Estimation of Key Dates and Stages in Rice Crops Using Dual-Polarization SAR Time Series and a Particle Filtering Approach

Caleb G. De Bernardis, Fernando Vicente-Guijalba, Tomas Martinez-Marin,
and Juan M. Lopez-Sanchez, *Senior Member, IEEE*

Abstract—Information of crop phenology is essential for evaluating crop productivity. In a previous work, we determined phenological stages with remote sensing data using a dynamic system framework and an extended Kalman filter (EKF) approach. In this paper, we demonstrate that the particle filter is a more reliable method to infer any phenological stage compared to the EKF. The improvements achieved with this approach are discussed. In addition, this methodology enables the estimation of key cultivation dates, thus providing a practical product for many applications. The dates of some important stages, as the sowing date and the day when the crop reaches the panicle initiation stage, have been chosen to show the potential of this technique.

Index Terms—Agriculture, multitemporal, particle filter, phenology, polarimetry, rice, synthetic aperture radar (SAR).

I. INTRODUCTION

RICE is one of the most important sources of food in the world, Asia being the largest producer with increasing importance in Africa and Latin America as well as pockets of production in Australia, Europe, and the U.S. To ensure a maximum yield, it is necessary to keep a continuous monitoring over fields. This would allow producers to have an accurate knowledge on the crop status and to apply a correct treatment at the precise moments. Phenology represents a measurement of crop evolution and it can be used as a control variable by the farmers [1], [2]. Traditionally, it is measured by means of visual inspection on ground but, due to clear limitations, different alternative ways of monitoring appear, such as those based on remote sensing satellite images [2]–[4].

In [5], the phenological stage estimation problem was treated in a dynamic context. A sequence of synthetic aperture radar (SAR) images, acquired by the German TerraSAR-X sensor at X-band, was used as input data to deduce the phenological stage. The methodology consisted of: 1) the generation of a transition model of the phenological evolution and 2) the

application of the extended Kalman filter (EKF) [6] approach to combine the observations and the derived model for estimation purposes. This filtering method uses a linearization of the transitions and assumes that the probability density function (pdf) of both observation and transition models are described by a Gaussian distribution in order to provide optimal estimations. However, the model is strongly nonlinear and the observations do not exhibit exactly a Gaussian distribution. For these reasons, another filtering method able to work under these conditions needs to be considered. The particle filtering (PF) approach [7] is proposed here because the estimation procedure is not affected by these limitations. It is a sequential Monte Carlo [8], [9] method based on approximating the posterior pdf of the state vector, based on all available information, by a set of samples (or "particles"). When a new input data is available the pdf given by the model is combined with the pdf of the observation to estimate the most likely state. Hence, the PF is proposed here as a convenient method for obtaining more accurate estimations of phenological stages than the EKF.

From the application point of view, different works in the literature are based on the analysis of normalized difference vegetation index (NDVI) curves. Phenological parameters are derived in correspondence to inflection points or local minimum/maximum points of the temporal signal [10], [11]. Despite the reduced observation space (only one observable), the time coordinate is exploited to estimate some stages, such as vegetation green-up or start of season, end of season, begin of brown-down, and end of brown-down. In our work, the transition model is presented in a state space defined by a set of variables (a detailed explanation is given in Section III) providing a continuous representation for the phenological stages. Therefore, the estimation of any date is feasible with this methodology.

This approach enables estimating the date of critical stages that are strongly related with the productivity, as the panicle initiation stage or the sowing date [12]. A tight control of nutrient inputs, especially nitrogen (N) fertilizer [13], is one of the key aspects in the improvement of yields in the world. While N fertilization maximizes grain yield, an overuse may actually have the opposite effect [14], [15]. For this reason, N fertilization is stopped at panicle initiation stage. The accurate prediction of a particular crop reaching this stage will enable to optimize production and reduce the impacts on the environment, which is becoming a serious issue as it happens in Asia [16]. Moreover, the access to an accurate sowing date estimate will provide the

required information to governments or relevant agencies to reduce the increase in pests and soil damages. Pests affect the crop yield and force farmers to use toxic products to combat the problem, and it can be dangerous to consumers. For instance, an alternative technique for breaking the life cycle of the insect pest is presented in [17]. The goal is to synchronize the planting date of crops in large areas. However, this study concludes with the need of a legislation control in the cycle of sowing to make it effective. On the other hand, dependence of yield with the sowing date is shown in [12] and [18], so the knowledge of sowing dates can be used to establish when is the optimum date to plant in certain area.

In this paper, the estimation of the current phenological stage of rice crops is addressed first, showing the improvements achieved by the PF instead of the EKF. In second place, this method provides a solution to predict or estimate any key date of the cultivation cycle. Due to their relevance, as described in the previous paragraph, the sowing date and the panicle initiation stage date have been chosen to test it.

An introduction on the PF theory is described in Section II, followed by the description of the methodology proposed for this particular problem in Section III. Then, results are presented in Section IV. Finally, conclusion and a discussion are addressed in Section V.

II. THEORY OF PARTICLE FILTER

Particle filter is a method based on Monte Carlo [8], [9] and recursive Bayesian sequential estimation [19], [20] that is very suitable for nonlinear and/or non-Gaussian applications [21]–[23].

Systems that have a dynamic behavior can be described mathematically by a set of inputs, outputs, and variables, using what is known as a state space [24]. The state-space approach is convenient when the process is nonlinear and non-Gaussian instead of using traditional time-series techniques [25]. The relevant information about a dynamic process is represented in an N-dimensional space. The state vector defines the position in the state space of the process at a precise moment. For example, in navigation tracking problems, the information could be defined by the distances to the boundaries and the orientation angle. Data acquired by noisy measurements define the resulting state vector and introduce an uncertainty, forcing us to work in statistical terms. Therefore, filtering techniques are necessary to obtain the optimal estimation.

The goal of Bayesian sequential estimation is to construct the posterior pdf, to know which is the most likely state, in a recursive way, by separating the process in two stages: prediction and update.

Prediction: The Chapman–Kolmogorov integral equation can be used to generate the prior pdf, if we have defined the transition model $p(X_k|X_{k-1})$, which will be explained in Section III-A, and the posterior pdf $p(X_{k-1}|Z_{1:k-1})$ is available. The prior pdf defines the most likely state at time k without introducing the observation.

Update: When the observation Z_k is available, the prior pdf is updated via Bayes' rule to obtain the posterior pdf. A complete expression is shown in [26].
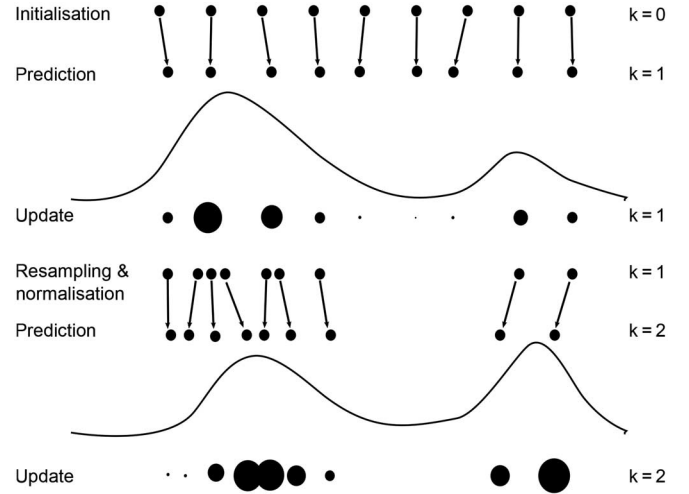

Fig. 1. Illustration of the stages of the PF algorithm for a pixel at 1-D.

Considering the Markovian assumption, the posterior pdf $p(X_k|Z_{1:k})$ can be obtained recursively from the pdf $p(X_{k-1}|Z_{1:k-1})$ calculated at a previous state $k-1$. In case the initial distribution pdf $p(X_0|Z_0) = p(x_0)$ is unknown, a uniform distribution over the whole state space can be considered.

PF is used to approximate the posterior pdf with a set of N samples when the prediction and update steps are not analytically tractable. Particles represent possible states over the state space. Each particle is defined by a vector (position) and a probability value (weight) based on the likelihood of an observation. In other words, the posterior pdf is represented by a set of N particles and their weights. In the *initialization* step, a set of N particles are distributed with an initial pdf that can be uniform (step 1). In the *prediction* step particles evolve in time according to the transition model $p(X_k|X_{k-1})$ to obtain the prior pdf. When an observation is available the *update* step is carried out, in which the weights of the particles are computed to generate the posterior pdf using (1) and, being N sufficiently large, it approximates the posterior pdf (step 3)

$$p(X_{0:k}|Z_{1:k}) \approx \sum_{i=1}^{N} \omega_k^i \delta(X_{0:k} - X_{0:k}^i) \qquad (1)$$

where δ is the delta-Dirac function, $p(X_{0:k}|Z_{1:k})$ is the true posterior pdf, $X_{0:k}^i$ is the ith simulated sample (particle), and ω_k^i is the weight of ith simulated sample (particle).

For a one-dimensional (1-D) process, a graphical evolution of the samples at each state is shown in Fig. 1. The particles are distributed over the horizontal axis represented by a circle, and the area indicates the weight of each one. At step k, the particles evolve using the transition model to obtain the prior pdf. When a new observation is available, weights are updated by the observation $p(Z_{k-1}|X_{k-1}^i)$ to obtain the posterior pdf.

As k increases, only a few particles will keep a significant weight, or in other words, the particles will degenerate with time. In order to resolve the degeneracy problem, it is necessary to implement a *resampling* step. In the literature, we can find a lot of resampling methods [27]. If the number of particles

Algorithm 1. RSR Method

Generate a random number $\triangle U^{(0)} \approx U[0, \frac{1}{M}]$
    for m = 1 to N **do**
        $i^{(m)} = \lfloor (\omega_n^{(m)} - \triangle U^{(m-1)}M \rfloor + 1$
        $\triangle U^{(m)} = \triangle U^{(m-1)} + \frac{i^{(m)}}{M} - \omega_n^{(m)}$
    end for

---

with high weight is below a threshold, resampling is employed to concentrate more particles where the posterior pdf is more likely. To measure the degeneracy, an estimate of the effective sample size is typically used as

$$N_{eff} = \frac{1}{\sum_{i=1}^{N}(\omega_k^i)^2}. \tag{2}$$

When N_{eff} is smaller than a previously defined threshold, resampling is applied. A small N_{eff} value means a large variance for the weights, hence more degeneracy. The objective of the resampling is to concentrate more particles, where there are high weights and drop only a few where weights are low. In this work, the residual-systematic resampling (RSR) [28] is used for this purpose. The pseudocode in Algorithm 1 shows the RSR algorithm for N input and M output particles, where ω_n^N is an array of weights. The $\lfloor \rfloor$ indicates a rounded down and $U[0, \frac{1}{M}]$ represents the uniform distribution between 0 and $\frac{1}{M}$.

The resulting set of samples is in fact an independent and identically distributed (i.i.d.) from the discrete posterior pdf $p(X_k|Z_{1:k})$. Therefore, the weights can now be reset to $\omega_k^i = 1/N_s$. In Fig. 1, the resampling and normalization steps are illustrated, showing the particles before and after being resampled. In addition resampling enables the adaption of the computational cost, i.e., the number of particles can be varied at each step, and the minimum number of particles can be used to keep the most important information from the pdf (hence reducing the computational cost), or it can be increased when necessary. However, if we reduce sharply the number of particles it is possible that the new acquisition pdf does not overlap the pdf given by the transition mode, meaning that the prediction and the observation are incoherent. To solve this issue, a uniform redistribution of the particles is necessary. To prevent this, it is important to adapt the number of particles after resampling.

The last step is the *normalization* of the weights before going to the prediction step again. A summary of all stages involved in the estimation process is shown in Table I.

III. METHODOLOGY

In this section, the methodology for our particular problem is presented. Point A describes the transition model employed in this work. In point B, the development of the PF in the framework of precision farming is shown, and finally, the approach to estimate different key dates and the phenological stage is exposed in Section III-C.

TABLE I
PF SEQUENCE

1) Initialisation	Generate samples for $i = 1, \ldots, N$ from initial distribution $p(x_0)$ if it is unknown, a uniform distribution is used (Set initial weight $\omega_0^i = 1/N$).
2) Prediction	Obtain the next state using the Transition Model (see section II.D) $x_{k+1} \propto p(x_{k+1}|x_k^i)$.
3) Update	Combine the observation and prediction on the basis of likelihood function. $\omega_{k+1}^i \propto \omega_k^i * p(Z_{k+1}|x_{k+1}^i)$.
4) Resampling	If the posterior pdf is represented by a small number of particles, $N_{eff} < N_{thrd}$, they are redistributed where the pdf is more likely. Set $\omega_0^i = 1/N$.
5) Normalisation	Normalise the weight $\omega_0^i = \frac{\omega_{k+1}^i}{\sum_{i=1}^{N} \omega_{k+1}^i}$. Go to step 2.
Definitions	
$p(x_0)$,	Prior distribution;
ω_k^i,	Weight of particle i at time k
N,	Number of particles
X_k^i,	State vector of particle i at time k
$p(Z_k|X_k)$,	Observation Model
$p(X_k|X_{k-1})$	Transition Model
N_{eff}	Number of effective particles, Eq. (2)
N_{thrs}	Threshold to make resampling.

A. Dynamic Model

A transition model, which characterizes the behavior over the state space, is required to apply the prediction step in a filtering procedure. It is a representation of the phenological evolution for any rice crop, and it can be employed by any dynamic filtering method. In this paper, the dynamic model defined in [5] is used.

From data provided by a polarimetric SAR sensor, it is possible to derive a set of polarimetric observables. In our case, we have decomposed the available dual-polarization data into three sets of polarimetric parameters: power terms, magnitude, and phase from the correlations between polarimetric channels, and parameters derived from the eigen decomposition of the coherency matrix. In this particular scenario, a total of 13 parameters are employed, as detailed in [5].

In order to reduce the number of these observables, but minimizing the loss of information, an orthogonal transformation is made based on a principal component analysis (PCA) [29], and a 3-D state space is defined. A time series of dual-polarization SAR images and the corresponding phenological ground truth are used to generate the dynamic model in the space state. The complete procedure is explained more detail in [5]. In this work, the estimation process is based on the exploitation of this previously generated model, but with a PF approach. In general, all parcels present a similar behavior, so for that reason they are combined to create an average model (or signature). Such a signature describes the temporal evolution of crops that follow the same management practices. The model must be built with homogeneous plots, since heterogeneities in the parcel means that some areas are not evolving correctly. In such cases, the signature in the space state is different from the signature of
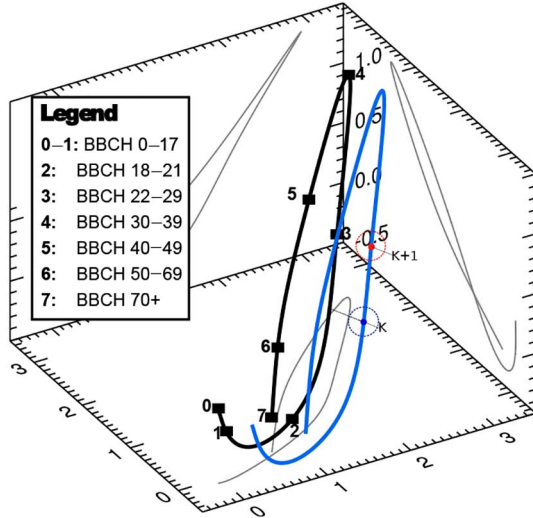
Fig. 2. Dynamic model applied to a particle at time k to obtain the new position at time $k+1$. Black line represents signature of rice on state space. Blue point represents particle position at time k. Blue line represents signature of rice shifted to the particle position. (Red point) Position of particle at time $k+1$ without noise. Red circle represents zone of possible positions when noise is added.

the model. This deviation can be used to alert the farmers of a possible problem in the crops.

It is important to remind that using a PCA transformation entails the loss of physical interpretation of the polarimetric data in the new state space. However, this study is focused on the final application (i.e., estimation of the current phenological stage) and not in the physical interpretation of the polarimetric data. It should be noticed that the definition of the state space employing such a tool provides us an effective mechanism to reduce the dimensions of the space while losing the minimum quantity of information.

Fig. 2 shows the phenological evolution of rice crops over the state space. The model is used to define the evolution of each particle from time $k-1$ to k. It is necessary to determine the way to apply it, due to the absence of an analytical model. First, each particle is projected to the closest state of the model (hereafter referred as closest state). Once we have a correspondence between both, the model is shifted to the particle position. Particles will evolve as their closest state would do. Therefore, the new position at time k is obtained by evolving over the shifted model (this is illustrated in Fig. 2). In case the temporal resolution of the model is in a daily basis, we have to increment as many steps as days from the previous acquisition to obtain the new state of the particle. The noise on the transition model was defined by a covariance matrix that depends of the variance of the input data (used when the model was generated). Once the new position is calculated, the state vector and the covariance matrix are multiplied to obtain a new state vector. The same process is applied to all particles to generate their prior pdf.

B. Filtering Approach

Initially, particles are distributed uniformly because the initial pdf is unknown. To generate the prior pdf the transition

model is applied for each particle, evolving to the new position k days later. When the observation is available the weights of the particles are computed using (3)

$$\omega_k^i \propto \omega_{k-1}^i \sum p^i \delta(Z_k - Z_k^i) \qquad (3)$$

where $p(Z_k|X_k^i) = \sum p^i \delta(Z_k - Z_k^i)$ represents the probability of particle X_k^i to generate an observation Z_k, and $p(X_k|X_{k-1})$ is used as proposal pdf [7]. In this case, the number of particles used to provide the estimations was $N = 1000$. With N below 1000 the representation of the posterior pdf's is deficient. In particular, some stages show a multimodal pdf which cannot be properly modeled due to the small number of samples, producing an incorrect estimation. If we increase N to values much larger than 1000 the estimations do not present any significant improvement but the computational burden is increased considerably. Consequently, a set of 1000 particles was selected as a good tradeoff.

The observations described in [5] are used in this work. However, instead of generating the covariance matrix employed in the EKF approach, the pdf is generated to characterize the observation.

The projection of the 3-D pdf over each axis of the state space is used to apply the methodology. Finally, when the posterior pdfs are available we can combine them to get the three-dimensional pdf again. The result is a set of particles located in a particular zone of the state space. Each particle will have an associated phenological stage which is determined by the projection of the state vector over the model. The particle with largest weight, i.e., the highest probability, shows the most likely phenological stage, and the standard deviation is given by the projection of the other particles. If the posterior pdf in (1) has degeneracy, i.e., only a few particles have a significant value whereas the rest of them have a low weight, we have to make a resampling. Finally, a normalization of the weight is necessary, and the sequence is repeated from the prediction step.

C. Application of PF

Using the previously introduced model and a set of observations, the PF method provides an estimation tool that can be applied for different purposes. We focus on the retrieval of the current phenological stage, the estimation of the sowing date, and prediction of the date at which the crop will reach the panicle initiation stage.

To have a proper description of the crop evolution, the general Biologische Bundesanstalt, Bundessortenamt, and Chemical industry (BBCH) scale for cereals [1], [30] is considered. The BBCH scale provides a numerical code to represent every growth stage along the life cycle of plants. This code ranges in a continuous way from 0 (associated with sowing) to 100 (associated with harvest). Therefore, every phenological stage along the cultivation period of cereals is associated with a BBCH code, which constitutes a convenient way to describe numerically the evolution of crops.

The evolution model contains all possible ranges of BBCH. The way to obtain the phenological stage from any state is by projecting its position over the closest state in the model. Using

only one input image the result depends on the observation precision. Instead, when a set of images is used, noise effects are reduced, wherewith the state is better determined, and the final estimate is improved.

To track phenological changes, particles evolve according to the transition model, generating a prior pdf. When a new observation is available, the pdf can be updated to provide the new distribution of particles. The projection of the new states over the model generates a set of possible phenological values (one for each particle). To provide a single solution, the maximum *a posterior* probability (MAP) estimate is used. The phenology is calculated using the particle with highest weight, and the variance is given by the estimates of the other particles. Each time a new observation is available, we can estimate the phenology of the crop by repeating this procedure.

Furthermore, each state of the model has a temporal stamp that represents the position along the temporal evolution of the crop over the state space. This temporal stamp information is exploited here to provide estimates of the date at which the crop reaches a specific stage. Due to their relevance in rice farming (explained in the Section I), we have chosen the date of sowing and the date of the panicle initiation stage for this study.

In order to know the time elapsed between two states, e.g., from $k - n$ to k, being k the state at the current observation, the difference between the two temporal stamps is calculated. Issues appear when $k - n$ is an unknown state, for instance when the observations available are only of later states. In such a case, the state is predicted from a signature (or trace) over the state space from $k - m$ to k, being $k - m$ the state defined by the first observation available and $k - m > k - n$. This trace is represented by a specific zone of the model.

To provide a solution for the sowing date, the model is used to complete the trace to the initial state to find the starting point. The amount of model used gives the elapsed time. As new data from the time series are incorporated, the trace is more representative and the fitting of the model is more accurate. This procedure allows us a more accurate estimation of the time elapse than when only one observation is available.

The anticipation of future events enables farmers to organize campaigns in an optimized way and maximize the yield. Once the current stage is identified in the transition model, it can be used to predict any future stages. Indirectly, this provides information about the time remaining before an event occurs. In such an approach, the temporal prediction is made in "open loop," so as new images are available they can be used to correct the estimation. Due to its importance for rice crops, an example of this procedure for the date of panicle initiation (BBCH 30) is considered here.

IV. RESULTS

This methodology was tested with rice crops located in Sevilla (Spain), using a stack of 11 HH/VV dual-polarization X-band images acquired by TerraSAR-X in 2009 at 30° of incidence angle, and with a resolution of 6.6 and 2.3 m in azimuth and ground range, respectively [4]. The dynamic model used was presented in [5] and is valid for this configuration mode and for the same crop management. In addition, ground truth
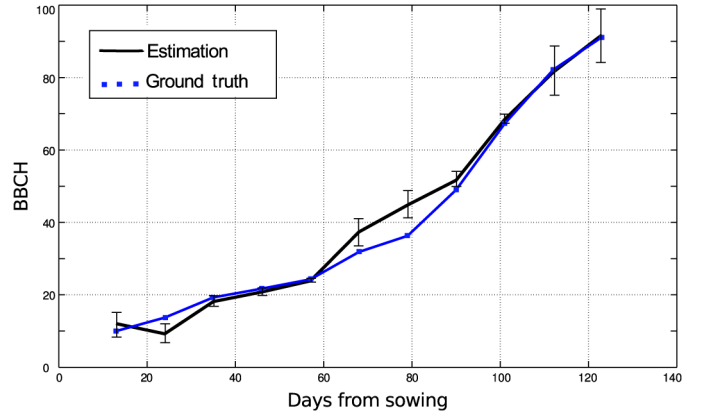

Fig. 3. BBCH estimates and ground truth values in the 2009 cultivation campaign versus days from sowing. Black line represents BBCH estimation. Blue square represents ground truth. Blue line represents ground truth linear interpolation.

data are available at parcel scale. First, current phenology is retrieved to show the advantages of employing a PF approach instead of the EKF. Then, the estimation of the date for two key events, i.e., sowing and panicle initiation stage, are presented to show the potential of this technique to be applied in precision farming.

A. Phenology Estimation

Fig. 3 shows the estimation results for one of the monitored parcels. The result is compared with the available ground truth. At any date, the estimate is taken from the most likely value of the posterior pdf defined in (1). There is a good agreement between estimates and ground data, but we can identify two zones in which estimates are clearly different from the ground truth data. As aforementioned, the first step of the estimation consists of a uniform distribution of particles over the state space (see initialization step in Table I), since prior information is not available. Consequently, the estimation during the first states presents larger errors. However, as soon as new images are incorporated the accuracy of the estimation increases, thanks to the convergence to the real value. The second zone in which results differ from ground truth occurs between BBCH 30 and 50. This phenological range presents the strongest nonlinear behavior in the model. At this point, after the posterior pdf's have been obtained (step 3 in Table I), the most likely value is projected over the model to provide an estimation of the current phenological stage. The projection in this zone of the model is more likely to provide an incorrect estimation because the stages before and after stage 4 are very close to each other (see Fig. 2).

To show the benefits involved when the PF approach is used, in contrast with the EKF method, both algorithms are tested for the phenological estimation process at all available parcels. To get a measurable difference between them, a confusion matrix is also calculated considering the phenology estimation as a classification problem. The confusion matrix shows the coincidences or agreements between the predictions (rows) and the ground truth values (columns) for a set

TABLE II
STATISTICS OF THE BBCH RESULTS WITH EIGHT PHENOLOGICAL INTERVALS USING PF

	0–15	16–27	28–34	35–49	50–59	60–84	85–100	Totals	User's Accuracy
0–15	9							6	100 %
16–27	2	14						16	87.5 %
28–34			4	1				5	80 %
35–49			2	6				8	75 %
50–59				2	2			4	50 %
60–84					1	6	2	9	66.6 %
85–100							4	4	100 %
Total	11	14	6	9	3	6	6	55	
Producer's Accuracy	81.8 %	100 %	66.6 %	66.6 %	66.6 %	100 %	66.6 %		Total:81.81%

Kappa: 0.7874

TABLE III
STATISTICS OF THE BBCH RESULTS WITH EIGHT PHENOLOGICAL USING EKF

	0–15	16–27	28–34	35–49	50–59	60–84	85–100	Totals	User's Accuracy
0–15	9	2						11	81.8 %
16–27	2	12	5	3				22	54.5 %
28–34		1	1	1				3	33.3 %
35–49				4	1			5	80 %
50–59				1	0			1	0 %
60–84					2	4	2	8	50 %
85–100						2	4	6	66.6 %
Total	11	15	6	9	3	6	6	55	
Producer's Accuaracy	81.8 %	80 %	16.6 %	44.4 %	0 %	66.6 %	66.6 %		Total:58.2%

Kappa: 0.5393

of phenological intervals. Along the main diagonal one finds the correct estimations (the estimation and the ground truth are in the same stage), so high values in the main diagonal with respect to off-diagonal positions mean that the estimation process is accurate. To interpret correctly the numbers in these tables, we provide here an example. In Table II, a total of 11 ground truth samples are in the BBCH range 0–15 (first column). Then, in the estimation process nine of them were correctly estimated (first row), and two were wrongly assigned to the next range of BBCH values (second row). To quantify the validity of the results, two different accuracy values are usually defined: producer's accuracy and user's accuracy. For a given set of samples in the ground truth data (a column), producer's accuracy provides the percentage of them that were classified correctly. On the other hand, for a set of samples in the estimates (a row) user's accuracy yields the percentage of the correctly classified. Finally, the kappa index reflects the difference between actual agreement and the agreement expected by chance, so high values (close to 1) mean results better than by chance alone (close to 0). Table II shows the results using the PF approach, and Table III the results using the EKF approach. In both cases, we use the same set of seven intervals of phenological stages. In the case of the PF, the kappa value improves 46% with respect to the value obtained with the EKF. This is mainly due to the zones in which the model is strongly nonlinear and a rapid trend change is present (i.e., around BBCH 40).

In such a situation, the linearization made by EKF technique is very poor to fit the evolution and cannot provide a good estimation, hence the error increases. Note that a value of kappa = 0.78 is regarded as a substantial agreement, whereas kappa = 0.53 corresponds to a moderate agreement [31].

In the approach presented here, as in [5], the observation noise may not be completely modeled. To test the consequences for both approaches (EKF and PF), a simulation in which the noise distribution was distorted in the input data has been carried out, generating a new set of simulated observations. After applying the PCA the most likely state keeps the same value, but new zones acquire relevance by the noise effects. Fig. 4 shows the projection of the observation pdf over one dimension (a state variable) and the same pdf with the simulated observation noise. With these conditions the PF can provide a more reliable estimation than the EKF. The reason is that the PF can consider all the potential zones by drawing particles in all potential states, but the EKF cannot work with multimodal pdf's. Fig. 5 shows the phenological estimation using the PF, the EKF, and the ground truth. When input data are noisy, and also the model has a strong nonlinearity, like around BBCH 40, the EKF is unable to follow the evolution to provide a good estimation.

Fig. 6 shows the estimation results at pixel level for a single parcel, using the BBCH ranges employed in the previous confusion matrices. The first three images show a completely
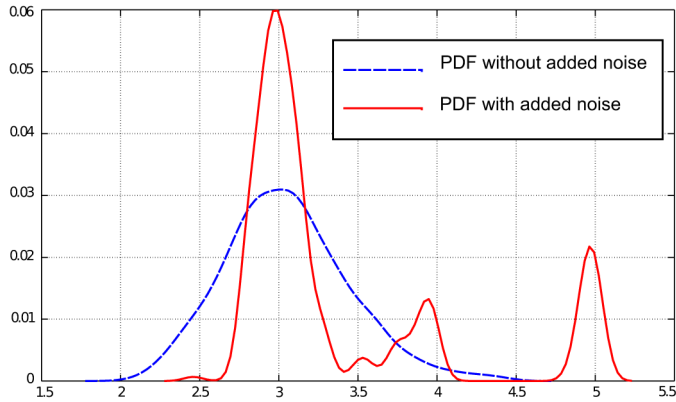
Fig. 4. Projection of the pdf of one observation over one dimension of state space. Blue line is the observation without noise added. Red line is the observation with noise added.
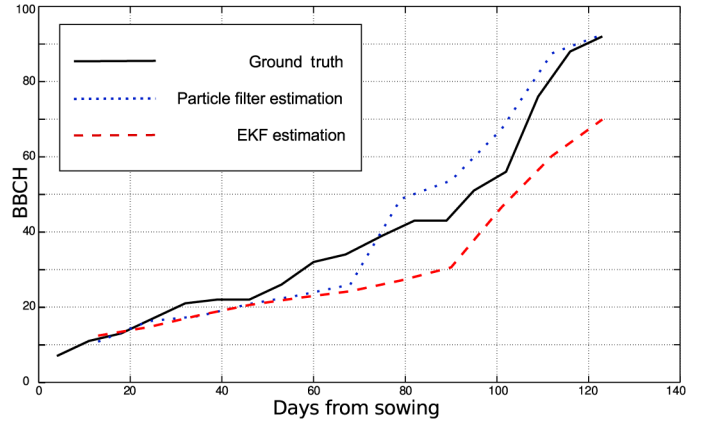

Fig. 5. Phenological estimation vs days regarding sowing when the observation has large distortion. Black line represents ground truth. Blue line represents particle filter estimation. Red line represents EKF estimation.

homogeneous estimation, since stages 1–27 are contained in a wide zone over the transition model (from stage 0 to 3 in Fig. 2). Afterward, there appear some heterogeneities, meaning that there exist pixels at diverse phenological stages within the same acquisition date. These pixels can evolve at different rates but all of them follows the same signature. These evolution differences make the prediction step less accurate but the observation step allows to minimize the impact over the estimation. In case the prediction is very inaccurate (a wider pdf), the estimation is limited by the observation accuracy. The last two phenological intervals, which dominate in the last four images, are the most mixed for the same date, as it was expected from their proximity in the state space (see stages 6 and 7 in Fig. 2).

Finally, in order to provide a visual insight of the potential of this approach, the estimation approach was applied to a large set of parcels (786 in total) for which information on the sowing date and harvest date was available. In Fig. 7, the values obtained at the time of the 11 acquisitions are represented in the form of colour maps. Although there was not ground truth available over these parcels for validation, and the model generation was carried out with information from only six parcels, this result shows a monotonic and spatially coherent evolution over the whole area, hence confirming its potential.

B. Estimation of the Date of Key Events

1) Sowing Date: Tables IV and V show the absolute errors (in days) that are made on the sowing date estimation for two different parcels. Series of up to 10 images are used to test this application. The acquisition dates, relative to the date of sowing, i.e., days after the sowing date (DaS), are indicated in the first row. The row index refers to the first image used in the estimation process, and the column index represents the last image used to compute the sowing date. For instance, for parcel A (Table IV) the element on row 1 and column 5 shows that the error is 1 day when the estimation is made with image 5 combined with the past information, given by the last four images. The main diagonal contains the estimation error when previous information is not used, i.e., with a single acquisition. In this situation, considering the results of both parcels, the error ranges

between 87 days in the worst scenario and 1 day in the best one (the closest state to the sowing day).

The noise of the observation is modeled by only taking into account the polarimetric speckle noise. Hence, other sources of noise are being dismissed. Zones with non modeled noise may produce larger deviations in the estimation, e.g., column 9. The PF allows us to reduce this effect significantly when a new image with low noise is available, for example the transition between column 9 and 10. Moreover, this improvement can be observed between images 4 and 5 on row 3, or images 6 and 7 on row 4 (Table IV). On the other hand, in some cases more than two images are necessary to guarantee an accurate estimation. The situation is shown in row 4. This is mainly due to the high nonlinearity and the large variance in this part of the model, and also because the observations exhibit their highest variance values. However, the results of parcel B (Table V) show much better estimates. For instance, estimates with only two images are much better for this parcel.

In this example, we have not made any assumption about the stage of the crop before the first observation is considered in the estimation. Therefore, in order to know the current phenological stage, when the first observation arrives, the observation is directly projected over the model (as explained in Section III-B). As aforementioned, these results, obtained by a direct projection, are contained in the main diagonal of Tables IV and V. In most cases, the induced errors are large because only the projection is used, so they depend on the accuracy of a single observation. Between DaS 65 and 88 (columns 5–7), the projection over the model is very precise and the result is accurate enough with only one observation. Improving the estimation using more observations is not possible due to the high variance of the model in these areas. In consequence, in this range of days, observations present a high weight in the filtering process, producing similar results employing a single observation. On the other hand, it is possible to reach a situation for which the errors are so large that we cannot recover. For these two particular parcels, this situation is found at image 8, from which the estimated dates are no longer precise. The error cause is the method employed to estimate the current phenological stage:

Thisarticlehasbeenacceptedforinclusioninafutureissueofthisjournal.Contentisfinalaspresented,withtheexceptionofpagination.

8

IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING
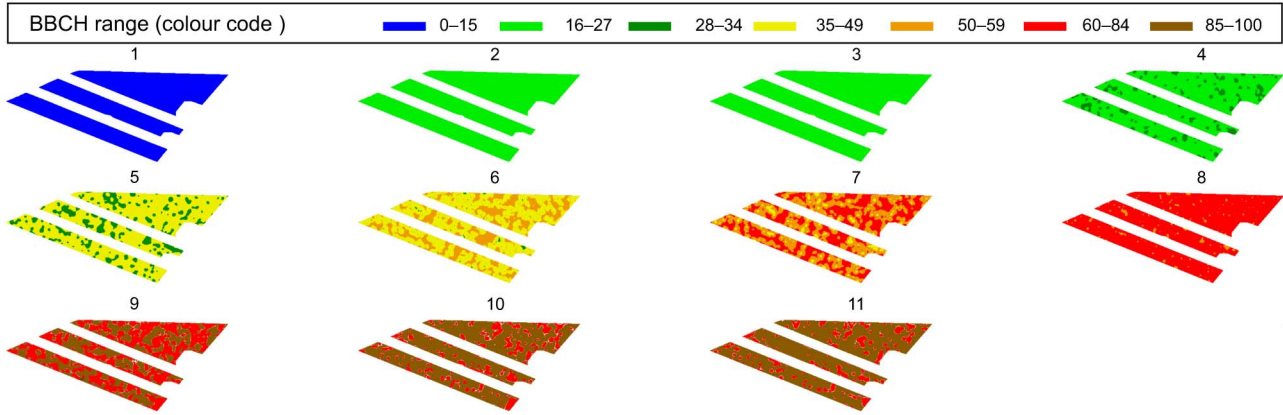
Fig. 6. Phenological estimation results obtained over the test parcel at pixel basis with the proposed methodology. The index indicates the images sequence. The first was acquired 12 days after sowing. The time between images is 11 days.
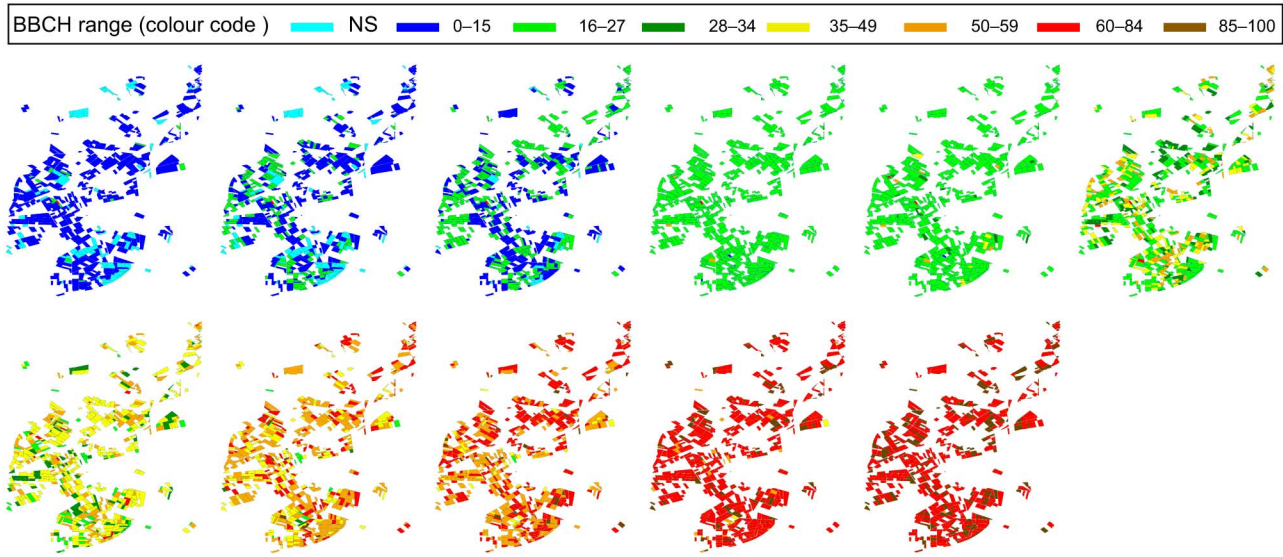

Fig. 7. Maps with the phenological estimation computed over 786 parcels in the area at 11 acquisition dates. The time between images is 11 days. NS means that the parcel was not sown yet at that date.

the projection fails to distinguish between starting states (1 and 2) and ending states (6 and 7) because they are very close in the model (see Fig. 2). An incorrect projection causes that when the next image is available (image 9) the prior pdf given by the transition model is not overlapped with the pdf given by the observation. Consequently, the posterior pdf cannot be defined and an uniform redistribution of the particles is necessary (step 1 in Table I). The same situation is repeated for images 9 and 10. This is the reason why it is not possible to have a good estimation when the first input data is after DaS 100. This issue can be solved by simply dividing the model into two sections to apply the projection. The first section, for instance, could be defined from state 0 to 3, and the second from state 4 to 7 (see Fig. 2). After projecting and applying the transition model, each section gives different prior pdf's. Considering the case in which there is an overlap between prior and observation pdf's, an accurate estimation starting at any image could be provided.

Finally, the improvement achieved in the estimation when a set of images is employed, instead of a single acquisition, can be observed by comparing elements of the main diagonal

TABLE IV
ERROR (IN DAYS) OF THE SOWING DATE ESTIMATION FOR PARCEL A

DaS	21	32	43	54	65	76	87	98	109	120
Image	1	2	3	4	5	6	7	8	9	10
1	2	5	4	1	1	3	3	9	8	1
2		6	5	6	1	3	3	8	8	1
3			30	34	1	4	3	7	8	1
4				32	21	21	8	8	8	1
5					2	4	4	7	8	1
6						4	3	7	8	1
7							4	9	9	1
8								65	78	84
9									77	83
10										81

with off diagonal elements. Extremely noisy input data could produce inaccurate estimates, as e.g., at image 4 in Table IV. The error in the estimation is about 32 days (row 4, column 4) when previous information is unknown, but using the proposed methodology the error decreases down to 1 day (row 1, column 4).

Thisarticlehasbeenacceptedforinclusioninafutureissueofthisjournal.Contentisfinalaspresented,withtheexceptionofpagination.

DE BERNARDIS *et al.*: ESTIMATION OF KEY DATES AND STAGES IN RICE CROPS

9

TABLE V
ERROR (IN DAYS) OF THE SOWING DATE ESTIMATION FOR PARCEL B

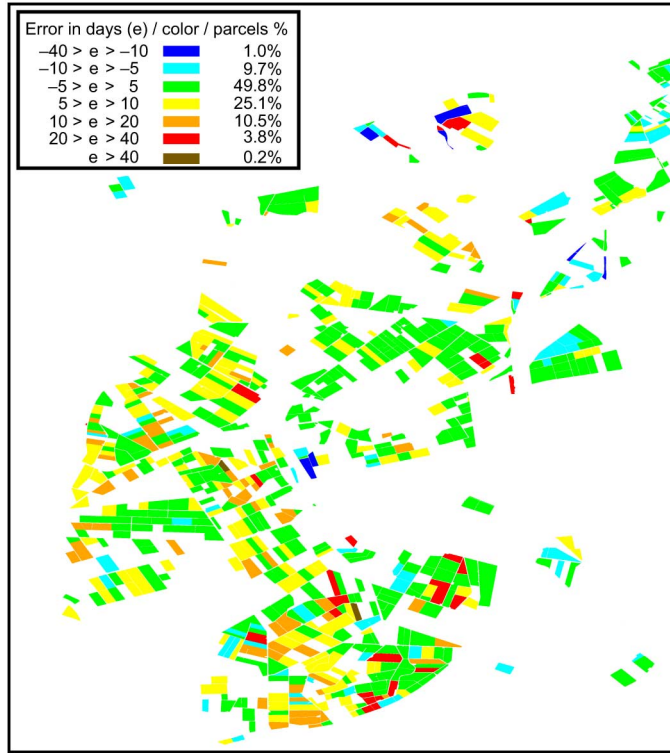| DaS	22	33	44	55	66	77	88	99	110	121
Image	1	2	3	4	5	6	7	8	9	10
1	1	2	1	2	2	6	1	0	8	1
2		3	1	2	2	6	1	0	9	1
3			3	2	2	6	1	1	8	1
4				12	4	2	2	1	9	1
5					5	8	0	0	9	1
6						7	1	0	8	1
7							2	0	9	1
8								65	78	86
9									76	83
10										87

Fig. 8. Map and histogram of the error in the estimation of the sowing day for a set of 786 parcels, obtained by using the first four SAR acquisitions.

To further evaluate the accuracy of the sowing date estimation process, the complete set of 786 parcels was also used for validation. The results are shown in Fig. 8. Date estimates are obtained by using the first four SAR images. Each color represents a range of error in days and the percentages of parcels in each range are presented also in the picture. Negative values mean that the estimated date is before the actual value, whereas positive values mean it corresponds to later dates. It must be emphasized that we are estimating the sowing date of 786 plots using a model that was created with information from only six plots [5]. Even so, we have obtained a very good accuracy: the absolute error is less than 5 days for 50% of the parcels and less that 10 days for 85% of them. If the number of plots used to build the model were increased, it is expected that the accuracy in the estimation process would increase because the average behavior of them would be better defined. The fact that these accuracies are achieved with so few samples is an evidence of the potential of the methodology.
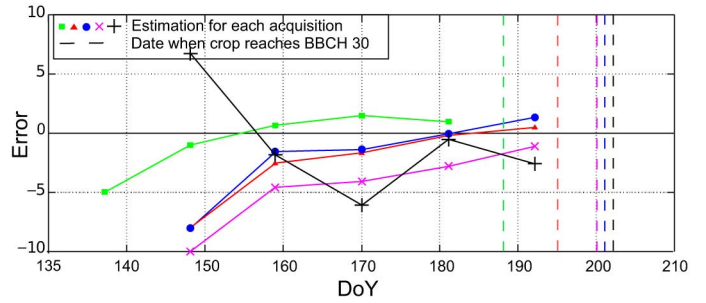

Fig. 9. Error (in days) in the estimation of the date in which rice reaches stage BBCH 30 (panicle initiation) for different parcels, represented by different colours. The vertical axis shows the estimation error in days, and the horizontal axis indicates the DoY when the acquisition was made. The vertical dashed lines represent the DoY in which each crop reaches BBCH 30 according to the ground truth.

2) Date of Panicle Initiation Stage: The final set of results corresponds to the predictions of the date in which crops reach the panicle initiation stage. Fig. 9 shows the results obtained for the six individual parcels with ground data, represented with different colours. The estimated error is depicted in the vertical axis and represents the difference in days between the date given by the ground truth and the estimated date. The horizontal axis shows the day of year (DoY) in which the acquisitions (and hence the estimations) were made. The vertical dashed lines indicate the dates in which each parcel reaches the BBCH 30 according to the ground truth. The first acquisitions available in the data set are around 10 days after the sowing date. A total of up to five images are employed to predict the moment in which stage 30 is reached. The first predicted dates (around 50 days before) exhibit the worst values, as expected, with errors greater than 5 days in most cases. Each time a new image is incorporated (every 11 days) the error decreases, even in the parcel represented in black, for which the nonmodeled noise has the strongest influence. All parcels show an improvement when temporal sequences are used in the estimation process. The results obtained from 40 days before the panicle initiation date onward show differences of only 3 days in average, and not larger than 5 days in most cases. Therefore, the estimation of the date associated with the panicle initiation stage shows that the PF is a robust method to predict dates of future phenological stages.

V. CONCLUSION AND DISCUSSION

In this paper, the advantage of the particle filter, over other filtering techniques like the EKF, is demonstrated in estimation applications relative to precision farming. It is much more robust to noisy measurements and nonlinear evolutions, making this method a better tool to obtain the required results. Two main applications have been tested: retrieval of current phenological stage and estimation of dates of particular crop events, like sowing and initiation of panicle. Phenology tracking has been improved with this technique, and the contribution of time series of data has been studied.

The sowing date was estimated with an error that is less than 1 day in parcels that are well characterized by the model. The estimation was also carried out over 786 parcels using a model

This article has been accepted for inclusion in a future issue of this journal. Content is final as presented, with the exception of pagination.

10 IEEE JOURNAL OF SELECTED TOPICS IN APPLIED EARTH OBSERVATIONS AND REMOTE SENSING

that was created with the information of only six parcels. In this case, we achieve errors lower than 5 days in the estimation process for 50% of the parcels, and less than 10 days for 85% of them.

This technique also enables the prediction of the date of future events. For instance, it has been shown that it is possible to get a prediction 40 days before the panicle initiation date with an error of only 3 days on average.

To ensure proper estimations, the model must be built with homogeneous fields. Plots producing a signature very different from the model can be identified, so this approach would provide an indicator that the plot is not developing properly and would help to alert farmers. According to our methodology, the model characterizes crops with the same management practices. To apply this approach to crops with different farming practices, it would be necessary to build a new model representing each typical crop development in the site.

Ongoing work is addressed to test the methodology over other types of crops. Moreover, results have to be evaluated as a function of the available revisit time. In case of exploiting polarimetric data, the influence of incidence angle and the available polarimetric space (i.e., dual-pol and compact-pol) will be analyzed. Such a study is of interest in view of the next availability of data from new missions like Sentinel-1 and ALOS-2. Finally, since the methodology allows us to combine input data from different sources, another promising line of research would consist in combining optical and SAR data in order to generate more accurate estimations.

REFERENCES

[1] U. Meier, "Growth stages of mono-and dicotyledonous plants," Biological Research Centre for Agriculture and Forestry, Tech. Rep., 2001 [Online]. Available: http://www.bba.de/veroeff/bbch/bbcheng.pdf

[2] T. Sakamoto, "A crop phenology detection method using time-series MODIS data," *Remote Sens. Environ.*, vol. 96, no. 3, pp. 366–374, 2005.

[3] J. Shang *et al.*, "Tracking crop phenological development of spring wheat using synthetic aperture radar (SAR) in Northern Ontario, Canada," in *Proc. IEEE 2nd Int. Conf. Agro-Geoinformat. (Agro-Geoinformatics'13)*, 2013, pp. 517–521.

[4] J. M. Lopez-Sanchez, S. R. Cloude, and J. D. Ballester, "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.

[5] F. Vicente-Guijalba, T. Martinez-Marin, and J. M. 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.

[6] R. E. Kalman and R. S. Bucy, "New results in linear filtering and prediction theory," *Trans. ASME, J. Basic Eng.*, vol. 83, pp. 95–107, 1961.

[7] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-Gaussian Bayesian tracking," *IEEE Trans. Signal Process.*, vol. 50, no. 2, pp. 174–188, Feb. 2002.

[8] A. Doucet, S. Godsill, and C. Andrieu, "On sequential Monte Carlo sampling methods for Bayesian filtering," *Statist. Comput.*, vol. 10, no. 3, pp. 197–208, 2000.

[9] M. H. Kalos and P. A. Whitlock, *Monte Carlo Method*. Hoboken, NJ, USA: Wiley, 2008.

[10] B. Tan *et al.*, "Vegetation phenology metrics derived from temporally smoothed and gap-filled MODIS data," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS'08)*, 2008, vol. 3, pp. III–593.

[11] M. Boschetti, D. Stroppiana, P. A. Brivio, and S. Bocchi, "Multi-year monitoring of rice crop phenology through time series analysis of MODIS images," *Int. J. Remote Sens.*, vol. 30, no. 18, pp. 4643–4662, 2009.

[12] H. Ozer, "Sowing date and nitrogen rate effects on growth, yield and yield components of two summer rapeseed cultivars," *Eur. J. Agron.*, vol. 19, no. 3, pp. 453–463, 2003.

[13] K. G. Cassman, A. Dobermann, D. T. Walters, and H. Yang, "Meeting cereal demand while protecting natural resources and improving environmental quality," *Annu. Rev. Environ. Resour.*, vol. 28, no. 1, pp. 315–358, 2003.

[14] Q. D. Pham, A. Abe, M. Hirano, S. Sagawa, and E. Kuroda, "Analysis of lodging-resistant characteristics of different rice genotypes grown under the standard and nitrogen-free basal dressing accompanied with sparse planting density practices," *Plant Prod. Sci.*, vol. 7, no. 3, pp. 243–251, 2004.

[15] R. M. Cu, T. W. Mew, K. G. Cassman, and P. S. Teng, "Effect of sheath blight on yield in tropical, intensive rice production system," *Plant Dis.*, vol. 80, no. 10, pp. 1103–1108, 1996.

[16] Z. L. Zhu and D. L. Chen, "Nitrogen fertilizer use in China–Contributions to food production, impacts on the environment and best management strategies," *Nutr. Cycl. Agroecosyst.*, vol. 63, no. 2–3, pp. 117–127, 2002.

[17] I. N. Oka, "Cultural control of the brown planthopper," in *Brown Planthopper: Threat to Rice Production in Asia*, Los Baños, Philippines: International Rice Research Institute (IRRI), 1979, pp. 359–369.

[18] M. A. Mannan, M. S. U. Bhuiya, S. M. A. Hossain, and M. I. M. Akhand, "Study on phenology and yielding ability of Basmati fine rice genotypes as influenced by planting date in aman season," *Bangladesh J. Agric. Res.*, vol. 34, no. 3, pp. 373–384, 2009.

[19] N. Bergman, "Recursive Bayesian estimation," Ph.D. dissertation, Linköping Studies Sci. Technol., Dept. Elect. Eng., Linköping Univ., 1999, vol. 579.

[20] M. Ghosh, N. Mukhopadhyay, and P. K. Sen, *Sequential Estimation*. Hoboken, NJ, USA: Wiley, 2011, vol. 904.

[21] S. K. Zhou, R. Chellappa, and B. Moghaddam, "Visual tracking and recognition using appearance-adaptive models in particle filters," *IEEE Trans. Image Process.*, vol. 13, no. 11, pp. 1491–1506, Nov. 2004.

[22] G. Catalin and S. Nedevschi, "Object tracking from stereo sequences using particle filter," in *Proc. IEEE 4th Int. Conf. Intell. Comput. Commun. Process. (ICCP'08)*, 2008, pp. 279–282.

[23] J. Welle, D. Schulz, T. Bachran, and A. B. Cremers, "Optimization techniques for laser-based 3D particle filter SLAM," in *Proc. IEEE Int. Conf. Robot. Autom. (ICRA'10)*, 2010, pp. 3525–3530.

[24] J. Durbin and S. J. Koopman, *Time Series Analysis by State Space Methods*. New York, NY, USA: Oxford Univ. Press, 2012, vol. 38.

[25] M. West and J. Harrison, *Bayesian Forecasting and Dynamic Models*. New York, NY, USA: Springer, 1997, ISBN 0-387-94725-6.

[26] Y.-C. Ho and R. Lee, "A Bayesian approach to problems in stochastic estimation and control," *IEEE Trans. Automat. Control*, vol. 9, no. 4, pp. 333–339, Oct. 1964.

[27] R. Douc and O. Cappé, "Comparison of resampling schemes for particle filtering," in *Proc. IEEE Int. Symp. Image Signal Process. Anal. (ISPA'05)*, 2005, pp. 64–69.

[28] M. Bolić, P. M. Djurić, and S. Hong, "Resampling algorithms for particle filters: A computational complexity perspective," *EURASIP J. Adv. Signal Process.*, vol. 2004, no. 15, pp. 2267–2277, 2004.

[29] L. Eklundh and A. Singh, "A comparative analysis of standardised and unstandardised principal components analysis in remote sensing," *Int. J. Remote Sens.*, vol. 14, no. 7, pp. 1359–1370, 1993.

[30] J. C. Zadoks, T. T. Chang, and C. F. Konzak, "A decimal code for the growth stages of cereals," *Weed Res.*, vol. 14, no. 6, pp. 415–421, 1974.

[31] A. J. Viera *et al.*, "Understanding interobserver agreement: The kappa statistic," *Fam. Med.*, vol. 37, no. 5, pp. 360–363, 2005.

Caleb G. De Bernardis was born in Mar del Plata, Buenos Aires, Argentina, in 1989. He received the Ingeniero Tecnico (B.S.) degree (*summa cum laude*) in sound and image engineering from the University of Alicante, Alicante, Spain, in 2010, and the Ingeniero (M.S.) degree in telecommunication engineering from the Technical University of Valencia (UPV), Valencia, Spain, in 2014. Currently, he is pursuing the Ph.D. degree at Signals, Systems and Telecommunications Group, University of Alicante.

He collaborated with the Nanophotonics Technology Centre (NTC), Valencia, Spain, in 2013.

Fernando Vicente-Guijalba was born in Elche, Spain, in 1981. He received the Ingeniero Tecnico (B.S.) degree in telecommunication engineering and the Ingeniero (M.S.) degree in sound and image engineering from the University of Alicante, Alicante, Spain, in 2006 and 2011, respectively.

Since 2011, he has been a Predoctoral Fellow with the Signals, Systems, and Telecommunications Group, University of Alicante. His research interests include dynamical systems analysis with applications in polarimetric and interferometric SAR methods.

Tomas Martinez-Marin received the Technical Engineering (B.S.) degree from the University of Alcalá (UAH), Madrid, Spain, in 1990, the M.S. and Ph.D. degrees in telecommunication engineering from the Technical University of Madrid (UPM), Madrid, Spain, in 1995 and 1999, respectively.

In 1990, he joined the University of Alcalá as an Assistant Professor. In 1997, he joined the European University of Madrid (UEM) as an Assistant Professor. Since 2000, he has been with the Department of Physics, System Engineering and Signal Theory, University of Alicante (UA), Alicante, Spain, where he is currently an Associate Professor. His research interests include reinforcement learning, optimal control, intelligent vehicles, and SAR filtering algorithms.

Juan M. Lopez-Sanchez (S'94–M'00–SM'05) was born in Alicante, Spain, in 1972. He received the Ingeniero (M.S) and Doctor Ingeniero (Ph.D.) degrees in telecommunication engineering from the Technical University of Valencia (UPV), Valencia, Spain, in 1996 and 2000, respectively.

From 1998 to 1999, he worked as a Predoctoral Grantholder with the Space Applications Institute, Joint Research Centre of the European Commission, Ispra, Italy. Since 2000, he has been leading the Signals, Systems, and Telecommunication Group, University of Alicante, Alicante, Spain, where he is a Full Professor since November 2011. He has coauthored more than 50 papers in refereed journals and more than 90 papers and presentations in international conferences and symposia. His research interests include microwave remote sensing for inversion of biophysical parameters, polarimetric and interferometric techniques, SAR imaging algorithms, and analytical and numerical models for multiple scattering problems.

Dr. Lopez-Sanchez was the Chair of the Spanish Chapter of the IEEE Geoscience and Remote Sensing Society, from 2006 to 2012. He was the recipient of the Indra Award for the best Ph.D. thesis about radar in Spain, in 2001.