Article:

Ran, H, Kang, S, Hu, X et al. (5 more authors) (2022) A framework to quantify uncertainty of crop model parameters and its application in arid Northwest China. Agricultural and Forest Meteorology, 316. 108844. ISSN 0168-1923

https://doi.org/10.1016/j.agrformet.2022.108844

A framework to quantify uncertainty of crop model parameters and its application in arid Northwest China

Hui Ran[a], Shaozhong Kang[b,*], Xiaotao Hu[a,*], Ning Yao[a], Sien Li[b], Wene Wang[a], Marcelo V. Galdos[c], Andrew J. Challinor[c]

[a]*Key Laboratory of Agricultural Soil and Water Engineering in Arid and Semiarid Areas, Ministry of Education, Northwest A&F University, Yangling 712100, China*

[b]*Center for Agricultural Water Research in China, China Agricultural University, Beijing 100083, China*

[c]*Institute for Climate and Atmospheric Science, School of Earth and Environment, University of Leeds, Leeds LS2 9JT, UK*

Corresponding author:

Prof. Shaozhong Kang

Tel: +86 10 62737611

E-mail: kangsz@cau.edu.cn

Prof. Xiaotao Hu

Tel: +86 29 87082117

E-mail: huxiaotao11@nwsuaf.edu.cn

Abstract

Crop modeling is affected by parameter uncertainty. We proposed a framework that integrates sensitivity, uncertainty and parameter calibration of crop models, to provide prediction intervals in place of single values for decision-makers to reduce management risks in agriculture. The framework includes four steps: 1) set prior distributions of parameters and collect measured data, 2) use Morris screening to find out sensitive parameters, 3) adopt Metropolis-Hastings within Gibbs algorithm to calculate posterior distributions of the sensitive parameters and model residual errors, and 4) analyze uncertainties propagation and their applications. The framework was firstly applied on 27 parameters of AquaCrop (version 6.1) on maize in four irrigation scenarios in arid Northwest China, given 5 time series and summary variables including canopy cover (CC), aboveground biomass (B_t), soil water content (SWC), daily evapotranspiration (ET) and final yield (Y) with 1458 measured data points of 27 irrigation treatment-year combinations from 2012 to 2015. The results showed that water stress parameters in AquaCrop were more sensitive in severe drought situations than in full irrigation conditions. The parameter uncertainty brought more variation to simulated final yield than simulated time series variables of maize in arid Northwest China. Model residual error was found to be the major contributor to overall prediction uncertainty, and interannual variation and severe water stress increased its contribution. Adding high-quality measured data of time series variables into MCMC iterations can make the estimated parameters more reliable and more biologically

significant. Medians of outputs using the framework were generally closer to the corresponding measurements when compared with the results of using trial and error method. Especially for SWC and Y, Nash–Sutcliffe coefficient (EF) improved from 0.364 to 0.739 and from 0.055 to 0.415, respectively. The framework is straightforward to be applied to other crop models that can be run in batches.

Key words: Morris method, Metropolis-Hastings within Gibbs, Markov Chain Monte Carlo (MCMC), Bayes' theorem, drought stress, AquaCrop

1. Introduction

Crop models are evolving from deterministic thought to uncertain theory in recent years, with a growing acknowledgement that the simulation results of crop models are greatly affected by various sources of uncertainty. The quantitative information on the reliability of crop model outputs should be carefully analyzed to provide a basis for decision-makers to conduct risk assessment in agricultural management. Wallach and Thorburn (2017) defined prediction uncertainty as to the sum of a bias plus a predictor uncertainty term in model structure, model parameters, and/or model inputs, and suggested that uncertainty assessment should be a standard part of crop models.

Parameter uncertainty is a major source of prediction uncertainty in crop modeling (Wallach et al., 2012). Inferential statistics, including Frequentist statistics and Bayesian statistics, is a major branch in statistics dealing with the problem of parameter uncertainty (Ellison, 2004). Frequentist statistics treats crop models as fixed with input variables known. Model parameters (usually calibrated values) are used to run a specific model. The outputs obtained are a set of fixed values. This approach ignores possible errors in the input data and in the crop model itself. In contrast, Bayesian statistics treats crop models as random and chooses parameters from distributions randomly. It can provide a coherent framework for dealing with uncertainty and is becoming increasingly popular for estimating crop model parameters (Wallach et al., 2019). Tremblay and Wallach (2004) showed that Bayesian methods can perform better for parameter estimation than the least squares method from Frequentist statistics when the ratio between measured data amount and number of parameters is low.

Markov Chain Monte Carlo (MCMC) methods, based on drawing values of parameters from approximated distributions and then correcting those draws to better approximate the target posterior distributions (Gelman et al., 2014), have been important in making Bayesian inference practical for quantifying parameter uncertainty. The Metropolis-Hastings algorithm, which generalizes the basic Metropolis algorithm, is one basic MCMC method. It builds up a chain of parameter vectors by taking a random walk with an acceptance/rejection rule to converge to the posterior distribution (Gelman et al., 2014). Gibbs sampling is another widely used MCMC algorithm that generates values for each parameter, in turn, using conditional probability distributions (Gilks et al., 1996). It has the advantage of not specifying proposal distribution but requires the knowledge of all the conditional distributions which is not always satisfied (Wallach et al., 2019). The Metropolis-Hastings within

Gibbs algorithm, combining Metropolis-Hastings and Gibbs sampling, is an alternative approach on the condition that some of the conditional posterior distributions in a model can be sampled directly and some cannot (Gelman et al., 2014).

In the Metropolis-Hastings within Gibbs algorithm, the Metropolis-Hastings part generates the chain for parameters, using measured data and the latest residual variances; and the Gibbs part generates the chain for residual variances, using measured data and the latest estimated parameters (Wallach et al., 2012). When using the Metropolis-Hastings within Gibbs algorithm, it is necessary to choose a starting value, a proposed distribution, a total number of iterations, and a number of discarded iterations. How to choose these elements is currently an area of active research for crop modelers (Wallach et al., 2019). The choice of starting value is generally not very critical but the choice of the proposed distribution matters (Gilks et al., 1995). Wallach et al. (2012) used the Metropolis-Hastings within Gibbs algorithm to assess the posterior distributions of 15 model parameters and the residual error variances of leaf area index, aboveground biomass, and yield simultaneously. Recently, Gao et al. (2020) applied the Metropolis-Hastings within Gibbs algorithm to estimate posterior distributions for 3 parameters of the phenology model in DSSAT-CERES-Rice under ten different environments. The likelihood function usually adopts a normal distribution (Wallach et al., 2012; Gao et al., 2020). The inverse gamma distribution was adopted in Wallach et al. (2012) and the inverse Wishart distribution in Gao et al.

(2020) to generate residual variances due to their properties of conjugate prior.

One challenge for the application of MCMC algorithms to crop models is the huge computation cost caused by a large number of parameters (Wallach et al., 2019). It is efficient to give priority to parameters that have a large impact on the model outputs to reduce the number of parameters. Sensitivity analysis is an efficient way to recognize influential parameters (Lamboni et al., 2009, 2011; Tan et al., 2017). The screening methods and the variance-based methods are widely used global sensitivity analysis methods. The Morris method, which is based on the computation of the absolute mean elementary effect of individual parameter changes on the model output, is the most commonly used screening approach (Wallach et al., 2019). It is very effective to identify a few influential factors among a large set of parameters (Lu et al., 2021). Studies have used the Morris method to distinguish the influential and non-influential parameters of AquaCrop under different climate-crop-soil combinations (Vanuytrecht et al., 2014; Lu et al., 2021). However, the results are not exactly the same due to different target output variables and various given conditions.

Although methods of estimating parameter distributions have been extensively studied, a framework combining sensitivity analysis and uncertainty quantification for assessing crop model parameters is still lacking. More specifically, some issues remain to be solved:

(i) Few such studies have focused on arid regions. How and to what extent the water stress, which often occurs in such an area, affects uncertainty in crop modeling is unclear.

(ii) Most such studies involve only a few parameters. The application of MCMC algorithms to crop models, including a large number of parameters and model residual errors, remains a challenge.

(iii) Uncertainty quantification based on measured time series variables (e.g., canopy cover, biomass accumulation, dynamic soil water content, and daily evapotranspiration) and summary variables (e.g., crop yield) simultaneously in crop modeling requires further study.

(iv) Sensitivity analysis and uncertainty quantification in different irrigation scenarios are lacking. Irrigation is likely to be a significant factor that affects the sensitivity and uncertainty of crop model parameters in arid climate, which should be taken into account.

The objective of this study is to develop a general framework to quantify uncertainty of crop model parameters in conjunction with model residual errors. The framework is firstly applied on 27 parameters of AquaCrop (version 6.1) on maize in four irrigation scenarios in arid Northwest China. Parameter uncertainties and their propagations to time series canopy cover (CC), biomass (B_t), soil water content (SWC), daily evapotranspiration (ET), and final yield (Y), are explored simultaneously. It is anticipated that this study can provide valuable insights into the application of MCMC algorithms to crop models with a large number of parameters, especially for uncertainty research under drought climates. The framework in this study is expected to be applied with other crop models and for the evaluation of other model outputs in other scenarios.

2. Materials and methods

2.1. Experiments and measurements

The measured data came from the same database described by Ran et al. (2017, 2018). The experiments that made up the database were conducted from 2012 to 2015 at the National Field Scientific Observation and Research Station on Efficient Water Use of Oasis Agriculture in Wuwei of Gansu Province, which is located in an arid region of Northwest China (37°52′ N, 102°50′ E, at 1581 m elevation). The mean annual precipitation is 164 mm, pan evaporation approximately 2000 mm, and groundwater table below 25 m (1955-2005, Li et al. (2015)). Maize in this region is grown between April and September with one harvest per year. The details of soil physical and chemical properties for the 0-100 cm soil layer are shown in Table S1 (see Supplementary material).

The summary of irrigation treatments and measured variables of maize from 2012 to 2015 is presented in Table S2. The measured data collected can be divided into two categories. The first group included daily ET measurements in a large fully irrigated field (300×200 m^2) equipped with an eddy covariance (EC) system for each year from 2012 to 2015. The second group included measurements of different plot irrigation treatments from 2012 to 2015, and the irrigation ranged from 3 to 7 applications per year. Overall, there were 27 irrigation treatment-year combinations in the database. In all cases, the application of fertilizers was enough to avoid nutrient stress. The measured data in the database was classified into four irrigation scenarios based on the irrigation and precipitation amount. They were Full Irrigation (S_{FI}), Deficit Irrigation (S_{DI}), Extreme Deficit Irrigation (S_{EDI}), and All Irrigation (S_{AI}) treatments (Table S2). The measured variables included in-season time series measurements of canopy cover (CC), aboveground biomass (B_t), soil water content (SWC), daily evapotranspiration (ET), and summary variable of final yield (Y).

2.2. Description of the AquaCrop model

AquaCrop is a water-driven dynamic crop model developed by FAO striking a balance among accuracy, simplicity, robustness and ease of use, and focuses on applications in arid regions where water is a key limiting factor in crop production (Steduto et al., 2009; Raes et al., 2009; Hsiao et al., 2009). The model consists of climate, crop, management, and soil modules. The core idea of AquaCrop is evolved from the crop water production function in FAO 33 (Doorenbos and Kassam, 1979), in which yield is calculated from evapotranspiration. To realize the processes of crop development and production, AquaCrop first simulates the daily green canopy cover (CC) including its expansion, ageing, and senescence. Then crop transpiration (T_r) is differentiated from soil evaporation (E_s) by using a "Kc-ET_0" approach based on CC. After that, daily aboveground biomass is calculated by multiplying T_r and the water productivity normalized for atmospheric demand and air CO_2 concentrations (named WP^*). Given the simulated biomass, crop yield is obtained with the help of the reference harvest index (HI_0) but the complex partitioning of biomass among various organs is avoided. AquaCrop actually does not simulate the process of phenology like other agronomic crop models. The phenology is specified as inputs in thermal time (growing degree days) or calendar days by the users. Soil water stress affects the development of CC and alters HI_0. AquaCrop also considers salt and nutritional stress, which is not the scope of this study. Our previous study has carefully described the calculation procedures in AquaCrop (Ran et al., 2018). More details can also be found in Raes et al. (2009, 2018).

In this study, AquaCrop plug-in program (version 6.0), in which the calculation procedures are identical to AquaCrop standard window program (version 6.1), was used because of its accessibility for iterative runs by R (R Core Team, 2013). AquaCrop was run in thermal time mode.

2.3. Description of the framework for quantifying uncertainties

Four steps are included in the framework for quantifying uncertainties (Figure 1).

Step 1. Initialization to set parameters of interest and collect measured data.

A total of 27 parameters in AquaCrop are set as target parameters (listed in Table 1), which are treated as random variables. Their nominal values are cited from Ran et al. (2018). The lower and upper boundaries of the prior distribution for each parameter are defined as ±30% of the nominal values. However, some parameters (e.g., T_{base} and T_{upper}) have their inherent ranges restricted in AquaCrop. Therefore, we have to modify the lower and upper boundaries for these parameters based on literature values and constraints in AquaCrop (Table 1).

The time series variables of CC, B_t, SWC and daily ET, and summary variable of Y are simultaneously used to calculate posterior distributions. Data of weather, management, and initial soil water content are model inputs.

Step 2. Sensitivity analysis to obtain sensitive parameters

To reduce subsequent computational cost, the Morris method is adopted to identify influential factors among the 27 parameters. Crop yield is set as the target variable because of its priority for all crop models. The principle is to calculate a sensitivity index for each parameter and to select parameters with absolute mean effect (μ^*) greater than 0.1 t ha^{-1}. The threshold value of 0.1 t ha^{-1} is determined based on Vanuytrecht et al. (2014) and Silvestro et al. (2017), which is considered to be a reasonable value for deviations in yield assessment studies.

The Morris method defines the elementary effect of the *k*th parameter for a set of parameter value scenarios $Z_i=(z_{i1},\dots, z_{ik-1}, z_{ik}, z_{ik+1},\dots, z_{iK})$ as (Morris, 1991):

$$d_k(Z_i)=\frac{f(z_{i1},\dots,z_{ik-1},z_{ik}+\Delta,z_{ik+1},\dots,z_{iK})-f(z_{i1},\dots,z_{ik-1},z_{ik},z_{ik+1},\dots,z_{iK})}{\Delta} \quad (1)$$

where $f(Z_i)$ is the model output (final yield in this study), $Z_i=(z_{i1},\dots,z_{iK})$ is the K-dimensional parameter vector, K is the number of parameters (K=27 in this study). Δ is a predetermined multiple of the grid spacing (i.e., size of grid jump).

Then the absolute mean and the standard deviation of the elementary effects of $d_k(Z_i)$ are calculated as (Morris, 1991; Campolongo et al., 2007):

$$\mu_k^* = \frac{\sum_{i=1}^{r} \left| d_k(Z_i) \right|}{r} \qquad (2)$$

$$\sigma_k = \sqrt{\frac{\sum_{i=1}^{r} \left[d_k(Z_i) - \frac{1}{r} \sum_{i=1}^{r} d_k(Z_i) \right]^2}{r}} \qquad (3)$$

where r is the number of trajectories. A high μ_k^* indicates a factor with an important influence on the output. σ_k estimates the ensemble of a factor's higher order effects, i.e. nonlinear effects and/or interfactor effects (Campolongo et al., 2007).

The Morris method is implemented with the help of *Morris* function of *sensitivity* package in R with levels=6, a jump Δ=3 (following Morris's recommendation of levels/2), and a number of trajectories r=500, which follows Wallach et al. (2019) and Lu et al. (2021). Thus a total number N=14000 of model evaluations is performed (N=r×(K+1)). Furthermore, sensitivity analysis is implemented for each of the four irrigation scenarios.

Step 3. Uncertainty quantification and posterior distributions.

The main idea for uncertainty quantification is to use the Markov Chain Monte Carlo (MCMC) algorithm under Bayes' theorem. The sensitive parameters (θ) and model residual variance (σ^2) of CC, B_t, SWC, ET, and Y are treated as random quantities and are assumed to be independent. All model residuals (ε) are assumed to be independent and identically distributed (*iid*) with normal distributions that have expectation 0 and variance σ_v^2. That is (Wallach et al., 2012),

$$M_{vi} = \hat{M}_{vi}(\theta) + \varepsilon_{vi} \qquad (4)$$

$$\varepsilon_{vi} \underset{iid}{\sim} N\left(0, \sigma_v^2\right) \qquad (5)$$

where M_{vi} and $\hat{M}_{vi}$ are the ith measured and simulated value of variable v, respectively, v stands for CC, B_t, SWC, daily ET, or Y in this study.

The basic equation for Bayesian parameter estimation is (Wallach et al., 2012):

$$P\left(\theta,\sigma^2 \middle| M\right) \propto P\left(M \middle| \theta,\sigma^2\right) P\left(\theta\right) P\left(\sigma^2\right) \quad (6)$$

where $P(\theta,\sigma^2|M)$ is posterior distribution given the observed values M. $P(M|\theta,\sigma^2)$ is a likelihood. $P(\theta)$ and $P(\sigma^2)$ are prior distributions for θ and σ^2, respectively. $P(\theta)$ is assumed a uniform distribution and the boundaries of each parameter are shown in Table 1. $P\left(\sigma^2\right)=\prod_{v=1}^{n} 1/\sigma_v{}^2$, which is a commonly used non-informative prior distribution (Wallach et al., 2012).

The specific approach used here is the Metropolis-Hastings within Gibbs algorithm (Wallach et al., 2012; Gao et al., 2020). It separates the estimation of parameters (in the Metropolis-Hastings step) from the estimation of the model residual error variance (in the Gibbs step).

(i) In the Metropolis-Hastings step, posterior distributions of parameters are calculated as:

$$P\left(\theta \middle| M,\sigma^2\right) \propto P\left(M \middle| \theta,\sigma^2\right) P\left(\theta\right) \quad (7)$$

The likelihood function is assumed to have a normal distribution as (Wallach et al., 2012):

$$P\left(M \middle| \theta,\sigma^2\right)=\prod_{v=1}^{n}\left(2\pi\sigma_v^2\right)^{-N_v/2} exp\left\{-\frac{\sum_{i=1}^{N_v}\left[M_{vi}-\hat{M}_{vi}\left(\theta\right)\right]^2}{2\sigma_v^2}\right\} \quad (8)$$

where n is the number of measured variables (n=5 in this study), N_v is the number of samples for each measured variable. A logarithmic transformation for the measured data of CC, B_t, SWC, ET, and Y is conducted separately to stabilize the variances in these series. Function *log1p* in R is applied to the measured data to prevent applying a logarithm to 0 values.

(ii) In the Gibbs step, the posterior distribution of residual variance is calculated using an inverse *gamma* distribution, i.e., σ^{-2} has a *gamma* distribution as:

$$P\left(\sigma^{-2}\middle|M,\theta\right)=\frac{1}{\Gamma\left(\alpha_v\right)\beta_v^{\alpha_v}}\left(\sigma^{-2}\right)^{\alpha_v-1}e^{-\sigma^{-2}/\beta_v} \tag{9}$$

where $\Gamma(\alpha_v)$ denotes the *gamma* function calculated at α_v.

The shape parameter (α_v) and scale parameter (β_v) of the *gamma* distribution for variable v are (Wallach et al., 2012):

$$\alpha_v = 2 + N_v / 2 \tag{10}$$

$$\beta_v = 2\Big/\sum_{i=1}^{N_v}\left[M_{vi}-\hat{M}_{vi}\left(\theta\right)\right]^2 \tag{11}$$

To realize the above two sub-steps, a Markov chain of values ($\theta^{(1)}$, $\sigma^{2(1)}$,…, $\theta^{(t)}$, $\sigma^{2(t)}$,…, $\theta^{(m)}$, $\sigma^{2(m)}$) is iterated by taking random steps in parameter spaces due to the lack of analytical expressions for their posterior distributions.

In the Metropolis-Hastings step, a proposal $\theta^{*(t+1)}|\theta^{(t)}$ is drawn from a multivariate normal distribution:

$$\theta^{*(t+1)}\middle|\theta^{(t)} \sim N\left(\theta^{(t)}, tune\times\Sigma\right) \tag{12}$$

where *tune* is a dynamic factor with initial value being set to 1, and it is multiplied by

2 if the rejection rate is <0.65 and divided by 2 if >0.85 to make sure the final acceptance rate of the proposed $\theta^{*(t+1)}|\theta^{(t)}$ around the recommended rate of 25% (Wallach et al., 2012); Σ is a diagonal matrix with the prior variances for each parameter on the diagonal, and the prior variance for uniform distribution is calculated as (upper boundary - lower boundary)^2/12.

After proposing a $\theta^{*(t+1)}$, an acceptance ratio $A(\theta^{*(t+1)}, \theta^{(t)})$ that decides whether to accept or reject the candidate is calculated as (Gelman et al., 2014):

$$\begin{aligned} A\left(\theta^{*(t+1)},\theta^{(t)}\right) &= min\left(1, \frac{P\left(M\left|\theta^{*(t+1)},\sigma^{2(t)}\right.\right)P\left(\theta^{*(t+1)}\right)P\left(\theta^{(t)}\left|\theta^{*(t+1)}\right.\right)}{P\left(M\left|\theta^{(t)},\sigma^{2(t)}\right.\right)P\left(\theta^{(t)}\right)P\left(\theta^{*(t+1)}\left|\theta^{(t)}\right.\right)}\right) \\ &= min\left(1, \frac{P\left(M\left|\theta^{*(t+1)},\sigma^{2(t)}\right.\right)}{P\left(M\left|\theta^{(t)},\sigma^{2(t)}\right.\right)}\right) \end{aligned} \tag{13}$$

Here, we have $P\left(\theta^{(t)}\left|\theta^{*(t+1)}\right.\right)=P\left(\theta^{*(t+1)}\left|\theta^{(t)}\right.\right)$ because the proposal distribution is multivariate normal. In addition, $P\left(\theta^{*(t+1)}\right)=P\left(\theta^{(t)}\right)$ because the prior distribution is the same uniform distribution.

Then generate a uniform random number $u\in[0,1]$. If $u<A\left(\theta^{*(t+1)},\theta^{(t)}\right)$, $\theta^{(t+1)}=\theta^{*(t+1)}$, otherwise, $\theta^{(t+1)}=\theta^{(t)}$.

In the Gibbs step, a value of $\sigma^{2(t+1)}$ is generated by sampling from the conditional distribution with the help of *rgamma* function in R and then taking the inverse.

To obtain stable results, three chains with three different starting points ($\theta^{(0)}$) for each of the four scenarios (a total of 12 chains) are run in parallel on a workstation (Intel(R) Xeon(R) CPU 2.20 GHz, 12 Kernels). Each chain is iterated 300,000 times to make the convergence toward the posterior distribution. The test basis for whether the convergence has occurred is the upper limit of the Gelman criterion, which should be below 1.1, as calculated by the R *coda* package (Plummer et al., 2006). Autocorrelation of the chains of each parameter is checked with the *acf* function of R. The first half of each chain is eliminated to remove the effect of starting value, and then the remaining vectors in each chain are combined to give a single chain of 450,000 vectors. Finally, the chains for different scenarios are thinned, keeping only one vector out of 1000 according to the results of *effectiveSize* function in *coda* package, to reduce autocorrelation, and 451 vectors of parameters are left. Uncertainties of T_{base} and T_{upper} are not considered in this study owing to phenology being specified by users rather than simulated by AquaCrop.

Step 4. Analysis and application.

After obtaining the parameter sensitivity using the Morris method, its dependence on target variables and irrigation scenarios is explored. Next, posterior parameter distributions are compared under the four irrigation scenarios to investigate the influence of measurements. In particular, the difference of posterior $Kc_{Tr,x}$ caused by whether involving measured daily ET during the MCMC iteration is investigated to emphasize the importance of measured data of intermediate variables on posterior parameter distributions (Figure 6). Then the posterior parameter distributions are used to calculate the distributions of CC, B_t, SWC, ET, and Y. The percentages of measured values that fell in different percentile ranges of corresponding predictions are calculated. After that, the contributions of parameter uncertainty (var_{parm}) and model residual variance (var_{model}) to overall prediction uncertainty (var_{pred}) for variable v are calculated using the following equations:

$$\begin{aligned} var_{parm} &= var\left\{\hat{M}_v\left(\theta_j\right)\right\} \\ var_{model} &= var\left\{\varepsilon_{vj}\right\} \\ var_{pred} &= var\left\{\hat{M}_v\left(\theta_j\right)+\varepsilon_{vj}\right\} \end{aligned} \qquad (14)$$

where θ_j is the jth parameter vector in the posterior distribution (Figure 3, j=1,...,451 in this study), ε_{vj} is a sample with one element drawn from $N(0, \sigma_{vj}^2)$ for each j, σ_{vj}^2 is the jth model residual variance in the posterior distribution (Figure 3). The difference between var_{pred} and the sum of var_{parm} and var_{model} is a measure of the interaction between the variance due just to parameter uncertainty and the variance due just to model residual error (Gao et al., 2020). Finally, the medians of model outputs are compared with the previous parameterization results using the trial and error method based on the same measurements in Ran et al. (2017, 2018).

2.4. Statistical analysis

The performance of the crop model was assessed using six statistical indices. They were regression coefficient through the origin (b_0), coefficient of determination (R^2), root mean square error (RMSE), normalized root mean square error (NRMSE), Nash–Sutcliffe model efficiency coefficient (EF), and Willmott's index of agreement (d). The formulas for these statistical indices can be found in Ran et al. (2020). The components of the framework, including data organization, crop model simulation, sensitivity analysis, uncertainty quantification, statistical analysis and plotting were programmed in R (please contact the corresponding author to access the code).

3. Results

3.1. Morris results based on final yield

Parameter sensitivity varied with different irrigation scenarios (Figure 2). The number of sensitive parameters, based on the criteria of μ^*>0.1 in this study, was 17, 18, 20, and 18 out of 27 for S_{FI}, S_{DI}, S_{EDI}, and S_{AI}, respectively. On the other hand, the major influential parameters demonstrated some similarities among different scenarios. Parameters with important influence (μ^*>0.1) for all scenarios were HI_0, WP^*, $Kc_{Tr,x}$, CGC, CC_x, GDD_{min}, T_{base}, fcsoil, wpsoil, CDC, p_{sen}, cc_0, T_{upper}, and p_{expl}. Parameters with negligible influence (μ^*<0.1) for all scenarios were Kc_{Trxle}, Ksatsoil, $cold_s$, and $root_s$. For S_{FI}, S_{DI}, and S_{AI}, the top three sensitive parameters were HI_0, WP^*, and $Kc_{Tr,x}$. For S_{EDI}, however, the difference was evident. The wpsoil was the most sensitive parameter. In addition, Parameters related to water stress, e.g., p_{sen}, p_{pol}, p_{sto}, and p_{sens}, became more sensitive in this scenario.

3.2. Posterior distributions for sensitive parameters

The acceptance rates of the proposed parameter vectors during the MCMC iterations for the four scenarios were 24.9%, 25%, 24.7%, and 24.9% (Table S3), which were close to the recommended rate of 25% (Wallach et al., 2012). Convergence diagnosis showed most values for each parameter as well as the multivariate value were below 1.1 (Table S3), which indicated that convergence toward the posterior parameter distribution had occurred. Some of the values were slightly above 1.1, but we examined graphs of each parameter versus iteration number and those indicated that it was reasonable to assume that all the Markov chains converged to the stationary distributions.

The shapes and the ranges of posterior parameter distributions were highly related to the four different irrigation scenarios (Figure 3). Some medians of posterior parameters in S_{EDI}, e.g. WP^* and $Kc_{Tr,x}$, were smaller than those in the other three scenarios, and some, e.g. fcsoil and GDD_{min}, were greater. On the other hand, the posterior parameter distributions showed some similarities among different scenarios. For example, the medians of HI_0, CGC, CC_x, and wpsoil were around 32%, 0.011, 85%, and 7.65% for the four scenarios, respectively (Table S4). In addition, the posterior distributions of WP^*, $Kc_{Tr,x}$, CGC, and CC_x were much narrower than the prior. The posterior distribution of HI_0 was very similar to the prior. The posterior residual variances of CC, B_t, SWC, ET, and Y can be recognized as inverse gamma distributions, which was in line with the properties of conjugate prior.

3.3. Prediction uncertainty

The simulated time series variables of CC, B_t, SWC, and ET using the posterior parameter distributions in the S_{EDI} scenario (only one treatment of 2013W3) are shown in Figure 4. Simulations of CC, B_t, SWC, and ET generally followed the trend of measured values, and they were better than the results in Ran et al. (2018) in this scenario. Most of the measured values were covered by or close to the 0th to 100th percentile band (Figure 4). However, for the other three scenarios (S_{FI}, S_{DI}, and S_{AI}), the 0th to 100th percentile band did not cover all measured values (Figure S1-S12).

The percentages of mean values of measured CC, B_t, SWC, and ET that fell in the 25th to 75th percentile band were 4%-16% in S_{FI}, S_{DI}, and S_{AI} scenarios, and they were 15%-62% in the 0th to 100th percentile band (Table 2).

The simulated summary variable of Y using the posterior parameter distributions in the four irrigation scenarios is shown in Figure 5. The 25th to 75th percentile band and of 0th to 100th percentile band were much wider than those of the simulated time series variables. The percentages of mean values of measured Y that fell in the 25th to 75th percentile band (33%-80%) and the 0th to 100th percentile band (89%-100%) were also much larger in S_{FI}, S_{DI}, and S_{AI} scenarios (Table 2). However, the single mean value of measured Y was outside the 0th to 100th percentile band in S_{EDI} scenario (Table 2, Figure 5).

The sum of the variance due just to parameters (var_{parm}) and the variance due just to model residual error (var_{model}) was almost equal to the total variance (var_{pred}) for each output variable (Table 3), which suggested that there was little interaction between the parameter uncertainty and the model residual error. For S_{FI}, S_{DI}, and S_{AI} scenarios, more than 95% of the total prediction uncertainty of CC, B_t, SWC, ET, and Y came from model residual errors. For S_{EDI} scenario, the ratios of parameter uncertainty and residual error uncertainty of the five target output variables to total prediction uncertainty varied from 0.1% to 24.1% and from 79.0% to 99.5%, respectively.

Medians of simulated CC, B_t, SWC, ET, and Y using the posterior parameter distributions were generally closer to the measured values, when they were compared to simulations of pre-calibrated AquaCrop using the trial and error method in Ran et al. (2017, 2018). In particular, EF of the medians for SWC and Y increased from 0.364 to 0.739 and from 0.055 to 0.415, respectively (Table 4).

4. Discussion

4.1. How to make prediction uncertainty a standard part of crop models?

Probability distributions of model outputs are useful because they give information about if the results are sufficiently reliable. Although a growing acknowledgement and characterization of uncertainty in crop model predictions is dominant in recent years, many crop models themselves, such as AquaCrop, DSSAT, etc., currently have no module to handle uncertainty. Recently, Gao et al. (2020) created an R version of the phenology model in DSSAT to study the parameter uncertainties instead of using DSSAT itself. How to quantify the uncertainty of a large number of parameters in crop models is still a big challenge so far.

The framework developed in this study integrating sensitivity and uncertainty algorithms can address this issue, provided that the target crop model can be run in batches with inputs and outputs. The framework uses R to modify the input files, invoke the execute program and read the output files of crop models. This process is repeated hundreds of thousands of times, which is necessary for MCMC iteration. This framework does not require the source code of crop models, therefore, it can theoretically be applied to any deterministic crop model which can be run in batches. The framework first uses the Morris screening to keep out non-sensitive parameters to reduce the number of parameters that need to be quantified for uncertainty. Then it adopts the Metropolis-Hastings within Gibbs algorithm to quantify the uncertainty of the remaining sensitive parameters.

The first application of the framework on AquaCrop for maize in Northwest China showed that posterior variances of parameters were generally much smaller than the prior (Table S4). Especially for parameters like WP^*, $Kc_{Tr,x}$, and fcsoil, they are dozens to hundreds of times lower than the prior variances. The propagation of parameter uncertainty to output variables is also quantified. The results demonstrate that the framework is successfully implemented on AquaCrop in arid Northwest China, and is straightforward to be applied on other crop models in other environments.

4.2. Why should sensitivity and uncertainty be conducted in given scenarios?

Our sensitivity analysis results are partially different from previous studies (Vanuytrecht et al., 2014; Silvestro et al., 2017; Lu et al., 2021), especially for the parameters of HI_0, WP^*, and $Kc_{Tr,x}$. The difference comes from the specific pre-defined ranges of these parameters. $Kc_{Tr,x}$ with a pre-defined range of 1.00-1.10, WP^*, 30-35 g m^{-2}, and HI_0, 46-50% in Vanuytrecht et al. (2014) and Lu et al. (2021). HI_0 with a range of 40-55% is pre-defined in Silvestro et al. (2017). The pre-defined ranges for these parameters are much wider in our study (Table 1).

Parameter sensitivities also vary with the target model output. Yield is the variable of interest in this study. If the target variable changes to final aboveground biomass,

WP^*, rather than HI_0, is the most sensitive parameter in the full irrigation scenario (data not shown). This is expected as the parameter of HI_0 affects yield formation rather than biomass accumulation processes.

In addition, the sensitivity analysis results of a same set of parameters and same target output in different irrigation scenarios are also different. For example, the sensitivity of water stress coefficients is generally low in full irrigation scenario while it is high in extreme deficit irrigation scenario (Figure 2). This is also expected because when the irrigation amount is enough to avoid water stress, the consequently related water stress coefficients would have no impact on the simulation of crop growth. Roux et al. (2014) also showed that uncertainty in a water stress parameter may lead to large uncertainty in water stress situations, but to little uncertainty in well-watered conditions. Furthermore, our results demonstrate that irrigation is a significant factor that affects the sensitivity and uncertainty of crop model parameters in arid climate, which should be carefully considered.

Although it is impossible to derive a list of sensitive parameters that are universally valid for all scenarios (Vanuytrecht et al., 2014; Lu et al., 2021), overlaps of influential parameter subsets under different irrigation scenarios can serve as a guide for calibrating AquaCrop in other environments. HI_0, WP^*, $Kc_{Tr,x}$ are generally the most sensitive parameters, and one needs to give priority to calibrate these three parameters when using AquaCrop. HI_0 determines how much biomass is allocated to crop yield. WP^* controls how much biomass is produced from transpiration. $Kc_{Tr,x}$

affects how much transpiration occurs which is the basis for biomass calculation. The most sensitive parameter in extreme deficit irrigation scenario is wpsoil. It indicates that how much water in the soil can be used by crops is critical in extreme drought conditions.

4.3. How important is the measured data to the posterior distribution?

When the measured daily ET data is not used during the MCMC iterations, the posterior distribution of $Kc_{Tr,x}$ is close to 1.4. After adding the measured daily ET, however, the posterior distribution value of $Kc_{Tr,x}$ is distributed around 1.17 (Figure 6), which is more biologically significant and is closer to the measured value. It suggests that 1) incorporating measured data of intermediate variables has a significant influence on the posterior distributions of the parameters that directly associated with these variables, and 2) adding intermediate measured data makes the estimated parameters more reliable. There should be sufficient interactions among the components of a system that, unless the detailed characteristics of these components can be specified independently, many representations may be equally acceptable (Beven and Freer, 2001). Generally, the number of estimated parameters should be substantially fewer than the number of observations to avoid overfitting and consequently poor predictive quality (Tremblay and Wallach, 2004). Our study, going a step further, indicates that high-quality measured data of intermediate time series variables needs to be incorporated into MCMC iteration to obtain reliable posterior parameter distributions for process-based crop models. Furthermore, the posterior distributions of the parameters should be carefully analyzed according to their biological meaning and need to be compared with their measured values (if available) to avoid overfitting and equifinality.

The measured data affects not only the posterior distributions of parameters but also the distributions of model outputs. For example, the distributions of CC, B_t, SWC, and ET of treatment 2013W3 in S_{EDI} (Figure 4) are much different from those in S_{AI} (Figure S9-S12, 2013W3). The reason is that the measured data used for the MCMC iteration in S_{EDI} involves only one treatment of 2013W3, while the iteration in S_{AI} scenario involves 27 treatments of four years and 2013W3 is just one of them. We also find that the distributions of CC, B_t, SWC, and ET of treatment 2013W3 in S_{EDI} are much closer to the measured values than in S_{AI}. It indicates that AquaCrop might have a problem to handle interannual variation.

Although the percentage of mean measured time series values that fall in the 25th to 75th percentile band in S_{FI}, S_{DI} and S_{AI} scenarios is relatively low, there are many intersections between error bars and confidence intervals (Figure S1-S12). Measurement error is also another important uncertainty source for crop modeling (Confalonieri et al., 2016), which is beyond the scope of this study.

4.4. What can we learn from the posterior distribution?

For some parameters (e.g., WP^*, $Kc_{Tr,x}$, CGC, and CC_x), the posterior distribution is much narrower than the prior, which allows us to considerably narrow the possible range of values. However, for other parameters (e.g., HI_0), the posterior distribution is similar to the prior. There is little information in the distribution that would allow us to reduce our initial uncertainty about these parameters.

The posterior distributions of WP^* for all the four scenarios are around 20 g m^{-2} (Table S4), which is close to the value of 20.9 g m^{-2} that was derived using measured biomass and EC data from 2012 to 2015 in Ran et al. (2018). On one hand, it indicates that the method to derive WP^* from the first derivative of the linear regression between measured biomass and the sum of normalized evapotranspiration in Ran et al. (2018), originated from Hsiao et al. (2009), can stand the proof. On the other hand, it seems feasible to obtain crop model parameters through algorithms. This does not necessarily mean that algorithms can replace measurements in model calibration. Parameters derived from measurements should be advocated in crop models, if available, to make the crop model outputs more reliable. We also found posterior distributions of soil parameters of fcsoil (close to 26%) and wpsoil (around 7%), two key parameters determining SWC, are away from their nominal values (30% and 10%). The use of the posterior values of fcsoil and wpsoil does improve the accuracy of SWC simulation when compared with using their nominal values (Table 4). It gives us new information that nominal values of fcsoil and wpsoil may be inaccurate, which needs further study.

The simulated results of AquaCrop are generally acceptable considering the multiple irrigation treatment-year combinations and the extremely arid climate in this study. However, model fitting to the measured data is partly unsatisfying. The decomposition of total prediction error into parameter uncertainty and model residual error shows that residual error makes the major contribution (Table 3). On one hand, it suggests that improving the internal calculation process of AquaCrop may lead to significant improvement in simulation accuracy rather than further calibrating the model parameters in this case. On the other hand, the reasons for the low contribution of model parameter uncertainty are different for different scenarios. For the full irrigation scenario (S_{FI}), the low contribution occurs given the bands of simulated yield are wide and cover all the measured values. It indicates that for crop models like AquaCrop that intend to adopt a same set of parameters for all the four years rather than different sets of parameters for each year, extra uncertainty will be brought into model residual error to weaken the contribution of parameter uncertainty. In addition, for the extreme deficit irrigation scenario (S_{EDI}), the low contribution occurs on the condition that the measured yield is outside the band of simulated yield. It suggests that severe water stress will bring greater variance to model residual error. The HI simulation that controls the yield formation process in AquaCrop under water stress conditions is poor and its improvement has been carefully studied in Ran et al. (2019).

4.5. Limits and future challenges?

Only one crop model, i.e. AquaCrop, is currently considered in the framework. More crop models should be involved in the future. Uncertainty of parameters related to phenology (e.g., T_{base} and T_{upper}) in AquaCrop are not considered in this study because phenology is specified as input by users. However, phenology is generally an important source of uncertainty (Gao et al., 2020), which should be considered in crop models. In addition, the framework only considers the parameter uncertainty, and the uncertainties of input data and model structure should be involved in this framework in the future. How to quantify the uncertainties of input, parameter, and model structure simultaneously is a great challenge in crop modeling, and research is lacking on this topic.

In this study, model residual errors are randomly generated from inverse gamma distributions under the assumption of independent and identically distributed with normal distributions. That means the covariances of residuals between variables are ignored. When covariances need to be considered, a common practice is to use inverse Wishart distribution to generate the variance-covariance matrix of residual errors (Gelman et al., 2014). However, the application of inverse Wishart distribution to handle residuals of multiple time series variables with nonhomogeneous data size is complicated and rarely reported in crop modelling.

The main purpose of the framework in this study is to quantify prediction uncertainty instead of finding the optimal parameters, although the results show that the framework does improve the accuracy of simulations when compared to the trial and error method. For example, the medians of SWC and Y are closer to the measurements, with EF improved from 0.364 to 0.739 and from 0.055 to 0.415, respectively (Table 4). However, it is computationally expensive if one uses the framework to find the optimal parameters. The Metropolis-Hastings within Gibbs algorithm tends to move to a smaller error at each iteration, but can also move to a larger error with a certain probability, and does not specifically look for the minimum error (Gao et al., 2020). Studies showed that other methods, e.g. PEST, are much faster at optimizing parameter values (Ma et al., 2020).

The purpose of sensitivity analysis in the framework is to quickly identify sensitive parameters among a large number of parameters to reduce subsequent computational costs. However, the Morris screening method is limited by the inability to quantify the source of variance. Adding the variance-based methods, like the new methodology in Lamboni et al. (2021), into the framework to perform dependent multivariate sensitivity is interesting. In addition, we set the boundaries of the parameters according to their biological meaning, add measured data of time series variables in the MCMC iterations, and compare the posterior distributions of the parameters to their measured values to avoid overfitting. However, one needs to carefully assess the overfitting issue when using the framework with one's own data.

In the framework, 300,000 iterations are needed to make the Markov chains converge given 17-19 parameters, 5 target variables, and 1458 measured data of 27 treatment-year combinations. Gao et al. (2020) found that 10,000 iterations are sufficient to make the chains converge for 3 parameters. More than 200,000 iterations were run for 15 parameters with 3 target variables in Wallach et al. (2012). The computation cost is the primary challenge for the application of the Metropolis-Hastings within Gibbs algorithm in the framework. Although we run 12 chains in parallel for the four irrigation scenarios, it takes an average of 9.19 days to complete these chains with 300,000 iterations of each. The algorithm needs to be further improved to reduce the time to convergence.

5. Conclusions

We have developed a framework that integrates sensitivity, uncertainty and parameter calibration of crop models, and have demonstrated its application to quantify the uncertainties of parameters and model residual errors of AquaCrop on maize in arid Northwest China. The framework has the ability to select sensitive parameters from a large number of parameters, and can also clarify the difference in sensitivity of a parameter between well-watered conditions and extreme water stress situations. For crop models like AquaCrop that intend to adopt similar parameters across years, interannual variation and severe water stress cause extra uncertainty of residual error, and thereby weaken the contribution of parameter uncertainty to the total prediction uncertainty. The different propagation of parameter uncertainty into output time series variables and summary variable is found using the framework. High-quality measured data of intermediate variables help the framework to obtain more reliable posterior parameter distributions. The framework can also improve the simulations of AquaCrop when comparing with using the trial and error method. It would be straightforward to use the framework on other crop models under other scenarios.

The framework only considers parameter uncertainty currently, but aims to quantify multiple uncertainties of input, parameter, and model structure simultaneously in the future, which is a great challenge in crop modeling and needs extensive study. In addition, adding the variance-based methods into the framework to perform dependent multivariate sensitivity is interesting. The overfitting issue should also be carefully assessed when one uses the framework.

Acknowledgements

The authors thank Prof. Ping Guo (CAU, China) for her valuable comments and suggestions on the manuscript. We thank Dr. Yu Hu (NWAFU, China) for her help in performing parallel runs using R. We thank Dr. Jintao Wang (IGDB, CAS, China) for kindly providing part of valuable measured data. We thank Dr. Yujing Gao (ZJU, China) for her advice on MCMC algorithm. The R code of this study is partially taken from Prof. Daniel Wallach's book of *Zebook* demo *ch10maizeB7Bayes_functionsMH*. We also want to express our sincere gratitude to the anonymous reviewers and editor for their insightful comments and kind help to improve this manuscript. This work is financially supported by the National Natural Science Foundation of China (grants 51809216 and 51790534), the Project Supported by Natural Science Basic Research Plan in Shaanxi Province of China (Program No. 2019JQ-646), the National Field Scientific Observation and Research Station on Efficient Water Use of Oasis Agriculture in Wuwei of Gansu Province (No. KF2021005), and the scholarship fund from China Scholarship Council (CSC, No. 202006305001).

References

[1] Beven, K. and Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. Journal of Hydrology, 249(1-4): 11-29.

[2] Campolongo, F., Cariboni, J. and Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. Environmental Modelling and Software, 22 (10): 1509-1518.

[3] Confalonieri, R., Bregaglio, S. and Acutis, M., 2016. Quantifying uncertainty in crop model predictions due to the uncertainty in the observations used for calibration. Ecological modelling, 328: 72-77.

[4] Doorenbos, J., Kassam, A.H., 1979. Yield Response to Water. FAO Irrigation and Drainage Paper 33. FAO, Rome, Italy, p. 257.

[5] Ellison, A.M., 2004. Bayesian inference in ecology. Ecology Letters, 7(6): 509-520.

[6] Gao, Y., Wallach, D., Liu, B., Dingkuhn, M., Boote, K. J., Singh, U., Asseng, S., Kahveci, T., He, J., Zhang, R., Confalonieri, R. and Hoogenboom, G., 2020. Comparison of three calibration methods for modeling rice phenology. Agricultural and Forest Meteorology, 280: 107785.

[7] Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A. and Rubin, D. B., 2014. Bayesian Data Analysis. CRC Press, Boca Raton, FL.

[8] Gilks, W.R., Richardson, S. and Spiegelhalter, D.J., 1995. Markov Chain Monte Carlo in Practice. Chapman and Hall, London.

[9] Gilks, W.R., Richardson, S. and Spiegelhalter, D.J., 1996. Introducing Markov chain Monte Carlo. In: W.R. Gilks, S. Richardson and D.J. Spiegelhalter (Editors), Markov chain Monte Carlo in practice. Chapman and Hall, London, pp. 1-19.

[10] Hsiao, T. C., Heng, L., Steduto, P., Rojas-Lara, B., Raes, D. and Fereres, E., 2009. AquaCrop-the FAO crop model to simulate yield response to water: III. Parameterization and testing for maize. Agronomy Journal, 101(3): 448-459.

[11] Lamboni, M. and Kucherenko, S., 2021. Multivariate sensitivity analysis and derivative-based global sensitivity measures with dependent variables. Reliability Engineering and System Safety, 212: 107519.

[12] Lamboni, M., Makowski, D., Lehuger, S., Gabrielle, B. and Monod, H., 2009. Multivariate global sensitivity analysis for dynamic crop models. Field Crops Research, 113(3): 312-320.

[13] Lamboni, M., Monod, H. and Makowski, D., 2011. Multivariate sensitivity analysis to measure global contribution of input factors in dynamic models. Reliability Engineering and System Safety, 96(4): 450-459.

[14] Li, S., Zhang, L., Kang, S., Tong, L., Du, T., Hao, X. and Zhao, P., 2015. Comparison of several surface resistance models for estimating crop evapotranspiration over the entire growing season in arid regions. Agricultural and Forest Meteorology, 208: 1-15.

[15] Lu, Y., Chibarabada, T.P., McCabe, M.F., De Lannoy, G.J.M. and Sheffield, J., 2021. Global sensitivity analysis of crop yield and transpiration from the

FAO-AquaCrop model for dryland environments. Field Crops Research, 269: 108182.

[16] Ma, H., Malone, R. W., Jiang, T., Yao, N., Chen, S., Song, L., Feng, H., Yu, Q. and He, J., 2020. Estimating crop genetic parameters for DSSAT with modified PEST software. European Journal of Agronomy, 115: 126017.

[17] Morris, M.D., 1991. Factorial Sampling Plans for Preliminary Computational Experiments. Technometrics, 33(2): 161-174.

[18] Plummer, M., Best, N., Cowles, K. and Vines, K., 2006. CODA: Convergence diagnosis and output analysis for MCMC. R News, 6(1): 7-11.

[19] Raes, D., Steduto, P., Hsiao, T.C. and Fereres, E., 2009. AquaCrop-the FAO crop model to simulate yield response to water: II. Main algorithms and software description. Agronomy Journal, 101(3): 438-447.

[20] Raes, D., Steduto, P., Hsiao, T.C. and Fereres, E., 2018. Reference Manual, Chapter 3- Calculation Procedures, AquaCrop. Land and Water Division. FAO, Rome, Italy, p. 151 version 6.0-6.1.

[21] Ran, H., Kang, S., Hu, X., Li, F., Du, T., Tong, L., Li, S., Ding, R. and Zhang, X., 2018. Parameterization of the AquaCrop model for full and deficit irrigated maize for seed production in arid Northwest China. Agricultural Water Management, 203: 438-450.

[22] Ran, H., Kang, S., Hu, X., Li, F., Du, T., Tong, L., Li, S., Ding, R., Zhou, Z. and Parsons, D., 2019. Newly developed water productivity and harvest index models for maize in an arid region. Field Crops Research, 234: 73-86.

[23] Ran, H., Kang, S., Hu, X., Li, S., Wang, W. and Liu, F., 2020. Capability of a solar energy-driven crop model for simulating water consumption and yield of maize and its comparison with a water-driven crop model. Agricultural and Forest Meteorology, 287: 107955.

[24] Ran, H., Kang, S., Li, F., Tong, L., Ding, R., Du, T., Li, S. and Zhang, X., 2017. Performance of AquaCrop and SIMDualKc models in evapotranspiration partitioning on full and deficit irrigated maize for seed production under plastic film-mulch in an arid region of China. Agricultural Systems, 151: 20-32.

[25] Roux, S., Brun, F. and Wallach, D., 2014. Combining input uncertainty and residual error in crop model predictions: A case study on vineyards. European Journal of Agronomy, 52: 191-197.

[26] Silvestro, P. C., Pignatti, S., Yang, H., Yang, G., Pascucci, S., Castaldi, F. and Casa, R., 2017. Sensitivity analysis of the Aquacrop and SAFYE crop models for the assessment of water limited winter wheat yield in regional scale applications. Plos One, 12(11).

[27] Steduto, P., Hsiao, T.C., Raes, D. and Fereres, E., 2009. AquaCrop-The FAO crop model to simulate yield response to water: I. Concepts and underlying principles. Agronomy Journal, 101(3): 426-437.

[28] Tan, J., Cui, Y. and Luo, Y., 2017. Assessment of uncertainty and sensitivity analyses for ORYZA model under different ranges of parameter variation.

European Journal of Agronomy, 91: 54-62.

[29] Team, R.C., 2013. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

[30] Tremblay, M. and Wallach, D., 2004. Comparison of parameter estimation methods for crop models. Agronomie, 24(6-7): 351-365.

[31] Vanuytrecht, E., Raes, D. and Willems, P., 2014. Global sensitivity analysis of yield output from the water productivity model. Environmental Modelling and Software, 51: 323-332.

[32] Wallach, D., Keussayan, N., Brun, F., Lacroix, B. and Bergez, J.E., 2012. Assessing the Uncertainty when Using a Model to Compare Irrigation Strategies. Agronomy Journal, 104(5): 1274-1283.

[33] Wallach, D., Makowski, D., Jones, J.W. and Brun, F., 2019. Working with Dynamic Crop Models: Methods, Tools and Examples for Agriculture and Environment (Third Edition). Working with Dynamic Crop Models (Third Edition). Academic Press.

[34] Wallach, D. and Thorburn, P.J., 2017. Estimating uncertainty in crop model predictions: Current situation and future prospects. European Journal of Agronomy, 88: A1-A7.

Table 1

Model parameters of interest for maize in the AquaCrop model that are treated as random variables.

Description and unit	Nominal value	Prior distribution	Abbreviation
Base temperature, °C	8	6.5-8.5[a]	T_{base}
Upper temperature, °C	30	29-31[a]	T_{upper}
Minimum growing degrees required for full biomass production, °C/day	12	8.4-15.6	GDD_{min}
Canopy size of the average seedling at 90% emergence, cm^2	6.5	4.55-8.45	cc_0
Shape factor describing root zone expansion	1.3	1.1-1.5[a]	$root_s$
Upper threshold for soil water depletion for leaf growth	0.14	0.098-0.182	p_{expu}
Lower threshold for soil water depletion for leaf growth	0.72	0.504-0.936	p_{expl}
Shape factor for Water stress coefficient for canopy expansion	2.9	2.03-3.77	p_{exps}
Soil water depletion threshold for stomatal control	0.5	0.35-0.65	p_{sto}
Shape factor for water stress coefficient for stomatal control	6	4.2-7.8	p_{stos}
Soil water depletion threshold for canopy senescence	0.5	0.35-0.65	p_{sen}
Shape factor for water stress coefficient for canopy senescence	2.7	1.89-3.51	p_{sens}
Soil water depletion threshold for failure of pollination	0.75	0.525-0.975	p_{pol}
Minimum air temperature below which pollination starts to fail (cold stress), °C	10	7-13	$cold_s$
Maximum air temperature above which pollination starts to fail (heat stress), °C	40	30-45[a]	$heat_s$
Maximum canopy cover, %	90	70-100[a]	CC_x
Crop coefficient when canopy is complete but prior to senescence	1.20	1.0-1.4[a]	$Kc_{Tr,x}$
Decline of crop coefficient as a result of ageing, nitrogen deficiency, etc., %/day	0.3	0.21-0.39	Kc_{Trxd}
Effect of canopy cover on reducing soil evaporation in late season stage	50	35-65	Kc_{Trxle}
Water productivity normalized for ET_0 and CO_2, g m^{-2}	20.9	14.63-27.17	WP^*

Reference harvest index, %	33.1	23-43	HI_0
Canopy growth coefficient	0.0115	0.008055-0.014959	CGC
Canopy decline coefficient	0.0052	0.003616-0.006716	CDC
Soil field capacity, %	30	26-34[a]	fcsoil
Soil saturated water content, %	41	37-45[a]	satsoil
Soil water content at permanent wilting point, %	10	6-14[a]	wpsoil
Soil saturated hydraulic conductivity, mm d^{-1}	500	350-650	Ksatsoil

The nominal values are cited from Ran et al. (2018). The lower and upper boundaries of prior distribution for each parameter is obtained by ±30% nominal value. The prior distribution is assumed to be a uniform distribution.

[a]means the lower and upper boundaries are modified based on literature values and regulations in AquaCrop.

Table 2

Percentages of mean values of measured data fell in the calculated 25th to 75th percentile band or 0th to 100th percentile band in the four irrigation scenarios.

Scenarios	Response variables	Cases in 25th to 75th percentile range (%)	Cases in 0th to 100th percentile range (%)
S_{FI}: Full Irrigation	CC	15	62
	B_t	6	26
	SWC	6	19
	ET	4	15
	Y	80	100
S_{DI}: Deficit Irrigation	CC	16	58
	B_t	4	23
	SWC	6	24
	ET	14	42
	Y	33	89
S_{EDI}: Extreme Deficit Irrigation	CC	25	92
	B_t	0	60
	SWC	0	56
	ET	0	53
	Y	0	0
S_{AI}: All Irrigation	CC	9	36
	B_t	6	17
	SWC	4	19
	ET	4	16
	Y	39	91

CC, B_t, SWC, ET and Y represent time series canopy cover, aboveground biomass, soil water content, daily evapotranspiration and final yield, respectively.

Table 3

The variance of parameter estimations (var_{parm}), model residual errors (var_{model}) and predictions (var_{pred}) derived for each response variable in the four irrigation scenarios.

Scenarios	Response variables	var_{parm}	var_{model}	var_{pred}
S_{FI}: Full Irrigation	CC	426.8 (1.8%)[(a)]	22946.0 (96.7%)	23739.2
	B_t	6.5 (2.7%)	238.0 (99.7%)	238.8
	SWC	277.6 (0.6%)	46561.8 (99.8%)	46635.6
	ET	3.5 (0.8%)	451.2 (98.1%)	460.0
	Y	7.8 (1.7%)	442.5 (96.6%)	457.9
S_{DI}: Deficit Irrigation	CC	1256.9 (0.3%)	503109.6 (100.1%)	502635.1
	B_t	8.1 (1.0%)	826.5 (100.2%)	825.0
	SWC	870.9 (0.5%)	188634.1 (99.6%)	189316.4
	ET	1.5 (3.9%)	37.6 (95.7%)	39.2
	Y	14.4 (0.6%)	2558.4 (99.9%)	2560.3
S_{EDI}: Extreme Deficit Irrigation	CC	38.2 (24.1%)	125.2 (79.0%)	158.5
	B_t	0.6 (11.0%)	4.8 (89.2%)	5.4
	SWC	233.3 (10.2%)	2006.4 (88.0%)	2280.6
	ET	0.3 (8.5%)	3.0 (92.6%)	3.3
	Y[(b)]	0.5 (0.1%)	528.6 (99.5%)	531.1
S_{AI}: All Irrigation	CC	1755.2 (0.3%)	512605.3 (99.4%)	515660.2
	B_t	8.8 (0.8%)	1054.2 (99.3%)	1061.3
	SWC	676.0 (0.3%)	201476.9 (99.8%)	201945.2
	ET	3.8 (0.6%)	642.8 (99.2%)	648.1
	Y	19.1 (0.4%)	4819.2 (99.3%)	4854.5

CC, B_t, SWC, ET and Y represent time series canopy cover, aboveground biomass, soil water content, daily evapotranspiration and final yield, respectively.

[(a)]The numbers in brackets represent the ratios of the variances of parameter estimations and model residual errors to the total variances of predictions.

[(b)]N_v instead of N_v-1 is used to calculate the variance of Y in S_{EDI} to avoid infinity value since there is only one measured value. N_v is the number of measurements.

Table 4

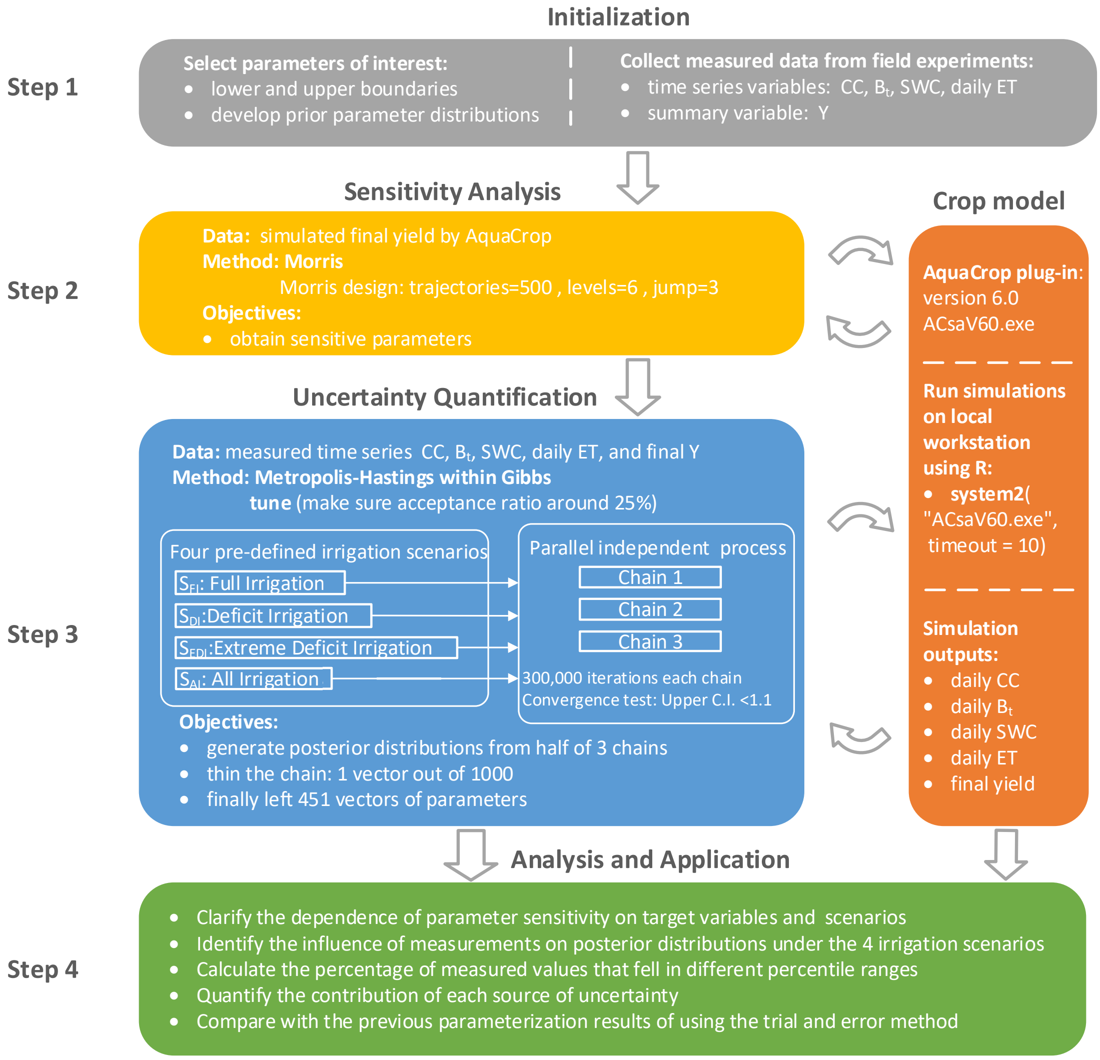
The goodness-of-fit between the medians of simulated canopy cover (CC), aboveground biomass (B_t), total soil water content in the 0–100 cm soil profile (SWC), daily evapotranspiration (ET), and yield (Y) using the 451 vectors of the posterior parameter distributions in the All Irrigation treatments scenario (S_{AI}) and the measured data, and its comparison with the result using the trial and error method.

Variable-Model	n[b]	b_0	R^2	RMSE	NRMSE	EF	d
CC-AquaCrop(median)	321	0.90	0.812	14.0	20.9	0.784	0.934
CC-AquaCrop[a]	321	0.96	0.818	12.9	19.3	0.811	0.947
B_t-AquaCrop(median)	219	0.97	0.931	1.742	16.8	0.931	0.982
B_t-AquaCrop[a]	219	1.05	0.929	1.972	19.1	0.903	0.977
SWC-AquaCrop(median)	178	0.98	0.744	24.1	11.1	**0.739**	0.923
SWC-AquaCrop[a]	178	1.09	0.736	33.1	15.2	**0.364**	0.854
ET-AquaCrop(median)	717	0.94	0.835	0.73	23.6	0.833	0.952
ET-AquaCrop[a]	717	0.96	0.825	0.75	24.4	0.822	0.952
Y-AquaCrop(median)	23	1.01	0.586	1.153	20.6	**0.415**	0.703
Y-AquaCrop[a]	23	1.12	0.496	1.466	26.2	**0.055**	0.681

[a]The simulated results of AquaCrop using the trial and error method based on the same measured data in this study are cited from Ran et al. (2017, 2018).

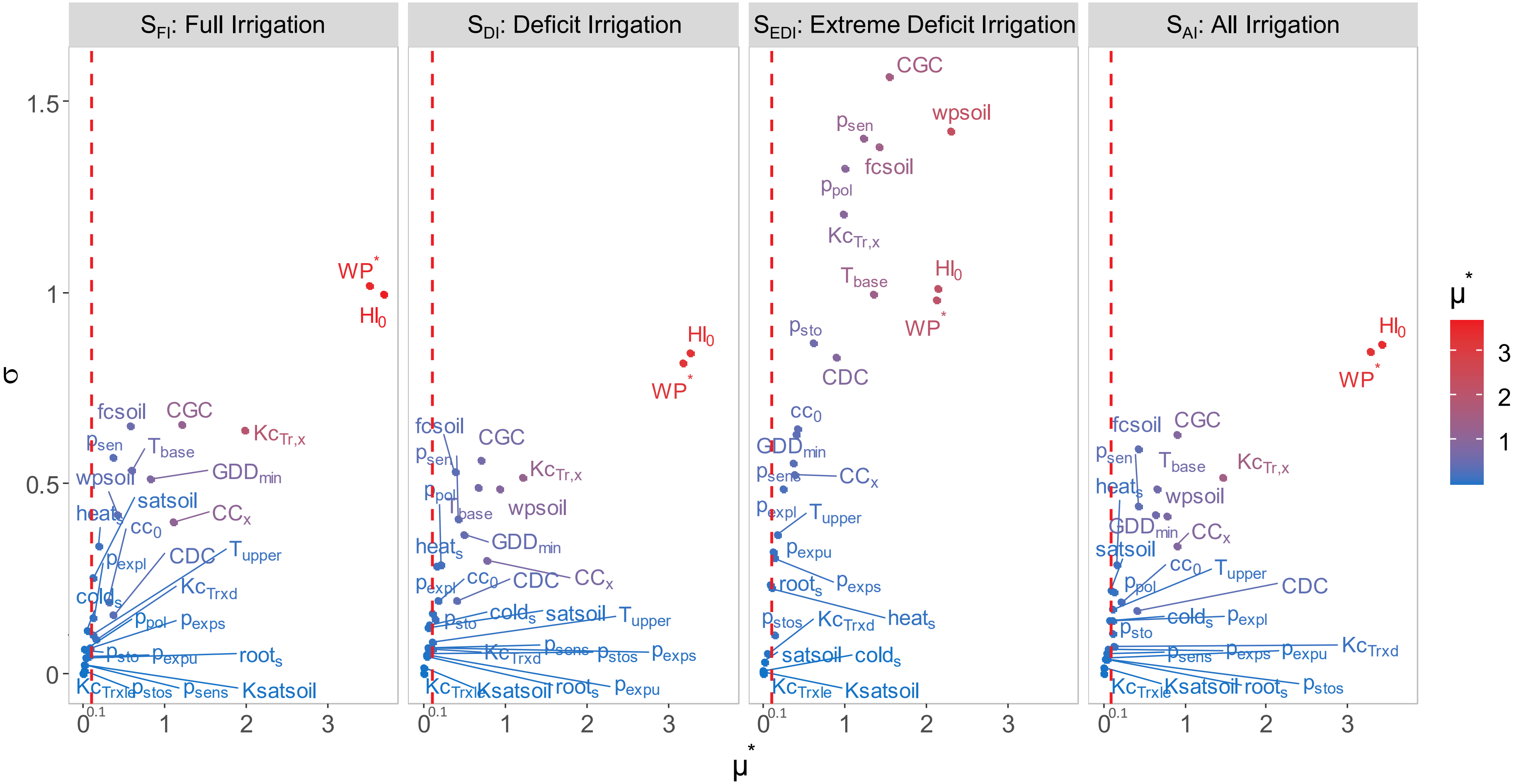
[b]n, b_0, R^2, RMSE, NRMSE, EF, and d represent the number of measured samples, regression coefficient through the origin, coefficient of determination, root mean square error, normalized root mean square error, Nash–Sutcliffe model efficiency coefficient, and Willmott's index of agreement, respectively. b_0, R^2, EF, and d are unitless. The unit of RMSE for CC, B_t, SWC, ET, and Y is %, t ha^{-1}, mm, mm d^{-1} and t ha^{-1}, respectively. The unit of NRMSE is %.

Figure 1



Framework for quantifying uncertainty of crop model parameters. CC, B_t, SWC, ET, and Y represent canopy cover, aboveground biomass, soil water content, daily evapotranspiration, and final yield, respectively. Upper C.I. represents the upper confidence limits of the potential scale reduction factor. The test of convergence is conducted by the *coda* package in R. *system2* is a function in R to invoke and run AquaCrop.

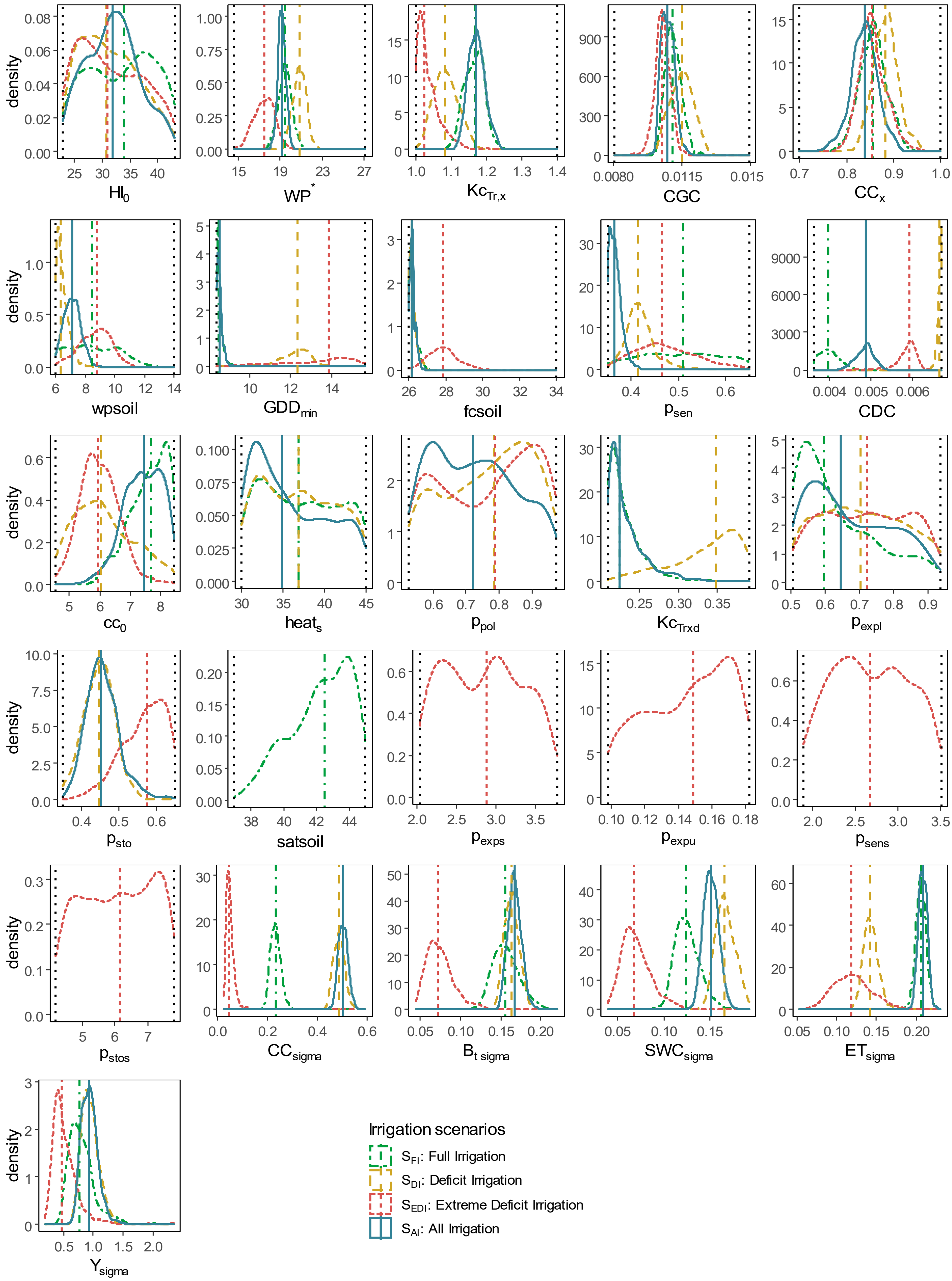
Figure 2



SFI: Full Irrigation
SDI: Deficit Irrigation
SEDI: Extreme Deficit Irrigation
SAI: All Irrigation
σ
μ*
WP*
HI0
CGC
fcsoil
wpsoil
psen
ppol
psto
KcTr,x
Tbase
CDC
CCx
cc0
GDDmin
Tupper
satsoil
heats
colds
pexpl
pexpu
pexps
psens
pstos
roots
KcTrxd
KcTrxle
Ksatsoil

Results of the Morris method obtained with AquaCrop for the four irrigation scenarios. μ^* and σ represent average Morris mean effects and the square root of variance. The method is implemented with a grid including 6 levels per factor, a jump equal to 3, and 500 trajectories. The meaning of parameter abbreviation is shown in Table 1.

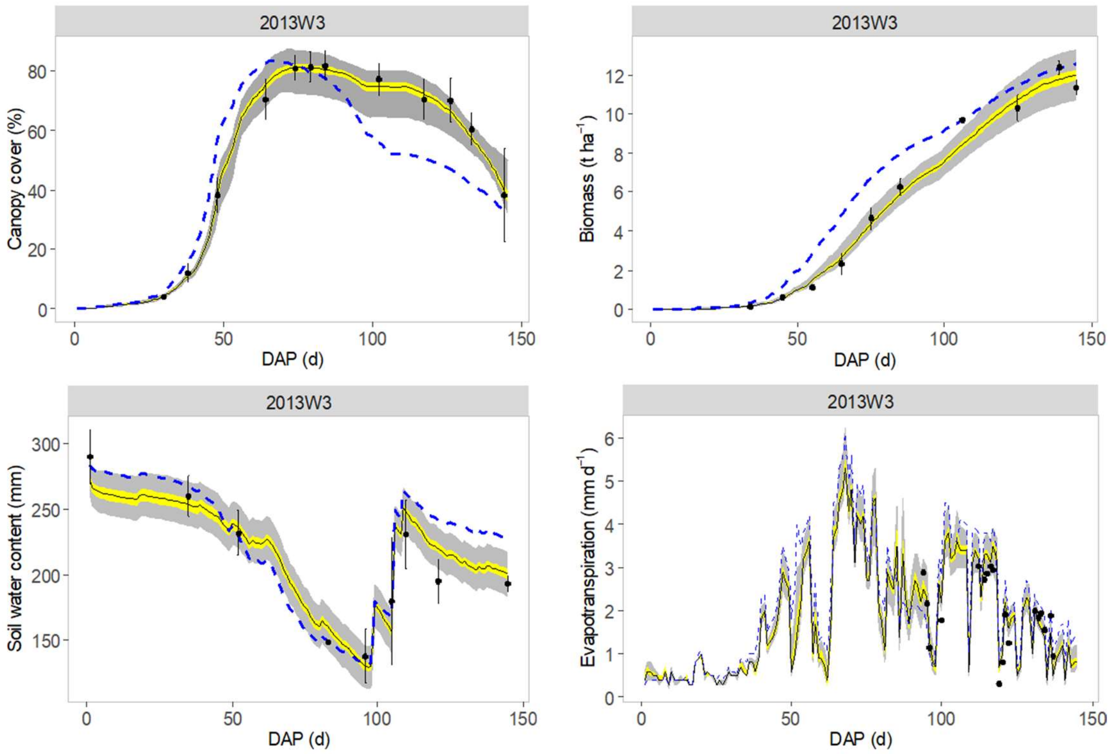
Figure 3



density
HI0
WP*
KcTr,x
CGC
CCx
wpsoil
GDDmin
fcsoil
psen
CDC
cc0
heats
ppol
KcTrxd
pexpl
psto
satsoil
pexps
pexpu
psens
pstos
CCsigma
Bt sigma
SWCsigma
ETsigma
Ysigma
Irrigation scenarios
SFI: Full Irrigation
SDI: Deficit Irrigation
SEDI: Extreme Deficit Irrigation
SAI: All Irrigation

Posterior distributions of the sensitive parameters and model residual standard error (sigma) for canopy cover (CC), biomass (B_t), soil water content (SWC), daily evapotranspiration (ET) and final grain yield (Y) in the four irrigation scenarios. The colored lines represent the corresponding medians. The black dotted lines represent the lower and upper boundaries of the prior distribution for each parameter. The sigma for each variable is generated from inverse gamma distribution after logarithmic transformation of the measured values.

Figure 4

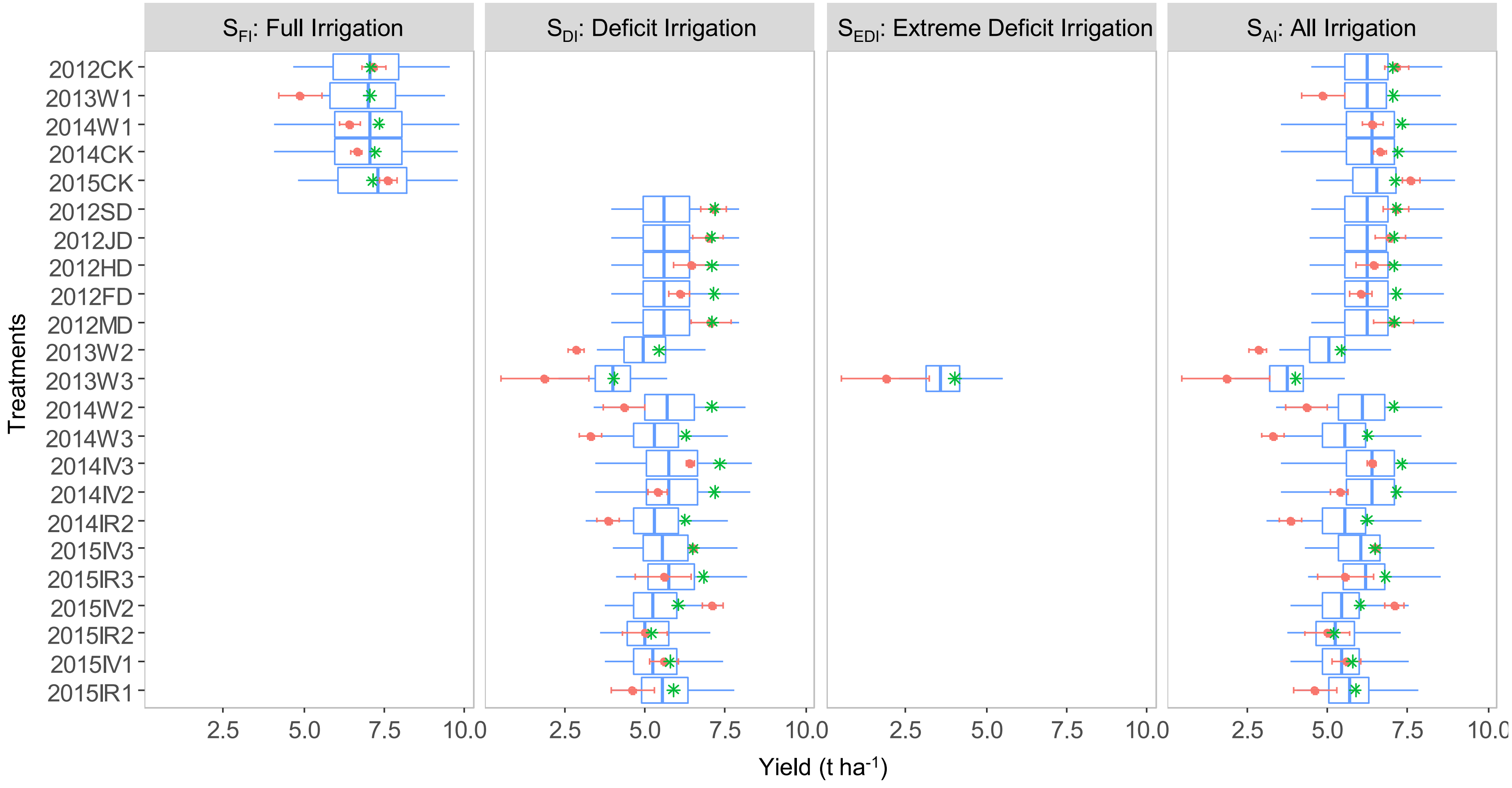


Simulated and measured canopy cover (CC), biomass (B_t), soil water content (SWC), and daily evapotranspiration (ET) for the treatment of 2013W3 in Extreme Deficit Irrigation scenario (S_{EDI}).

DAP is days after planting.

Black dots with error bars represent measured values with ±1 standard deviation. Solid black lines represent the medians of the simulations. Gray areas indicate the 0th to 100th percentile band and yellow areas the 25th to 75th percentile band of the values simulated by AquaCrop with 451 vectors of the posterior parameter distribution. The blue dash line represents the calibrated results of AquaCrop with the same measured data of this study using the trial and error method in Ran et al. (2017, 2018).

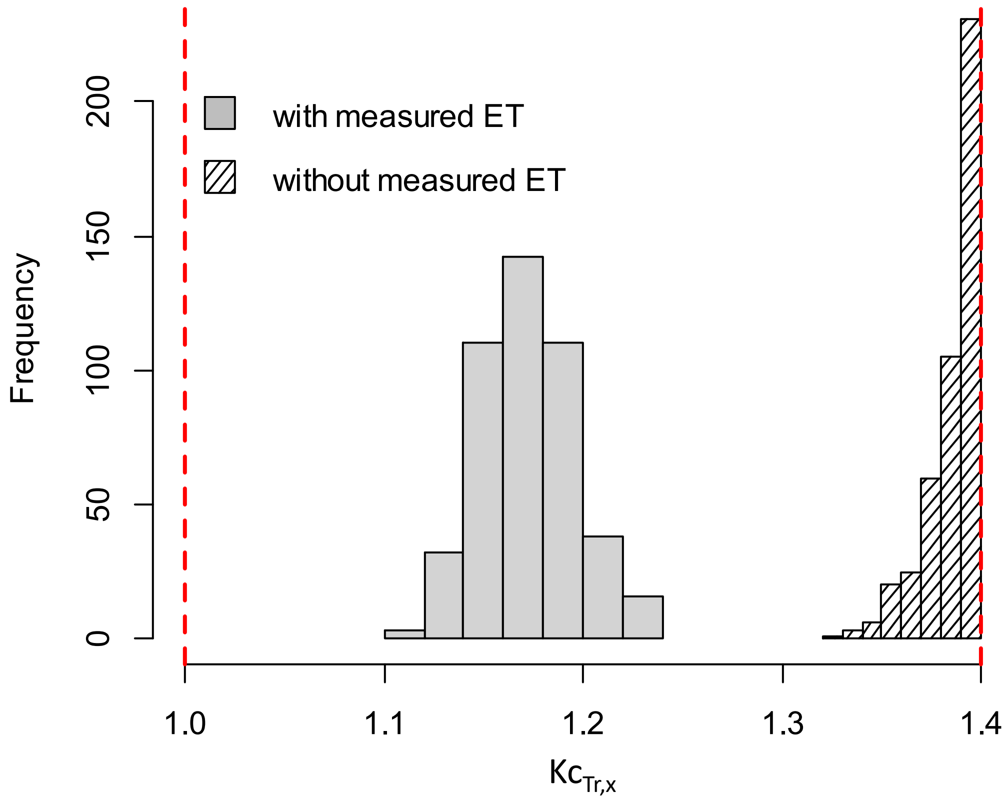
Figure 5



S_{FI}: Full Irrigation
S_{DI}: Deficit Irrigation
S_{EDI}: Extreme Deficit Irrigation
S_{AI}: All Irrigation
Treatments
2012CK
2013W1
2014W1
2014CK
2015CK
2012SD
2012JD
2012HD
2012FD
2012MD
2013W2
2013W3
2014W2
2014W3
2014IV3
2014IV2
2014IR2
2015IV3
2015IR3
2015IV2
2015IR2
2015IV1
2015IR1
2.5
5.0
7.5
10.0
Yield (t ha^{-1})

Measured and simulated final grain yield (Y) with the AquaCrop model in the four irrigation scenarios. Red points with error bars represent measured yield and its ±1 standard deviation. Blue boxplots represent the 0th, 25th, 50th, 75th, and 100th percentile range of the values simulated by AquaCrop with 451 vectors of the posterior parameter distributions. Green points represent the calibrated results of AquaCrop with the same measured data of this study using the trial and error method in Ran et al. (2018).

Figure 6



The posterior distribution for parameter of $Kc_{Tr,x}$ with or without involving measured daily ET during the MCMC iteration in the All Irrigation treatments scenario (S_{AI}). Red dash lines mean the lower and upper boundaries of the prior distribution for each parameter.