

# Modeling and Prediction of Time-Varying Environmental Data Using Advanced Bayesian Methods

**Mansouri Majdi, Dumont Benjamin, Destain Marie-France.**

*Département des Sciences et Technologies de l'Environnement, GxABT,  
Université de Liège,  
2 Passage des Déportés, 5030 Gembloux, Belgium  
mfdestain@ulg.ac.be  
tel : +32 (0)81622164  
fax : +32(0)81622167*

## **ABSTRACT**

The problem of state/parameter estimation represents a key issue in crop models which are nonlinear, non-Gaussian and include a large number of parameters. The prediction errors are often important due to uncertainties in the equations, the input variables, and the parameters. The measurements needed to run the model (input data), to perform calibration and validation are sometimes not numerous or known with some uncertainty.

In these cases, estimating the state variables and/or parameters from easily obtained measurements can be extremely useful. In this work, we address the problem of modeling and prediction of leaf area index and soil moisture (LSM) using state estimation. The performances of various conventional and state-of-the-art state estimation techniques are compared when they are utilized to achieve this objective. These techniques include the extended Kalman filter (EKF), unscented Kalman filter (UKF), particle filter (PF), and the more recently developed technique variational Bayesian filter (VF).

The objective of this work is to extend the state and parameter estimation techniques (i.e., EKF, UKF, PF and VF) to better handle nonlinear and non-Gaussian processes without a priori state information, by utilizing a time-varying assumption of statistical parameters. In this case, the state vector to be estimated  $z_k$  at any instant  $k$  is assumed to follow a Gaussian model, where the expectation  $\mu_k$  and the covariance matrix  $\lambda_k$  are both random. The randomness of the expectation and the covariance of the state/parameter vector are assumed here to further capture the uncertainty of the state distribution. One practical choice of these distributions can be a Gaussian distribution for the expectation and a multi-dimensional Wishart distribution for the covariance matrix. The assumption of random mean and random covariance of the state leads to a probability distribution covering a wide range of tail behaviors, which allows discrete jumps in the state variables,  $z_k$

The results of the comparative studies show that the PF provides a higher accuracy than the EKF, which is due to the limited ability of the EKF to handle highly nonlinear processes. The results also show that the PF provides a significant improvement over the UKF. This is because the covariance is propagated through linearization of the underlying non-linear model, when the state transition and observation models are highly non-linear. The results also indicate that the VF provides a significant improvement over the PF because, unlike the PF which depends on the choice of sampling distribution used to estimate the posterior distribution, the VF yields an optimum choice of the sampling distribution, which also accounts for the observed data. The original data were issued from experiments carried out on a silty soil in Belgium, with a wheat crop, during two consecutive years, the seasons 2008-09 and 2009-10.

## I. INTRODUCTION

Parameter and states estimation in nonlinear environmental systems is an important issue in diagnosis, measurement and modeling. However, due to the difficulty of, or cost associated with, obtaining these measurements, state and/or parameter estimators are often used to overcome this problem. Crop models such as EPIC [1], WOFOST [2], DAISY[3], STICS [4], and SALUS [5] are dynamic non-linear models that describe the growth and development of a crop interacting with environmental factors (soil and climate) and agricultural practices (crop species, tillage type, fertilizer amount, etc. ). They are developed to predict crop yield and quality or to optimize the farming practices in order to satisfy environmental objectives, as the reduction of nitrogen lixiviation. More recently, crop models are used to simulate the effects of climate changes on the agricultural production. Nevertheless, the prediction errors of these models may be important due to uncertainties in the estimates of initial values of the states, in input data, in the parameters, and in the equations. The measurements needed to run the model are sometimes not numerous, whereas the field spatial variability and the climatic temporal fluctuations over the field may be high. The degree of accuracy is therefore difficult to estimate, apart from numerous repetitions of measurements. For these reasons, the problem of state/parameter estimation represents a key issue in such nonlinear and non-Gaussian crop models including a large number of parameters, while measurement noise exists in the data.

Several state estimation techniques have been developed and used in practice. These techniques include the extended Kalman filter (EKF), unscented Kalman filter (UKF), particle filter (PF), and more recently the variational Bayesian filter (VF). The classical Kalman Filter (KF) was developed in the 1960s [6], and has been widely used in various engineering and science applications, including communications, control, machine learning, neuroscience, and many others. In the case where the model describing the system is assumed to be linear and Gaussian, the KF provides an optimal solution [7, 8, 9, 10]. The KF has also been formulated in the context of Takagi-Sugeno fuzzy systems to handle nonlinear models, which can be described as a convex set of multiple linear models [11, 12, 13]. It is known that the KF is computationally efficient; however, it is limited by the non-universal linear and Gaussian modeling assumptions. To relax these assumptions, the extended Kalman filter (EKF) [7, 8, 14, 15, 16] and the unscented Kalman filter (UKF) [17, 18, 19] have been developed. In extended Kalman filtering, the model describing the system is linearized at every time sample (in order to estimate the mean and covariance matrix of the state vector), and thus the model is assumed to be differentiable. Unfortunately, for highly nonlinear or complex models, the EKF does not usually provide a satisfactory performance. On the other hand, instead of linearizing the model to approximate the mean and covariance matrix of the state vector, the UKF uses the unscented transformation to improve the approximation of these moments. In the unscented transformation, a set of samples (called sigma points) are selected and propagated through the nonlinear model, which provides more accurate approximations of the mean and covariance matrix of the state vector, and thus more accurate state estimation.

Other state estimation techniques use a Bayesian framework to estimate the state and/or parameter vector [20]. This approach relies on computing the probability distribution of the unobserved state given a sequence of the observed data in addition to a state evolution model. Consider an observed data set  $y$  which is generated from a model defined by a set of unknown state variables and/or parameters  $z$ . The beliefs about the data are completely expressed via the parametric probabilistic observation model  $P(y|z)$ . The learning on uncertainty or randomness of a process is solved by constructing a distribution  $P(z|y)$ , called the posterior distribution, which quantifies our belief about the system after obtaining the measurements. According to Bayes theorem, the posterior can be expressed as

$$P(z|y) = \frac{P(y|z)P(z)}{P(y)}$$

where  $P(y|z)$  is the conditional distribution of the data given the vector  $z$ , which is called the likelihood function,  $P(z)$  is the prior distribution which quantifies our belief about  $z$  before obtaining the measurements, and  $P(y)$  is the distribution of the data. Unfortunately, for most nonlinear systems and non-Gaussian noise observations, closed-form analytic expressions of the posterior distribution of the state vector are untractable ([21]). To overcome this drawback, a nonparametric Monte Carlo sampling based method called Particle Filtering (PF) has been proposed by Doucet and Tadic [25]. The latter method presents several advantages since: (i) it can account for the constraint of small number of data samples, (ii) the online update of the filtering distribution and its compression are simultaneously performed, and (iii) it yields an optimal choice of the sampling distribution over the state variable by minimizing the Kullback-Leibler (KL) divergence.

Recently, a variational filtering (VF) has been proposed for solving the nonlinear parameter estimation problem encountered in crop models. Mansouri et al. ([35,36]) used a Bayesian sampling method for modeling and prediction of nonlinear environmental system, where the nonlinear observed system was assumed to progress according to a probabilistic state space model. In this investigation, the state vector to be estimated (at any time instant) was assumed to follow a Gaussian model, where the expectation and the covariance matrix are constants.

Each of the above state estimation techniques has its advantages and disadvantages. For example, the variational filter can be applied to large parameter spaces, has better convergence properties, and is easier to implement than the particle filter. Both of them can provide improved accuracy over the EKF.

The general objective of this chapter is compare the performances of the EKF, UKF, PF, and VF when used to monitor and model an environmental process called LMS, which represents the temporal evolution of three state variables implied in the wheat crop growth and development: the leaf-area index (LAI), the soil moisture of the topsoil layer (0-30 cm) (HUR1), and the soil moisture of the subsoil layer (30-50 cm) (HUR2). This comparative study will assess the accuracy and convergence of the proposed techniques, as well as the effect of the size of the parameter space (i.e., number of estimated parameters) on the performances of the estimation techniques. Some practical challenges, however, can affect the accuracy of estimated states and/or parameters. Beside the existence of many parameters, it is necessary to consider the presence of noise in the measured data, and the restricted availability of some measured data samples. Consequently, the objectives of this chapter is two-fold: i) we study the accuracy and convergence of EKF, UKF, PF, and VF techniques, ii) we investigate the effect of the above challenges on the performances of these techniques.

The rest of the chapter is organized as follows. In Section II, a statement of the problem addressed in this chapter is presented, followed by description of variational Bayesian filtering. Then, in Section III, the performances of the various state estimation techniques are compared through their application to estimate the state variables and model parameters of the LSM process. Finally, some concluding remarks are presented in Section IV.

## II. MATERIAL AND METHODS

### A. Problem Formulation

Here, the estimation problem of interest is formulated for a general system model. Let a nonlinear state space model be described as follows:

$$\begin{cases} \dot{x} = g(x, u, \theta, w) \\ y = l(x, u, \theta, v) \end{cases} \quad (1)$$

where,  $x \in \mathbb{R}^n$  is a vector of the state variables,  $u \in \mathbb{R}^p$  is a vector of the input variables (which can be changed as desired),  $\theta \in \mathbb{R}^q$  is an unknown parameter vector,  $y \in \mathbb{R}^m$  is a vector of the measured variables,  $g$  and  $l$  are nonlinear differentiable functions, and  $w \in \mathbb{R}^n$  and  $v \in \mathbb{R}^m$  are process and measurement noise, which quantify randomness in the process and errors in the measurements, respectively.

Discretizing the state space model (1), the discrete model can be written as follows:

$$\begin{cases} x_k = f(x_{k-1}, u_{k-1}, \theta_{k-1}, w_{k-1}) \\ y_k = h(x_k, u_k, \theta_k, v_k) \end{cases}, \quad (2)$$

which describes the state variables at some time step ( $k$ ) in terms of their values at a previous time step ( $k-1$ ). Note that in equation (2). The process and measurement noise vectors have the following properties:

$$E[w_k] = 0, \quad E[w_k w_k^T] = Q_k, \quad E[v_k] = 0, \quad E[v_k v_k^T] = R_k.$$

The function  $f$  is used to predict the value of the state vector at some time step ( $k$ ) given its value at the previous time step ( $k-1$ ), and the function  $h$  relates the measured vector ( $y_k$ ) to the state vector ( $x_k$ ) at the same time step. Also, defining the augmented vector,  $u_k$  is the vector of input variables,  $\theta_k$  is a parameter vector (assumed to be known),  $y_k$  is the vector of the measured variables,  $w_k$  and  $v_k$  are respectively model and measurement noise vectors, and the matrices,  $Q_k$  and  $R_k$ , represent the covariance matrices of the process and measurement noise vectors, respectively. We assume that the error terms  $w_k$  and  $v_k$  have normal distributions with zero expectation, and that they are mutually independent.

Since we are interested to estimate the state vector  $x_k$ , as well as the parameter vector  $\theta_k$ , let's assume that the parameter vector is described by the following model:

$$\theta_k = \theta_{k-1} + \gamma_{k-1} \quad (3)$$

where  $\gamma_{k-1}$  is white noise. In other words, the parameter vector model (3) corresponds to a stationary process, with an identity transition matrix, driven by white noise. We can define a new state vector that augments the two vectors together as follows:

$$z_k = \begin{bmatrix} x_k \\ \theta_k \end{bmatrix} = \begin{bmatrix} f(x_{k-1}, u_{k-1}, \theta_{k-1}, w_{k-1}) \\ \theta_{k-1} + \gamma_{k-1} \end{bmatrix} \quad (4)$$

where  $z_k$  is assumed to follow a Gaussian model as  $z_k \sim N(\mu_k, \lambda_k)$ , and where at any time  $k$  the expectation  $\mu_k$  and the covariance matrix  $\lambda_k$  are both random. Also, defining the augmented vector,

$$\mathcal{E}_{k-1} = \begin{bmatrix} w_{k-1} \\ \gamma_{k-1} \end{bmatrix} \quad (5)$$

the model (2) can be written as:

$$\begin{aligned} z_k &= \mathfrak{F}(z_{k-1}, u_{k-1}, \mathcal{E}_{k-1}) \\ y_k &= \mathfrak{R}(z_k, u_k, v_k) \end{aligned} \quad (6)$$

### B. Time-variant evolution systems

Instead of the kinematic parametric model [28, 29, 30] which is usually used in estimation problems, we employ a time-variant evolution systems (TVES) [31]. This model is more appropriate to practical non-linear and non-gaussian situations where no a priori information on the state value is available. The state variable  $z_k$  at instant  $k$  is assumed to follow a Gaussian model, where the expectation  $\mu_k$  and the precision matrix  $\lambda_k$  are both random. Gaussian distribution for the expectation and Wishart distribution for the precision matrix form a practical choice for the filtering implementation. The hidden state  $z_k$  is extended to an augmented state  $\alpha_k = (z_k, \mu_k, \lambda_k)$ , yielding a hierarchical model as follows,

$$\begin{cases} \mu_k \sim N(\mu_k | \mu_{k-1}, \bar{\lambda}) \\ \lambda_k \sim W_n(\bar{S} | \bar{n}) \\ z_k \sim N(\mu_k, \lambda_k) \end{cases} \quad (7)$$

where the fixed hyperparameters  $\bar{\lambda}$ ,  $\bar{S}$ , and  $\bar{n}$  are respectively the random walk precision matrix, the degrees of freedom and the precision of the Wishart distribution. Note that assuming random mean and covariance for the state  $z_k$  leads to a prior probability distribution covering a wide range of tail behaviors allowing discrete jumps in the state variable.

In fact, the marginal state distribution is obtained by integrating over the mean and precision matrix:

$$p(z_k | z_{k-1}) = \int p(z_k | \mu_k, \lambda_k) p(\mu_k, \lambda_k | x_{k-1}) d\mu_k \lambda_k \quad (8)$$

where the integration with respect to the precision matrix leads to the known class of scale mixture distributions introduced by Barndorff-Nielsen ([21]). Low values of the degrees of freedom  $\bar{n}$  reflects the heavy tails of the marginal distribution  $p(z_k | z_{k-1})$ .

Particle filtering methods offer a number of significant advantages over other conventional methods. However, since they use the prior distribution as the importance distribution [11], the latest data observation is not considered and not taken into account when evaluating the weights of the particles. Even this choice of the importance sampling distribution has computational advantages, it can cause filtering divergence. In cases where the likelihood distribution is too small compared to the prior distribution, very few particles will have significant weights. Hence, a better proposal distribution that takes the latest observation data into account is needed. The VF yields an optimal choice of the sampling distribution over the state variable by minimizing the Kullback-Leibler (KL) divergence. In addition, compared to the particle filtering method, the computational cost and the memory requirements are dramatically reduced by the variational approximation in the prediction phase. In fact, the expectations involved in the computation of the predictive distribution have closed forms, avoiding the use of Monte Carlo integration. Next, we present the variational Bayesian filter algorithm.

### C. Variational Bayesian Filter

The variational approach consists in approximating  $p(\alpha_k | z_{1:k})$  by a separable distribution  $q(\alpha_k) = q(x_k)q(\mu_k)q(\lambda_k)$  that minimizes the Kullback-Leibler divergence (KLD) between the true filtering distribution and the approximate distribution,

$$D_{KL}(q||p) = \int q(\alpha_k) \log \frac{q(\alpha_k)}{p(\alpha_k | \alpha_{1:k})} d\alpha_k \quad (9)$$

The minimization is subject to constraint  $\int q(\alpha_k) d\alpha_k = 1$ . The Lagrange multiplier method used in [31,32] shows that the updated separable approximating distribution  $q(\alpha_k)$  has the following form:

$$\begin{aligned} q(z_k) &\propto p(y_k | z_k) N(z_k | \langle \mu_k \rangle, \langle \lambda_k \rangle) \\ q(\mu_k) &\propto N(\mu_k | \mu_k^*, \lambda_k^*) \\ q(\lambda_k) &\propto W_n(\lambda_k | S_k^*) \\ q(\mu_k | \mu_{k-1}) &\propto N(\langle \mu_k^p \rangle, \langle \lambda_k^p \rangle) \end{aligned} \quad (10)$$

where  $\langle \cdot \rangle$  denotes the expectation operator relative to the distribution  $q$ . The parameters are iteratively updated according to the following scheme:

$$\begin{aligned}
\boldsymbol{\mu}_k^* &= \boldsymbol{\lambda}_k^{*-1} (\langle \boldsymbol{\lambda}_k \rangle \langle \mathbf{z}_k \rangle + \boldsymbol{\mu}_k^p \boldsymbol{\lambda}_k^p) \\
\boldsymbol{\lambda}_k^* &= \langle \boldsymbol{\lambda}_k \rangle + \boldsymbol{\lambda}_k^p \\
n^* &= \bar{n} + 1 \\
\mathbf{S}_k^* &= (\langle \mathbf{z}_k \mathbf{z}_k^T \rangle - \langle \mathbf{z}_k \rangle \langle \boldsymbol{\mu}_k \rangle^T + \langle \boldsymbol{\mu}_k \rangle \langle \mathbf{z}_k \rangle^T + \langle \boldsymbol{\mu}_k \boldsymbol{\mu}_k^T \rangle + \bar{\mathbf{S}}^{-1})^{-1} \\
\boldsymbol{\mu}_k^p &= \boldsymbol{\mu}_{k-1}^* \\
\boldsymbol{\lambda}_k^p &= (\boldsymbol{\lambda}_{k-1}^{*-1} + \bar{\boldsymbol{\lambda}}^{-1})
\end{aligned} \tag{11}$$

In fact, taking into account the separable approximate distribution at time  $k-1$ , the predictive distribution is written,

$$\begin{aligned}
p(\boldsymbol{\alpha}_{k-1} | z_{1:k-1}) &\propto \int p(\boldsymbol{\alpha}_k | \boldsymbol{\alpha}_{k-1}) q(\boldsymbol{\alpha}_{k-1}) d\boldsymbol{\alpha}_{k-1} \\
&\propto p(z_k, \boldsymbol{\lambda}_k | \boldsymbol{\mu}_k) q_p(\boldsymbol{\mu}_k)
\end{aligned} \tag{12}$$

The exponential form solution, which minimizes the Kullback-Leibler divergence between the predictive distribution  $p(\boldsymbol{\alpha}_k | z_{1:k-1})$  and the separable approximate distribution  $q_{k|k-1}(\boldsymbol{\alpha})$ , yields Gaussian distributions for the state and its mean and Wishart distribution for the precision matrix:

$$\begin{aligned}
q_{k|k-1}(z_k) &\propto N(\langle \boldsymbol{\mu}_k \rangle_{q_{k|k-1}}, \langle \boldsymbol{\lambda}_k \rangle_{q_{k|k-1}}) \\
q_{k|k-1}(\boldsymbol{\mu}_k) &\propto N(\boldsymbol{\mu}_{k|k-1}^*, \boldsymbol{\lambda}_{k|k-1}^*) \\
q_{k|k-1}(\boldsymbol{\lambda}_k) &\propto N_n(V_{k|k-1}^*, n_{k|k-1}^*)
\end{aligned} \tag{13}$$

where the parameters are updated according to the same iterative scheme as and the expectations are exactly computed as follows:

$$\begin{aligned}
\boldsymbol{\mu}_k^p &= \boldsymbol{\mu}_{k-1}^* \\
\boldsymbol{\lambda}_k^p &= (\boldsymbol{\lambda}_{k-1}^{*-1} + \bar{\boldsymbol{\lambda}}^{-1}) \\
\boldsymbol{\mu}_{k|k-1}^* &= \boldsymbol{\lambda}_{k|k-1}^{*-1} (\langle \boldsymbol{\lambda}_k \rangle_{k|k-1} \langle \mathbf{z}_k \rangle_{k|k-1} + \boldsymbol{\mu}_k^p \boldsymbol{\lambda}_k^p) \\
\boldsymbol{\lambda}_{k|k-1}^* &= \langle \boldsymbol{\lambda}_k \rangle_{k|k-1} + \boldsymbol{\lambda}_k^p \\
n_{k|k-1}^* &= \bar{n} + 1 \\
\mathbf{S}_k^* &= (\langle \mathbf{z}_k \mathbf{z}_k^T \rangle_{q_{k|k-1}} - \langle \mathbf{z}_k \rangle_{q_{k|k-1}} \langle \boldsymbol{\mu}_k \rangle_{q_{k|k-1}}^T + \langle \boldsymbol{\mu}_k \rangle_{q_{k|k-1}} \langle \mathbf{z}_k \rangle_{q_{k|k-1}}^T + \langle \boldsymbol{\mu}_k \boldsymbol{\mu}_k^T \rangle_{q_{k|k-1}} + \bar{\mathbf{S}}^{-1})^{-1}
\end{aligned} \tag{14}$$

and the predictive expectations of the target state are now evaluated by the following expressions:

$$\begin{aligned}
\langle z_k \rangle_{q_{k|k-1}} &= \langle \mu_k \rangle_{q_{k|k-1}} \\
\langle z_k z_k^T \rangle_{q_{k|k-1}} &= \langle \lambda_k \rangle_{q_{k|k-1}}^{-1} + \langle \mu_k \rangle_{q_{k|k-1}} \langle \mu_k \rangle_{q_{k|k-1}}^T
\end{aligned} \tag{15}$$

In the next Section, these state estimation techniques (EKF, UKF, PF, and VF) are used to estimate the states variables (the leaf-area index LAI, the volumetric water content of the layer 1, HUR1 and the volumetric water content of the layer 2, HUR2) as well as the model parameters of a LSM process.

### III. SIMULATIONS RESULTS ANALYSIS

#### A. Crop model

The original data were issued from experiments carried out on a silty soil in Belgium, with a wheat crop (*Triticum aestivum* L., cultivar Julius), during 2 consecutive years, the crop seasons 2008-09 and 2009-10. The experimental blocks were prepared on two soil types (loamy and sandy loam), corresponding to the agro-environmental conditions of the Hesbaye region in Belgium. The measurements were the results of four repetitions by date, nitrogen level, soil type and crop season. Each repetition was performed on a small block (2 m × 6 m) within the original experiment as a complete randomised block distribution, spread over the field within each soil type, to ensure measurement independence. A wireless microsensors network (eKo pro series system, Crossbow) was used to continuously characterize the soil (water content, suction, temperature at two depths: 20 and 40 cm) and the atmosphere (radiation, temperature, relative humidity, wind speed) within the vegetation. Pluviometry data were also acquired in the experimental field. The plant characteristics (LAI and biomass) were also measured using reference techniques at regular intervals (2 weeks) along the crop seasons. The measurements were performed since the middle of February (around Julian day 410) till harvest. During the season 2008-2009, yields were quite high and close to the optimum of the cultivar. This was mainly explained by the good weather conditions and a sufficient nitrogen nutrition level. The season 2009-2010 was known to induce deep water stresses, and thus characterized by yield losses.

The model for which the methods were tested is Mini-STICS model. The equations are presented in [33], and the parameters given at Table 1. The dynamic equations indicate how each state variable evolves from one day to the next as a function of the current values of the state variables, of the explanatory variables, and of the parameters value. Encoding these equations over time allows one to eliminate the intermediate values of the state variables and relate the state variables at any time to the explanatory variables on each day.

In the first step we are interested to compare the estimation performances of EKF, UKF, PF and VF in estimating three state variables of the mini-STICS model: the leaf-area index LAI, the volumetric water content of the layer 1, HUR1 and the volumetric water content of the layer 2, HUR2. Based on the model equations described in [33], the mathematical model of the LAI and soil moisture (called in the rest of the document LSM model) is given by:

$$\begin{aligned}
LAI(t) &= f_1(LAI(t-1) + \theta) \\
HUR1(t) &= f_2(HUR1(t-1) + \theta) \\
HUR2(t) &= f_3(HUR2(t-1) + \theta)
\end{aligned} \tag{16}$$



Where  $t$  is the time,  $f_{1:3}$  are the corresponding model functions, and  $\theta$  is the vector of parameters driving the simulations (Table 1). LAI is the leaf area index and HUR1 (resp. HUR2) is the volumetric water content of the layer 1 (resp. the layer 2). Discretizing the model (15) using a sampling interval of  $\Delta t$  (one day), it can be written as,

$$\begin{aligned} LAI_k &= [f_1(\theta)]\Delta t + LAI_{k-1} + w_{k-1}^1 \\ HUR1_k &= [f_2(\theta)]\Delta t + HUR1_{k-1} + w_{k-1}^2 \\ HUR2_k &= [f_3(\theta)]\Delta t + HUR2_{k-1} + w_{k-1}^3 \end{aligned} \quad (17)$$

where  $w_{j \in \{1, \dots, 3\}}^j$  is a process Gaussian noise with zero mean and known variance  $\sigma_{\gamma_j}^2$ .

Table 1. Model parameters ([34])

Parameter name	Meaning	True value
ADENS (-)	Parameter of compensation between stem number and plant density	-0.8
BDENS (plants $m^{-2}$ )	Maximum density above which there is competition between plants	1.25
CROIRAC (cm degree $day^{-1}$ )	Growth rate of the root front	0.25
DLAIMAX ( $m^2$ leaves $m^{-2}$ soil degree days $^{-1}$ )	Maximum rate of the setting up of LAI	0.0078
EXTIN (-)	Extinction coefficient of photosynthetic active radiation in the canopy	0.9
KMAX (-)	Maximum crop coefficient for water requirements	1.2
LVOPT (cm root $cm^{-3}$ s)	Optimum root density	0.5
PSISTO (bars)	Absolute value of the potential of stomatal closing	10
PSISTURG (bars)	Absolute value of the potential of the beginning of decrease in the cellular extension	4
RAY ON (cm)	Average radius of roots	0.02
TCMIN ( $^{\circ}C$ )	Minimum temperature of growth	6
TCOPT ( $^{\circ}C$ )	Optimum temperature of growth	32
ZPENTE (cm)	Depth where the root density is 1/2 of the surface root density for the reference profile	120
ZPRLIM (cm)	Maximum depth of the root profile for the reference profile	150

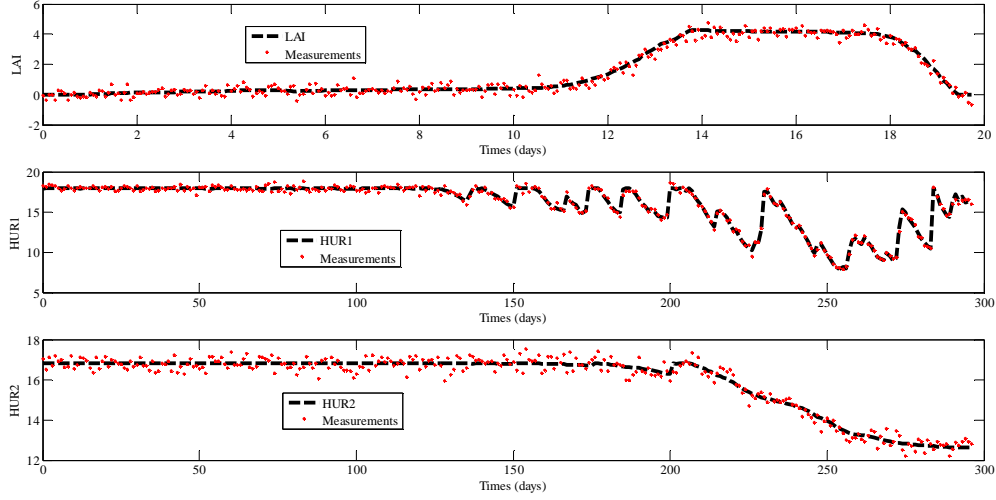
## B. Sampling data generation

To obtain original dynamic data, the model was first used to simulate the temporal responses  $LAI_k$ ,  $HUR1_k$ ,  $HUR2_k$  on basis of the recorded climatic variables of the crop season “2008-2009”. The sampling time used for discretization was 1 day.

Moreover, to characterize the ability of the different approaches to estimate both the states and the parameters at same time, “true” parameter values were chosen (Table 1). The advantage of working by simulation rather than on real data is that the true parameter values are known. It is thus possible to calculate the quality of the estimated parameters and the predictive quality of the adjusted model for each method. The drawback is that the generality of the results is hard to know. The results may depend on the details of the model, on the way the data are generated and on the specific data that are used. The simulated values, assumed to be noise free, are shown in Figure 1. The evolution of LAI during the

wheat's lifecycle presents the three expected phases, growth, stability, and senescence. Daily variations of shallow ground water show fluctuations that were damped in the subsoil layer 30 – 50 cm.

These simulated states were then contaminated with zero mean Gaussian errors, i.e., the measurement noise  $v_{k-1} \sim N(0, \sigma_v^2)$ , where  $\sigma_v^2 = 0.1$ .



**Figure 1** Simulated LSM data used in estimation: state variables (LAI leaf area index, HUR1 volumetric water content of the layer 1; HUR2 volumetric water content of the layer 2).

### C. Comparative Study 1: Estimation of State Variables from Noisy Measurements

In the first scenario, it is assumed that the model parameters are known and the objective is to estimate only the state variables, LAI, HUR1, and HUR2. Therefore, we consider the state vector that we wish to estimate as:

$$z_k = x_k = [LAI_k \ HUR1_k \ HUR2_k] \quad (18)$$

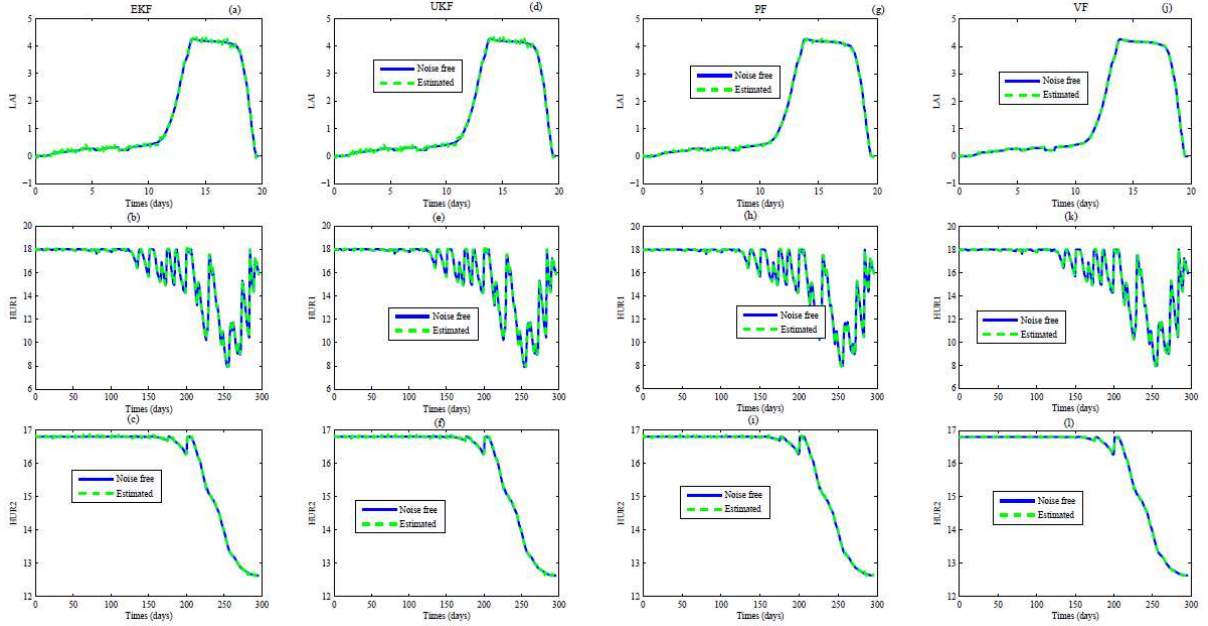
Eventually, to perform comparison between the techniques, the estimation root mean square errors (RMSE) criteria are used and calculated on the states (with respect to the noise free data)

$$RMSE = \sqrt{E((x - \hat{x})^2)} \quad (19)$$

where  $x$  (resp.  $\hat{x}$ ) is the true parameter/state (resp. the estimated parameter/state).

The simulation results of estimating the three states  $LAI_k$ ,  $HUR1_k$  and  $HUR2_k$  using EKF, UKF, PF, and VF are shown in Figures 2(a,b,c), Figures 2(d,e,f), Figures 2(g,h,i), Figures 2(j,k,l) respectively. Also, the estimation root mean square errors (RMSE) for the estimated states are presented in Table 2. It can be observed from Figure 3 and Table 2 that EKF resulted in the worst performance of all

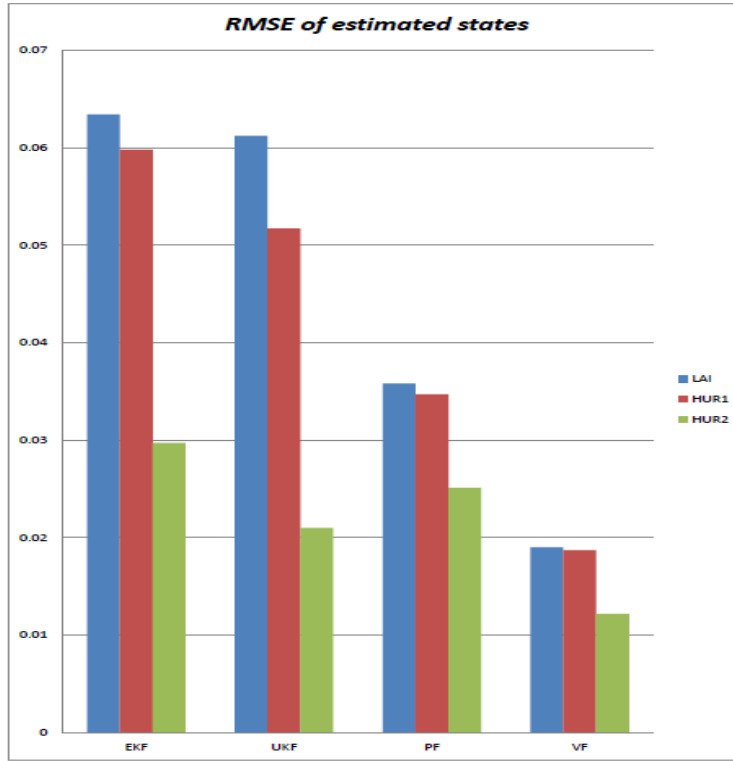
estimation techniques, which is expected due to the limited ability of EKF to accurately estimate the mean and covariance matrix of the estimated states through linearization of the nonlinear process model. The results also show that the PF provides a significant improvement over the UKF. This is because the covariance is propagated through linearization of the underlying non-linear model, when the state transition and observation models are highly non-linear. VF provides a significant improvement over the PF, which is due to the fact that the VF yields an optimal choice of the sampling distribution,  $p(\alpha_k | \alpha_{k-1}, y_k)$ , by minimizing a KL divergence criterion that also utilizes the observed data  $y_k$ .



**Figure 2 Estimation of state variables using various state estimation techniques (comparative study 1).**

*Table 2. Root mean square errors (RMSE) of estimated states*

Technique	RMSE		Mean at steady state	
	LAI m2leavesm- 2soil	HUR1 (%)	HUR2 (%)	ADENS (-)
EKF	0.0634	0.0598	0.0297	-0.8
UKF	0.0612	0.0517	0.0201	-0.8
PF	0.0358	0.0347	0.0251	-0.8
VF	0.0190	0.0187	0.0122	-0.8



**Figure 3** Histogram showing the RMSE of estimated states using EKF, UKF, PF and VF

#### **D. Comparative Study 2: Simultaneous Estimation of State Variables and Model Parameters**

The model (17) assumes that the parameters are fixed and/or have been determined previously. However, the model involves several parameters that are usually not exactly known, or that have to be estimated. Estimating these parameters, to completely define the model, usually requires several experiment setups, which can be expensive and challenging in practice. In a second step, in this work, we propose to use a Bayesian approach that can considerably simplify the task of modeling compared to the conventional experimentally intensive methods. Let's thus consider that some of the parameters have to be estimated to improve the simulations, by example the *ADENS*, *DLAIMAX*, *PSISTURG*, *EXTIN*, and *LVOPT* parameters. *ADENS* is the parameter of compensation between stem number and plant density, *DLAIMAX* is the maximum rate of the setting up of LAI, *PSISTURG* is the absolute value of the potential of the beginning of decrease in the cellular extension, *EXTIN* is the extinction coefficient of photosynthetic active radiation in the canopy, and *LVOPT* is the optimum root density. To estimate these parameters, the following equations that describe their evolution are also needed:

$$ADENS_k = ADENS_{k-1} + \gamma_{k-1}^1$$

$$DLAIMAX_k = DLAIMAX_{k-1} + \gamma_{k-1}^2$$

$$PSISTURG_k = PSISTURG_{k-1} + \gamma_{k-1}^3$$

$$\begin{aligned}
EXTIN_k &= EXTIN_{k-1} + \gamma_{k-1}^4 \\
LVOPT_k &= LVOPT_{k-1} + \gamma_{k-1}^5
\end{aligned} \tag{20}$$

where  $\gamma_{j \in \{1, \dots, 5\}}^j$  is a process Gaussian noise with zero mean and known variance  $\sigma_\gamma^2$ . Combining (19) and (20), one obtains:

$$\begin{aligned}
f_1 : LAI_k &= [g_1(\theta_{k-1})]\Delta t + LAI_{k-1} + w_{k-1}^1 \\
f_2 : HUR1_k &= [g_2(\theta_{k-1})]\Delta t + HUR1_{k-1} + w_{k-1}^2 \\
f_3 : HUR2_k &= [g_3(\theta_{k-1})]\Delta t + HUR2_{k-1} + w_{k-1}^3 \\
f_4 : ADENS_k &= ADENS_{k-1} + \gamma_{k-1}^1 \\
f_5 : DLAIMAX_k &= DLAIMAX_{k-1} + \gamma_{k-1}^2 \\
f_6 : PSISTURG_k &= PSISTURG_{k-1} + \gamma_{k-1}^3 \\
f_7 : EXTIN_k &= EXTIN_{k-1} + \gamma_{k-1}^4 \\
f_8 : LVOPT_k &= LVOPT_{k-1} + \gamma_{k-1}^5
\end{aligned} \tag{21}$$

where  $f_{k \in \{1, \dots, 8\}}$  are some nonlinear functions and where  $w = (w_1, w_2, w_3)^T$  and  $\gamma = (\gamma_1, \gamma_2, \gamma_3)^T$  are respectively the measurement and process noise vector, which quantify randomness at both levels. In other words, we are forming the augmented state:  $z_k = (x_k, \theta_k)^T$  which is the vector that we wish to estimate. It can be given by a 8 by 1 matrix:

$$\begin{aligned}
x_k(1, :) &\rightarrow LAI_k \\
x_k(2, :) &\rightarrow HUR1_k \\
x_k(3, :) &\rightarrow HUR2_k \\
\theta_k(1, :) &\rightarrow ADENS_k \\
\theta_k(2, :) &\rightarrow DLAIMAX_k \\
\theta_k(3, :) &\rightarrow PSISTURG_k \\
\theta_k(4, :) &\rightarrow EXTIN_k \\
\theta_k(5, :) &\rightarrow LVOPT_k
\end{aligned} \tag{22}$$

In the following, we denote  $w = (w_1 w_2 w_3)^T$ , and  $\gamma = (\gamma_1 \gamma_2 \gamma_3 \gamma_4 \gamma_5)^T$ , respectively the measurement and process noise vectors, which quantify (i) errors in the measurements and (ii) randomness in the process. The idea here is that, if a dynamic model structure is available, the model parameters can be estimated using one of state estimation technique, namely EKF, UKF, PF, and VF. To characterize the ability of the different approaches to estimate both at same time, the states and the parameters, we have chosen true parameter values and then tested each technique to see how well it could retrieve these true parameter values given the data. It was thus possible to calculate the quality of the estimated parameters and the predictive quality of the adjusted model for each method.

To go further in the analysis, the effect of the number estimated parameters on the estimation performances are investigated. Five cases summarized below are considered. In all cases, it is assumed that three states ((LAI, HUR1, and HUR2)) are measured.

- i) Case 1: the three states (LAI, HUR1 and HUR2) along with the parameter ADENS be estimated.

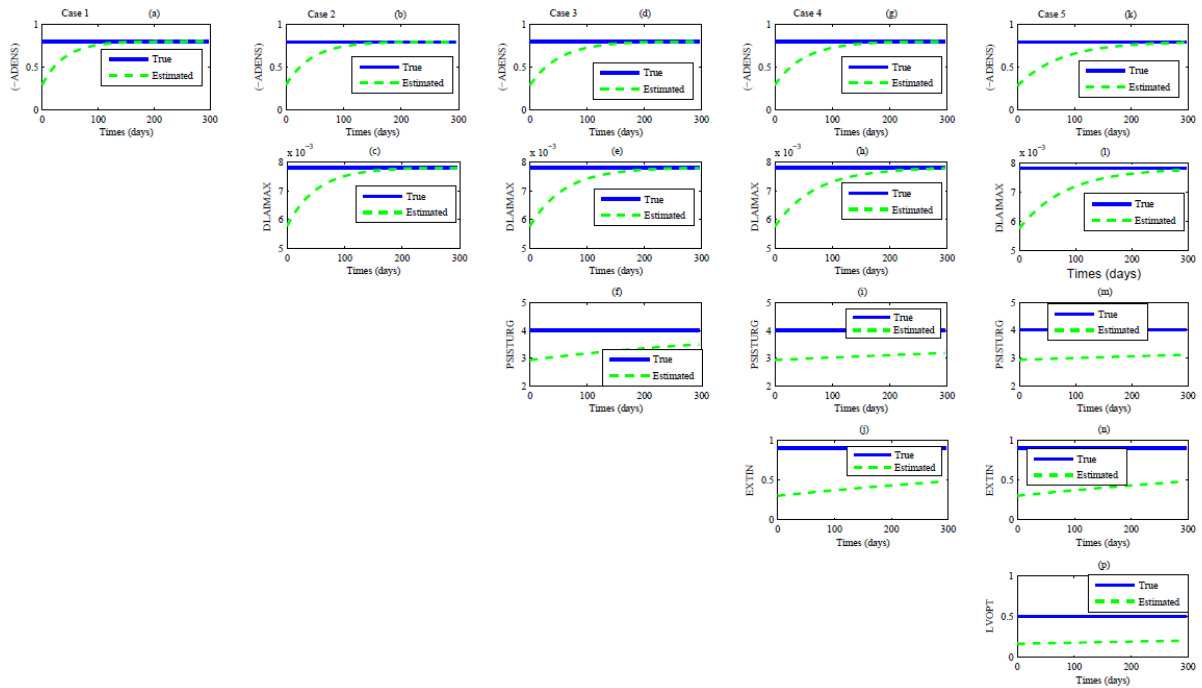
- ii) Case 2: the three states (LAI, HUR1 and HUR2) along with the parameters ADENS, and DLAIMAX will be estimated.
- iii) Case 3: the three states (LAI, HUR1 and HUR2) along with the parameters ADENS, DLAIMAX and PSISTURG will be estimated.
- iv) Case 4: the three states (LAI, HUR1 and HUR2) along with the parameters ADENS, DLAIMAX, PSISTURG, and EXTIN will be estimated.
- v) Case 5: the three states (LAI, HUR1 and HUR2) along with the parameters ADENS, DLAIMAX, PSISTURG, EXTIN and LVOPT will be estimated.

Here, we are interested in examining the effect of the number of estimated parameters on the estimation performances of EKF, UKF, PF and VF and in estimating the states and parameters of the LSM process model, during the first crop season 2008-2009 (unstressed growth data). The estimation of the state variables and parameter(s) for these five cases were performed using the four state estimation techniques, EKF, UKF, PF, and VF, and the estimation results for the model parameters using these techniques are shown in Figures 4, 5, 6, and 7, respectively. For example, Figure 3(a) shows the estimation of the parameter DLAIMAX in case 1, Figures 4(b,c) show the estimation of the parameters DLAIMAX and ADENS in case 2, and Figures 4(d,e,f) show the estimation of all five parameters ADENS, DLAIMAX, PSISTURG, EXT IN, and LV OPT in case 5. Also, Tables 3 to 7 compare the performances of the four estimation techniques for the five cases. For example, for case 1, Table 3 compares the estimation mean square errors for the three state variables LAI, HUR1 and HUR2 (with respect to the noise-free data) and the mean of the estimated parameter DLAIMAX at steady state (i.e., after convergence of parameter(s)). Tables 4 to 7 present similar comparisons for cases 2-5, respectively. Moreover, Figures 8 to 12 present histograms comparing the estimation RMSE for the three state variables (LAI, HUR1 and HUR2), respectively, using the various state estimation techniques.

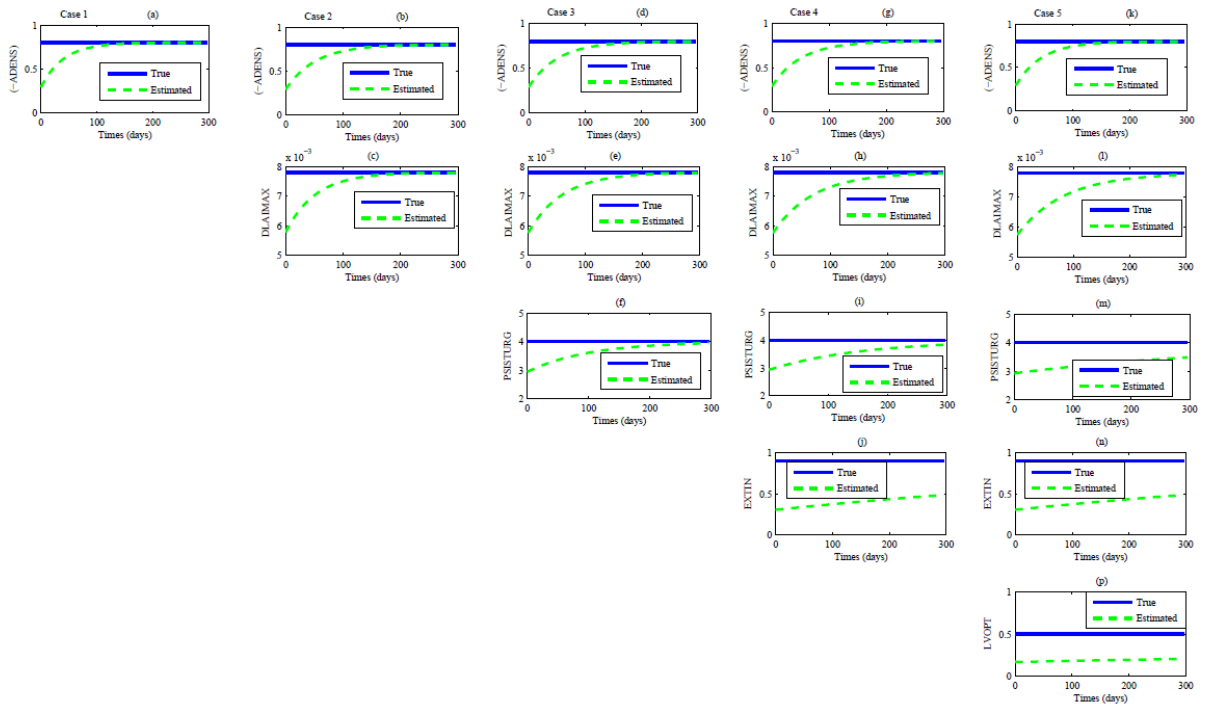
It can be seen from the results presented in Tables 3-7 that in all cases, the PF outperforms UKF and EKF (i.e., provides smaller RMSE for the state variables), and that the VF shows a significant improvement over all other techniques. These results confirm those obtained in the first comparative study, where only the state variables are estimated. The advantages of the VF over the PF (and the PF over the the UKF and EKF) can also be seen through their abilities to estimate the model parameters. For example, EKF could perfectly estimate one parameter in case 1 (see Figure 4(a)), but it took longer to estimate a second parameter in case 2 (see Figures 4(b,c)), and it could not converge for the third parameter in case 3 (see Figure 4(d,e,f)), where it is used to estimate three parameters. While, UKF could estimate one parameter in case 1 (see Figure 5(a)) and two parameters in case 2 (see Figure 5(b,c)), but it took longer to estimate a third parameter in case 3 (see Figures 5(d,e,f)), and it could not converge for the fourth and the fifth parameters in cases 4 and 5 (see Figures 5(g,h,i,j) and 5(k,l,m,n,p)), where it is used to estimate all five parameters. The PF, on the other hand, could estimate all parameters in all cases 1-5, even though it took longer to converge in case 5, where all five parameters are estimated (see Figure 6). The VF, however, could estimate all parameters in all five cases (see Figure 7), and converged faster than all other techniques. These advantages of the VF are due to the fact it provides an optimum choice of the sampling distribution used to approximate the posterior density function, which also accounts for the observed data.

The results also show that the number of estimated parameters affect the estimation accuracy of the estimated state variables. In other words, for all estimation techniques, the estimation RMSE of LAI, HUR1 and HUR2 increases from the first comparative study (where only the state variables are estimated) to case 1 (where only one parameter, DLAIMAX, is estimated) to case 5 (where all five parameters, ADENS, DLAIMAX, PSISTURG, EXT IN, and LV OPT , are estimated). For example, the RMSEs obtained using EKF for LAI in the first comparative study and cases 1-5 of the second comparative study are 0.0634, 0.0814, 0.105, 0.131, 0.187 and 0.221, respectively, which increase as the

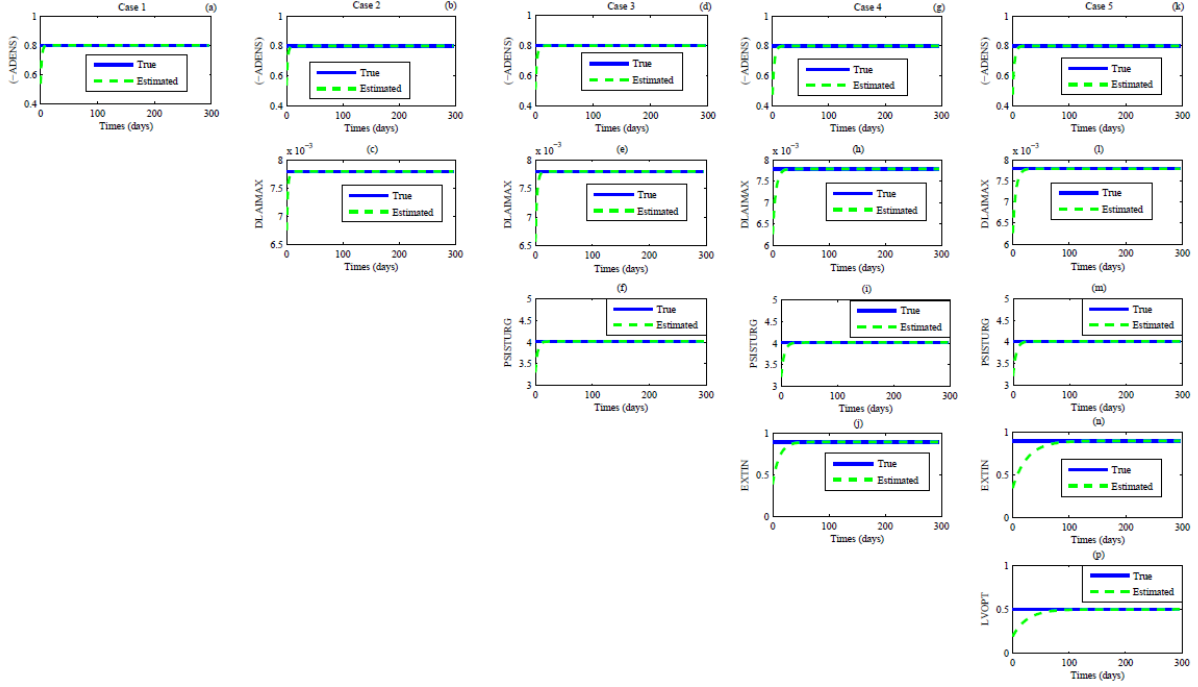
number of estimated parameters increases (refer to Tables 3-7). This observation is valid for the other state variables HUR1 and HUR2 and for all other estimation techniques, UKF, PF and VF.



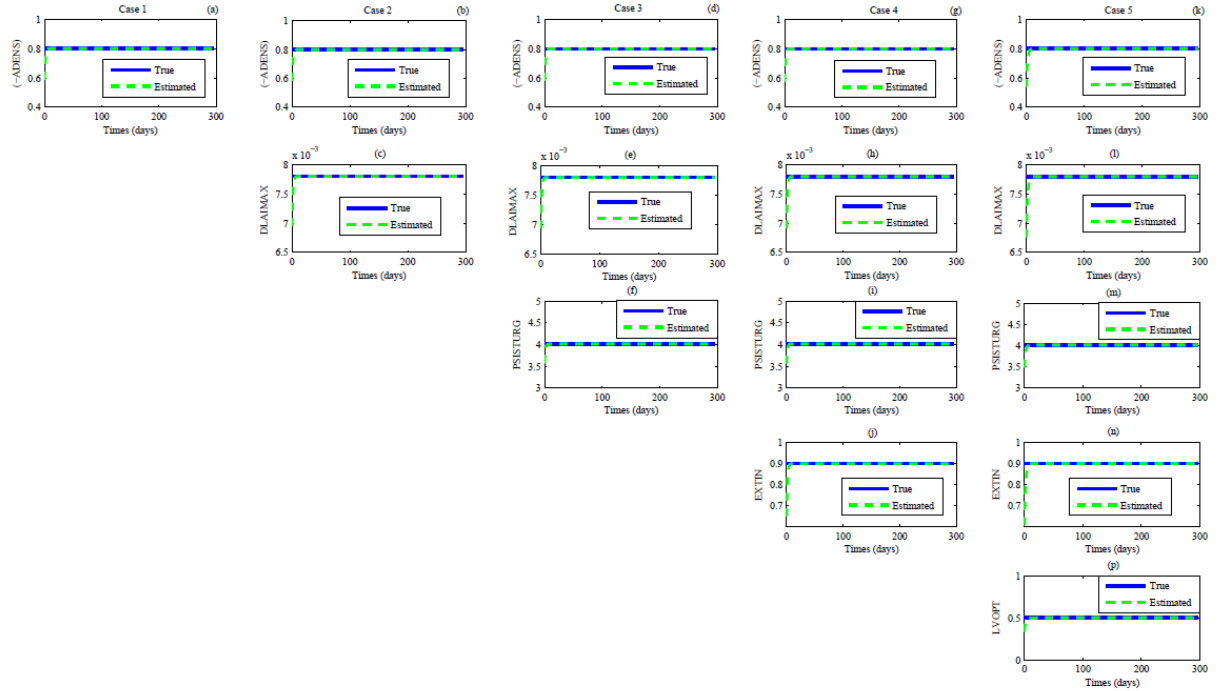
**Figure 4** Estimation of the LSM model parameters using EKF for all cases - case 1: (a), case 2: (b),(c), case 3: (d),(e),(f), case 4: (g),(h),(i),(j), case 5: (k),(l),(m),(n),(o),(p).



**Figure 5 Estimation of the LSM model parameters using UKF for all cases - case 1: (a), case 2: (b),(c), case 3: (d),(e),(f), case 4: (g),(h),(i),(j), case 5: (k),(l),(m),(n),(o),(p).**



**Figure 6 Estimation of the LSM model parameters using PF for all cases - case 1: (a), case 2: (b),(c), case 3: (d),(e),(f), case 4: (g),(h),(i),(j), case 5: (k),(l),(m),(n),(o),(p).**





**Figure 7 Estimation of the LSM model parameters using VF for all cases - case 1: (a), case 2: (b),(c), case 3: (d),(e),(f), case 4: (g),(h),(i),(j), case 5: (k),(l),(m),(n),(p).**

*Table 3. Root mean square errors (RMSE) of estimated states and mean of estimated parameter - case 1*

Technique	RMSE			Mean at steady state
	LAI m <sup>2</sup> leaves m <sup>-2</sup> soil	HUR1 (%)	HUR2 (%)	ADENS (-)
EKF	0.0814	0.0841	0.0411	-0.8
UKF	0.0758	0.0798	0.0358	-0.8
PF	0.0511	0.0581	0.0315	-0.8
VF	0.0315	0.0317	0.0278	-0.8

*Table 4. Root mean square errors of estimated states and mean of estimated parameters - case 2*

Technique	RMSE			Mean at steady state	
	LAI m <sup>2</sup> leaves m <sup>-2</sup> soil	HUR1 (%)	HUR2 (%)	ADENS (-)	DLAIMAX (m <sup>2</sup> leavesm <sup>-2</sup> soildegree days <sup>-1</sup> )
EKF	0.105	0.119	0.0493	-0.8	0.0078
UKF	0.0944	0.104	0.0423	-0.8	0.0078
PF	0.0798	0.091	0.0412	-0.8	0.0078
VF	0.0548	0.0651	0.0298	-0.8	0.0078

*Table 5. Root mean square errors (RMSE) of estimated states and mean of estimated parameters - case 3*

Technique	RMSE			Mean at steady state		
	LAI m <sup>2</sup> leavesm <sup>-2</sup> soil	HUR1 (%)	HUR2 (%)	ADENS (-)	DLAIMAX (m <sup>2</sup> leaves m <sup>-2</sup> soil degree days <sup>-1</sup> )	PSISTURG (bars)
EKF	0.131	0.142	0.061	-0.8	0.0078	Did not converge (DNC)
UKF	0.125	0.134	0.054	-0.8	0.0078	4
PF	0.117	0.1023	0.045	-0.8	0.0078	4
VF	0.094	0.0758	0.0354	-0.8	0.0078	4

*Table 6. Root mean square errors (RMSE) of estimated states and mean of estimated parameters - case 4*

Technique	RMSE			Mean at steady state			
	LAI m <sup>2</sup> leavesm <sup>-2</sup> soil	HUR1 (%)	HUR2 (%)	ADENS (-)	DLAIMAX (m <sup>2</sup> leaves m <sup>-2</sup> soil degree days <sup>-1</sup> )	PSISTURG (bars)	EXTIN (bars)
EKF	0.187	0.202	0.112	-0.8	0.0078	DNC	DNC
UKF	0.165	0.189	0.092	-0.8	0.0078	DNC	DNC
PF	0.141	0.149	0.085	-0.8	0.0078	4	0.9
VF	0.113	0.093	0.0614	-0.8	0.0078	4	0.9

Table 7. Root mean square errors (RMSE) of estimated states and mean of estimated parameters - case 5

Technique	RMSE			Mean at steady state				
	LAI m <sup>2</sup> leavesm <sup>-2</sup> soil	HUR1 (%)	HUR2 (%)	ADENS (-)	DLAIMAX (m <sup>2</sup> leaves m <sup>-2</sup> soil degree days <sup>-1</sup> )	PSISTURG (bars)	EXTIN (bars)	LVOPT (-)
EKF	0.221	0.235	0.178	-0.8	0.0078	DNC	DNC	DNC
UKF	0.204	0.214	0.196	-0.8	0.0078	DNC	DNC	DNC
PF	0.149	0.158	0.103	-0.8	0.0078	4	0.9	0.5
VF	0.125	0.102	0.096	-0.8	0.0078	4	0.9	0.5

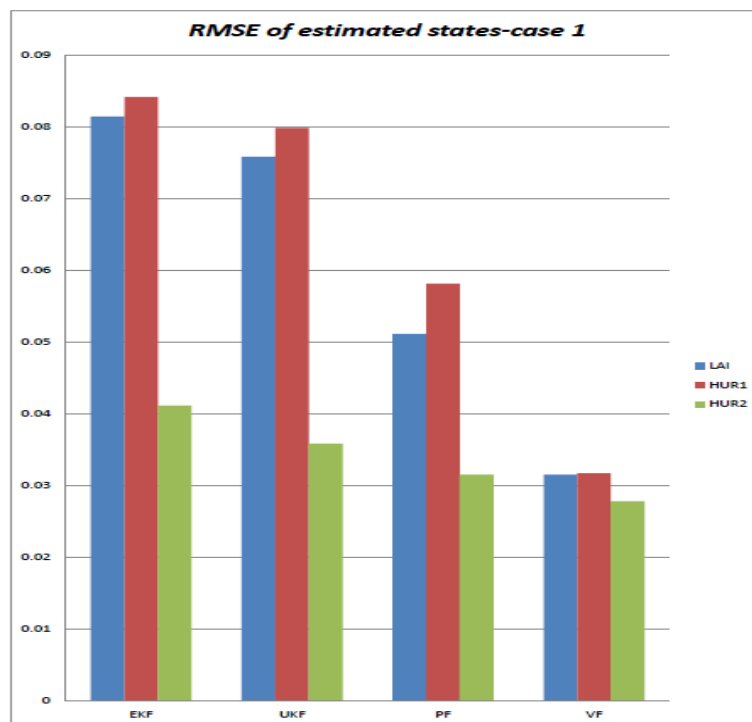


Figure 8 Histogram showing the RMSE of estimated states for case 1 using EKF, UKF, PF and VF

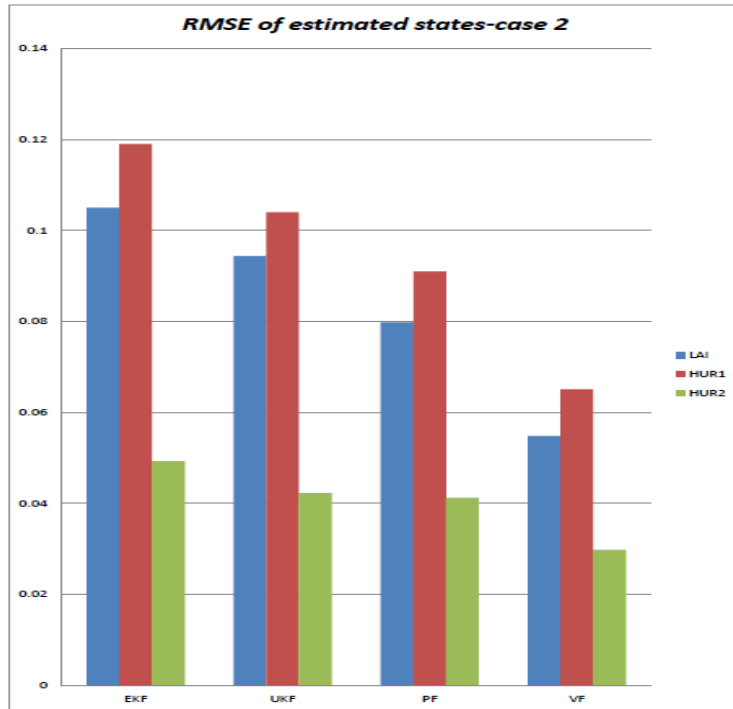


Figure 9 Histogram showing the RMSE of estimated states for case 2 using EKF, UKF, PF and VF

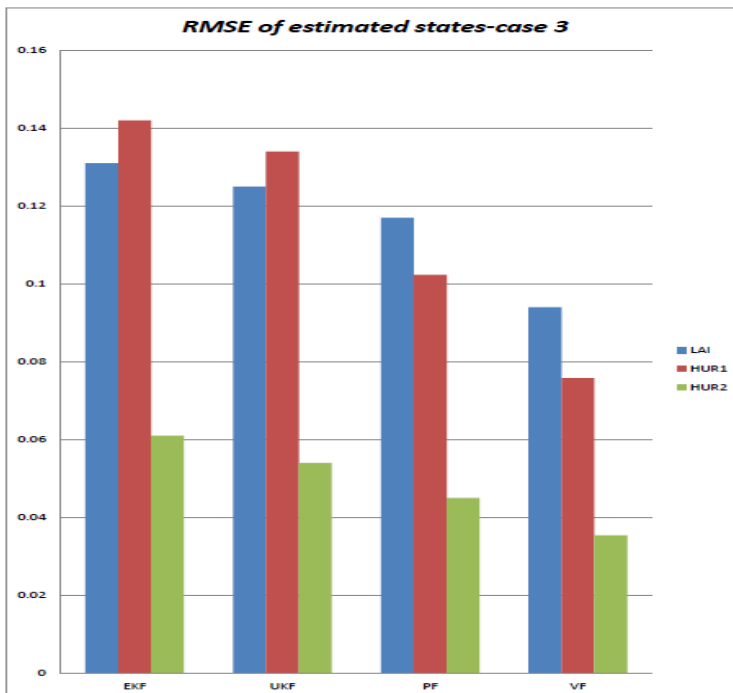


Figure 10 Histogram showing the RMSE of estimated states for case 3 using EKF, UKF, PF and VF

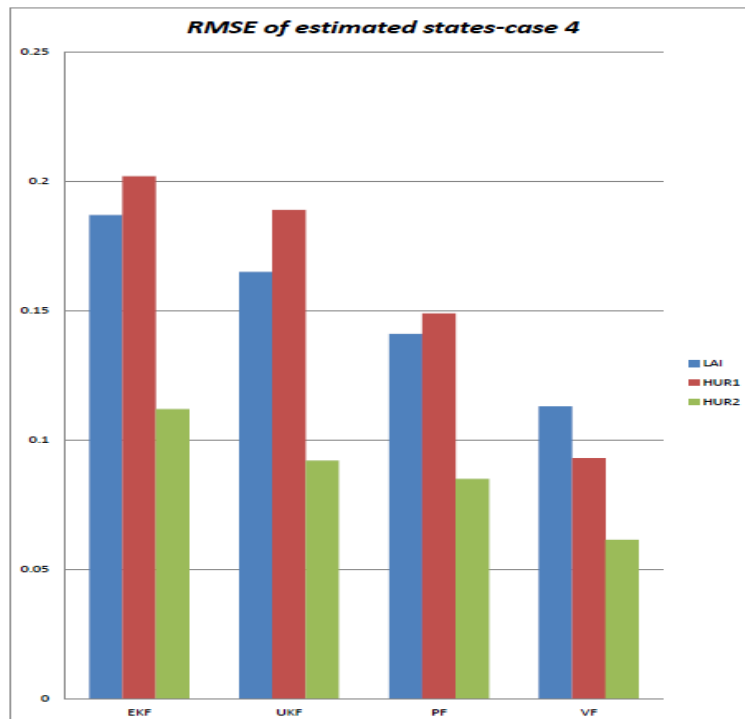


Figure 11 Histogram showing the RMSE of estimated states for case 4 using EKF, UKF, PF and VF

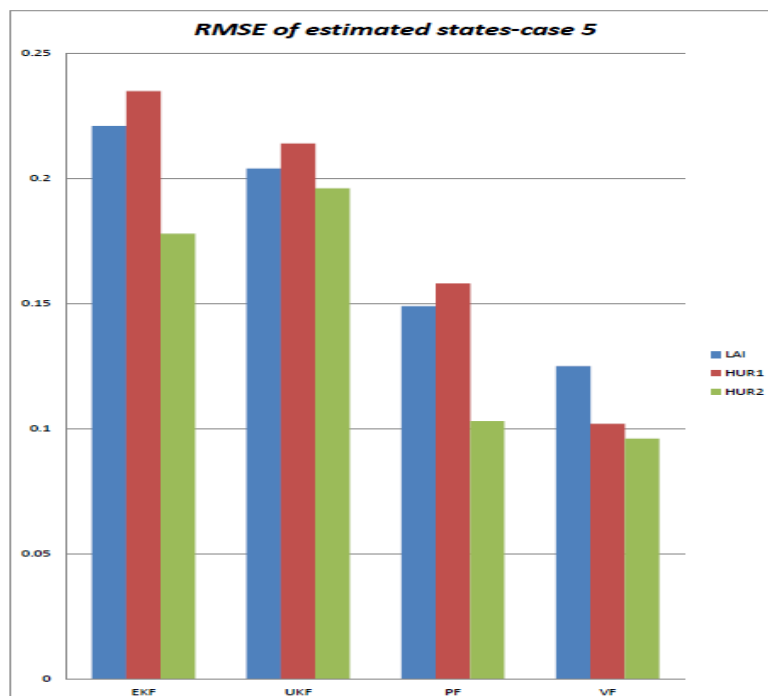


Figure 12 Histogram showing the RMSE of estimated states for case 5 using EKF, UKF, PF and VF

### E. The effect of driving variables

Eventually, to achieve the research, the effect of the driving variables are determined. The five studied cases presented at step 2 are applied on the climatic weather database of the crop season 2009-2010. This procedure was furthermore developed since the PSISTURG parameter was optimised at cases 2 and 3, while the climatic season 2008-2009 was known not to induce stresses. We applied the different algorithms described above, EKF, UKF, PF and VF to simulate the responses of LAI, HUR1 and HUR2 as functions of time, in the second crop season 2009-2010 (season with deep water stresses). And respectively the above techniques are used for estimating the five model parameters, ADENS, DLAIMAX, PSISTURG, EXTIN, and LVOPT. From Tables 8 to 12, we can show that variational Bayesian filtering algorithm outperforms the classical algorithms, and demonstrate the performance and the good behavior of the proposed algorithm when the growing season is varied. Moreover, Figures 13 to 17 present histograms comparing the estimation RMSE for the three state variables (LAI, HUR1 and HUR2), respectively, using the four state estimation techniques EKF, UKF, PF and VF.

Table 8. Root mean square errors (RMSE) of estimated states and mean of estimated parameter - case 1

Technique	RMSE			Mean at steady state	
	LAI m <sup>2</sup> leavesm- 2soil	HUR1 (%)	HUR2 (%)	ADENS (-)	
EKF	0.0939	0.0901	0.0461	-0.8	
UKF	0.0714	0.0745	0.0387	-0.8	
PF	0.0542	0.0531	0.0342	-0.8	
VF	0.0341	0.0357	0.0232	-0.8	

Table 9. Root mean square errors of estimated states and mean of estimated parameters - case 2

Technique	RMSE			Mean at steady state	
	LAI m <sup>2</sup> leavesm- 2soil	HUR1 (%)	HUR2 (%)	ADENS (-)	DLAIMAX (m <sup>2</sup> leaves m-2 soil degree days-1)
EKF	0.1502	0.1589	0.0723	-0.8	0.0078
UKF	0.122	0.1325	0.0618	-0.8	0.0078
PF	0.0988	0.0957	0.0592	-0.8	0.0078
VF	0.0638	0.0557	0.0364	-0.8	0.0078

Table 10. Root mean square errors (RMSE) of estimated states and mean of estimated parameters - case 3

Technique	RMSE			Mean at steady state		
	LAI m <sup>2</sup> leavesm- 2soil	HUR1 (%)	HUR2 (%)	ADENS (-)	DLAIMAX (m <sup>2</sup> leaves m-2 soil degree days-1)	PSISTURG (bars)
EKF	0.213	0.204	0.141	-0.8	0.0078	DNC
UKF	0.183	0.175	0.115	-0.8	0.0078	4
PF	0.1431	0.1433	0.0945	-0.8	0.0078	4
VF	0.0847	0.0859	0.0594	-0.8	0.0078	4

Table 11. Root mean square errors (RMSE) of estimated states and mean of estimated parameters - case 4

Technique	RMSE			Mean at steady state			
	LAI <small>m<sup>2</sup>leavesm- 2soil</small>	HUR1 (%)	HUR2 (%)	ADENS (-)	DLAIMAX <small>(m<sup>2</sup> leaves m-2 soil degree days-1)</small>	PSISTURG (bars)	EXTIN (bars)
EKF	0.2161	0.235	0.155	-0.8	0.0078	DNC	DNC
UKF	0.191	0.195	0.124	-0.8	0.0078	DNC	DNC
PF	0.1435	0.1454	0.096	-0.8	0.0078	4	0.9
VF	0.0853	0.087	0.070	-0.8	0.0078	4	0.9

Table 12. Root mean square errors (RMSE) of estimated states and mean of estimated parameters - case 5

Technique	RMSE			Mean at steady state				
	LAI <small>m<sup>2</sup>leavesm- 2soil</small>	HUR1 (%)	HUR2 (%)	ADENS (-)	DLAIMAX <small>(m<sup>2</sup> leaves m-2 soil degree days-1)</small>	PSISTURG (bars)	EXTIN (bars)	LVOPT (-)
EKF	0.226	0.255	0.168	-0.8	0.0078	DNC	DNC	DNC
UKF	0.196	0.212	0.135	-0.8	0.0078	DNC	DNC	DNC
PF	0.145	0.1464	0.0975	-0.8	0.0078	4	0.9	0.5
VF	0.085	0.0902	0.0814	-0.8	0.0078	4	0.9	0.5

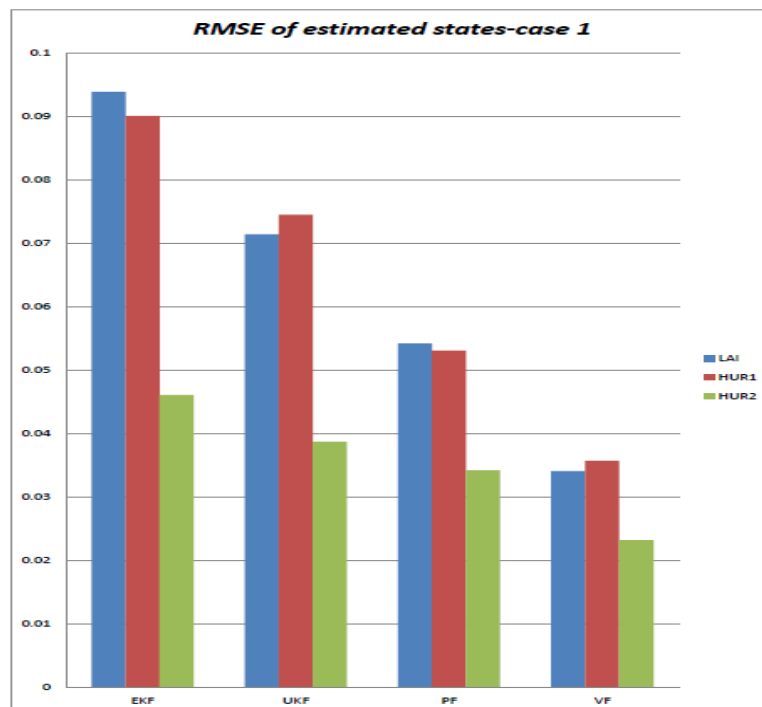


Figure 13 Histogram showing the RMSE of estimated states for case 1 using EKF, UKF, PF and VF

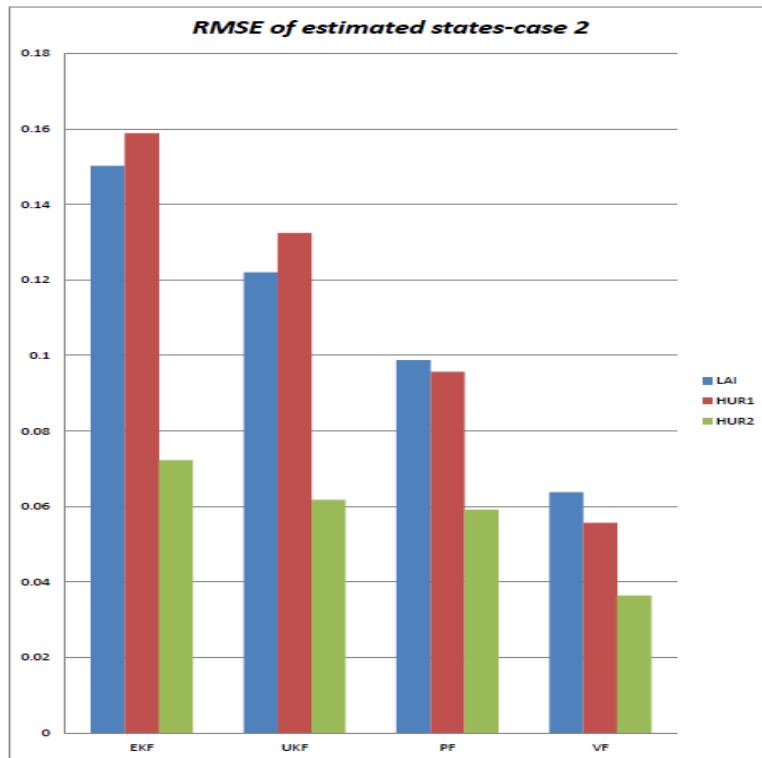


Figure 14 Histogram showing the RMSE of estimated states for case 2 using EKF, UKF, PF and VF

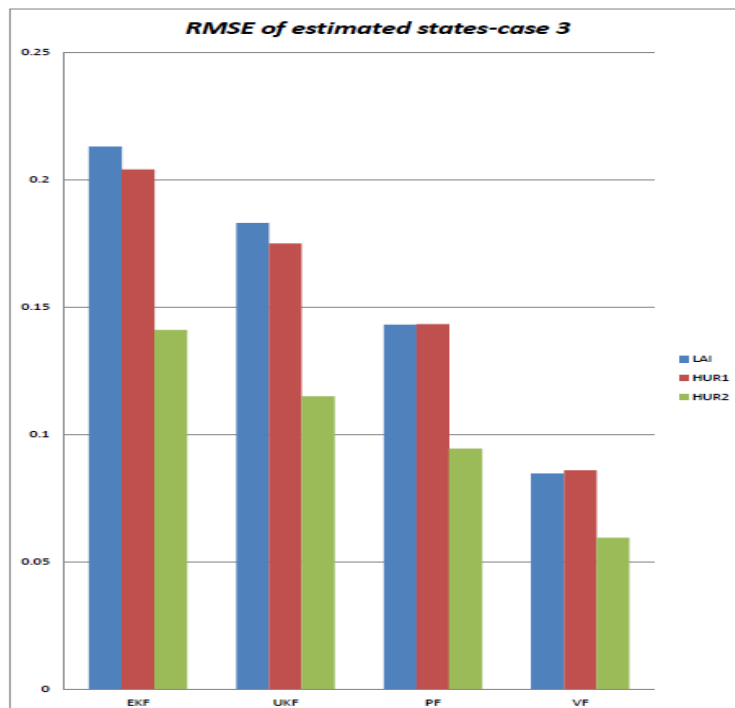


Figure 15 Histogram showing the RMSE of estimated states for case 3 using EKF, UKF, PF and VF

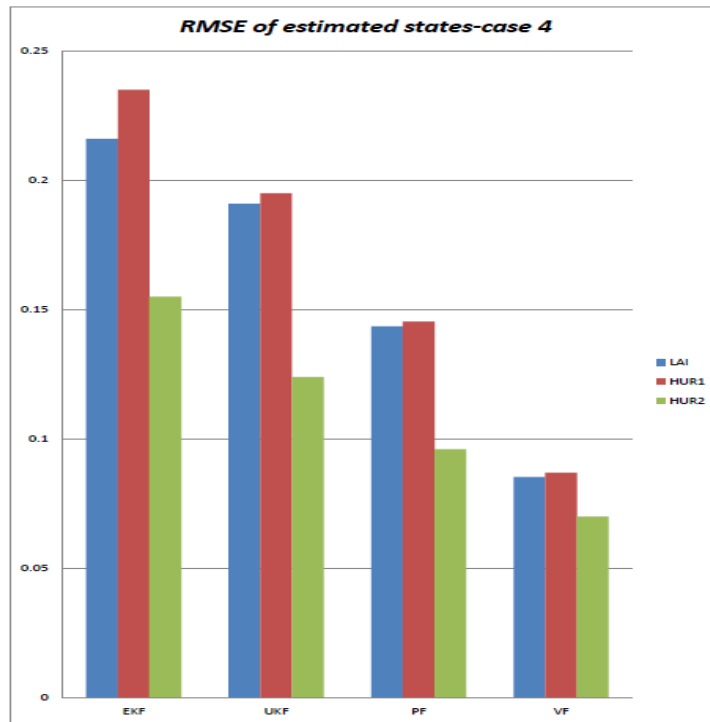


Figure 16 Histogram showing the RMSE of estimated states for case 4 using EKF, UKF, PF and VF

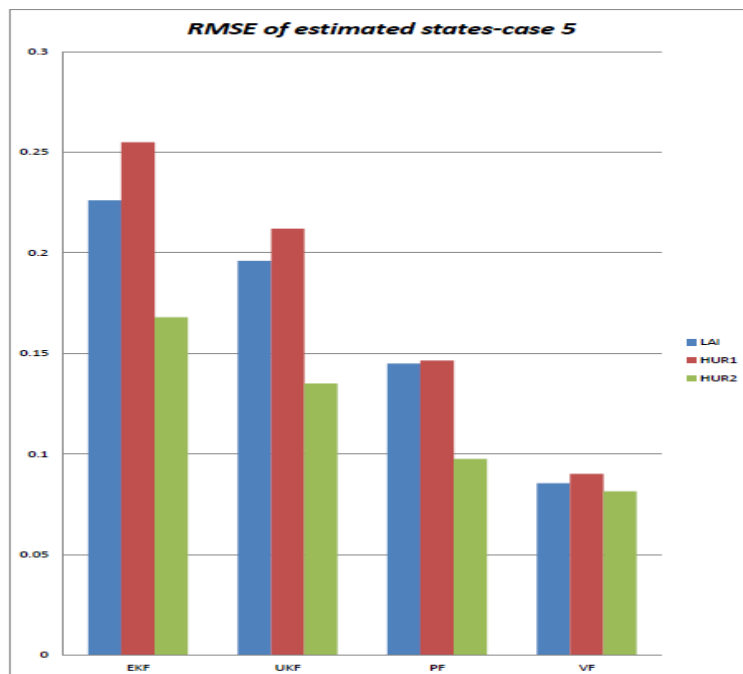


Figure 17 Histogram showing the RMSE of estimated states for case 5 using EKF, UKF, PF and VF



## IV. CONCLUSIONS

In this chapter, state estimation techniques are used to predict simultaneously three state variates (Leaf area index (LAI) and soil moisture of two soil layers, HUR1 and HUR2) for a winter wheat crop model and several parameters. Various state estimation techniques, which include the extended Kalman filter (EKF), unscented Kalman filter (UKF), particle filter (PF), and variational Bayesian filter (VF) are compared as they are used to achieve this objective.

These techniques (i.e., EKF, UKF, PF and VF) have been extended to better handle nonlinear and non-Gaussian processes with no a priori information on the state, by utilizing a time-varying assumption of statistical parameters. The time-varying assumption of statistical parameters is assumed here to further capture the uncertainty of the state distribution.

Two comparative studies have been conducted to compare the estimation performances of these four estimation techniques. In the first comparative study, EKF, UKF, PF and VF are used to estimate the three state variables, LAI, HUR1, and HUR2. In the second comparative study, the state variables and the model parameters are simultaneously estimated, and the effect of number of estimated parameters on the performances of the four estimation techniques is investigated.

The simulation results of both comparative studies show that the PF provides a higher accuracy than the EKF and the UKF due to the limited ability of the EKF and the UKF to deal with highly nonlinear process models. The results also show that the VF provides a significant improvement over the PF. This is because, unlike the PF which depends on the choice of sampling distribution used to estimate the posterior distribution, the VF yields an optimum choice of the sampling distribution, which also utilizes the observed data. The results of the second comparative study show that, for all techniques, estimating more model parameters affects the estimation accuracy as well as the convergence of the estimated states and parameters. The VF, however, still provides advantages over other methods with respect to estimation accuracy as well convergence.

## V. REFERENCES

- [1] J. Williams, C. Jones, J. Kiniry, and D. Spanel, "The Epic crop growth model," *Trans. ASAE*, vol. 32, no. 2, pp. 497–511, 1989.
- [2] C. Diepen, J. Wolf, H. Keulen, and C. Rappoldt, "Wofost: a simulation model of crop production," *Soil use and management*, vol. 5, no. 1, pp. 16–24, 1989.
- [3] S. Hansen, H. Jensen, N. Nielsen, and H. Svendsen, *NPo-research, A10: DAISY: Soil Plant Atmosphere System Model*. Miljøstyrelsen, 1990.
- [4] N. Brisson, B. Mary, D. Ripoche, M. Jeuffroy, F. Ruget, B. Nicoulaud, P. Gate, F. Devienne-Barret, R. Antonioletti, C. Durr et al., "Stics: a generic model for the simulation of crops and their water and nitrogen balances. i. theory, and parameterization applied to wheat and corn," *Agronomie*, vol. 18, no. 5-6, pp. 311–346, 1998.
- [5] B. Basso and J. Ritchie, "Impact of compost, manure and inorganic fertilizer on nitrate leaching and yield for a 6-year maize-alfalfa rotation in michigan," *Agriculture, ecosystems & environment*, vol. 108, no. 4, pp. 329–341, 2005.
- [6] R. E. Kalman, "A new approach to linear filtering and prediction problem," *Trans. ASME, Ser. D, J. Basic Eng.*, vol. 82, pp. 34–45, 1960.

- [7] D. Simon, *Optimal State Estimation: Kalman, H1, and Nonlinear Approaches*. John Wiley and Sons, 2006.
- [8] M. Grewal and A. Andrews, *Kalman Filtering: Theory and Practice using MATLAB*. John Wiley and Sons, 2008.
- [9] V. Aidala, "Parameter estimation via the kalman filter," *IEEE Trans. on Automatic Control*, vol. 22, no. 3, pp. 471–472, 1977.
- [10] L. Matthies, T. Kanade, and R. Szeliski, "Kalman filter-based algorithms for estimating depth from image sequences," *International Journal of Computer Vision*, vol. 3, no. 3, pp. 209–238, 1989.
- [11] G. Chen, Q. Xie, and L. Shieh, "Fuzzy kalman filtering," *Journal of Information Science*, vol. 109, pp. 197–209, 1998.
- [12] D. Simon, "Kalman filtering of fuzzy discrete time dynamic systems," *Applied Soft Computing*, vol. 3, pp. 191–207, 2003.
- [13] H. Nounou and M. Nounou, "Multiscale fuzzy kalman filtering," *Engineering Applications of Artificial Intelligence*, vol. 19, pp. 439–450, 2006.
- [14] S. Julier and J. Uhlmann, "New extension of the kalman filter to nonlinear systems," *Proceedings of SPIE*, vol. 3, no. 1, pp. 182–193, 1997.
- [15] L. Ljung, "Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems," *IEEE Trans. on Automatic Control*, vol. 24, no. 1, pp. 36–50, 1979.
- [16] Y. Kim, S. Sul, and M. Park, "Speed sensorless vector control of induction motor using extended kalman filter," *IEEE Trans. on Industrial Applications*, vol. 30, no. 5, pp. 1225–1233, 1994.
- [17] E. Wan and R. V. D. Merwe, "The unscented kalman filter for nonlinear estimation," *Adaptive Systems for Signal Processing, Communications, and Control Symposium*, pp. 153–158, 2000.
- [18] R. V. D. Merwe and E. Wan, "The square-root unscented kalman filter for state and parameter-estimation," *IEEE International Conference on Acoustics, Speech, and Signal Processing*, vol. 6, pp. 3461–3464, 2001.
- [19] S. Sarkka, "On unscented kalman filtering for state estimation of continuous-time nonlinear systems," *IEEE Trans. Automatic Control*, vol. 52, no. 9, pp. 1631–1641, 2007.
- [20] M. Beal, "Variational algorithms for approximate bayesian inference," Ph.D. dissertation, Gatsby Computational Neuro- science Unit, University College London, 2003.
- [21] J. Kotecha and P. Djuric, "Gaussian particle filtering," *IEEE Trans. on Signal Processing*, vol. 51, no. 10, pp. 2592–2601, 2003.
- [22] G. Storvik, "Particle filters for state-space models with the presence of unknown static parameters," *IEEE Trans. On Signal Processing*, vol. 50, no. 2, pp. 281–289, 2002.
- [23] A. Doucet and V. Tadić, "Parameter estimation in general state-space models using particle methods," *Annals of the institute of Statistical Mathematics*, vol. 55, no. 2, pp. 409–422, 2003.
- [24] G. Poyiadjis, A. Doucet, and S. Singh, "Maximum likelihood parameter estimation in general state-space models using particle methods," in *Proc of the American Stat. Assoc*, 2005.
- [25] M. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, "A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking," *Signal Processing, IEEE Transactions on*, vol. 50, no. 2, pp. 174–188, 2002.
- [26] M. Mansouri, H. Snoussi, and C. Richard, "A nonlinear estimation for target tracking in wireless sensor networks using quantized variational filtering," *Proc. 3rd International Conference on Signals, Circuits and Systems*, pp. 1–4, 2009.
- [27] B. Balaji and K. Friston, "Bayesian state estimation using generalized coordinates," *Proc. SPIE 8050*, 2011.
- [28] P. Djurić, M. Vemula, and M. Bugallo, "Target tracking by particle filtering in binary sensor networks," *IEEE Transactions on Signal Processing*, vol. 56, no. 6, pp. 2229–2238, 2008.
- [29] A. Chhetri, D. Morrell, and A. Papandreou-Suppappola, "Energy efficient target tracking in a sensor network using non-myopic sensor scheduling," in *Proc. 8th International Conference on Information Fusion*, 2005.

- [30] J. Yick, B. Mukherjee, and D. Ghosal, "Analysis of a prediction-based mobility adaptive tracking algorithm," in Proceedings of the IEEE Second International Conference on Broadband Networks (BROADNETS), Boston, 2005.
- [31] J. Vermaak, N. Lawrence, and P. Perez, "Variational inference for visual tracking," in Conf. Computer Vision and Pattern Recognition, Jun. 2003.
- [32] A. Corduneanu and C. Bishop, "Variational bayesian model selection for mixture distribution," in Artificial Intelligence and Statistics, 2001.
- [33] M. Tremblay and D. Wallach, "Comparison of parameter estimation methods for crop models," *Agronomie*, vol. 24, no. 6-7, pp. 351–365, 2004.
- [34] F. BARET, "Contribution au suivi radiometrique de cultures de cereales," Ph.D. dissertation, 1986.
- [35] Mansouri, M., Dumont, B. and Destain, M.-F. "Modeling and Prediction of Nonlinear Environmental System Using Bayesian Methods", *Computers & Electronics in Agriculture*, 92, 16-31, 2013.
- [36] Mansouri, M., Dumont, B. and Destain, M.-F. "Bayesian methods for predicting LAI and soil Moisture", Proceedings of the 12th International Conference on Precision Agriculture (ICPA), 15-18 July Indianapolis (USA), 2011.

## VI. ADDITIONAL READING SECTION

- [1] V. Aidala. Parameter estimation via the kalman filter. *IEEE Trans. on Automatic Control*, 22(3):471–472, 1977.
- [2] J. Aitchison and S.D. Silvey. Maximum-likelihood estimation of parameters subject to restraints. *The Annals of Mathematical Statistics*, 29(3):813–828, 1958.
- [3] J. Aitchison and S.D. Silvey. Maximum-likelihood estimation procedures and associated tests of significance. *Journal of the Royal Statistical Society. Series B (Methodological)*, pages 154–171, 1960.
- [4] B.W. Andrews, T.M. Yi, and P.A. Iglesias. Optimal noise filtering in the chemotactic response of *escherichia coli*. *PLoS computational biology*, 2(11):e154, 2006.
- [5] J. Lin, "Divergence measures based on the shannon entropy," *Information Theory, IEEE Transactions on*, vol. 37, no. 1, pp. 145–151, 1991.
- [6] M.S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp. A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking. *Signal Processing, IEEE Transactions on*, 50(2):174–188, 2002.
- [7] J. Burbea and C. Rao. On the convexity of some divergence measures based on entropy functions. *Information Theory, IEEE Transactions on*, 28(3):489–495, 1982.
- [8] A. Corduneanu and C.M. Bishop. Variational bayesian model selection for mixture distribution. In *Artificial Intelligence and Statistics*, 2001.
- [9] D.D. Dorfman and E. Alf. Maximum-likelihood estimation of parameters of signal-detection theory and determination of confidence intervals–rating-method data. *Journal of Mathematical Psychology*, 6(3):487–496, 1969.
- [10] A. Ervadi-Radhakrishnan and E. O. Voit. Controllability of non-linear biochemical systems. *Mathematical Biosciences*, 196:99–123, 2005.
- [11] Georg Fritz, Christiane Koller, Korinna Burdack, Larissa Tetsch, Ina Haneburger, Kirsten Jung, and Ulrich Gerland. Induction kinetics of a conditional ph stress response system in *escherichia coli*. *Journal of Molecular Biology*, 393(2):272–286, 2009.
- [12] J.C. Geromel. Optimal linear filtering under parameter uncertainty. *IEEE Trans. on Signal Processing*, 47(1):168–175, 1999.
- [13] M. Grewal and A. Andrews. *Kalman Filtering: Theory and Practice using MATLAB*. John Wiley and Sons, 2008.
- [14] S.P. Han. A globally convergent method for nonlinear programming. *Journal of optimization theory and applications*, 22(3):297–309, 1977.

- [15] AJ Haug. A tutorial on bayesian estimation and tracking techniques applicable to nonlinear and non-gaussian processes. MITRE Corporation, McLean, 2005.
- [16] J. Holtz. The representation of ac machine dynamics by complex signal flow graphs. *Industrial Electronics, IEEE Transactions on*, 42(3):263–271, 1995.
- [17] H. Jeffreys and H. Jeffreys. An invariant form for the prior probability in estimation problems. *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences*, 186(1007):453–461, 1946.
- [18] B.H. Juang and LR Rabiner. A probabilistic distance measure for hidden Markov models. *AT&T Bell Laboratories technical journal*, 64(2):391–408, 1985.
- [19] R. E. Kalman. A new approach to linear filtering and prediction problem. *Trans. ASME, Ser. D, J. Basic Eng.*, 82:34–45, 1960.
- [20] J.H. Lee and N.L. Ricker. Extended kalman filter based nonlinear model predictive control. *Industrial & engineering chemistry research*, 33(6):1530–1541, 1994.
- [21] A. Leshem. Maximum likelihood separation of constant modulus signals. *IEEE Trans. on Signal Processing*, 48(10):2948–2952, 2000.
- [22] Z.F. Li, M.R. Osborne, and T. Prvan. Numerical algorithms for constrained maximum likelihood estimation. *Anziam Journal*, 45(1):91–114, 2003.
- [23] L. Matthies, T. Kanade, and R. Szeliski. Kalman filter-based algorithms for estimating depth from image sequences. *International Journal of Computer Vision*, 3(3):209–238, 1989.
- [24] N. Meskin, H. Nounou, M. Nounou, A. Dattta, and E. R. Dougherty. Intervention in biological phenomena modeled by s-systems. *IEEE Transactions on Biomedical Engineering*, 58(5):1260–1267, 2011.
- [25] N. Meskin, H. Nounou, M. Nounou, A. Dattta, and E. R. Dougherty. Parameter Estimation of Biological Phenomena Modeled by S-Systems: An Extended Kalman Filter Approach. In *Proc. IEEE Conference on Decision and Control and European Control Conference, Orlando, FL.*, to appear, 2011.
- [26] H. Nounou and M. Nounou. Multiscale fuzzy kalman filtering. *Engineering Applications of Artificial Intelligence*, 19:439–450, 2006.
- [27] M.R. Osborne. Fisher’s method of scoring. *International Statistical Review*, pages 99–117, 1992.
- [28] R. Sibson. Information radius. *Probability Theory and Related Fields*, 14(2):149–160, 1969.
- [29] G. Storvik. Particle filters for state-space models with the presence of unknown static parameters. *IEEE Trans. on Signal Processing*, 50(2):281–289, 2002.
- [30] G.A. Terejanu. Unscented kalman filter tutorial. In *Workshop on Large-Scale Quantification of Uncertainty, Sandia National Laboratories*, pages 1–6, 2009.
- [31] N. V. Torres. Modelization and experimental studies on the control of the glycolytic-glucogenolytic pathway in rat liver. *Molecular and Cellular Biochemistry*, 132(2), 1994.
- [32] A. Vajesta and R. Schmitz. An experimental study of steady-state multiplicity and stability in an adiabatic stirred reactor. *AIChE Journal*, 3:410–419, 1970.
- [33] J. Vermaak, N.D. Lawrence, and P Perez. Variational inference for visual tracking. In *Conf. Computer Vision and Pattern Recognition*, June 2003.
- [34] E. O Voit and J. Almeida. Decoupling dynamical systems for pathway identification from metabolic profiles. *Bioinformatics*, 20(11):1670–1681, 2004.
- [35] Z. Zhang. Parameter estimation techniques: A tutorial with application to conic fitting. *Image and vision Computing*, 15(1):59–76, 1997.
- [36] Y. Bard, *Nonlinear parameter estimation*. Academic Press New York, 1974, vol. 477.
- [37] J. Beck and K. Arnold, *Parameter estimation in engineering and science*. Wiley New York, 1977, vol. 8.
- [38] A. Tarantola, *Inverse problem theory and methods for model parameter estimation*. Society for Industrial Mathematics, 2005.
- [39] P. Brown, V. Pietra, S. Pietra, and R. Mercer, “The mathematics of statistical machine translation: Parameter estimation,” *Computational linguistics*, vol. 19, no. 2, pp. 263–311, 1993.

[40] W. Cash, "Parameter estimation in astronomy through application of the likelihood ratio," *The Astrophysical Journal*, vol. 228, pp. 939–947, 1979.

[41] D. Rife and R. Boorstyn, "Single tone parameter estimation from discrete-time observations," *Information Theory, IEEE Transactions on*, vol. 20, no. 5, pp. 591–598, 1974.

## KEY TERMS AND DEFINITIONS

States and parameters estimation:

State and parameter estimation is the process of obtaining the estimators of state variables and model parameters based on a dataset.

Extended Kalman filter (EKF):

EKF is an estimation technique that is applicable to nonlinear and non-Gaussian models. As the Kalman filter is applicable to linear systems, the EKF can be viewed as an extension of the Kalman filter that is applied to a linearized version of the nonlinear model.

Unscented Kalman filter (UKF):

UKF is an estimation technique that uses a deterministic sampling technique known as the unscented transform to approximate the mean and the covariance of the state and parameter vector.

Particle filter (PF):

PF, which is based on Bayesian estimation, is a sequential Monte Carlo state estimation method for nonlinear and non-Gaussian systems. The EKF and UKF algorithms do not always provide a satisfactory performance, especially for highly nonlinear processes as model linearization does not necessarily provide good estimates of the mean of the state vector and the covariance matrix of the estimation error, which are used in state estimation. These issues are addressed by the PF.

Variational filter (VF):

VF, which is based on Bayesian estimation, is state estimation method for nonlinear and non-Gaussian systems. It has been proposed recently to enhance state estimation because VF yields an optimal choice of the sampling distribution by minimizing a Kullback-Leibler (KL) divergence criterion. In fact, variational calculus leads to a simple Gaussian sampling distribution whose parameters (which are estimated iteratively) also utilize the observed data, which provides more accurate and computationally efficient computation of the posterior distribution.

Environmental data:

Environmental data is that which is based on the measurement of environmental pressures, the state of the environment and the impacts on ecosystems.

Leaf Area Index:

Leaf Area Index is a method of measurement of leaf area in a given section of ground area expressed in terms of half the square meters of leaf per square meter of ground.

Soil moisture:

Soil moisture is the volumetric water content in the soil.