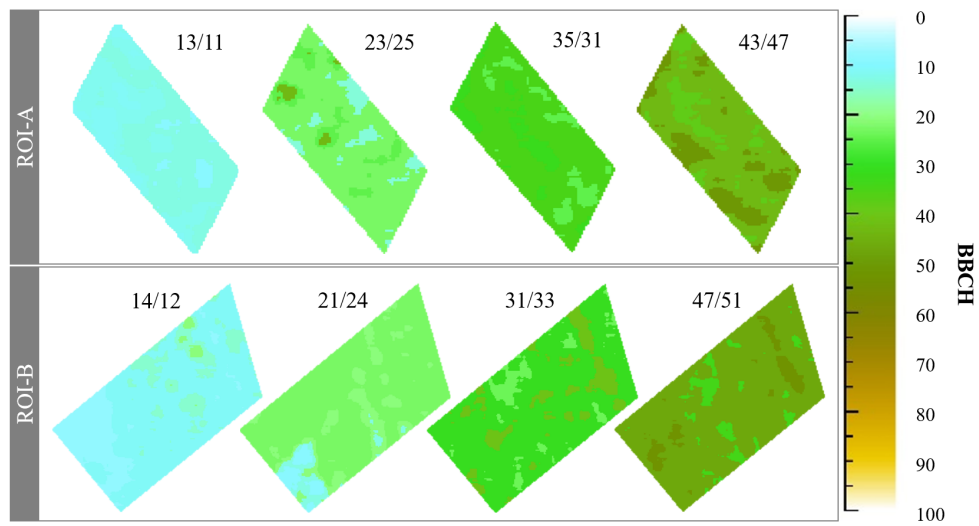


Figure 12: Phenological stage estimation results of the proposed algorithm obtained over two different Region Of Interest (ROI) located in the Ipsala 2014 dataset. The growth stages are given as estimated/measured BBCH stage.

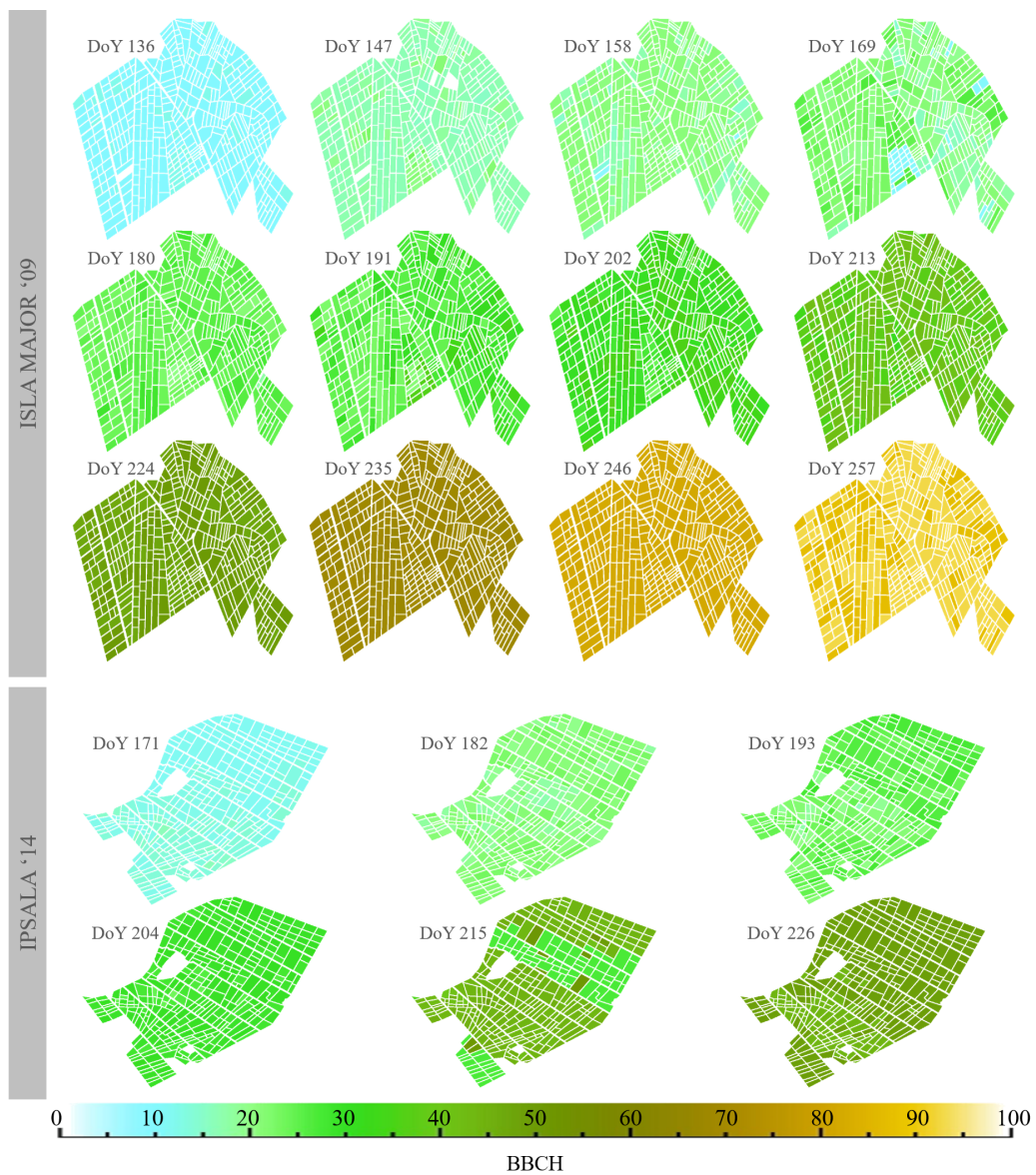


Figure 13: Growth maps with the phenological stage estimation with the BBCH scale in two different areas exploiting their temporal behavior. The date of the images is given as Day of Year (DoY).

## References

- Atzberger, C. (2013, feb). Advances in Remote Sensing of Agriculture: Context Description, Existing Operational Monitoring Systems and Major Information Needs. *Remote Sensing* 5(2), 949–981.
- Blatman, G. and B. Sudret (2010, nov). Efficient computation of global sensitivity indices using sparse polynomial chaos expansions. *Reliability Engineering & System Safety* 95(11), 1216–1229.
- Bouman, B. (1991, sep). Crop parameter estimation from ground-based X-band (3-cm wave) radar backscattering data. *Remote Sensing of Environment* 37(3), 193–205.
- Bouman, B., D. van Kraalingen, W. Stol, and H. van Leeuwen (1999, feb). An agroecological modeling approach to explain ERS SAR radar backscatter of agricultural crops. *Remote Sensing of Environment* 67(2), 137–146.
- Dupacova, J. (2013). *Stochastic Optimization: Algorithms and Applications*. Springer US.
- Erten, E., J. M. Lopez-Sanchez, O. Yuzugullu, and I. Hajnsek (2017, dec). Retrieval of agricultural crop height from space: A comparison of SAR techniques. *Remote Sensing of Environment* 187, 130–144.
- Erten, E., C. Rossi, and O. Yuzugullu (2015, jul). Polarization impact in TanDEM-X data over vertical-oriented vegetation: The paddy-rice case study. *IEEE Geoscience and Remote Sensing Letters* 12(7), 1501–1505.
- GRI SP (Global Rice Science Partnership) (2011). *Rice Almanac* (4th edition ed.). Los Banos (Philippines): International Rice Research Institute.
- Inoue, Y., T. Kurosui, H. Maeno, S. Uratsuka, T. Kozu, K. Dabrowska-Zielinska, and J. Qi (2002). 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 Sensing of Environment* 81(2-3), 194–204.
- Inoue, Y. and E. Sakaiya (2013, mar). Relationship between X-band backscattering coefficients from high-resolution satellite SAR and biophysical variables in paddy rice. *Remote Sensing Letters* 4(3), 288–295.
- Inoue, Y., E. Sakaiya, and C. Wang (2014, jan). Capability of C-band backscattering coefficients from high-resolution satellite SAR sensors to assess biophysical variables in paddy rice. *Remote Sensing of Environment* 140, 257–266.
- Karam, M. and A. Fung (1989, nov). Leaf-shape effects in electromagnetic wave scattering from vegetation. *IEEE Transactions on Geoscience and Remote Sensing* 27(6), 687–697.

- Karam, M., A. Fung, and Y. Antar (1988). Electromagnetic wave scattering from some vegetation samples. *IEEE Transactions on Geoscience and Remote Sensing* 26(6), 799–808.
- Karam, M. A. and A. K. Fung (1988, jun). Electromagnetic scattering from a layer of finite length, randomly oriented, dielectric, circular cylinders over a rough interface with application to vegetation. *International Journal of Remote Sensing* 9(6), 1109–1134.
- Koay, J.-Y., C.-P. Tan, K.-S. Lim, S. B. bin Abu Bakar, H.-T. Ewe, H.-T. Chuah, and J.-A. Kong (2007, sep). Paddy fields as electrically dense media: Theoretical modeling and measurement comparisons. *IEEE Transactions on Geoscience and Remote Sensing* 45(9), 2837–2849.
- Koppe, W., M. L. Gnyp, C. Hutt, Y. Yao, Y. Miao, X. Chen, and G. Bareth (2013, apr). Rice monitoring with multi-temporal and dual-polarimetric TerraSAR-x data. *International Journal of Applied Earth Observation and Geoinformation* 21, 568–576.
- Lancashire, P. D., H. Bleiholder, T. V. D. Boom, P. Langeluddeke, R. Stauss, E. Weber, and A. Witzemberger (1991, dec). A uniform decimal code for growth stages of crops and weeds. *Annals of Applied Biology* 119(3), 561–601.
- Lim, K.-S., C.-P. Tan, J. Y. Koay, V. Koo, H. T. Ewe, Y.-C. Lo, and A. Ali (2007). Multitemporal C-Band Radar Measurement on Rice Fields. *PIERS Online* 3(1), 44–47.
- Lin, Y.-C. and K. Sarabandi (1999, may). Retrieval of forest parameters using a fractal-based coherent scattering model and a genetic algorithm. *IEEE Transactions on Geoscience and Remote Sensing* 37(3), 1415–1424.
- Lopez-Sanchez, J. M., S. R. Cloude, and J. D. Ballester-Berman (2012, jul). Rice Phenology Monitoring by Means of SAR Polarimetry at X-Band. *IEEE Transactions on Geoscience and Remote Sensing* 50(7), 2695–2709.
- Lopez-Sanchez, J. M., F. Vicente-Guijalba, J. D. Ballester-Berman, and S. R. Cloude (2014, may). Polarimetric Response of Rice Fields at C-Band: Analysis and Phenology Retrieval. *IEEE Transactions on Geoscience and Remote Sensing* 52(5), 2977–2993.
- Marelli, S. and B. Sudret (2014). *UQLab: A framework for uncertainty quantification in Matlab*, Chapter 257, pp. 2554–2563.
- Marelli, S. and B. Sudret (2015). UQLab user manual - Polynomial Chaos Expansion. Technical report, Chair of Risk, Safety & Uncertainty Quantification, ETH Zurich. Report - UQLab-V0.9-104.
- Mulla, D. J. (2013, apr). Twenty five years of remote sensing in precision agriculture: Key advances and remaining knowledge gaps. *Biosystems Engineering* 114(4), 358–371.

- Ribbes, F. (1999, jan). Rice field mapping and monitoring with RADARSAT data. *International Journal of Remote Sensing* 20(4), 745–765.
- Rossi, C. and E. Erten (2015, feb). Paddy-Rice Monitoring Using TanDEM-X. *IEEE Transactions on Geoscience and Remote Sensing* 53(2), 900–910.
- Sakamoto, T., M. Yokozawa, H. Toritani, M. Shibayama, N. Ishitsuka, and H. Ohno (2005, jun). A crop phenology detection method using time-series MODIS data. *Remote Sensing of Environment* 96(3-4), 366–374.
- Shao, Y., X. Fan, H. Liu, J. Xiao, S. Ross, B. Brisco, R. Brown, and G. Staples (2001, jun). Rice monitoring and production estimation using multitemporal RADARSAT. *Remote Sensing of Environment* 76(3), 310–325.
- Smith, D. K. (1995, sep). Operations Research: Deterministic Optimization Models. *Journal of the Operational Research Society* 46(9), 1154–1155.
- Sobol, I. M. (1993). Sensitivity estimates for nonlinear mathematical models. *Mathematical modelling and computational experiment*.
- Stiles, J. and K. Sarabandi (2000). Electromagnetic scattering from grassland. i. a fully phase-coherent scattering model. *IEEE Transactions on Geoscience and Remote Sensing* 38(1), 339–348.
- Sudret, B. (2008, jul). Global sensitivity analysis using polynomial chaos expansions. *Reliability Engineering & System Safety* 93(7), 964–979.
- Thomas J. Santner, Brian J. Williams, W. I. N. (2013). *Design and Analysis of Computer Experiments*. Springer New York.
- Toan, T. L., F. Ribbes, L.-F. Wang, N. Floury, K.-H. Ding, J. A. Kong, M. Fujita, and T. Kurosu (1997). Rice crop mapping and monitoring using ERS-1 data based on experiment and modeling results. *IEEE Transactions on Geoscience and Remote Sensing* 35(1), 41–56.
- Tsang, L., K.-H. Ding, G. Zhang, C. Hsu, and J. A. Kong (1995, may). Backscattering enhancement and clustering effects of randomly distributed dielectric cylinders overlying a dielectric half space based on Monte-Carlo simulations. *IEEE Transactions on Antennas and Propagation* 43(5), 488–499.
- Ulaby, F., C. Allen, G. Eger, and E. Kanemasu (1984, jan). Relating the microwave backscattering coefficient to leaf area index. *Remote Sensing of Environment* 14(1-3), 113–133.
- Vapnik, V. N. (1995). *The nature of statistical learning theory*. Springer-Verlag.

- Vicente-Guijalba, F., T. Martinez-Marin, and J. M. Lopez-Sanchez (2014, jun). Crop Phenology Estimation Using a Multitemporal Model and a Kalman Filtering Strategy. *IEEE Geoscience and Remote Sensing Letters* 11(6), 1081–1085.
- Wang, L., J. Kong, K. Ding, T. Le Toan, F. Ribbes, and N. Floury (2005). Electromagnetic scattering model for rice canopy based on Monte Carlo simulation. *Progress In Electromagnetics Research* 52, 153–171.
- Yonezawa, C., M. Negishi, K. Azuma, M. Watanabe, N. Ishitsuka, S. Ogawa, and G. Saito (2012, sep). Growth monitoring and classification of rice fields using multitemporal RADARSAT-2 full-polarimetric data. *International Journal of Remote Sensing* 33(18), 5696–5711.
- Yuzugullu, O., E. Erten, and I. Hajnsek (2015, jun). Rice growth monitoring by means of X-band co-polar SAR: Feature clustering and BBCH scale. *IEEE Geoscience and Remote Sensing Letters* 12(6), 1218–1222.
- Yuzugullu, O., E. Erten, and I. Hajnsek (2017, jan). Estimation of rice crop height from x- and c-band PolSAR by metamodel-based optimization. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing* 10(1), 194–204.