Agricultural Systems 103 (2010) 256–264

Contents lists available at ScienceDirect

Agricultural Systems

journal homepage: www.elsevier.com/locate/agsy

Influence of likelihood function choice for estimating crop model parameters using the generalized likelihood uncertainty estimation method

Jianqiang He, James W. Jones *, Wendy D. Graham, Michael D. Dukes

University of Florida, Agricultural and Biological Engineering, Museum Road, Gainesville, FL 32611, United States

ARTICLE INFO

Article history:
Received 6 January 2009
Received in revised form 14 December 2009
Accepted 27 January 2010
Available online 7 March 2010

Keywords:
Parameter estimation
GLUE
Likelihood function
CERES-Maize
DSSAT
Sweet corn

ABSTRACT

Proper estimation of model parameters is required for ensuring accurate model predictions and good model-based decisions. The generalized likelihood uncertainty estimation (GLUE) method is a Bayesian Monte Carlo parameter estimation technique that makes use of a likelihood function to measure the closeness-of-fit of modeled and observed data. Various likelihood functions and methods of combining likelihood values have been used in previous studies. This research was conducted to determine the effects of using previously reported likelihood functions in a GLUE procedure for estimating parameters in a widely-used crop simulation model. A factorial computer experiment was conducted with synthetic measurement data to compare four likelihood functions and three methods of combining likelihood values using the CERES-Maize model of the Decision Support System for Agrotechnology Transfer (DSSAT). The procedure used an arbitrarily-selected parameter set as the known "true parameter set" and the CERES-Maize model to generate true output values. Then synthetic observations of crop variables were randomly generated (four replicates) by using the simulated true output values (dry yield, anthesis date, maturity date, leaf nitrogen concentration, soil nitrate concentration, and soil moisture) and adding a random observation error based on the variances of corresponding field measurements. The environmental conditions were obtained from a sweet corn (*Zea* mays L.) experiment conducted in 2005 in northern Florida. Results showed that the method of combining likelihood values had a strong influence on parameter estimates. The combination method based on the product of the likelihoods associated with each set of observations reduced the uncertainties in posterior distributions of parameter estimates most significantly. It was also found that the likelihood function based on Gaussian probability density function was the best among those tested. This combination accurately estimated the true parameter values, suggesting that it can be used when estimating CERES-Maize model parameters for real experiments.

1. Introduction

Proper estimation of model parameters is required for ensuring accurate model predictions (Makowski et al., 2002). Modeling of complex environmental systems generally involves the indirect identification of model components or parameters by posing an inverse problem. Often, such inverse problems involve multiple parameters and observations that are only indirectly related to the parameters of interest, or which may be at different scales to the variables and parameters used in distributed predictions. There are many methods for estimating parameters using inverse modeling methods. Bayesian approaches can be used to estimate parameters using two types of information, a sample of data and prior information about parameter values. Results from a Bayesian method are probability distributions of parameter values and predicted outputs (Makowski et al., 2006a).

Bayesian methods are becoming increasingly popular for estimating parameters of complex mathematical models (Campbell et al., 1999). The Generalized Likelihood Uncertainty Analysis (GLUE) methodology (Beven and Binley, 1992), one such Bayesian method, allows information from different types of observations to be combined to estimate probability distributions of parameter values and model predictions (Lamb et al., 1998). Many parameter sets are generated from specified prior distributions of parameters and then used to simulate outputs by Monte Carlo simulation. The performance of each parameter set in predicting observed model states is evaluated via a likelihood measure that is used to weight the predictions from the different parameter sets. The GLUE method transforms the problem of searching for an optimum parameter set into a search for sets of parameter values that would give reliable simulations for a range of model inputs (Candela et al., 2005).

Parameters estimated using any inverse modeling approach are uncertain and subject to equifinality (e.g. Beven and Binley, 1992; Beven and Freer, 2001; Beven, 2006). Equifinality refers to the situation where the likelihood values are equal for two or more

* Corresponding author. Tel.: +1 352 392 1864; fax: +1 352 392 4092.
E-mail address: jimj@ufl.edu (J.W. Jones).

0308-521X/$ - see front matter
doi:10.1016/j.agsy.2010.01.006

parameter values, and one cannot select a best one from these two or more values. It can be argued on grounds of physical theory that there may be sufficient interactions among the components of a system that, unless the detailed characteristics of these components are specified independently, many representations may be equally acceptable. This is particularly true of those parameters to which the model is not sensitive in a particular environment. For this reason, a sensitivity analysis is needed to select parameters to which the model is sensitive for the range of experiments being used before attempting to estimate them using inverse modeling methods. One implication of equifinality is that the uncertainty associated with the use of models might be wider than is usually considered.

As with any model parameter estimation method, the GLUE method requires the definition of some measure of goodness-of-fit or likelihood. Beven and Binley (1992) pointed out that various likelihood measures might be appropriate in a given application. For example, Romanowicz et al. (1994, 1996) used a likelihood measure based on an autocorrelated Gaussian error model; Beven and Binley (1992) used a likelihood measure based on inverse error variance with a shaping factor *N*; Freer et al. (1996) suggested a likelihood measure based on Nash and Sutcliffe efficiency criterion; and Keesman and Van Straten (1989, 1990) used a likelihood measure based on scaled maximum absolute residuals. However, Stedinger et al. (2008) criticized the use of an arbitrary likelihood function. The choice of a likelihood function is critical and needs to be a reasonable description of the distribution of model errors for the statistical inference and resulting uncertainty and prediction intervals to be valid. If an arbitrary likelihood measure is adopted that does not reasonably reflect the distribution of model errors, then GLUE may generate arbitrary results without statistical validity that should not be used in scientific work.

With multiple observations and multiple types of observations, likelihood values for each observation must be combined into an overall value for each candidate parameter set (Beven and Binley, 1992). Available methods of combining likelihood values include multiplication (e.g. Beven and Binley, 1992), weighted addition (Zak et al., 1997), pseudomaximum likelihood measure (Van Straten, 1983), fuzzy union, fuzzy interaction, and weighted fuzzy combination (Aronica et al., 1998). Beven and Freer (2001) and Beven and Binley (1992) suggested that when a likelihood approach is being considered, the choice of method of combining likelihood values is subjective. However, when the GLUE method is used with a model for the first time, it is important to make sure that the choice of likelihood measure and combination method can produce reliable model parameters.

The CERES-Maize model (Jones and Kiniry, 1986; Ritchie, 1998; Hoogenboom et al., 2003) is a maize (*Zea mays* L.) crop growth model in the cropping system model (CSM) that is in the Decision Support System for Agrotechnology Transfer (DSSAT) (Jones et al., 2003; Tsuji et al., 1998). The DSSAT-CSM incorporates all crops as modules using a single soil model. Hereafter, CERES-Maize will be used to refer to the model used in this study. This model has many parameters that characterize crop and soil processes, a number of which usually need to be estimated using field experiments. Over the years, a number of methods have been used to estimate parameters for the DSSAT models, including the simplex method (Grimm et al., 1993), simulated annealing (Mavromatis et al., 2002), sequential search software (Hunt et al., 1993), and even visual methods. Each of these methods has its own advantages and limitations. Our main reasons for selecting the GLUE method for this study were that it can help us understand uncertainties in the parameters and how those uncertainties affect predictions and it is relatively simple and straightforward to implement.

In using this method for the first time with the CERES-Maize model, the question arises as to how much the different likelihood measures and combination methods influence the results of parameter estimation. The objective of the study was to answer this question for this widely-used crop model. We evaluated the influence of four different likelihood functions and three combination methods in GLUE on the parameter estimates for the CERES-Maize model.

2. Materials and methods

2.1. CERES-Maize model

Crop growth and development are simulated by the CERES-Maize model in DSSAT V4.0 (Hoogenboom et al., 2003) with a daily time step from planting to maturity using physiological process relationships that describe the responses of maize to soil and environmental conditions. Potential growth is dependent on photosynthetically active radiation and its interception, whereas actual biomass production on any day is constrained by suboptimal temperatures, soil water deficits, and nitrogen deficiencies (Ritchie and Godwin, 1989; Ritchie, 1998).

There are four types of input data to the model: weather, plant, soil, and management. The weather input data are daily sum of global radiation (MJ m^{-2}), daily minimum and maximum air temperatures (°C), and daily sum of precipitation (mm). Plant parameters and physiological characteristics are given in the form of genetic coefficients, which describe physiological processes such as development, photosynthesis, and growth for individual crop varieties in response to soil, weather, and management during a season. Soil inputs describe the physical, chemical, and morphological properties of the soil surface and each soil layer within the root zone. The management information includes planting density, row spacing, planting depth, irrigation, application of fertilizer, etc. (Ritchie, 1998).

2.2. Soil parameters and genetic coefficients

There are usually many parameters and inputs in complex crop simulation models. Each of these parameters and inputs is subject to errors. Ideally, one would directly measure all inputs and parameters, but this is not possible in many cases (Bechini et al., 2006). Furthermore, uncertainties in some parameters are likely to cause more variations in simulated results than others. Thus, a common strategy is to select a subset of parameters to estimate using sensitivity analysis, and fixing the others to their nominal values (Makowski et al., 2006a,b; Monod et al., 2006; Wallach et al., 2001). Through a global sensitivity analysis with one-at-a-time (OAT) method (Morris, 1991), He (2008) selected the most sensitive genetic and soil parameters (Table 1) relative to their influence on CERES-Maize model predictions of dry matter yield and cumulative nitrogen leaching for the growing conditions in this study. Other parameters or inputs may be important for different environmental and management conditions, and the sensitivity analysis would need to be repeated for other experiments.

The selected soil parameters (SLLL, SDUL, and SSAT) define soil water holding capacity and influence the amount of available water in the soil profile on a day to day basis. Parameters SLRO and SLDR influence the amount of soil water runoff and water drained from the soil profile. Parameter SLPF represents the effect of other limiting soil factors that reduce crop growth. Genetic coefficients P1 and P5 control the phenological development of the crop through their effects on anthesis and maturity dates. Coefficient PHINT influences both phenological development and yield. See Jones and Kiniry (1986) and Ritchie (1998) for more details regarding these parameters in the CERES-Maize model.

Table 1
Soil parameters and genetic coefficients for the CERES-Maize model in DSSAT that were estimated in this study.

Parameter	Definition	Unit
SLLL	Lower limit of soil water available to plants	m^3/m^3
SDUL	Drained upper soil water limit	m^3/m^3
SSAT	Saturated soil water content	m^3/m^3
SLRO	Soil water runoff curve number	–
SLDR	Soil water drainage rate	–
SLPF	Growth reduction/fertility factor	–
P1	Thermal time from seedling emergence to the end of the juvenile phase (expressed in degree days above a base temperature of 8 °C) during which the plant is not responsive to changes in photoperiod	°C d
P5	Thermal time from silking to physiological maturity (expressed in degree days above a base temperature of 8 °C)	°C d
PHINT	Phylochron interval; the interval in thermal time (degree days) between successive leaf tip appearances	°C d

2.3. Field environmental conditions

The synthetic parameter estimation experiment was based on a sweet corn field experiment conducted at the Plant Science Research and Education Unit, the University of Florida in the spring of 2005. The research unit is located near Citra (29.4094°N, 82.1777°W, 21 m above sea level), Marion County, Florida, USA. The variety of sweet corn planted was 'Saturn SH2'. Although there were two treatments in this field in 2005, the high-nitrogen treatment (422 kg N ha^{-1}) was used in this study. The experimental plot was about 1.82 ha. Irrigation was scheduled based on daily evapotranspiration and water balance in the soil profile. The weather data were obtained from the Florida Automated Weather Network (FAWN) weather station located at Citra. See He (2008) for more details regarding the field experiment.

Six types of field measurement data were collected from the high-nitrogen treatment, including dry yield (kg ha^{-1}), anthesis date (ADAT, days after planting), maturity data (MDAT, days after planting), leaf nitrogen concentration (TKN, %), soil nitrate concentration (mg kg^{-1}) and soil moisture (cm^3 cm^{-3}) in four soil layers (0–15 cm, 15–30 cm, 30–60 cm, and 60–90 cm). These data were used to estimate measurement standard deviations (Table 2) for generating the synthetic observations for the GLUE procedure. Other field data, such as the dates and methods of planting, tillage, irrigation, fertigation, pesticide, and herbicide application, and harvest were also collected for model inputs.

2.4. Synthetic data generation

Although real data from our experiments could have been used to compare methods (e.g., see He et al., 2009; Casanova et al., 2006, 2007), we chose to use synthetic data generated by the CERES-Maize model. This was done so that we could compare estimated parameters obtained using the different GLUE methods with "true" values of the parameters. Thus in this study, a parameter set, which was arbitrarily selected from the parameters estimated by He (2008), was chosen as the known "true parameter set" (see the right-most column in Table 3). The CERES-Maize model was run with this parameter set for the environmental conditions and management in the 2005 field experiment. Model outputs (dry matter yield, anthesis date, maturity date, leaf nitrogen concentration, soil nitrate concentration, and soil moisture) were tabulated from the model-simulations on dates when real measurements were taken in the field study (He, 2008). Four synthetic observations of each variable were then generated by adding a Gaussian random error with zero mean and a standard deviation corresponding to that of the actual field measurements (Table 2) using the following equation:

$$X = p \times \sigma + \mu \tag{1}$$

where p is the standard normal error, σ the standard deviation of the appropriate field measurements, and μ is the "true" value of the variable obtained from model simulation outputs generated from the "true parameter set".

We assumed that errors in simulated output variables of the maize model (e.g., sweet corn yield, anthesis date, maturity date, etc.) follow normal distributions, although there were not enough measurements in this field experiment to test for normality. In He et al. (2009), residual errors of these output variables were analyzed after estimating parameters. These residual errors all followed Gaussian distributions. Table 2 shows the synthetic observations generated for yield, anthesis date, and maturity date. There were also four replicates of each of the temporally-variable synthetic variables (leaf TKN concentrations, soil nitrate, and moisture concentrations), each at four depths and on five dates, and therefore they are not shown in this table for the sake of brevity.

Table 2
Generated synthetic measurement data.

2005	Synthetic observation		Generated replicates			
	Mean[a]	STDEV[b]	1	2	3	4
Yield (kg ha^{-1})	3451	269	3769	3447	3595	3259
ADAT (days)	59	3	57	60	59	62
MDAT (days)	94	4	97	91	91	93

[a] Mean of the synthetic observation was the simulation results with the "true parameter set".
[b] STDEV was the standard deviation of the synthetic observation, which was the real standard deviation of field experiment observation.

2.5. Generalized likelihood uncertainty estimation method

2.5.1. GLUE implementation

The GLUE procedure as described by Beven and Binley (1992) was used in this study and is summarized in the following steps:

(1) *Develop prior parameter distributions*. The soil parameters and genetic coefficients contained in the database of the DSSAT model (Hoogenboom et al., 2003) were analyzed. The means and variances of parameters and covariances among the parameters were calculated, and normality tests were conducted to see whether the parameters followed normal distributions (He, 2008). Based on these results a multivariate normal distribution was used for all parameter prior distributions except for SLPF for which a uniform distribution was used.

(2) *Generate random parameter sets from the prior parameter distributions*. A MATLAB (2004) program, 'mvnrnd.m', was used to generate N multivariate normal realizations of parameter sets with each set containing all parameters shown in Table 1 except for SLPF, which was generated with a uniform distribution. From the point of view of Monte Carlo sampling in the GLUE method, more parameter sets lead to more stable results. The number of parameter set realizations to gener-

Table 3
Means and standard deviations (STDEV) of posterior distributions derived from likelihood functions and combination methods in the first round of GLUE.[a]

	Prior distribution		Under $C1$		Under $C2$		Under $C3$		True[b]
	Mean	STDEV	Mean	STDEV	Mean	STDEV	Mean	STDEV	
L1 likelihood function									
P1	225.10	67.83	119.95	54.01	116.12	6.27	120.99	55.00	95.12
P5	763.59	98.80	663.78	86.61	634.83	21.92	666.22	85.22	572.04
PHINT	41.17	4.01	40.06	3.65	42.73	0.88	40.17	3.64	39.57
SLDR	0.463	0.192	0.498	0.162	0.661	0.003	0.500	0.159	0.739
SLRO	73.00	11.56	76.36	10.57	84.21	0.96	76.60	10.51	89.45
SDUL	0.263	0.100	0.195	0.113	0.106	0.001	0.191	0.113	0.104
SLLL	0.138	0.084	0.107	0.079	0.070	0.019	0.106	0.078	0.060
SSAT	0.388	0.094	0.327	0.104	0.304	0.006	0.324	0.104	0.319
SLPF	0.962	0.114	0.847	0.077	0.915	0.023	0.849	0.076	0.931
L2 likelihood function									
P1	225.10	67.83	124.43	57.53	117.16	4.56	126.54	59.72	95.12
P5	763.59	98.80	667.57	85.14	638.48	15.92	671.35	83.40	572.04
PHINT	41.17	4.01	40.21	4.16	42.88	0.64	40.44	4.15	39.57
SLDR	0.463	0.192	0.510	0.175	0.662	0.002	0.515	0.172	0.739
SLRO	73.00	11.56	76.38	10.73	84.37	0.70	76.83	10.57	89.45
SDUL	0.263	0.100	0.198	0.111	0.107	0.001	0.191	0.111	0.104
SLLL	0.138	0.084	0.105	0.080	0.073	0.014	0.101	0.079	0.060
SSAT	0.388	0.094	0.331	0.103	0.305	0.004	0.325	0.103	0.319
SLPF	0.962	0.114	0.845	0.078	0.918	0.017	0.846	0.077	0.931
L3 likelihood function									
P1	225.10	67.83	120.87	54.96	114.03	8.37	122.18	56.30	95.12
P5	763.59	98.80	664.39	86.10	627.52	29.22	667.11	84.52	572.04
PHINT	41.17	4.01	40.08	3.84	42.43	1.18	40.23	3.83	39.57
SLDR	0.463	0.192	0.501	0.166	0.660	0.004	0.504	0.163	0.739
SLRO	73.00	11.56	76.38	10.60	83.89	1.28	76.69	10.51	89.45
SDUL	0.263	0.100	0.197	0.112	0.106	0.001	0.192	0.112	0.104
SLLL	0.138	0.084	0.107	0.079	0.063	0.026	0.104	0.079	0.060
SSAT	0.388	0.094	0.328	0.104	0.302	0.008	0.324	0.103	0.319
SLPF	0.962	0.114	0.847	0.077	0.907	0.030	0.849	0.077	0.931
L4 likelihood function									
P1	225.10	67.83	122.08	56.99	115.91	6.54	124.20	59.20	95.12
P5	763.59	98.80	654.41	86.14	634.10	22.86	658.20	84.42	572.04
PHINT	41.17	4.01	39.37	4.21	42.70	0.92	39.60	4.21	39.57
SLDR	0.463	0.192	0.499	0.175	0.661	0.003	0.504	0.172	0.739
SLRO	73.00	11.56	74.72	10.79	84.18	1.00	75.17	10.64	89.45
SDUL	0.263	0.100	0.196	0.110	0.106	0.001	0.189	0.110	0.104
SLLL	0.138	0.084	0.104	0.079	0.069	0.020	0.100	0.079	0.060
SSAT	0.388	0.094	0.326	0.102	0.304	0.006	0.320	0.102	0.319
SLPF	0.962	0.114	0.827	0.079	0.914	0.024	0.829	0.079	0.931

[a] "$L1$ likelihood function" means deriving posterior distribution using likelihood function $L1$, "Under $C1$" uses the method of likelihood value combination $C1$, etc. (Eqs. (5)–(8) and (11)–(13)).
[b] "True" means the initially selected "true parameter set".

ate depends on the number required for the estimated parameters and their variances to stabilize. Since the prior distributions obtained from the DSSAT database in this study were very broad, only a limited number of parameter sets had significant likelihood values that could be used to derive posterior distributions, even though a large number of parameter sets were generated (30,000). The overwhelming majority of the generated parameter sets were eliminated since their likelihood values were near zero. One may need to run the GLUE procedure two or more times in order to ensure that a sufficient number of realizations are used to obtain smooth posterior distributions for the selected parameters. In this paper, two runs of GLUE were conducted in sequence.

(3) *Run the model with the random parameter sets*. The model was run for each parameter set using MATLAB programs developed in this study. The standard CERES-Maize soil input file 'soil.sol' and genetic input file 'MZCER040.cul', were changed to simulate each random parameter set in sequence. Model outputs (dry matter yield, anthesis date, maturity date, leaf nitrogen concentration, soil nitrate concentration, and soil moisture) for each parameter set were tabulated for use in the GLUE likelihood calculations.

(4) *Calculate the likelihood values*. The generated observations (O, four replicates each for each variable) were used along with the corresponding simulated outputs to compute the likelihood value, $L(\theta_i|O)$, for each of the N generated parameter vectors θ_i. Then, the probability p_i of each parameter set was computed with the following equation:

$$p(\theta_i) = \frac{L(\theta_i|O)}{\sum_{j=1}^{N} L(\theta_i|O)} \tag{2}$$

where $p(\theta_i)$ is probability or likelihood weight of the ith parameter set θ_i, $L(\theta_i|O)$ is the likelihood value of parameter set θ_i, given observations O.

(5) *Construct posterior distribution and statistics*. The pairs of parameter sets and probabilities, $(\theta_i, p_i), i = 1, \ldots, N$, were used to construct empirical posterior distributions and to compute the means and variances of the selected parameters using the following equations:

$$\underset{post}{\widehat{\mu}}(\theta) = \sum_{i=1}^{N} p(\theta_i) \cdot \theta_i \tag{3}$$

$$\underset{post}{\widehat{\sigma}}^2(\theta) = \sum_{i=1}^{N} p(\theta_i) \cdot (\theta_i - \underset{post}{\widehat{\mu}})^2 \tag{4}$$

where $\underset{post}{\widehat{\mu}}(\theta)$ and $\underset{post}{\widehat{\sigma}}^2(\theta)$ are the estimated mean, variance of the posterior distribution of parameters θ; $p(\theta_i)$ is the probability of the ith parameter set θ_i calculated by Eq. (2); and N is the number of random parameter sets.

2.5.2. Likelihood functions

Four likelihood functions were chosen and investigated using the same model outputs. The four likelihood functions identified as $L1$, $L2$, $L3$, and $L4$, are as follows:

$$L1: \quad L[\theta_i|O] = \prod_{j=1}^{M} \frac{1}{\sqrt{2\pi\sigma_o^2}} \exp\left(-\frac{(O_j - P(\theta_i))^2}{2\sigma_o^2}\right), \quad (i = 1, 2, 3, \ldots N) \tag{5}$$

$$L2: \quad L[\theta_i|O] = \frac{1}{\sqrt{2\pi\sigma_o^2}} \cdot \exp\left(-\frac{(\overline{O} - P(\theta_i))^2}{2\sigma_o^2}\right), \quad (i = 1, 2, 3, \ldots N) \tag{6}$$

$$L3: \quad L[\theta_i|O] = \exp\left(-\frac{MSE_i}{\min(MSE)}\right) \quad (i = 1, 2, 3, \ldots N) \tag{7}$$

$$L4: \quad L(\theta_i|O) = \exp\left(-\frac{MSE_i}{2\sigma_o^2}\right), \quad (i = 1, 2, 3, \ldots N) \tag{8}$$

where θ_i is the ith parameter set; $P_j(\theta_i)$ the jth type of model output under parameter set θ_i; O the observation; O_j the jth observation of O; σ_o^2 the variance of model errors, assumed to be the variances of observations for this study; $\overline{O}$ the mean value of the observations; MSE_i the mean square model prediction error for the ith parameter set; $\min(MSE)$ the minimum value of MSE_i; N the number of parameter sets; and M is the number of observation replicates. The mean square model prediction error and mean value of the observation replicates were calculated with following equations:

$$MSE_i = \frac{1}{M} \sum_{j=1}^{M} (P_j(\theta_i) - O_j)^2 \tag{9}$$

$$\overline{O} = \frac{1}{M} \sum_{j=1}^{M} O_j \tag{10}$$

The first likelihood measure $L1$ is the maximum likelihood function used by Makowski et al. (2002), while $L2$ is a variant of $L1$ that uses the mean value of the observations instead of calculating the product of likelihood values of each replicate of the observations. Likelihood measure $L3$ is based on minimum mean square error and was used by Wang et al. (2005). $L3$ is a special case of $L4$ in which the minimum MSE produced from all parameter sets is used as an estimation of model error variance (Wang et al., 2005).

2.5.3. Methods of combining likelihood values

Another variation of the GLUE procedure that can influence results is the method of combining likelihood values from different types of observations. In this study, observations were made for dry grain yield, anthesis date, maturity date, leaf TKN concentration, soil nitrate concentration, soil volumetric moisture, dry grain yield, anthesis date (ADAT), and maturity date (MDAT). Three methods to combine likelihood values of different types of observations, identified as $C1$, $C2$, and $C3$ respectively, were investigated in this study:

Table 4
Mean values and standard deviations (STDEV) of posterior distributions derived from different likelihood functions and likelihood combinations in the second round of GLUE.

Parameter	*L1C2*				*L2C2*				*L3C2*				*L4C2*				True[c]
	Max-L[a]	Mean	STDEV	ARE[b] (%)	Max-L	Mean	STDEV	ARE (%)	Max-L	Mean	STDEV	ARE (%)	Max-L	Mean	STDEV	ARE (%)	
P1	100.71	100.723	0.266	5.9	111.313	111.124	0.985	16.8	102.739	102.845	0.531	8.1	105.173	105.006	1.581	10.4	95.12
P5	603.692	603.714	0.466	5.5	641.578	641.497	0.829	12.1	652.502	650.713	8.040	13.8	624.394	624.525	0.505	9.2	572.04
PHINT	42.276	42.277	0.015	6.8	42.375	42.387	0.099	7.1	41.388	41.360	0.289	4.5	42.125	42.115	0.106	6.4	39.57
SLDR	0.666	0.666	0.000	9.9	0.657	0.657	0.000	11.1	0.662	0.662	0.002	10.4	0.658	0.658	0.001	10.9	0.739
SLRO	86.229	86.228	0.021	3.6	86.797	86.733	0.346	3.0	86.892	86.880	0.171	2.9	87.337	87.195	0.488	2.5	89.45
SDUL	0.105	0.105	0.000	1.5	0.106	0.106	0.000	1.8	0.106	0.106	0.000	2.6	0.104	0.104	0.000	0.6	0.104
SLLL	0.065	0.065	0.000	7.5	0.063	0.063	0.000	3.6	0.062	0.062	0.000	2.4	0.059	0.060	0.001	1.4	0.060
SSAT	0.306	0.306	0.000	4.1	0.299	0.300	0.002	6.0	0.295	0.295	0.002	7.5	0.307	0.306	0.001	4.0	0.319
SLPF	0.916	0.916	0.000	1.6	0.883	0.883	0.004	5.1	0.875	0.875	0.004	6.0	0.907	0.911	0.014	2.2	0.931
Mean	–	–	–	5.2	–	–	–	7.4	–	–	–	6.5	–	–	–	5.3	–

[a] "Max-L" means parameter set that had the maximum likelihood values.
[b] "ARE" means absolute relative error between mean values of posterior distributions and the "true parameter set" value.
[c] "True" means the initially selected "true parameter set".

Table 5
Mean values and standard deviations (STDEV) of model outputs in the second round of GLUE.

	Outputs Unit	Yield kg ha^{-1}	Anthesis date Days	Maturity date Days	Average
*L*1*C*2	Mean	3325	60	97	–
	STDEV	2.45	0.00	0.00	–
	ARE (%)[a]	3.66	1.69	3.19	2.85
*L*2*C*2	Mean	3678	62	100	–
	STDEV	26.03	0.22	0.22	–
	ARE (%)	6.58	5.01	6.33	5.97
*L*3*C*2	Mean	3767	60	99	–
	STDEV	57.97	0.12	0.51	–
	ARE (%)	9.17	1.67	5.20	5.34
*L*4*C*2	Mean	3802	60	98	–
	STDEV	63.12	0.23	0.11	–
	ARE (%)	10.16	1.74	4.27	5.39
True	Mean	3451	59	94	–
	STDEV	269	3	4	–

[a] "ARE" means absolute relative error between mean values of model outputs and the "true measurement" resulted from the initial "true parameter set".

$$C1: \quad L_{combined}[\theta_i] = \frac{\sum_{k=1}^{K} L_k[\theta_i|O_k]}{K} \tag{11}$$

$$C2: \quad L_{combined}[\theta_i] = \prod_{k=1}^{K} L_k[\theta_i|O_k] \tag{12}$$

$$C3: \quad L_{combined}[\theta_i] = \left[\sum_{k=1}^{K} \frac{1}{K} \cdot L_k[\theta_i|O_k]^2\right]^{1/2} \tag{13}$$

where $L_{combined}[\theta_i]$ is the combined likelihood value of *i*th parameter set θ_i; $L_k[\theta_i|O_k]$ indicates the *k*th type likelihood value conditioned with the *i*th parameter set θ_i and the *k*th observation type O_k; and K is the number of observation types.

Method *C*1 (equation 11) is a special case of the weighted addition method for likelihood value combination (Zak et al., 1997), where the weighting coefficients of all terms are equally set as $1/K$. Method *C*2 (equation 12) is the method of Bayes' multiplication (e.g. Beven and Binley, 1992; Romanowicz et al., 1994, 1996). Method *C*3 (equation 13) is a special case of the aggregated function suggested by Wang et al. (2005), where the weighting coefficients are all set to $1/K$.

2.5.4. Computer experiment design

A 4×3 factorial (four likelihood functions and three methods of likelihood combination) computer experiment was designed resulting in 12 methods of likelihood value calculations. The CERES-Maize model was run (N = 30,000 times) with the random parameter sets generated from the prior parameter distributions. Subsequently, posterior distributions were derived for each parameter using each of the 12 methods. Then, the mean values and standard deviations of these posterior distributions were compared with the prior distributions. The likelihood combinations that failed to significantly reduce uncertainties in input parameters were eliminated and the remaining strategies were used in a second round of GLUE simulations. The probability of generating the correct parameter sets was very low after the first round of GLUE and the posterior distributions were not smooth.

In the second round of GLUE procedure, 10,000 new parameter sets were generated based on the posterior distributions from the first round. This procedure ensured that a large number parameter sets contributed to the posterior distribution. A multivariate normal distribution was again assumed for the second round based on the pattern of posterior distribution of first round of GLUE. Then the mean values, maximum likelihood parameters, and standard deviations of the new posterior distributions were compared with the "true parameter set" values. The best likelihood function and method of combining likelihood values were selected as the combination that had the closest mean values to the "true parameter set" and lowest variances or uncertainties in the posterior distributions of model input parameters and predicted yield and phenology dates.

3. Results and discussion

3.1. Comparison of likelihood combination methods

The mean values and standard deviations of the 12 posterior distributions derived from four different likelihood functions and three methods of combining likelihood values are summarized in Table 3. The method of combining likelihood values had a very strong influence on the corresponding posterior distributions,

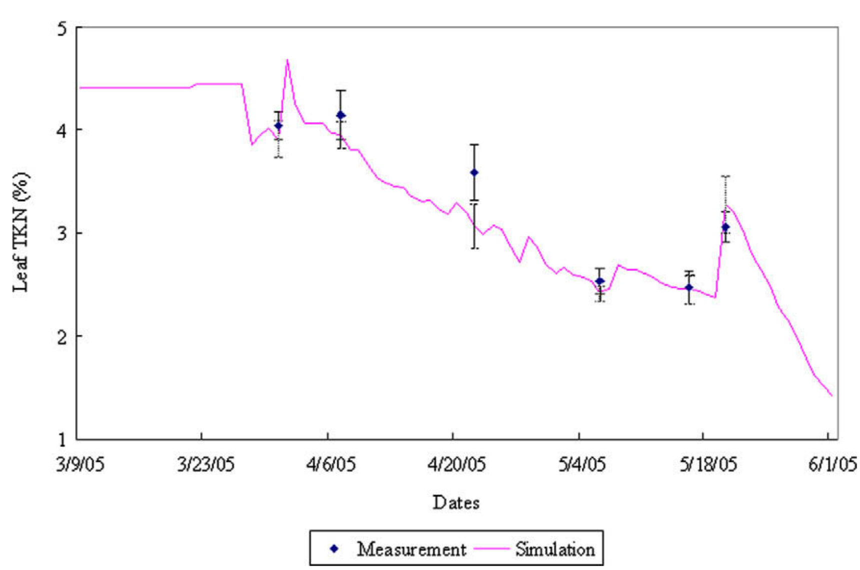

Fig. 1. Measured and simulated leaf TKN concentrations with the parameter set that had the highest likelihood value under *L*1*C*2 in the second round of GLUE during the sweet corn season. The solid and dashed error bars show the standard deviations of synthetic measurements and simulations, respectively.

particularly the standard deviations of parameter estimates. For example, under combination methods *C*1 and *C*3, the standard deviations of most input parameters did not decrease much relative to their prior values, and even increased in some cases. When using the *L*1 likelihood function, the standard deviation of SDUL increased to 0.113 under *C*1 and *C*3; in the prior distribution, it was 0.100. A similar trend occurred for the soil parameter SSAT. The same results were observed for likelihood functions *L*2, *L*3, and *L*4 when combinations *C*1 and *C*3 were used. These results show that the combination methods *C*1 and *C*3 failed to efficiently reduce the parameter uncertainties after 30,000 runs.

The combination methods *C*1 and *C*3 failed to eliminate parameter sets that simultaneously had very good predictions for some variables and very poor predictions for others. For example, one parameter set had a likelihood value of 0.09 in predicting the dry matter yield, but only 0.0001 in predicting the maturity date, meaning that this parameter set was better in predicting yield but was not adequate for predicting maturity date. Under *C*1 and *C*3, the combined likelihood values of this parameter set were high enough (0.0451 and 0.0636, respectively) to contribute to the mean parameter estimate. However under *C*2, the combined likelihood value of this parameter set was near zero and neglected when deriving the posterior distribution. If this parameter set was selected as a member of the posterior distribution parameter sets, the ranges of parameters provided reasonably good corn yield predictions, but poor predictions of maturity dates. If many such parameter sets were retained in the posterior distribution set, the posterior distribution would be broad and the uncertainties of parameters may not be significantly reduced relative to their prior distributions, or they could even be increased as shown in Table 3.

For combination method *C*2, the uncertainties of input parameters were all dramatically reduced under all likelihood functions. For example, the uncertainty in the estimated soil parameter SDUL was reduced to 0.001 (Table 3). In addition, the mean value of SDUL under *L*1*C*2 was 0.106, while the "true value" of SDUL was 0.104, an error of only 0.002. The *C*2 method was the most discriminating in eliminating unsatisfactory parameter sets. When *C*2 was used,

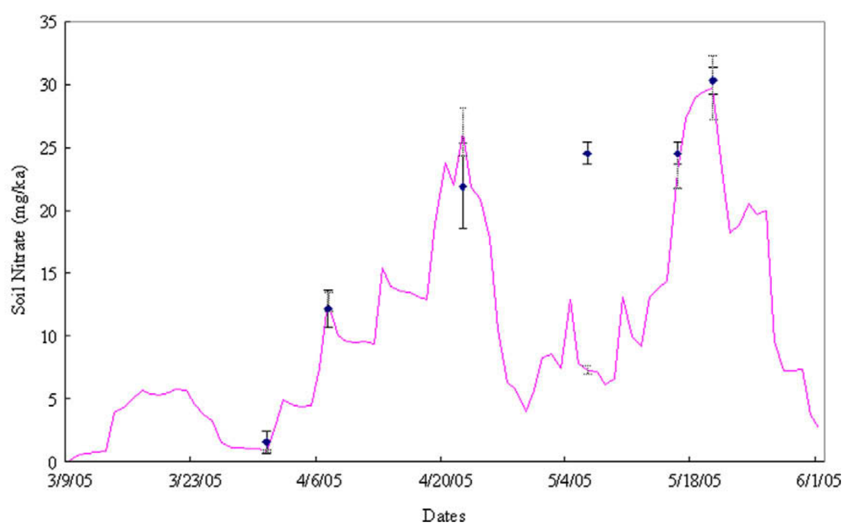

Fig. 2. Measured and simulated soil nitrate of Layer 1 (0–15 cm) with the parameter set that had the highest likelihood value under *L*1*C*2 in the second round of GLUE during the sweet corn season. The solid and dashed error bars show the standard deviations of synthetic measurements and simulations, respectively.

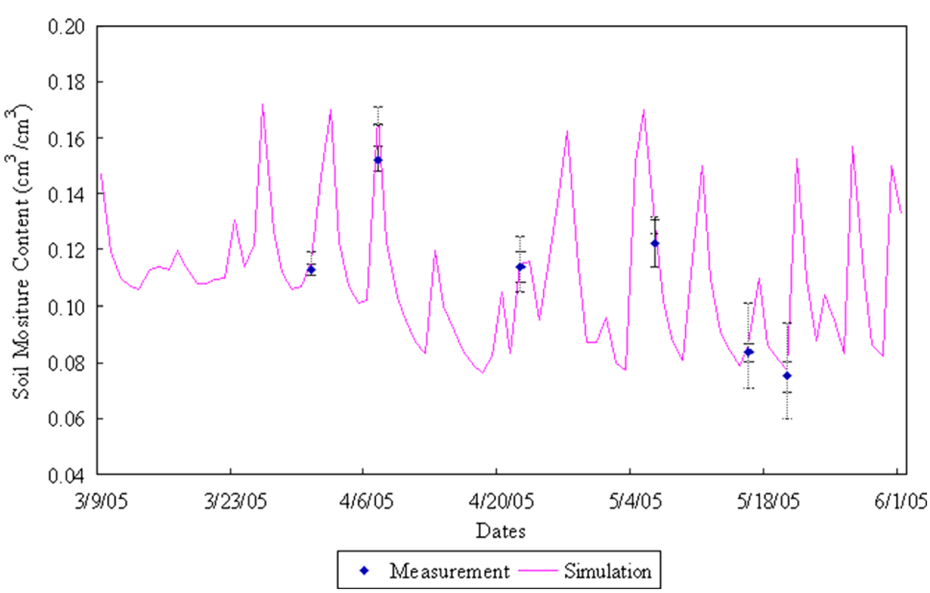

Fig. 3. Measured and simulated soil moisture of Layer 1 (0–15 cm) with the parameter set that had the highest likelihood value under *L*1*C*2 in the second round of GLUE during the sweet corn season. The solid and dashed error bars show the standard deviations of synthetic measurements and simulations, respectively.

posterior distribution differences were quite small regardless of the likelihood function that was used. For example, the mean value of the prior distribution of P1 was 225.1. The mean values of the posterior distributions were 116.12, 117.16, 114.03, and 115.91 under *L*1, *L*2, *L*3, and *L*4, respectively. These results were not surprising since the forms of these likelihood functions were all variants based on the probability density function of the normal distribution. We did not test likelihood functions that are based on other probability distributions in this study.

It was found that the combination method *C*2 was the most efficient one among the three tested methods. However, it was difficult to determine which likelihood function was the best one in estimating the parameter values. Thus, the methods *L*1*C*2, *L*2*C*2, *L*3*C*2, and *L*4*C*2 were used in a second round of analysis, while the other methods that used either *C*1 or *C*3 were eliminated to simplify the study.

3.2. Comparison of likelihood functions

The mean values and standard deviations of posterior distributions obtained in the second round of GLUE analysis are summarized in Table 4. The method *L*1*C*2 had the lowest average absolute relative error (ARE) value of 5.2% between the mean values of posterior distribution and the initially selected "true parameter set". In addition, the ARE values of input parameters were all lower than 10% for the *L*1*C*2 method. Methods *L*4*C*2, *L*3*C*2, and *L*2*C*2 ranked second through fourth, respectively, with average ARE values of 5.3%, 6.5%, and 7.4%. Though, method *L*4*C*2 also had a small average ARE value, it had higher standard deviation (STDEV) values relative to *L*1*C*2. For example, the STDEV value for P1 was only 0.266 under *L*1*C*2, but it was 1.581 under *L*4*C*2. Similar results were observed for all other parameters except for the soil parameter SDUL.

Under each likelihood function and combination method, the parameter set that had the highest likelihood values had parameter values close to the "true parameter set". After two rounds of GLUE, the uncertainties in parameter estimates were reduced to a very low level. Thus it was concluded that the LlC2 method was the best one for minimizing both ARE values and standard deviations of posterior distributions, i.e. likelihood function *L*1 was the best one among the four likelihood functions evaluated.

3.3. Comparison of model outputs

To further evaluate the reliability of the likelihood function and method of combing likelihood values, model outputs from different posterior distributions (Table 4) were also compared in Table 5. Outputs from *L*1*C*2 most precisely matched the synthetic field observations with the lowest average ARE value of 2.85%. The estimated anthesis and maturity dates (60 and 97 days) under *L*1*C*2 were within the range of synthetic observations presented Table 2. However, the estimated maturity dates under *L*2*C*2, *L*3*C*2, and *L*4*C*2 were all greater than the 97 days that was the maximum generated observed maturity date listed in Table 2. The simulated mean value of dry grain yield under *L*4*C*2 was also greater than the maximum observation of yield in Table 2. Based on model output comparisons, *L*1*C*2 produced outputs closest to the "true observations". Since leaf TKN concentration, soil nitrate and moisture concentration were both temporally and/or spatially variable, they are compared in Figs. 1–3. In these figures, the parameter set that had the highest likelihood value under *L*1*C*2 was used to simulate the daily results shown by the solid line. The points are mean values of the generated synthetic data of leaf TKN, soil nitrate concentration of layer 1 (0–15 cm), and soil moisture content of layer 1 (0–15 cm), which represent the measured means of field observations. The solid and dashed error bars about the points show measurement and simulation standard deviations, respectively. In general, the simulations using the highest likelihood parameters identified using the *L*1*C*2 method matched the measurements very well, except for sporadic outliers in leaf TKN and soil nitrate concentrations (Figs. 1 and 2).

4. Summary and conclusions

Although many likelihood functions and methods of likelihood combination have been suggested in the literature, arbitrary selection of likelihood function is likely to produce unreliable parameter estimates. In this study, it was found that the likelihood function and method of combing likelihood values had a very strong influence on parameter estimation results for the CERES-Maize model.

The likelihood function that was directly derived from the probability density function of normal distribution (*L*1) performed best in this study and produced model outputs from the posterior distribution that were closest to the "true measurements" obtained from the initial "true parameter set", with an average ARE value of 2.85%. The likelihood combination method using the mathematical product (*C*2) efficiently reduced the uncertainties in posterior parameter distributions. This method avoided the extreme random parameter sets that simultaneously predicted some outputs very well and other outputs very poorly. The methods based on mathematical addition and mean square failed to efficiently reduce the uncertainties, since they did not eliminate extreme parameter sets. The best likelihood function and method of combining likelihood values for this study was determined to be *L*1*C*2 (Eqs. (5) and (12)). The approach for combining likelihood values from different types of measurements was more important than the choice of likelihood function in this study. Although these results are specific to the data used to estimate the parameters and the model that we used, our study clearly demonstrated the need to avoid using arbitrary relationships as criteria for estimating model parameters.

Acknowledgements

This paper was partially supported by the Office of Natural Resources Management and Office of Agriculture in the Economic Growth, Agriculture, and Trade Bureau of the US Agency for International Development, under terms of Grant No. LAG-G-00-97-00002-00. Special appreciation is expressed to the Plant Science Research & Education Unit, University of Florida. The authors thank Dr. M. Annable, Dr. G. Hochmuth, Dr. J. Judge, and D. Burch of the Department of Agricultural and Biological Engineering, University of Florida for their help on this project.

References

Aronica, G., Hankin, B.G., Beven, K.J., 1998. Uncertainty and equifinality in calibrating distributed roughness coefficients in a flood propagation model with limited data. Advances in Water Resources 22 (4), 349–365.

Bechini, L., Bocchi, S., Maggiore, T., Confalonieri, R., 2006. Parameterization of a crop growth and development simulation model at sub model component level. An example for winter wheat (*Triticum aestivum* L.). Environmental Modelling & Software 21, 1042–1054.

Beven, K., 2006. A manifesto for the equifinality thesis. Journal of Hydrology 320, 18–36.

Beven, K., Binley, A., 1992. The future of distributed models: model calibration and uncertainty prediction. Hydrological Processes 6, 279–298.

Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modeling of complex environmental systems using the GLUE methodology. Journal of Hydrology 249, 11–29.

Candela, A., Noto, L.V., Aronica, G., 2005. Influence of surface roughness in hydrological response of semiarid catchments. Journal of Hydrology 313, 119–131.

Campbell, E.P., Fox, D.R., Bates, B.C., 1999. A Bayesian approach to parameter estimation and pooling in nonlinear flood event models. Water Resource Research 35, 211–220.

Casanova, J.J., Judge, J., Jang, M.Y., 2006. Vertical Distribution of Moisture in a Growing Sweet Corn Canopy During MicroWEX-4. Circular No. 1492. University of Florida, IFAS, Gainesville, Fla. <http://edis.ifas.ufl.edu/AE395>.

Casanova, J.J., Yan, F., Judge, J., Tien, K.J., Miller, L.W., Lanni, O., 2007. Field Observations During the Fifth Microwave, Water, and Energy Balance Experiment (MicroWEX-5): from March 9 through May 26, 2006. Circular No. 1514. University of Florida, IFAS, Gainesville, Fla. <http://edis.ifas.ufl.edu/AE407>.

Freer, J., Beven, K.J., Ambroise, B., 1996. Bayesian estimation of uncertainty in runoff prediction and the value of data: an application of the GLUE approach. Water Resources Research 32 (7), 2161–2173.

Grimm, S.S., Jones, J.W., Boote, K.J., Hesketh, J.D., 1993. Parameter estimation for predicting flowering dates of soybean cultivars. Crop Science 33, 137–144.

He, J., 2008. Best Management Practice Development with the CERES-Maize Model for Sweet Corn Production in North Florida. Dissertation, Agricultural and Biological Engineering Department, University of Florida, Gainesville, FL.

He, J., Dukes, M.D., Jones, J.W., Graham, W.D., 2009. Applying GLUE for estimating CERES-Maize genetic and soil parameters for sweet corn production. Transactions ASABE 52 (6), 1907–1921.

Hoogenboom, G., Jones, J.W., Porter, C.H., Wilkens, P.W., Boote, K.J., Batchelor, W.D., Hunt, L.A., Tsuji, G.Y. (Eds.), 2003. Decision Support System for Agrotechnology Transfer Version 4.0: Overview, vol. 1. University of Hawaii, Honolulu, HI.

Hunt, L.A., Pararyasingham, S., Jones, J.W., Hoogenboom, G.D., Imamura, T., Ogoshi, R.M., 1993. GENCALC – software to facilitate the use of crop models for analyzing field experiments. Agronomy Journal 85, 1090–1094.

Jones, J., Hoogenboom, G., Porter, C., Boote, K., Batchelor, W., Hunt, L., Wilkens, P., Singh, U., Gijsman, A., Ritchie, J., 2003. The DSSAT model cropping system model. European Journal of Agronomy 18 (3–4), 235–265.

Jones, C.A., Kiniry, J.R., 1986. CERES-Maize: A Simulation Model of Maize Growth and Development. Texas A&M University Press, College Station, Texas.

Keesman, K., van Straten, G., 1989. Identification and prediction propagation of uncertainty in models' with bound noise. International Journal of Control 49, 2259–2269.

Keesman, K., van Straten, G., 1990. Set membership approach to identification and prediction of lake eutrophication. Water Resources Research 26 (11), 2643–2652.

Lamb, R., Beven, K., Myrabo, S., 1998. Use of spatially distributed water table observations to constrain uncertainty in a rainfall–runoff model. Advances in Water Resources 22 (4), 305–317.

Makowski, D., Hillier, J., Wallach, D., Andrieu, B., Jeuffroy, M.H., 2006a. Parameter estimation for crop models. In: Wallach, D., Makowski, D., Jones, J. (Eds.), Working with Dynamic Crop Models. Elsevier, Amsterdam, pp. 101–150.

Makowski, D., Naud, C., Jeuffroy, M.H., Barbottin, A., Monod, H., 2006b. Global sensitivity analysis for calculating the contribution of genetic parameters to the variance of crop model predictions. Reliability Engineering and System Safety 91, 1142–1147.

Makowski, D., Wallach, D., Tremblay, M., 2002. Using a Bayesian approach to parameter estimation: comparison of the GLUE and MCMC methods. Agronomie 22, 191–203.

MATLAB, 2004. Version 7.0.0. MathWorks, Inc.

Mavromatis, T., Boote, K.J., Jones, J.W., Wilkerson, G.G., Hoogenboom, G., 2002. Repeatability of model genetic coefficients derived from soybean performance trials across different states. Crop Science 42, 76–89.

Monod, H., Naud, C., Makowski, D., 2006. Uncertainty and sensitivity analysis for crop models. In: Wallach, D., Makowski, D., Jones, J. (Eds.), Working with Dynamic Crop Models. Elsevier, Amsterdam, pp. 55–100.

Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. Technometrics 33 (2), 161–174.

Ritchie, J.T., Godwin, D.C., 1989. Description of soil water balance. In: Virnami, S.M., Tandon, H.L.S., Alagarswamy, G. (Eds.), Modeling the Growth and Development of Sorghum and Pearl Millet. ICRISAT, Research Bulletin 12, Pradesh, India, pp. 14–16.

Ritchie, J.T., 1998. Soil water balance and plant water stress. In: Tsuji, G.Y., Hoogenboom, G., Thornton, P.K. (Eds.), Understanding Options for Agricultural Production. Kluwer Academic Publishers, Dordrecht, The Netherlands, pp. 41–54.

Romanowicz, R., Beven, K.J., Tawn, J., 1994. Evaluation of predictive uncertainty in non-linear hydrological models using a Bayesian approach. In: Barnett, V., Turkman, K.F. (Eds.), Statistics for the Environment II. Water Related Issues. Wiley, New York, pp. 297–317.

Romanowicz, R., Beven, K.J., Tawn, J., 1996. Bayesian calibration of flood inundation models. In: Anderson, M.G., Walling, D.E., Bates, P.D. (Eds.), Flood Plain Processes. Wiley, Chichester.

Stedinger, J.R., Vogel, R.M., Lee, S.U., Batchelder, R., 2008. Appraisal of the generalized likelihood uncertainty estimation (GLUE) method. Water Resources Research 44, W00B06.

Tsuji, G.Y., Hoogenboom, G., Thornton, P.K. (Eds.), 1998. Systems Approach for Sustainable Agricultural Development: Understanding Options for Agricultural Production. Kluwer Academic Publishers in Cooperation with ICASA, pp. 1–6.

Van Straten, G., 1983. Maximum likelihood estimation of parameter and uncertainty in phytoplankton models. In: Beck, M.B., van Straten, G. (Eds.), Uncertainty and Forecasting of Water Quality. Springer-Verlag, Berlin, pp. 157–171.

Wang, X., He, X., Williams, J.R., Izaurralde, R.C., Atwood, J.D., 2005. Sensitivity and uncertainty analyses of crop yields and soil organic carbon simulated with EPIC. Transaction of ASAE 48 (3), 1041–1054.

Wallach, D., Goffinet, B., Bergez, J.E., Debaeke, P., Leenhardt, D., Aubertot, J.N., 2001. Parameter estimation for crop models: a new approach and application to a corn model. Agronomy Journal 93, 757–766.

Zak, S.K., Beven, K.J., Reynolds, B., 1997. Uncertainty in the estimation of critical loads: a practical methodology. Soil, Water and Air Pollution 98, 297–316.