State and parameter estimation of the AquaCrop model for winter wheat using sensitivity informed particle filter

Tianxiang Zhang[1,a,b], Jinya Su[1,c], Cunjia Liu[b,*], Wen-Hua Chen[b]

[a] School of Automation and Electrical Engineering, University of Science and Technology Beijing, Beijing 100083, China
[b] Department of Aeronautical and Automotive Engineering, Loughborough University, Loughborough, LE11 3TU, U.K.
[c] School of Computer Science and Electronic Engineering, University of Essex, Colchester, CO4 3SQ, U.K.

Abstract

Crop models play a paramount role in providing quantitative information on crop growth and field management. However, its prediction performance degrades significantly in the presence of unknown, uncertain parameters and noisy measurements. Consequently, simultaneous state and parameter estimation (SSPE) for crop model is required to maximize its potentials. This work aims to develop an integrated dynamic SSPE framework for the AquaCrop model by leveraging constrained particle filter, crop sensitivity analysis and UAV remote sensing. Both Monte Carlo simulation and one winter wheat experimental case study are performed to validate the proposed framework. It is shown that: (i) the proposed framework with state/parameter bound and parameter sensitivity information outperforms conventional particle filter and constrained particle filter in both state and parameter estimation in Monte Carlo simulations; (ii) in real-world experiment, the proposed approach achieves the smallest root mean squared error for canopy cover estimation among the three algorithms by using day forward-chaining validation method.

Keywords: Particle filter; Sensitivity analysis; Machine learning; Multispectral image; Unmanned Aerial Vehicle

1. Introduction

Crop simulation models, providing quantitative crop growth information during the crop life-cycle, play a paramount role in sustainable agriculture management. It contributes to intelligent irrigation, nutrient management, and yield prediction before harvest, which directly promote agriculture sustainability and food security [1]. However, the prediction performance of crop model degrades significantly in real-life applications due to the presence of unknown and uncertain system parameters. In this regard, a timely and reliable Simultaneous State and Parameter Estimation (SSPE) for crop model is highly desirable to realize its full potentials. Recently, the integration of crop models and remote sensing information is drawing ever-increasing research interest in precision agriculture, where the problem is usually addressed by using various data assimilation techniques (optimization approaches at large) [2].

Crop models are able to quantitatively simulate crop physiological process at a daily basis. Due to their practical

* Corresponding author (Cunjia Liu).
Email addresses: T.Zhang@lboro.ac.uk (Tianxiang Zhang[1]), j.su@essex.ac.uk (Jinya Su[1]), C.Liu5@lboro.ac.uk (Cunjia Liu), W.Chen@lboro.ac.uk (Wen-Hua Chen)
[1] Tianxiang Zhang and Jinya Su are joint first authors contributing equally to this work.

usability, a number of crop models have been developed recently from different principles such as WOFOST, DASSAT, STICS and AquaCrop model. Unlike other crop models (i.e. light driven or carbon driven), AquaCrop model is water-driven model, which is simple, robust and accurate [3]. This model has been widely applied in precision agriculture practices such as crop monitoring, intelligent irrigation management and yield prediction before harvest [4]. Therefore, the AquaCrop crop model is adopted to demonstrate the proposed SSPE framework in this study.

Regarding remote sensing information, images of various spatial/spectral resolutions can be captured by sensing platforms such as satellite, manned-aircraft and Unmanned Aerial Vehicles (UAV) [1]. Among them, UAV remote sensing is drawing increasing research interests and has become an important supplement to conventional platforms [5]. This is mainly due to its attractive characteristics including a relatively affordable cost, a high spatial and user-defined temporal resolution, and a good flexibility [6]. It has also been widely applied in a large number of applications such as crop stress monitoring (e.g. disease, weed, drought), crop state estimation (e.g. canopy cover, biomass, leaf area index) and crop parameter inference [2, 7].

In this study, UAV remote sensing is to derive canopy cover (CC) of the AquaCrop model. CC is defined as the ratio of plant leaves projected to the horizontal surface to the total ground area [8] and is one of the most important state variables in the AquaCrop model. The calculation of CC value is formulated as an image classification problem, which is addressed by the random forest classifier. Image pixels are segmented into wheat and non-wheat pixels, based on which the proportion of wheat pixels is calculated as the CC value. It is shown in [9] that this machine learning based approach outperforms threshold based approaches [10] and is therefore adopted in this study.

In addition, state estimation problem can be found in many applications such as crop state estimation, hazardous target tracking, hydrological parameter inference [11, 12]. This problem is usually addressed by Kalman Filter (KF) or its nonlinear variants such as Extended Kalman Filter (EKF) and Unscented Kalman Filter (UKF). The SSPE problem, in comparison to the standard state estimation, poses new research challenges such as poor observability. Because unknown or uncertain system parameters should also be estimated along with unknown states by using partial noisy measurements. Due to the high non-linearity involved in SSPE problem, the generic nonlinear filtering approach is usually adopted such as Particle Filtering (PF). PF is a sequential Bayesian approach by Monte Carlo sampling, and is particularly suitable for non-linear and non-Gaussian filtering. In this approach, a large number of particles are drawn to approximate the probability density function of states and parameters [13, 14, 15, 16] rather than only propagating the mean and variance. This distribution can provide confidence information which is not possible in point estimation (e.g. optimization based approaches).

It should be noted that in real-life agricultural applications, various types of extra information (or background knowledge) is usually available. For instance, many unknown or uncertain parameters in the AquaCrop model have physical meanings and therefore are with upper and lower bounds. This parameter bound information, if taken into account properly, can further improve the estimation performance of particle filter [17, 18, 19]. It is also discovered in this study that the sensitivity of various parameters in crop models may vary significantly in different crop growth stages, that is, a parameter being sensitive in stage A may become insensitive in stage B and vice versa [20]. As a result, a static parameter modelling error term is insufficient to capture this dynamic sensitivity behaviour and alternative approaches should be sought. Therefore, this work aims to develop an improved particle filter framework for SSPE of the AquaCrop model, which can accommodate these extra information (including parameter bound information and parameter sensitivity information) for better estimation performance. The improved particle filter is compared against the conventional PF and constrained PF by using both Monte Carlo (MC) simulations and real-life experiment. To be more exact, the main contributions are summarized as follows:

(1) A sequential particle filter with parameter bound and sensitivity information is drawn to integrate AquaCrop model and UAV remote sensing measurements so that the posterior distribution of both states and uncertain parameters can be estimated in near real time.

(2) MC simulations and an experimental case study are performed to validate the developed framework against conventional particle filter and constrained particle filter.

2. Problems Formulation

The core problem in this study can be formulated as a sequential state and parameter estimation (SSPE) problem by integrating AquaCrop model and remote sensing observations. Since the AquaCrop model is non-analytical, conventional particle filter cannot be used and an open access source AquaCrop model is selected (AquaCrop-OS) [4]. This model can be modelled as a discrete-time dynamic state-space model, and satisfies a Markov process where the future states at $k+1$ step is only associated with the states at k step [21]. The compact system including state dynamics and observation model can be represented by (1).

$$\begin{cases} x_{k+1} = F(x_k, \theta, u_k) + \nu_k, \\ Y_k = G(x_k, \theta, u_k) + \eta_k, \text{with } \eta_k \sim N(0, \sigma) \end{cases} \tag{1}$$

where x_{k+1} represents the canopy dynamic states in crop model at $k+1$ time step. θ and u_k denote the selected parameters and forcing data (e.g. weather data, fixed parameters). Y_k represents the observations at k time step. ν_k and η_k are independent, representing the model process noise and measurement noise, which are with zero means and proper covariances, respectively. $F(.)$ and $G(.)$ are non-linear functions relating the relevant variables. The SSPE problem in this study is to estimate the state x_k and unknown parameter θ based on the available measurements $Y_{1,\cdots,k}$ up to day k.

3. Methodology

In this section, some key components of the developed framework are introduced including Sobol sensitive analysis and the improved PF algorithm for SSPE problem of AquaCrop model.

3.1. Crop parameter sensitivity analysis

In this section, crop parameter sensitivity analysis is considered. This is because the parameter sensitivity in crop models may vary significantly in different crop growth stages. As a result, new parameter error modelling instead of a static one should be designed to capture the dynamic behaviour of parameter sensitivity.

3.1.1. Sobol sensitivity analysis

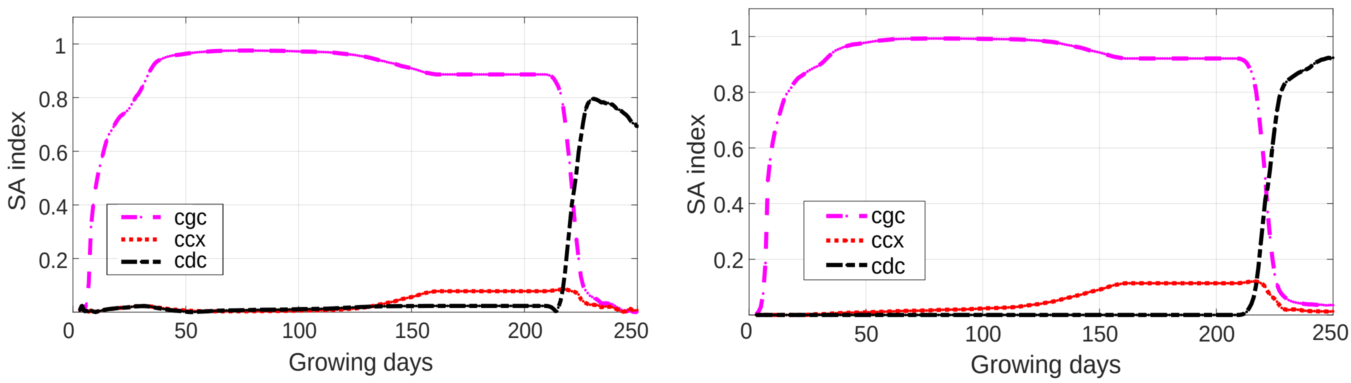

Figure 1: First order (left) and total order (right) SA index of three parameters over time.

Sensitive analysis (SA) is one effective tool to quantitatively analyse the uncertain factors (parameters or driving variables) on model outputs and identify the most sensitive ones [20]. The Sobol method is a variance based approach decomposing the model output variance into contributions associated with each input parameter. For AquaCrop model, it can evaluate the contribution of separate parameters and interactions to the model outputs (e.g. CC or biomass) [22]. Crop parameters *cgc*, *ccx* and *cdc* can determine canopy growth gradient, maximum canopy cover and canopy decline gradient at growing stages. An approximate canopy cover growth can be simulated once these three parameters are confirmed. As a consequence, these three parameters are selected for sensitive analysis and estimated in SSPE. Details of Sobol method is referred to [22, 23].

3.1.2. SA results

The sensitivity analysis results by Sobol analysis are depicted in Fig 1 including the first order index (left plot) and the total order index (right). It follows from Fig 1 that both indices share the same sensitivity trend. In particular, *cgc* has a high sensitivity during the whole wheat growing period. However, *ccx* starts its influence when the canopy is saturated; and *cdc* plays a significant role in crop degrading period. First order sensitive index can be taken into account in the proposed PF framework by adaptively adjusting the parameter modelling error (via variance).

3.2. PF framework with parameter bound and sensitivity information

In this section, the improved particle filter based SSPE framework is discussed, which can accommodate the parameter bound and sensitivity information. The SSPE framework for AquaCrop is shown in Fig 2.

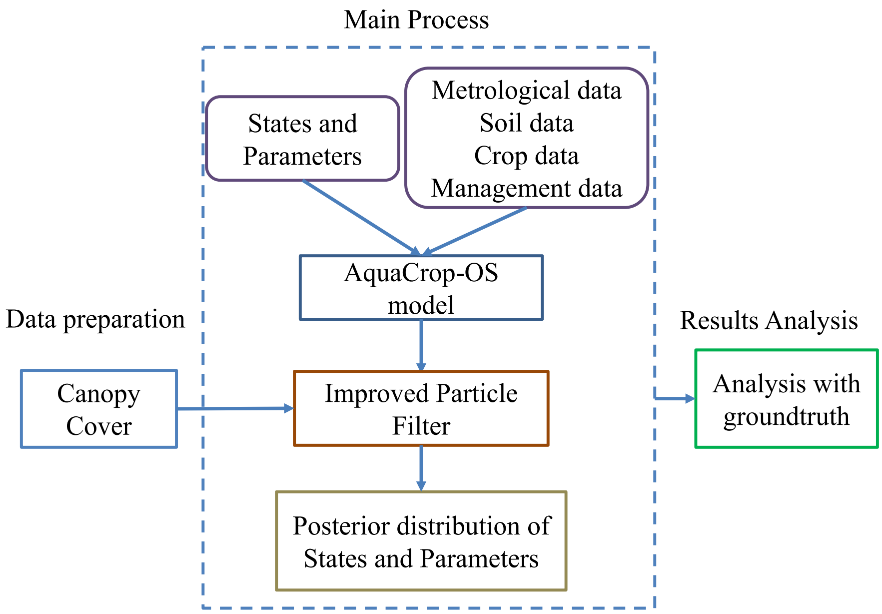

Figure 2: Framework of the proposed particle filter for state and parameter estimation of the AquaCrop model.

3.2.1. Recursive AquaCrop model

As one open-sourced model, the AquaCrop-OS can be easily integrated with different algorithms for various applications [3, 4, 24]. However, this crop model need to be revised recursively before realizing its full potentials so that the main relationship between crop states (e.g. biomass, canopy cover) and other properties (e.g. weather, crop and soil parameters, field management) is reformulated using crop water balance principle. Thus, our new program is able to generate and record new crop state in real-time, meanwhile all parameters and states can be easily updated with the advert of new observations. Consequently, this model can complete the crop simulation process by using different filtering algorithms in a recursive way.

3.2.2. Improved particle filter

Particle filter, one of recursive filtering methods, can be modified to solve the SSPE problem [13]. In conventional particle filter, a series of particle samples with corresponding weights are used to model the posterior state distribution [11], where the weight is calculated based on Bayesian equation fusing prior information and observations. However, in some practical scenarios, constrained particle filter is usually utilised as physical principles and process restrictions can be taken into account as constraints (as additional knowledge/information) so that effective particles can be increased to generate a better posterior distribution [25].

In the proposed problem, the parameters to be estimated θ_k are augmented with the original states x_k to become the augmented states X_k,

$$X_{k+1} = \begin{bmatrix} x_{k+1} \\ \theta_{k+1} \end{bmatrix} \tag{2}$$

A Gaussian random walk is assumed for the parameters

$$\theta_{k+1} = \theta_k + \mu_k \tag{3}$$

where $\mu_k \sim N(0, \alpha)$ is a Gaussian distribution with zero mean and a pre-defined covariance α. Consequently, by substituting Eq.2 and Eq.3 into Eq.1, the dynamic model with augmented state X_{k+1} can be rewritten as

$$\begin{cases} X_{k+1} = F^{'}(X_k, u_k) + \nu_k^{'} \\ Y_k = G(X_k, u_k) + \eta_k \end{cases} \tag{4}$$

where the modified function $F^{'}(.)$ and modified noise $\nu_k^{'}$ are given by

$$F^{'}(X_k, u_k) = \begin{bmatrix} F(x_k, \theta_k, u_k) \\ \theta_k \end{bmatrix}, \ \nu_k^{'} = \begin{bmatrix} \nu_k \\ \mu_k \end{bmatrix}. \tag{5}$$

It should be noted that the selection of μ_k (i.e. its variance) is paramount for the filtering performance. It follows from [26] that μ_k reflects the intensity of process noise and the size of sampling range. If a small variance term is chosen, it is difficult to converge to the truth values timely. And if a too large variance is chosen, more invalid particles will be generated, impairing algorithm effectiveness. In order to reduce the adverse effect caused by the inappropriately selected covariance in SSPE problem, the process noise is related to the sensitivity analysis result, which is defined as

$$\mu_k = H(index_k) \tag{6}$$

where $index_k$ is time-series parameter sensitive index at day k, and $H(.)$ represents the process error function. $H(.)$ is designed so that a high sensitivity value leads to a smaller error covariance and a low sensitivity value results in a larger error covariance. It will be shown that this strategy can significantly improve the effectiveness of the particle filter.

In real-life agriculture applications, some states and parameters have physical properties and as a result certain constraints information (e.g. bound information) is usually available [25]. These constraints can be represented by certain inequality function $g(X_k) <= 0$. The probability conditional on X_k can be defined as $p(D_k|X_k)$. In practical implementation, the particles can be accepted if they satisfy the constraints and be rejected if the constraints are violated.

It follows from [17] that the posterior distribution of $X_{0:k+1}$ with constraints information $D_{1:k+1}$ can be derived according to the Bayesian recursion once measurement $Y_{1:k+1}$ is available, given by

$$p(X_{0:k+1}|Y_{1:k+1}, D_{1:k+1}) = \frac{p(Y_{k+1}|X_{k+1})p(D_{k+1}|X_{k+1})p(X_{k+1}|X_k)p(X_{0:k}|Y_{1:k}, D_{1:k})}{p(Y_{k+1}, D_{k+1}|Y_{1:k}, D_{1:k})} \tag{7}$$

where $p(X_{0:k+1}|Y_{1:k+1}, D_{1:k+1})$ describes the posterior distribution at time $k+1$ from the posterior $p(X_{0:k}|Y_{1:k}, D_{1:k})$ at time k. $p(X_{k+1}|X_k)$ denotes the crop model function and $p(Y_{k+1}|X_{k+1})$ expresses the likelihood function. $p(D_{k+1}|X_{k+1})$ is hard constraint related probability.

Our proposed PF applies enough particle samples to approximate the posterior probability density function (PDF), where each particle represents a specific state X^i_{k+1} with a proper probability weight w^i_{k+1}. The posterior PDF of states and parameters could be approximated by N particles and their corresponding weights, given by

$$p(X_{0:k+1}|Y_{1:k+1}, D_{1;k+1}) \approx \sum_{i=1}^{N} w^i_{k+1}\delta(X_{0:k+1} - X^i_{0:k+1}) \tag{8}$$

where N means the particle number and δ is the Dirac delta function. $p(X_{0:k+1}|Y_{1:k+1}, D_{1;k+1})$ is the truth posterior PDF, $X^i_{0:k+1}$ is the i-th particle with related weight w^i_{k+1}.

According to sequential importance sampling principle (particle weight selection), X_{k+1} from proposal distribution

$q(X^i_{0:k+1}|Y_{1:K+1}, D_{1;k+1})$ can be generated and assigned with the corresponding weights according to

$$w^i_{k+1} \propto \frac{p(X^i_{0:k+1}|Y_{1:K+1}, D_{1;k+1})}{q(X^i_{0:k+1}|Y_{1:K+1}, D_{1;k+1})} \tag{9}$$

According to [13, 27], the proposal distribution could be factorised as

$$q(X_{0:k+1}|Y_{1:k+1}, D_{1;k+1}) = q(X_{k+1}|X_{0:k}, Y_{1:k+1}, D_{1;k+1})q(X_{0:k}|Y_{1:k}, D_{1;k}) \tag{10}$$

By inputting Eq. 10 and Eq. 7 into the importance weights Eq. 9, particles weights can be updated

$$w^i_{k+1} \propto \frac{w^i_k p(Y_{k+1}|X^i_{k+1})p(D_{k+1}|X^i_{k+1})p(X^i_{k+1}|X^i_k)}{q(X^i_{k+1}|X^i_{0:k}, Y_{1:k+1}, D_{1;k+1})} \tag{11}$$

where $q(X^i_{k+1}|X^i_{0:k}, Y_{1:k+1}, D_{1;k+1})$ means the posterior probability density function affecting the particle filter results. In this study, the proposal distribution is assumed to be the prior information $q(X_{k+1}|X_k, Y_{k+1}, D_{k+1}) = p(X_{k+1}|X_k)$, thus the above equation can be simplified as

$$w^i_{k+1} \approx w^i_k p(Y_{k+1}|X^i_{k+1})p(D_{k+1}|X^i_{k+1}) \tag{12}$$

Considering the hard constraint property, the constraint probability condition on X_{k+1} can be written as

$$p(D_{k+1}|X^i_{k+1}) = \begin{cases} 1, & if \quad g(X_{k+1}) <= 0 \\ 0, & otherwise \end{cases} \tag{13}$$

Consequently, the particle weights can be rewritten by considering the constraints information

$$w^i_{k+1} = \begin{cases} \propto w^i_k p(Y_{k+1}|X^i_{k+1}), & if \quad g(X_{k+1}) <= 0 \\ 0, & otherwise \end{cases} \tag{14}$$

Assuming that the measurement noise follows a Gaussian distribution with zero mean and a covariance R, the likelihood function and updated particle weight are given by [11, 16]

$$P(Y_{k+1}|X^i_{k+1}) = \frac{1}{\sqrt{2\pi}\sqrt{R_{k+1}}} \exp[-\frac{(Y_{k+1} - G(X^i_{k+1}))^2}{2R_{t+1}}] \tag{15}$$

$$w^i_{k+1} = \frac{w^i_{k+1}}{\sum_{i=1}^N w^i_{k+1}}. \tag{16}$$

Due to the particle degeneracy problem, a useful measure of effective number can be defined as

$$N_{eff} = \frac{1}{\sum_{i=1}^N (w^i_{k+1})^2}. \tag{17}$$

To attenuate the particle degeneracy problem, resampling strategy is adopted [28]. In particular, Metropolis resampling is employed due to its reliability and a low computation cost. The weight of each particle will be $1/N$

after resampling process. The posterior distribution of filtered states and parameters is in the form of particles, thus the updated particles after resampling will become the initial particles for the next evolutionary iteration until all available measurements are assimilated into AquaCrop model. The above steps are summarized in Algorithm 1.

Algorithm 1: Particle filter with constraints and sensitivity information

Require X_k^i,w_k^i

Initialization: Generate N samples from the prior PDF $P(X_k)$ with equal weight $1/N$

For: $i = 1 : N$

    Parameter process error covariance function $\mu_{k+1} = H(index_{k+1})$

    Draw new particles $X_{k+1} \sim p(X_{k+1}|X_k^i)$

    Update particle weight w_{k+1}^i according to hard constraints information by Eq. 14

End for

Particle weight satisfy $\sum_{i=1}^{N} w_{k+1}^i = 1$

Resampling using Metropolis resampling method

Ensure: new samples X_{k+1}^i,$w_{k+1}^i = 1/N$

3.3. Validation method for field experiment

In this study, two validation methods are adopted for the proposed algorithm including MC simulation in a simulated environment and field validation based on an experimental case study. In particular, due to the lack of groundtruth crop parameter data for field experiment owning to the logistic issues, the validation of field experiment is only based on CC estimation performance instead of both CC and crop parameters estimation. A modified Day Forward-Chaining (DFC) method [29, 30], as shown in Fig 3, is used for temporal validation in the field experiment. This is because forward chaining method can avoid the problems that an arbitrary selection of the hold-out test set may lead to a biased estimate, especially for the case with a limited number of observations.

In the modified DFC of Fig 3, each measurement denotes one fold with 12 folds in total in the field case study. Index N in red cells denotes the first few folds used for training; in other words, the measurements on these first few days are for state and parameter estimation using the proposed particle filter. In real-life agriculture management

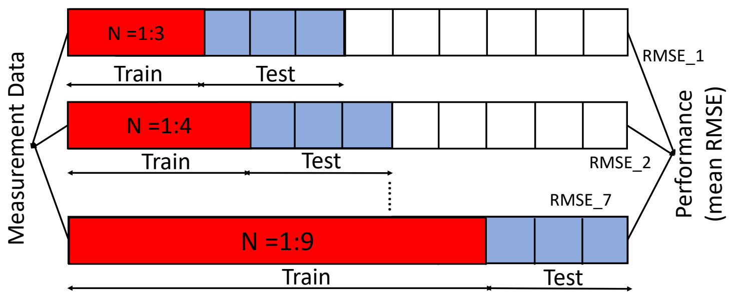

Figure 3: Day Forward-Chaining validation illustration for field case study.

practice, the decision is usually made by analysing the forecasting information in a number of future dates. Following this line of thought, the three folds following the training folds are used for testing. This means that the estimated state and parameters using the folds in red are inserted into the AquaCrop model to generate the prediction for folds in blue so that performance can be evaluated. The folds used for training and testing are illustrated in Fig 3.

4. Systematic settings

In this section, various validation approaches including Monte Carlo simulation and a real-world experiment case study are implemented to test the improved particle filter against the conventional particle filters with constraints or sensitivity information.

4.1. Monte Carlo simulation settings

MC simulations are firstly adopted to evaluate the SSPE performance for AquaCrop model, in particular for parameter estimation. Three parameters strongly associated with dynamic state canopy cover (CC) are selected. Thus, a four dimensional state and parameters vector is defined.

$$X = [cgc, ccx, cdc, CC]^T.$$

The states and parameters with their bound information and definitions are displayed in Table 1. In MC simulation, the default crop parameter values (being constant in its life-cycle) and model-generated CC in the AquaCrop simulation model are set to be the groundtruth. The noisy observations for model comparisons are derived from groundtruth CC by adding a Gaussian measurement noise with zero mean and a variance of 0.05^2. The time period of the AquaCrop model is consistent with the experiment from 08/Oct/2018 to 06/June/2019 under the same treatment and the measurement interval is 10 days. 50 MC simulations with random initials and random measurement noises are performed to test the robustness of all three methods including the conventional PF, the PF with constraints and the improved PF with constraints and sensitivity information.

Table 1: Selected state and parameters definition with bounds information for MC simulation and field experiment.

Variables	Prior information	Physical meaning
cgc	(0.005,0.02)	Canopy growth coefficient
ccx	(0.82,0.98)	Maximum canopy cover fraction
cdc	(0,0.02)	Canopy decline coefficient
CC	(0,1)	Canopy cover

4.2. Experimental evaluation

In addition to MC simulation, experimental verification is also performed. There is one case study (winter wheat) conducted from 2018-2019 to validate the proposed method. The key model state, wheat canopy cover, is extracted from multi-spectral images as below.

4.2.1. Experiment wheat field and UAV aerial imaging

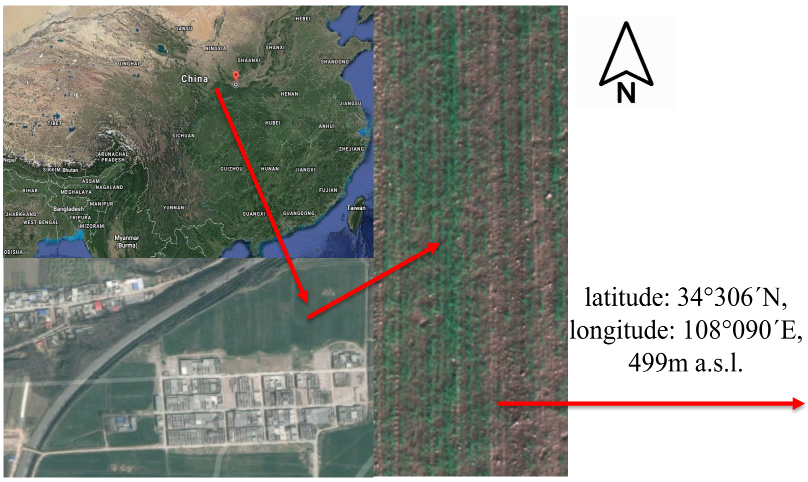

Figure 4: Geographic details of the study area.

The experiment site is located at Caoxinzhuang experiment field, which belongs to Northwest Agriculture and Forestry University (see Fig 4 for the location) [31]. The climate in this area is characterized by semi-humidity and semi-aridity, with an average annual temperature of 12.9^oC. In this study, the cultivar Xiaoyan22 (winter wheat) is adopted, which was developed by Northwest A&F university. In addition, some key information that is required in AquaCrop model, such as meteorological data and basic soil data can be downloaded from National Meteorological Information Center (http://data.cma.cn) and national Earth system Science Data Sharing Infrastructure (http://www.geodata.cn).

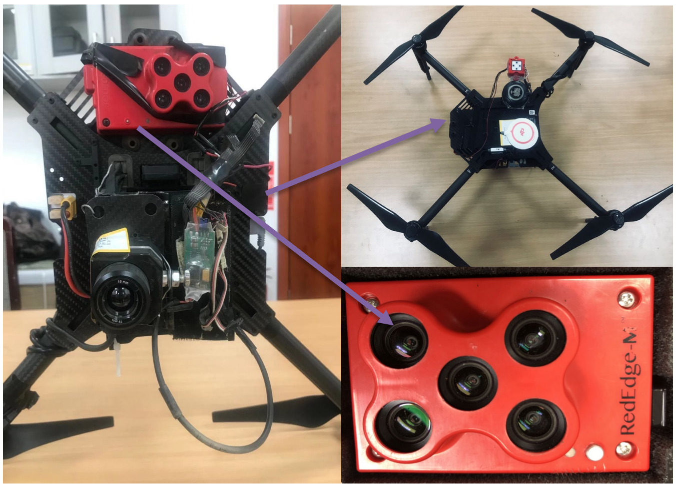

Figure 5: UAV camera system: DJI M100 Quadrotor UAV (left), GPS (upper right) and RedEdge camera(lower right)

In this study, UAV remote sensing images are preferred due to its high spatial/spectral resolutions. In particular, commercial DJI M100 Quadrotor (DJI Company, Shenzhen, China) and RedEdge camera (MicaSense Company, Seattle, USA) with five multi-spectral bands was integrated as the UAV sensing system (see Fig 5). RedEdge camera, in comparison to conventional RGB camera, has extra Rededge and NIR bands, providing extra spectral information for better classification performance [7]. The weight, dimensions and image resolution of RedEdge camera are 135g, $5.9cm \times 4.1cm \times 3.0cm$ and 1280×960 pixels, respectively.

In each flight, the RedEdge camera was fixed on the UAV, pointing vertically downwards to the wheat canopy. Flight altitude was set to be about 20 meters above ground with a ground image spatial resolution of about 1.2 cm/pixel. Image overlap and sidelap were set to be up to 75 % for an accurate orthomosaic generation. Reflectance calibration panel was always imaged at 1m height before each flight to account for camera and reflectance characteristics, and environmental variations [1, 31]. After data collection, Pix4DMapper, a commercial photogrammetry software, was then used to process the raw images to generate the caliberated orthomosaic images for each band. The overall process includes initial processing, Point Cloud and Mesh generation and orthomosaic generation, where more details are referred to the existing studies [7, 31]. Finally, a total of twelve multi-spectral images were collected covering winter wheat key developmental stages (please refer to Table. 2 for the specific imaging times) including tillering stage, green-up stage, jointing stage, anthesis stage and grain filling stage [32].

4.2.2. CC calculation

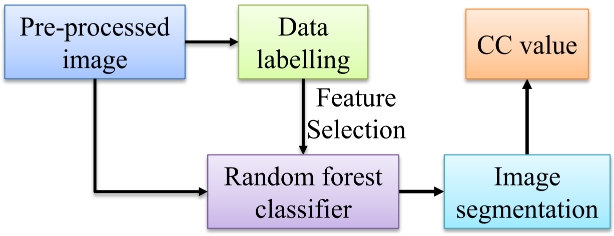

Figure 6: Steps for canopy cover calculation by using remote sensing images and random forest classifier.

In this study, the CC calculation can formulated as a wheat/non-wheat two-class classification problem so that the wheat pixel proportion can be calculated for the region of interest. The overall process is displayed in Fig 6, which include several components such as data labelling, random forest classifier and CC value calculation. One typical example for the data on 16/Dec/2018 is presented, where these steps are detailed in the following subsections.

Supervised classification depends on data labelling. In this study, wheat and non-wheat pixels are directly labelled according to on-site experiment and UAV RGB color image, where the RGB image is generated by using Red-Green-Blue bands of the multispectral image. The labelled sample image is displayed in Fig 7, where wheat pixels (wheat), non-wheat (others) and unlabelled pixels (un) are represented in different colours. Moreover, all available five bands including Blue (B), Green (G), Red (R), RedEdge and Near-infrared (NIR) bands are selected as the features for supervised classification. Spectral comparison between wheat and non-wheat pixels is referred to [1].

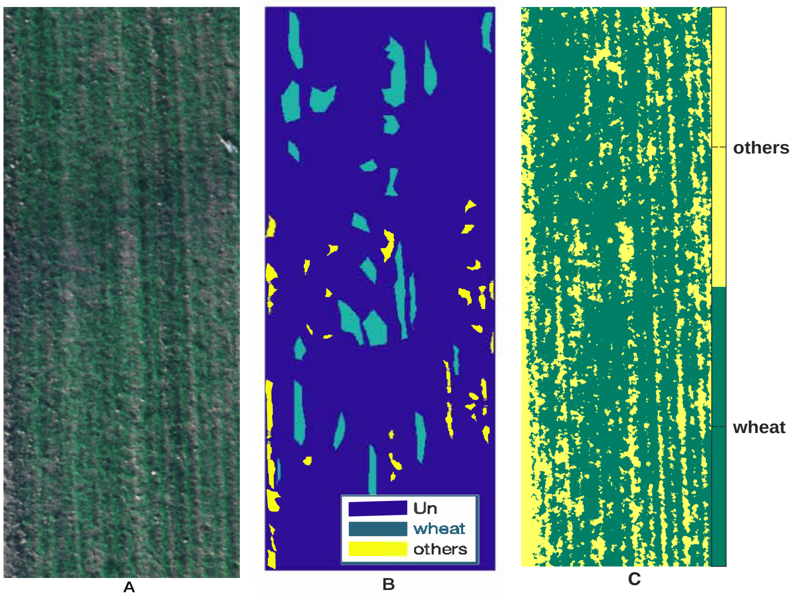

Figure 7: Survey data on 16/Dec/2018: A. RGB color image generated by Pix4DMapper; B. labelled image for supervised classification; C. segmented image by the random forest classifier.

A classifier is then required to perform the classification task so that new aerial images can be automatically classified for CC calculations. In this study, random forest classifier is implemented due to its good performance in terms of accuracy and robustness and a relatively low computation load, where the hyper-parameters are further automatically tuned by using Bayesian optimization [31]. Random forest algorithm has been previously used for wheat canopy segmentation in previous studies [1], which show that an accuracy of 99% can be achieved in selected labelled dataset. The detailed algorithm is omitted due to the lack of space and is referred to [31, 33].

The labelled image is split into training and testing data with a proportion of 70% and 30%. The trained random forest classifier is then applied to the original example image, where the classification map is displayed in the right plot of Fig 7. Then CC value can be calculated by $CC = N_{wp}/(N_{wp} + N_{nwp})$, where N_{wp} and N_{nwp} represent the number of wheat and non-wheat pixels in the region of interest. All CC values over time are displayed in Table 2 by following the above steps.

Table 2: Canopy cover values over time

Date	CC Value	Date	CC Value
11/11/2018	0.362	09/12/2018	0.7972
16/12/2018	0.8296	30/12/2018	0.8973
03/03/2019	0.9361	25/03/2019	0.9494
30/03/2019	0.9649	05/04/2019	0.9874
15/04/2019	0.9775	23/04/2019	0.9893
27/04/2019	0.9686	02/05/2019	0.9321

4.2.3. Experiment settings

Measurement noise of the experiment data can be estimated by using the algorithm in [1], where the covariance value is set to be 0.0021. The prior information of state and parameters as well as other settings of the improved PF algorithm remain the same as MC simulation (see Table 1).

5. Results and discussion

This section demonstrates a comparative estimation result using various PF methods. For MC simulation, the SSPE performance is evaluated by the root mean squared error (RMSE) of all MC runs. While in experimental validation, due to the absence of parameters groundtruth, error analysis is only tested on CC.

5.1. Results of MC simulation

MC analysis with random initial values and various measurements is first performed for the three methods. Mean and variance can be calculated from all particles at each observation day for one MC simulation run. For RMSE comparison, the estimated parameters on the last day of each MC run are used to calculate the RMSE value. In addition to parameter comparison, for CC, the total error is obtained by using the mean RMSE of all MC runs, where the RMSE of each run can be calculated by all CC estimations against the groundtruth at all observation dates. In particular, in MC simulations, all parameters to be estimated are constant in one certain local field and the default parameters are set to be groundtruth for performance evaluation [3].

Table 3: RMSE of 50 Monte Carlo simulations using different particle filter methods.

Parameters	Proposed PF(error)	Constrained PF(error)	Conventional PF
cgc	0.000592(-60.7%)	0.000833(-44.6%)	0.001506
cdc	0.010725(-31.5%)	0.035224(+125%)	0.015655
ccx	0.002034(-71.8%)	0.008303(+15%)	0.007217
CC	0.014137(-16.4%)	0.016027(-5.2%)	0.016907

The RMSE value of each method is shown in Table 3. Meanwhile, the error is also displayed, where $'-'$ means the error reduction and $'+'$ denotes the error increment in comparison to RMSE value using conventional PF. The RMSE error using different methods is defined by the following formula.

$$E_{PPF} = \frac{RMSE_{PPF} - RMSE_{PF}}{RMSE_{PF}} * 100\%; E_{CPF} = \frac{RMSE_{CPF} - RMSE_{PF}}{RMSE_{PF}} * 100\% \tag{18}$$

where E_{PPF} and E_{CPF} denote the parameters error using the proposed PF and constrained PF in comparison to using conventional PF. $RMSE_{PPF}$, $RMSE_{CPF}$ and $RMSE_{PF}$ represent the RMSE of estimated parameters and state CC using the proposed method, constrained PF and conventional PF, respectively.

It follows from Table 3 that the RMSE value of the proposed method is much smaller than the constrained PF and the conventional PF in terms of all parameters and canopy cover over 50 MC simulations. For cdc estimation, the result is not as good as other parameter estimation, the possible reason is that, as shown in Fig 1, CC is not sensitive to cdc parameter for most of the growing period. However, one can still see that the proposed PF and the constrained PF result in 16.4% and 5.2% improvement over the conventional PF in terms of canopy cover RMSE, which is significant in error percentage. The relationship between the parameters cdc, cgc, ccx and CC is complex and generally nonlinear, and the weightings of different parameters on CC are also diverse and time-varying (since the parameter sensitivity

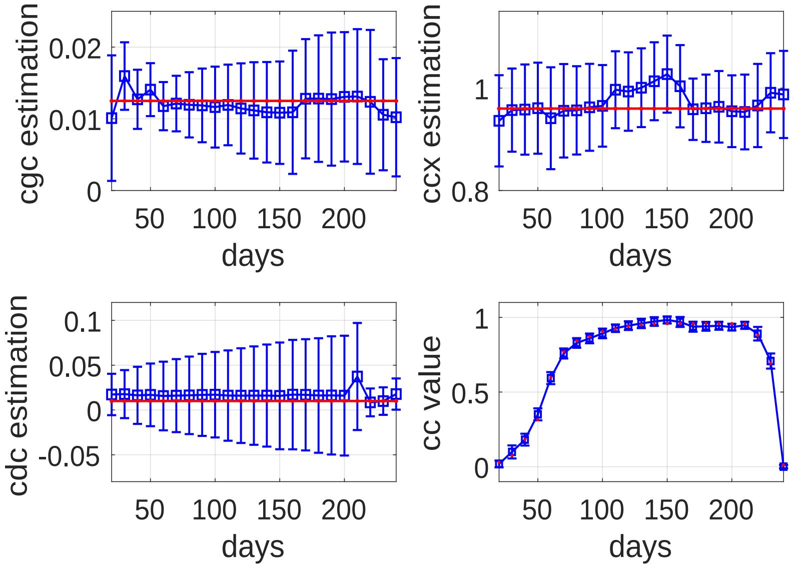

Figure 8: Error bar of the conventional PF on parameters and CC estimation from one MC simulation with groundtruth (red line/points).

is time-varying). As a result, the CC estimation performance may improve even some parameters estimation being poorer (by comparing the constrained PF and conventional PF). But our proposed PF still results in (significantly) better performance over the conventional PF for both CC and model parameters estimation. Consequently, it can be summarised that the proposed method can improve the SSPE estimation performance on both parameter and canopy cover whereas the constrained PF can only marginally improve the estimation performance on *cgc* and canopy cover.

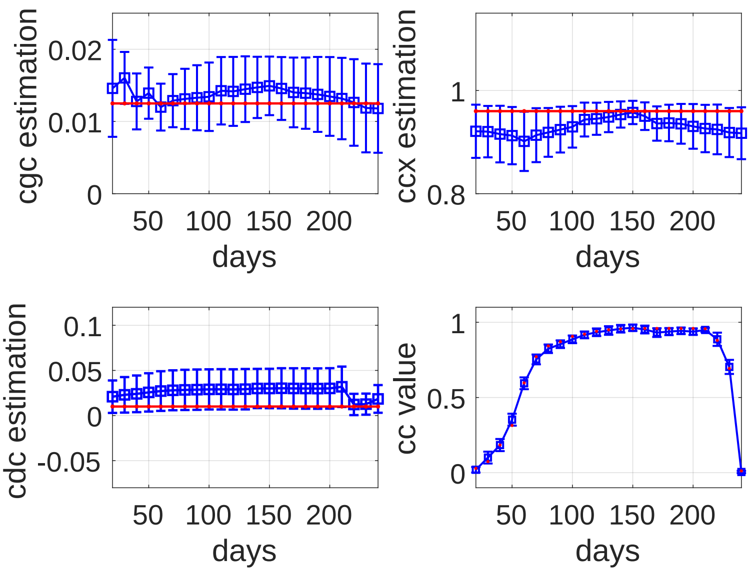

Figure 9: Error bar of the constrained PF on parameters and CC estimation from one MC simulation with groundtruth (red line/points).

An error bar of one MC simulation is given in Fig 8–Fig 10 to evaluate the estimation performance by uncertainty

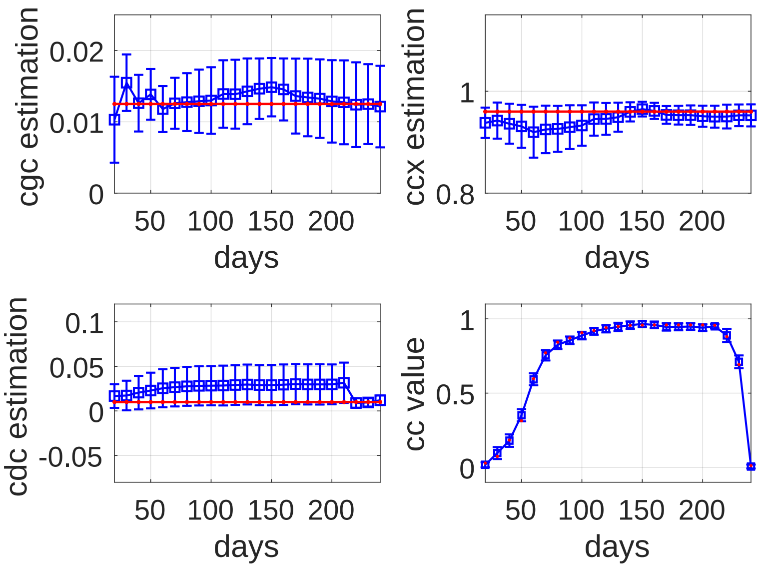

Figure 10: Error bar of the improved PF with constraints and sensitivity information on parameters and CC estimation from one MC simulation with groundtruth (red line/points).

analysis, where mean value and self-defined 1.5 times the standard deviation of all generated particles at each observation day are in blue, and the red line and points denote the groundtruth parameter and CC, respectively. It can be seen from Fig 10 that all estimated parameters are close to the groundtruth after sensitive period. In contrast, some parameters do not converge well by using other PF methods due to the absence of constraints and sensitivity information. Furthermore, it can also be seen from CC results that all estimated CC are all closer to the truth CC by using the proposed PF. In addition to mean value, it can be visually seen that the uncertainties by the proposed method are the smallest among these three methods, mainly due to the sensitivity information making the measurements more efficient in the process of estimation. Therefore, the proposed method achieves the best performance on parameter and CC estimation in terms of stability and accuracy.

In addition, the time-series 3D histogram of state and parameter estimation distribution of one MC run is also displayed in Fig 11. It can be seen that PF can provide posterior distribution of each parameter and state instead of point estimation, and therefore it can provide estimation confidence. The confidence rule is that the less spread the distribution is, the more reliable the estimation is. As is shown in Fig 11.C, the state and parameters distribution with small variance can take effects on sensitive period and thereafter. Consequently, in comparison to the conventional PF and constrained PF, the proposed PF with both constraints and sensitivity information achieve the best estimation with concentrated distributions.

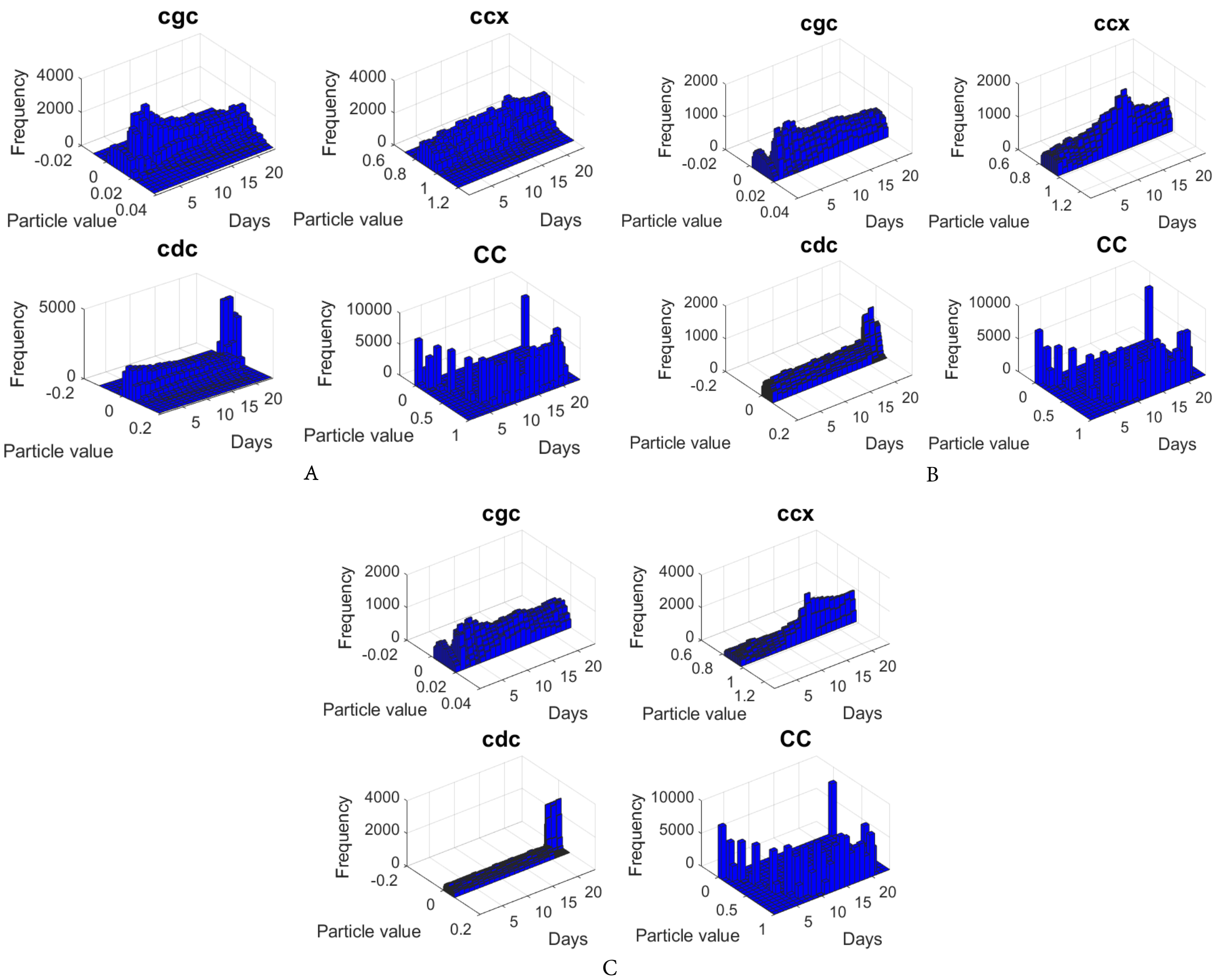

Figure 11: Time-series 3D histogram of one MC run by different methods: A. conventional PF; B. constrained PF; C. proposed PF.

5.2. Results of experimental validation

By using the DFC validation method, the first few folds are used for training and the following three ones are for testing. Testing data at each run derives a RMSE value against canopy cover observation value. The final validation performance is evaluated in terms of mean RMSE of all runs to test the algorithm robustness. In addition to the aforementioned three SSPE approaches, the default parameter-based canopy cover is also simulated for the field experiment using the default parameter values (cgc=0.0111; ccx=0.9051; cdc=0.0300) and noisy CC. The comparative results with error percentage against conventional PF are summarized in Table. 4. It follows from Table 4 that: (1) all SSPE approaches significantly outperform the default parameter based one (i.e. without parameter estimation); (2) our proposed method achieves the smallest RMSE among these four methods.

Table 4: Mean RMSE of different methods by DFC validation.

State	Proposed PF(error)	Constrained PF(error)	Conventional PF	Default Parameter PF(error)
CC	0.0656 (-4.8%)	0.0691(+0.2%)	0.0689	0.0758(+10.0%)

It follows that the experimental validation performance is not as good as MC simulation among three methods. It may be due to the lack of measurements at certain key sensitive stages. Meanwhile, the performance of constrained PF is worse than the conventional PF. The possible reason is that the constrained PF reduces the number of particles and the measurement is not enough to get a better estimation on CC value. Still it can be concluded that all three methods with parameter estimation are capable of solving the SSPE problem for the AquaCrop model, however, the estimation performance of the proposed method considering both constraints and sensitivity information is the best.

6. Conclusions and future work

This paper introduces an improved particle filter framework to integrate UAV multispectral images into AquaCrop model so that state and parameter estimation performance can be improved for the AquaCrop model. Machine learning classifier is applied to UAV multispectral image to calculate canopy cover value for winter wheat. Then particle filter is drawn to assimilate canopy cover information and AquaCrop model information in deriving posterior distributions for state and parameters. Notably, crop sensitivity information is accommodated in the improved particle filter in addition to model parameter bound information. Both Monte Carlo simulations and experimental case study are conducted to assess the performance of the improved particle filter against conventional and constrained particle filter. Monte Carlo simulation shows that the proposed method yields the best performance on state and parameter estimation. The proposed approach also obtains accurate canopy cover estimation in experiment in term of root mean square error. Consequently, the proposed approach provides one alternative to the existing particle filter methods for the simultaneous state and parameter estimation problem.

The AquaCrop model is a very useful model for crop management (e.g. crop growth monitoring, irrigation decision). However, some key crop model parameters should be estimated for local farmlands in order to accurately reflect the local behaviour so that its full potentials can be realized. This study achieves this objective by developing a state and parameter estimation algorithm by using particle filter along with parameter sensitivity information. The developed algorithm generally outperforms the conventional approaches, and therefore results in better simulation performance. As a result, this study provides a better management model to the local farmers, so that they can manage their fields in a more precise and sustainable manner. Therefore, local farmers can potentially benefit from an increased productivity while with a reduced input (e.g. water resources). Although the results are promising, there is still much room for further improvement, the number of observations need to be increased and obtained at crop sensitive stage in real-world experiment; in addition, more state information (e.g. biomass, yield) can be collected to further evaluate the algorithm performance.

Acknowledgements

This work was supported by Science and Technology Facilities Council (STFC) under Newton fund with grant number ST/N006852/1.

References

[1] T. Zhang, J. Su, C. Liu, and W.-H. Chen, "Bayesian calibration of aquacrop model for winter wheat by assimilating uav multi-spectral images," *Computers and Electronics in Agriculture*, vol. 167, p. 105052, 2019.

[2] X. Jin, L. Kumar, Z. Li, H. Feng, X. Xu, G. Yang, and J. Wang, "A review of data assimilation of remote sensing and crop models," *European Journal of Agronomy*, vol. 92, pp. 141–152, 2018.

[3] P. Steduto, T. C. Hsiao, D. Raes, and E. Fereres, "Aquacrop–the fao crop model to simulate yield response to water: I. concepts and underlying principles," *Agronomy Journal*, vol. 101, no. 3, pp. 426–437, 2009.

[4] T. Foster, N. Brozović, A. Butler, C. Neale, D. Raes, P. Steduto, E. Fereres, and T. C. Hsiao, "Aquacrop-os: An open source version of fao's crop water productivity model," *Agricultural water management*, vol. 181, pp. 18–22, 2017.

[5] J. Su, D. Yi, B. Su, Z. Mi, C. Liu, X. Hu, X. Xu, L. Guo, and W.-H. Chen, “Aerial visual perception in smart farming: Field study of wheat yellow rust monitoring,” *IEEE Transactions on Industrial Informatics*, 2020.

[6] B. Bansod, R. Singh, R. Thakur, and G. Singhal, “A comparision between satellite based and drone based remote sensing technology to achieve sustainable development: A review,” *Journal of Agriculture and Environment for International Development (JAEID)*, vol. 111, no. 2, pp. 383–407, 2017.

[7] J. Su, C. Liu, X. Hu, X. Xu, L. Guo, and W.-H. Chen, “Spatio-temporal monitoring of wheat yellow rust using uav multispectral imagery,” *Computers and electronics in agriculture*, vol. 167, p. 105035, 2019.

[8] A. C. Fiala, S. L. Garman, and A. N. Gray, “Comparison of five canopy cover estimation techniques in the western oregon cascades,” *Forest ecology and management*, vol. 232, no. 1-3, pp. 188–197, 2006.

[9] E. Hamuda, M. Glavin, and E. Jones, “A survey of image processing techniques for plant extraction and segmentation in the field,” *Computers and Electronics in Agriculture*, vol. 125, pp. 184–199, 2016.

[10] J. Su, M. Coombes, C. Liu, Y. Zhu, X. Song, S. Fang, L. Guo, and W.-H. Chen, “Machine learning-based crop drought mapping system by uav remote sensing rgb imagery,” *Unmanned Systems*, vol. 8, no. 01, pp. 71–83, 2020.

[11] Z. Jiang, Z. Chen, J. Chen, J. Liu, J. Ren, Z. Li, L. Sun, and H. Li, “Application of crop model data assimilation with a particle filter for estimating regional winter wheat yields,” *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 7, no. 11, pp. 4422–4431, 2014.

[12] M. Hutchinson, C. Liu, and W.-H. Chen, “Source term estimation of a hazardous airborne release using an unmanned aerial vehicle,” *Journal of Field Robotics*, vol. 36, no. 4, pp. 797–817, 2019.

[13] M. S. Arulampalam, S. Maskell, N. Gordon, and T. Clapp, “A tutorial on particle filters for online nonlinear/non-gaussian bayesian tracking,” *IEEE Transactions on signal processing*, vol. 50, no. 2, pp. 174–188, 2002.

[14] Z. Zhu, Z. Meng, T. Cao, Z. Zhang, and Y. Dai, “Particle filter-based robust state and parameter estimation for nonlinear process systems with variable parameters,” *Measurement Science and Technology*, vol. 28, no. 6, p. 065003, 2017.

[15] H. Moradkhani, K.-L. Hsu, H. Gupta, and S. Sorooshian, “Uncertainty assessment of hydrologic model states and parameters: Sequential data assimilation using the particle filter,” *Water resources research*, vol. 41, no. 5, 2005.

[16] H. Li, Z. Chen, W. Wu, Z. Jiang, B. Liu, and T. Hasi, “Crop model data assimilation with particle filter for yield prediction using leaf area index of different temporal scales,” in *2015 Fourth International Conference on Agro-Geoinformatics (Agro-geoinformatics)*, pp. 401–406, IEEE, 2015.

[17] C. Liu, B. Li, and W.-H. Chen, "Particle filtering with soft state constraints for target tracking," *IEEE Transactions on Aerospace and Electronic Systems*, vol. 55, no. 6, pp. 3492–3504, 2019.

[18] R. López-Negrete, S. C. Patwardhan, and L. T. Biegler, "Constrained particle filter approach to approximate the arrival cost in moving horizon estimation," *Journal of Process Control*, vol. 21, no. 6, pp. 909–919, 2011.

[19] N. Amor, N. Bouaynaya, P. Georgieva, R. Shterenberg, and S. Chebbi, "Eeg dynamic source localization using constrained particle filtering," in *2016 IEEE Symposium Series on Computational Intelligence (SSCI)*, pp. 1–8, IEEE, 2016.

[20] H. Xing, X. Xu, Z. Li, Y. Chen, H. Feng, G. Yang, and Z. Chen, "Global sensitivity analysis of the aquacrop model for winter wheat under different water treatments based on the extended fourier amplitude sensitivity test," *Journal of integrative agriculture*, vol. 16, no. 11, pp. 2444–2458, 2017.

[21] A. Kanso, M.-C. Gromaire, E. Gaume, B. Tassin, and G. Chebbo, "Bayesian approach for the calibration of models: application to an urban stormwater pollution model," *Water Science and Technology*, vol. 47, no. 4, pp. 77–84, 2003.

[22] X.-Y. Zhang, M. Trame, L. Lesko, and S. Schmidt, "Sobol sensitivity analysis: a tool to guide the development and evaluation of systems pharmacology models," *CPT: pharmacometrics & systems pharmacology*, vol. 4, no. 2, pp. 69–79, 2015.

[23] J. Nossent, P. Elsen, and W. Bauwens, "Sobolsensitivity analysis of a complex environmental model," *Environmental Modelling & Software*, vol. 26, no. 12, pp. 1515–1525, 2011.

[24] T. Zhang, J. Su, C. Liu, and W.-H. Chen, "Bayesian calibration of aquacrop model," in *Chinese Control Conference, the 37th*, IEEE, 2018.

[25] Z. Zhao, B. Huang, and F. Liu, "A particle filter based on a constrained sampling method for state estimation," in *2012 15th International Conference on Information Fusion*, pp. 816–823, IEEE, 2012.

[26] K. Berntorp and S. Di Cairano, "Process-noise adaptive particle filtering with dependent process and measurement noise," in *2016 IEEE 55th Conference on Decision and Control (CDC)*, pp. 5434–5439, IEEE, 2016.

[27] C. De Bernardis, F. Vicente-Guijalba, T. Martinez-Marin, and J. Lopez-Sanchez, "Particle filter approach for real-time estimation of crop phenological states using time series of ndvi images," *Remote Sensing*, vol. 8, no. 7, p. 610, 2016.

[28] T. Li, M. Bolic, and P. M. Djuric, "Resampling methods for particle filtering: classification, implementation, and strategies," *IEEE Signal Processing Magazine*, vol. 32, no. 3, pp. 70–86, 2015.

[29] C. Bergmeir and J. M. Benítez, "On the use of cross-validation for time series predictor evaluation," *Information Sciences*, vol. 191, pp. 192–213, 2012.

[30] A. Perotti, G. Boella, and A. d. Garcez, "Runtime verification through forward chaining," in *Runtime Verification*, pp. 185–200, Springer, 2015.

[31] J. Su, C. Liu, M. Coombes, X. Hu, C. Wang, X. Xu, Q. Li, L. Guo, and W.-H. Chen, "Wheat yellow rust monitoring by learning from multispectral uav aerial imagery," *Computers and electronics in agriculture*, vol. 155, pp. 157–166, 2018.

[32] A. Ozturk, A. Unlukara, A. Ipek, and B. Gurbuz, "Effects of salt stress and water deficit on plant growth and essential oil content of lemon balm (melissa officinalis l.)," *Pak. J. Bot*, vol. 36, no. 4, pp. 787–792, 2004.

[33] C. Thornton, F. Hutter, H. H. Hoos, and K. Leyton-Brown, "Auto-weka: Combined selection and hyperparameter optimization of classification algorithms," in *Proceedings of the 19th ACM SIGKDD international conference on Knowledge discovery and data mining*, pp. 847–855, ACM, 2013.