Agricultural Systems 135 (2015) 31–47

Contents lists available at ScienceDirect

Agricultural Systems

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

Parameter and uncertainty estimation for maize, peanut and cotton using the SALUS crop model

K.A. Dzotsi [a,*], B. Basso [b], J.W. Jones [a]

[a] Agricultural and Biological Engineering Department, University of Florida, P.O. Box 110570, Gainesville, FL, United States
[b] Department of Geological Sciences, W.K. Kellogg Biological Station, Michigan State University, East Lansing, MI, United States

A R T I C L E I N F O

Article history:
Received 13 September 2013
Received in revised form 18 November 2014
Accepted 9 December 2014
Available online

Keywords::
Bayesian parameter estimation
Crop modeling
DSSAT
Markov Chain Monte Carlo
Metropolis-Hastings
SALUS

A B S T R A C T

The generic and simple version of SALUS (System Approach to Land Use Sustainability) crop model was recently integrated in the DSSAT (Decision Support System for Agrotechnology Transfer) cropping system model to provide an alternative approach to more complex crop models without the need for a detailed parameterization.

A previous uncertainty and sensitivity analysis of the model (SALUS-Simple) established that accurate estimation of 15 of the 20 crop parameters required for predicting crop performance under water limitation was necessary to achieve reliable simulations. The present study used a Markov Chain Monte Carlo-based Bayesian stepwise approach for estimating crop parameters in SALUS-Simple using limited, end-of-season data (limited data case) and detailed in-season data (detailed data case). Independent testing were performed using data distributed with DSSAT version 4.5.

Results of the detailed data case indicated that the estimated parameters resulted in smaller deviations between simulated and measured variables and in posterior parameter distributions with smaller variances. Independent testing showed that maize growth simulations (based on both data cases) were in good agreement with observations while peanut and cotton growth was simulated with mixed success. SALUS-Simple predictions using parameters estimated in the limited data case were concordant with observations for end-of-season biomass and yield, but simulations of in-season growth were degraded relative to the use of parameters estimated in the detailed data case.

We conclude that the use of a sequential approach reduced compensation errors and, the use of a range of data types combined with a higher ratio between the number of data points and the number of estimated parameters significantly reduced uncertainties associated with the estimated parameters. Furthermore, model predictions based on mean parameter values can be regarded as reliable estimators of the expected values of the distributions of model predictions when an average prediction rather than a distribution is needed. Results from this study highlighted the principle that parameters estimated based on end-of-season data do not guarantee accurate prediction of in-season growth even if a Bayesian approach is used. The ability of the SALUS-Simple model to be parameterized or adapted for simulating canopy-level potential production of annual plants based on limited data is promising. Further testing of the model will help establish its response to different soils, climates and crops.

Published by Elsevier Ltd.

1. Introduction

Scientific interest in using crop models at regional and global scales for large-scale assessment of agricultural systems has drawn attention to generic, simplified crop models. A number of large-scale crop modeling studies have raised concerns about knowledge gaps in calibrating crop models at regional scales and accurately modeling multi-crop coverage for adequate comparison with national agricultural statistics (Adam et al., 2011; Faivre et al., 2004; Rosenzweig et al., 2013). The major benefit of a generic model for annual crops and grasses at this scale is to provide a consistent framework for representing cropping system components present in a spatial grid cell with the same model and set of parameters. A primary advantage associated with a simple model is the relatively small number of parameters required to characterize a crop and the ability to use literature or limited crop data to provide reasonable estimates of the parameters. In addition, recent model intercomparison studies (Asseng et al., 2013) have shown that alternative approaches can assist with identifying strengths and weaknesses in different models, provide a more realistic explanation of causes of simulation errors, promote model improvement, and contribute to improved model-based recommendations.

* Corresponding author. Tel.: +1 352 392 1864x263; fax: +1 352 392 4092.
E-mail address: kofikuma@ufl.edu (K.A. Dzotsi).

http://dx.doi.org/10.1016/j.agsy.2014.12.003
0308-521X/Published by Elsevier Ltd.

The generic and simple version of the SALUS (System Approach to Land Use Sustainability) crop model was recently integrated in the DSSAT (Decision Support System for Agrotechnology Transfer) cropping system model to provide an alternative approach to more complex crop models in DSSAT without the need for estimation of detailed crop parameters (Dzotsi et al., 2013). The typical crop described in the model (SALUS-Simple) has an annual cycle with its leaf area index (LAI) curve described by an increase and a decline with thermal time during the season, and its total dry matter estimated using the radiation use efficiency (RUE) approach (Monteith, 1977). The ability of a model to simulate crops and grasses with a relatively small number of parameters makes it a good candidate for large-scale modeling studies in which within-field variations are not critical due to aggregation of model outputs (Bondeau et al., 2007; Challinor et al., 2004; Lobell and Burke, 2010; Stehfest et al., 2007). Parameterization of crops in SALUS-Simple is a major component of the model development and provides the opportunity of assessing the ability of the model to correctly represent within- and between-species variations as well as interactions occurring in the soil-plant-atmosphere continuum. The model potentially can be parameterized for simulating the characteristics of crop maturity groups based on literature and limited data, and therefore can contribute to reducing uncertainties in estimating detailed parameters for specific cultivars using aggregated crop data.

An uncertainty and sensitivity analysis demonstrated that the SALUS model is sensitive to crop parameters involved in the accumulation of thermal time, the representation of leaf area index (LAI), and the calculation of plant dry matter (Dzotsi et al., 2013). That initial study demonstrated strong relationships between some parameters and specific model outputs, which suggested that an adaptive approach could be used to estimate crop parameters with the goal of reducing compensation errors that tend to occur during model calibration (Wallach, 2011). This paper demonstrates such a stepwise approach for maize, peanut and cotton while accounting for the limited availability of data that affect most large-scale crop modeling studies and the existence of uncertain prior information about the parameters. Our stepwise procedure emphasizes a Bayesian parameter estimation approach but integrates a frequentist approach to reduce compensation errors. A frequentist approach considers a parameter to have a true, fixed value that can be estimated by obtaining a random sample of data from an experiment (Ellison, 1996). This approach is solely based on available data and may lead to over-fitting of crop parameters to data without accounting for other sources of uncertainty affecting the parameters. Frequentist approaches also exhibit a greater risk of producing data-, site-specific parameter estimates because they do not provide a mean of incorporating existing knowledge that may have been accumulated about the parameter from other sources into the parameter estimation procedure. Since the true value of the parameter is unknown, a confidence interval constructed about the estimated parameter value in a frequentist approach captures the true value in *k*% of all possible samples. A more transferrable result would be an interval that contains *k*% of all possible parameter values (Ellison, 1996), or even better a probability distribution that quantifies the uncertainty about the parameter given the data: this is the typical methodology of a Bayesian approach (Makowski et al., 2002).

A Bayesian approach integrates several sources of information about the parameters treated as random variables for which posterior distributions can be derived by combining prior knowledge with observed data. Before using data to estimate parameters, uncertainty ranges of parameters can be defined, for example based on values found in the literature or expert knowledge. A Bayesian approach further associates a likelihood function to the data, which is the probability of observing the data conditional upon the parameter set. In theory, the posterior parameter distribution can be calculated using Bayes theorem. However, in practice, this is generally complicated by high numbers of model parameters and crop model non-linearities. One of the most recognized numerical methods originally used by physicists is the Metropolis algorithm (Metropolis et al., 1953), later generalized as the Metropolis–Hastings (MH) algorithm (Hastings, 1970) which is a Markov Chain Monte Carlo (MCMC) approach because it relies on the current value of the parameter to determine the next sample in the sequence. The Bayesian approach based on the MCMC-MH algorithm has been increasingly popular in the literature and applications in various fields including hydrology (Bates and Campbell, 2001), forestry (Ceglar et al., 2011), large-scale crop modeling (Iizumi et al., 2009), astrophysics (Putze et al., 2010), and field scale crop modeling (Makowski et al., 2002) have been reported. A comprehensive description and analysis of the MH algorithm can be found in Chib and Greenberg (1995).

This paper pursues the following objectives related to the SALUS-Simple model development and testing: i) develop and test parameter sets for maize, peanut and cotton, and estimate associated uncertainties using available data in DSSAT version 4.5; ii) investigate the effect of data detail on the accuracy of estimated parameters.

2. Materials and methods

2.1. The SALUS-Simple crop model

Parameters were estimated for maize (*Zea mays* L.), peanut (*Arachis hypogaea* L.) and cotton (*Gossypium hirsutum* L.) using the SALUS-Simple crop model (Basso et al., 2006; Dzotsi et al., 2013) integrated in the Decision Support System for Agrotechnology Transfer (DSSAT, Jones et al., 2003). SALUS-Simple is a generic crop model that derives from ALMANAC (Agricultural Land Management Alternatives with Numerical Assessment Criteria; Kiniry et al., 1992) and was designed to simulate a wide range of annual plant species and the impact of different environments. Within-species variations may be confined to maturity groups. A detailed description of the SALUS-Simple model for maize, peanut and cotton can be found in Dzotsi et al. (2013). While the model can be parameterized for specific cultivars, it was not designed to capture detailed differences among plant cultivars. Total crop growth duration is expressed as cumulative degree-days from planting to maturity (*TTMature*, Table 1) and the relative thermal time is the fraction of *TTMature* such that 0 represent planting and 1 corresponds to maturity. Timing of germination and emergence are predicted using crop parameters that describe the cumulative degree-days required for the occurrence of these events. The beginning of leaf senescence, which marks the end of vegetative growth, is a crop parameter expressed in terms of relative thermal time. Leaf area index is simulated directly using a sigmoid function whose shape is controlled by three crop parameters (Table 1). Total dry matter is calculated using the RUE approach (Monsi and Saeki, 1953; Monteith, 1977). A relative thermal time-dependent function is used to dynamically partition total dry matter between roots and aboveground plant parts while yield is considered to be a fixed fraction of aboveground dry matter. The current version of SALUS-Simple uses 20 parameters to model a crop in water-limited conditions (Table 1).

2.2. Parameter and uncertainty estimation cases

A total of 13 crop parameters were estimated (Table 1 and Fig. 1) following two distinct procedures to study the effect of limited availability of data on the parameters and uncertainty estimates. The case of availability of a detailed dataset (hereafter denoted by detailed