A Theoretical and Real World Evaluation of Two Bayesian Techniques for the Calibration of Variety Parameters in a Sugarcane Crop Model

J SEXTON[1], Y EVERINGHAM[1, 2], G INMAN-BAMBER[1]

[1]*College of Science, Technology and Engineering, James Cook University, Townsville, 4811*

[2]*Centre for Tropical Environmental & Sustainability Science*

Available at: http://dx.doi.org/10.1016/j.envsoft.2016.05.014

Highlights

- We evaluate two Bayesian methods of calibrating sugarcane varieties in a crop model
- Variety parameters can be estimated using limited biomass and sucrose yield data
- We were able to calibrate differences between parameters of two pre-defined varieties
- MCMC calibration estimates of variety parameter values were physically meaningful
- Bayesian calibration can be used to routinely update crop models for new varieties

Abstract

Process based agricultural systems models allow researchers to investigate the interactions between variety, environment and management. The 'Sugar' module in the **A**gricultural **P**roductions **S**ystems s**IM**ulator (APSIM-Sugar) currently includes definitions for 14 sugarcane varieties, most of which are no longer commercially grown. This study evaluated the use of two Bayesian approaches to calibrate sugarcane varieties in APSIM-Sugar: Generalized Likelihood Uncertainty Estimation (GLUE) and Markov Chain Monte Carlo (MCMC). Both GLUE and MCMC calibrations were able to accurately simulate green biomass and sucrose yield in both a theoretical and real world evaluation. In the theoretical evaluation GLUE and MCMC parameter estimates accurately reflected differences between two pre-defined sugarcane varieties. We found that the MCMC approach can be used to calibrate varieties in APSIM-Sugar based on yield data. With appropriate variety definitions, APSIM-Sugar could be used for early risk assessment of adopting new varieties.

Keywords: APSIM; Sugarcane; GLUE; MCMC; Bayesian; Calibration

1. Introduction

Australian sugarcane varieties have changed over the last 15 years, yet their representation in crop growth models has not. The Australian sugar industry is constantly developing new varieties to improve sucrose yields, pest and disease resistance and tolerance of abiotic stresses such as drought and water logging. Diseases such as orange rust (Magarey et al., 2001) have greatly affected prominent varieties like Q124, requiring new disease resistant varieties for commercial applications. Most varieties commonly grown in 1999 have been superseded by new varieties (Fig. 1). Despite this crop models such as the **A**gricultural **P**roduction **S**ystems s**IM**ulator (APSIM) (Holzworth et al., 2014) do not offer options for these new varieties (APSIM-Sugar Version 7.5 r3124) (Keating et al., 1999). This has limited the ability of modellers to explore sugarcane varietal differences through simulation. Yet, the ability to simulate varietal differences has been well explored for a range of other commercial crops. For example, varietal simulations have been used to identify ideal varieties for given environments in crops such as peanuts (Putto et al., 2013), wheat (Laurila et al., 2012), and rice (Aggarwal et al., 1997). Varietal simulations have also been used to investigate likely sources of genetic gains in soybeans (Boote et al., 2001) and to link crop models with genetic models (Chapman, 2008).

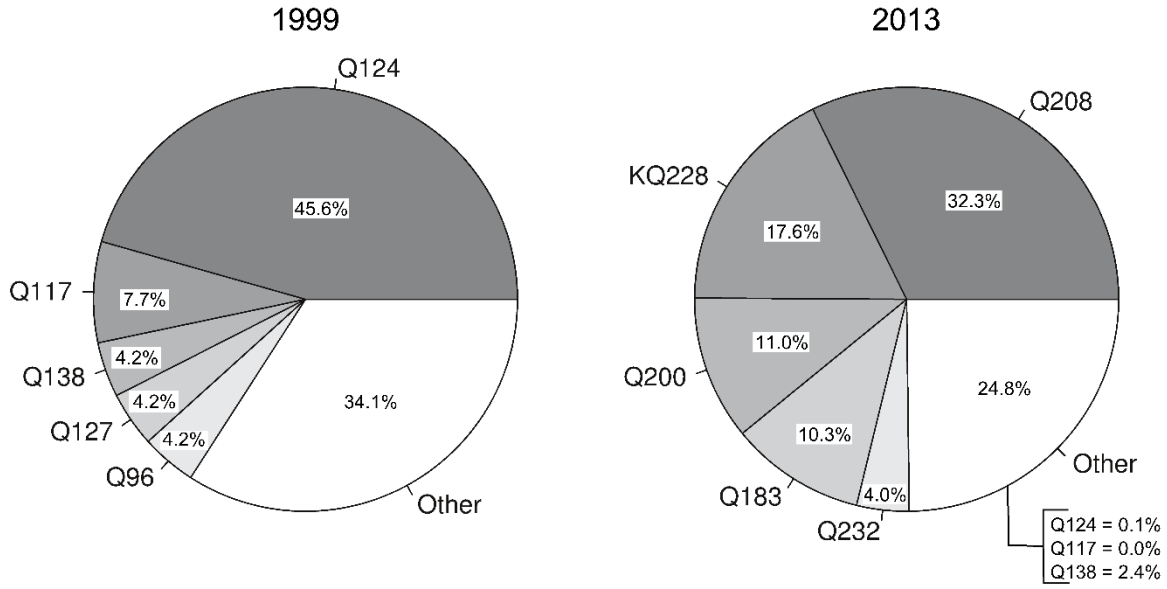

Fig. 1. Percentage of hectares grown by variety in Australia (QLD and NSW) for 1999 and 2013. Values downloaded from QCANESelect™ provided by Sugar Research Australia (accessed 11-09-2014). Data were collected as part of the SPIDNet database (Lethbridge and Cox, 2010).

Dynamic, process-based, crop models are advantageous as they simulate the underlying physiological processes required to grow a crop (Lisson et al., 2005). Cropping-system simulators

model crop growth in response to environments, soils, abiotic stresses, varieties and management decisions. Many such models are available for sugarcane including APSIM-Sugar, DSSAT-Canegro (Singels et al., 2008), QCane (Liu and Bull, 2001), AUSCANE (Jones et al., 1989) and STICS-Sugarcane (Brisson et al., 2003). These models are widely used as decision support tools to help identify industry best practices as well as helping farmers develop site specific management plans. DSSAT-Canegro and APSIM-Sugar are two of the most widely used models. For example APSIM-Sugar has been used in Australia for irrigation scheduling (Everingham et al., 2002), investigating nitrogen best management practices (Skocaj et al., 2013; Stewart et al., 2006; Thorburn et al., 2010), and climate change impact studies (Biggs et al., 2013; Webster et al., 2009). Canegro in the DSSAT (Decision Support System for Agrotechnology Transfer) environment has been used for similar studies largely in South Africa (Bezuidenhout et al., 2002; Jones et al., 2014; McGlinchey and Inman-Bamber, 1996; van der Laan et al., 2011) but also for climate change impacts in Swaziland (Knox et al., 2010) and Australia and Brazil (Singels et al., 2013). To date little research has considered varietal effects on such decision support research in sugarcane.

Canegro (DSSAT V4.5) represents 13 sugarcane varieties using 22 variety specific trait parameters while APSIM contains 14 sugarcane varieties described using 14 parameters. Some parameters representing physiological traits such as leaf area and number of green leaves are easy to measure. However, accurately obtaining parameters for other traits (e.g. transpiration efficiency) can be expensive and time consuming. An alternative approach is to statistically estimate parameters that represent difficult to measure traits from knowledge about traits that are routinely or more easily measured such as biomass, cane yield or sucrose yield. Various methods have been used to estimate crop model parameters such as sensitivity analysis, least squares regression and maximum likelihood regression (Makowski et al., 2006). Recently more complex statistical approaches such as genetic algorithms (Mitchell, 1999) and Bayesian statistical approaches such as Generalized Likelihood Uncertainty Estimation (GLUE; Beven and Binley, (1992)) and Markov Chain Monte Carlo (MCMC; Gelman et al. (1997)) have found widespread application.

Bayesian statistical parameter estimation techniques like GLUE and MCMC allow researchers to use various sources of data and prior knowledge of parameter values. GLUE has become widely used in a range of crop models because of its computational simplicity (Makowski et al., 2002). GLUE has been used effectively for parameterizing generic crop models (Wang et al., 2005), models for maize (He et al., 2010), wheat (Mo and Beven, 2004), Cotton (Pathak et al., 2012) and sugarcane (Marin et al., 2011). MCMC algorithms have been used to estimate crop model variety parameters for rice (Iizumi et al., 2009, 2011), maize (Tao et al., 2009), wheat (Dumont et al., 2014; Tao and Zhang, 2013) and soybeans (Archontoulis et al., 2014) but have not been applied to sugarcane crop models.

In a recent thematic issue on agricultural systems modelling and software published in the journal of Environmental Modelling and Software, Holzworth et al. (2015) identified the need for more objective and reproducible model calibration and validation as a way forward for models in a growing agricultural domain. The use of Bayesian statistical calibration techniques supply model developers with a reproducible method for calibrating model parameters while simultaneously quantifying the uncertainty in those parameter values. While some modelling environments such as DSSAT provide inbuilt tools for this purpose (Jones et al., 2011), others such as APSIM do not have an operationally available calibration method (Archontoulis et al., 2014). The availability of a reproducible calibration methodology also simplifies the documentation of model calibrations in an industry where problems with model documentation has been a long-standing issue (Holzworth et al. 2015).

By applying both GLUE and MCMC calibration techniques in APSIM-Sugar this paper aims to evaluate both qualitatively and quantitatively the advantages and disadvantages of these two common Bayesian parameter estimation methods. The use of statistical techniques for parameter estimation would provide a much needed systematic method for updating variety specific trait parameters while quantifying uncertainty about those parameter values. Keeping APSIM-Sugar up-to-date with new commercially released varieties will also allow crop modellers to investigate the impact of interactions between variety, environment and farm management practices and will help improve the model as a decision support tool.

2. Theory

2.1. Agricultural Production System sIMulator

The APSIM-Sugar module simulates biomass accumulation on a daily time step (Keating et al., 1999). This requires detailed information on climate parameters such as rainfall, temperature and solar radiation. Biomass accumulation is driven largely by radiation (in the 300 nm to 3000 nm range) through radiation use efficiency (RUE) and is divided into five live pools (leaf, cabbage, structural stem, roots and sucrose). Biomass accumulation is limited by high or low temperature, excess or deficit water stresses and nitrogen deficit effects on RUE (Singels, 2013). This allows APSIM to simulate differences between potential and attainable yields and hence benchmark production based on yield actually obtained (Inman-Bamber, 2013). From biomass, key productivity measures such as cane yield, sugar yield and sucrose content of cane (measured as commercial cane sugar in Australia), are simulated.

Soil dynamics are a core concept of the APSIM modelling framework (Holzworth et al., 2014). Within the APSIM framework, specific modules are used to simulate soil water and nitrogen balances. This requires detailed information on soil properties. Further modules simulate farm management and the growth of specific crops, such as sugarcane (APSIM-Sugar; Keating et al., 1999). In order to simulate the growth of the specific crop, the underlying biophysical processes have to be represented by a range of model parameters.

Parameters in APSIM-Sugar that represent the biophysical processes of sugarcane growth are divided into generic plant and ratoon parameters and variety specific plant and ratoon parameters. Table 1 lists parameters that represent key traits such as canopy development (parameters 1, 2, 12, 13 and 14), biomass partitioning (parameters 3-8) and phenological stages (parameters 9-11). Some traits are represented by a combination of parameters. For example parameters *leaf_size* (area of fully expanded leaf) and *leaf_size_no* (leaf position along stalk) represent canopy development by describing inflection points of a Gompertz curve (Inman-Bamber, 2013) while *sucrose_fraction_stalk* and *stress_factor_stalk* reflect how stress affects the partitioning of assimilates into sucrose (Keating et al., 1999).

Although parameters 1-14 listed in Table 1 are classed as variety parameters in APSIM-Sugar, current definitions of varieties in APSIM-sugar differ only in (i) fully expanded area of leaves and (ii) partitioning of biomass to sucrose in the stalk. Flowering in sugarcane is sporadic and phenology trait parameters for flowering, although available are not implemented (Keating et al., 1999; Sexton and Everingham, 2014). Transpiration efficiency in sugarcane has recently been shown to vary between

varieties (Jackson et al., 2014) but the corresponding parameter in APSIM-Sugar (*transp_eff_cf*) remains a constant for different varieties. The RUE parameter in APSIM (*rue*) is generally considered a species constant. However, in the DSSAT-Canegro model RUE has a corresponding variety parameter '*parcemax*' based on photosynthetically active radiation ranging from 400 nm to 2000 nm (Marin et al., 2014; Marin et al., 2011). Apart from Marin et al. (2014 and 2011) there is no published evidence for genetic variation in this trait which is difficult to measure and has a profound effect on crop productivity.

Table 1. APSIM-Sugar parameters: Parameters 1 to 14 are considered variety specific in the APSIM-Sugar model. Parameters 1, 2, 12, 13 and 14 represent morphological traits such as the leaf size profile, and number of green leaves, while others represent developmental stages (parameters 9, 10 and 11), physiological traits such as RUE (parameter 16) or partitioning between sucrose and biomass (parameters 3 and 4). Some traits are described using pairs of variables such as the leaf size profile which is described using *leaf_size* (parameter 1) and *leaf_size_no* (parameter 2).

Parameter	Description	Units	Values (Variety Q117)
1 *leaf_size*	Area of each leaf	mm^2	1500 55000 55000
2 *leaf_size_no*	Leaf number from top leaf	leaf	1 14 20
3 *cane_fraction*	Fraction of accumulated biomass partitioned to cane	$g\ g^{-1}$	0.70
4 *sucrose_fraction_stalk*	Fraction accumulated biomass partitioned to sucrose	$g\ g^{-1}$	1 0.55
5 *stress_factor_stalk*	Stress factor for sucrose accumulation	nil	0.2 1.0
6 *sucrose_delay*	Sucrose accumulation delay	g	0
7 *min_sstem_sucrose*	Minimum stem biomass before partitioning to sucrose commences	g	800
8 *min_sstem_sucrose_redn*	Reduction to minimum stem sucrose under stress	g	10
9 *tt_emerg_to_begcane*	Thermal time required from emergence to start stalk growth	°C day	1900
10 *tt_begcane_to_flowering*	Thermal time required from start of stalk growth to start of flowering	°C day	6000
11 *tt_flowering_to_crop_end*	Thermal time from flowering to crop death	°C day	2000
12 *green_leaf_no*	Green leaf number	leaves	13.0
13 *tillerf_leaf_size*	Expansion factor applied to *leaf_size* due to tillering	nil	1 1 1.5 1 1
14 *tillerf_leaf_size_no*	Leaf number from top leaf	leaf	1 4 10 16 26
15 *transp_eff_cf*	Intrinsic transpiration efficiency coefficient	$g\ kPa\ g^{-1}$	0.0080
16 *rue*	Radiation use efficiency	$g\ MJ^{-1}$	1.8 (plant) 1.65 (ratoon)
17 k_L	Root water extraction coefficients	d^{-1}	Varies with soil

2.2. Bayesian statistical parameter estimation

Bayes' rule (Eq. (1)) relates prior belief about parameter values before observing any data (prior probability; $P(\boldsymbol{\theta})$), to posterior beliefs (posterior probability; $P(\boldsymbol{\theta}|\mathbf{Y})$) after observations are made through a sampling distribution known as the likelihood function ($P(\mathbf{Y}|\boldsymbol{\theta})$) (Gelman et al., 1997). That is, Bayes' rule defined as:

$$P(\boldsymbol{\theta}|\mathbf{Y}) \propto P(\boldsymbol{\theta})\, P(\mathbf{Y}|\boldsymbol{\theta}) \tag{1}$$

where

$\boldsymbol{\theta}$ = a vector of p unknown parameter values: $\boldsymbol{\theta} = [\theta_1, \theta_2, \ldots, \theta_p]$;

$\mathbf{Y}$ = a vector of O observations: $\mathbf{Y} = [y_1, y_2, \ldots, y_O]$;

$P(\boldsymbol{\theta})$ = the joint prior distribution of unknown parameters;

$P(\mathbf{Y}|\boldsymbol{\theta})$ = the likelihood function and

$P(\boldsymbol{\theta}|\mathbf{Y})$ = the joint posterior distribution of unknown parameters,

provides a systematic method of updating our beliefs of the parameter probability as more observations are made.

The likelihood function $P(\mathbf{Y}|\boldsymbol{\theta})$ describes the probability of the data $\mathbf{Y}$ given the parameters $\boldsymbol{\theta}$, dependent on the probability distribution of model errors – the difference between observed and simulated values (Makowski et al., 2006). In the case of parameter estimation the parameters are the unknown quantity of interest and the likelihood function is generally defined as $L(\boldsymbol{\theta}|\mathbf{Y})$, that is the likelihood of parameter set $\boldsymbol{\theta}$ given observations $\mathbf{Y} = [y_1, y_2, \ldots, y_o]$ with model error variance (σ^2). The choice of likelihood function should reflect the actual distribution of model errors (He et al., 2010; Makowski and Wallach, 2002; Stedinger et al., 2008).

The most widely used likelihood functions for both GLUE and MCMC are based on the Gaussian distribution such that:

$$L(\boldsymbol{\theta}|\mathbf{Y}) = \prod_{o=1}^{O} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{[y_o - \hat{y}_o(\boldsymbol{\theta})]^2}{2\sigma^2}\right) \tag{2}$$

where

y_o is the o^{th} observation and

$\hat{y}_o(\boldsymbol{\theta})$ is the o^{th} simulated value using parameter set $\boldsymbol{\theta}$,

assuming that model errors (residuals) are normally distributed (Dumont et al., 2014; He et al., 2010; Iizumi et al., 2009; Jones et al., 2011; Makowski et al., 2006; Makowski et al., 2002; Marin et al., 2011;

Pathak et al., 2012). Model error variance (σ^2) can be estimated by the variance in the observations or from a well calibrated example (Pathak et al., 2012; Wang et al., 2005). Furthermore error variance may be specified for each observation (σ_o^2) if variances for each are known (Iizumi et al., 2009).

Model errors are often assumed to be normally distributed without reporting the validity of the assumption (Jones et al., 2011; Marin et al., 2011; Iizumi et al., 2009; Pathak et al., 2012). In practice observations often used to calibrate crop models such as biomass or leaf area index (LAI) may have variances that change with the size of the observation (Wallach et al., 2011) and are therefore not normally distributed. The assumption of normality in such cases may be addressed by performing the calibration using appropriately transformed (e.g. log transformed) observations (Wallach et al., 2011; Kennedy and O'Hagan 2001). The likelihood based on a normal distribution (Eq. (2)) was used in this study as it has been used effectively in similar crop model calibration studies such as Marin et al. (2011).

The main advantage of the Bayesian approach to parameter estimation is that prior knowledge of parameter values is taken into account to describe a posterior probability based on observations. This allows researchers to estimate parameters from different data sources and quantify the uncertainty in model outputs due to parameter uncertainty (Makowski et al., 2006). In practice, Bayesian calibration of crop models generally use independent uniform prior distributions of parameter values as little information is available on parameter covariance. Uniform priors have been used in the Bayesian calibration of crop models for wheat (Dumont et al., 2014), soybeans (Archontoulis et al., 2014) and sugarcane (Marin et al., 2011). Prior distributions should be based on published data or wide enough to be relatively sure of capturing the true parameter value (Beven and Binley, 1992).

2.2.1 Generalized Likelihood Uncertainty Estimation

Generalized Likelihood Uncertainty Estimation or "GLUE" (Beven and Binley, 1992; Beven and Freer, 2001; Stedinger et al., 2008) is a Monte Carlo approach whereby a large sample of parameter sets [$\boldsymbol{\theta}_1$, $\boldsymbol{\theta}_2$, …, $\boldsymbol{\theta}_Q$] are chosen from a defined prior distribution. A goodness-of-fit likelihood is calculated for each parameter set and used to produce weighted posterior probability densities for each parameter set which in turn are used to approximate the posterior distribution for each parameter. A simplistic overview can be outlined in four steps:

1. A prior probability distribution is developed for each parameter

2. A Monte Carlo approach is used to generate a large number (Q) of random samples of parameter sets from the prior distribution
3. Calculate likelihood $L(\boldsymbol{\theta}_i|\mathbf{Y})$ for each parameter set using an explicit likelihood function
4. Calculate the posterior probability density $p(\boldsymbol{\theta})$ of each parameter set (He et al., 2010; Jones et al., 2011) as:

$$p(\boldsymbol{\theta}_i) = \frac{L(\boldsymbol{\theta}_i|\mathbf{Y})}{\sum_{i=1}^{Q} L(\boldsymbol{\theta}_i|\mathbf{Y})} \quad (3)$$

Posterior probability densities can then be used to define the empirical posterior distribution and the distribution mean:

$$\hat{\boldsymbol{\mu}}_{post} = \sum_{i=1}^{Q} p(\boldsymbol{\theta}_i)\cdot\boldsymbol{\theta}_i \quad (4)$$

and variance:

$$\hat{\boldsymbol{\sigma}}_{post}^{2} = \sum_{i=1}^{Q} p(\boldsymbol{\theta}_i)\cdot(\boldsymbol{\theta}_i - \hat{\boldsymbol{\mu}}_{post})^2 \quad (5)$$

for each parameter.

An additional step generally applied in the GLUE process is to remove "non-behavioural" parameter sets from the posterior distribution (Beven and Freer, 2001; Montanari, 2005; Nott et al., 2012). That is, after calculating the likelihood of each parameter set (step 3), parameter sets with a likelihood below a subjective threshold are discarded and the posterior probability density (step 4) is calculated based on this reduced set. While this extra step has been used in the calibration of crop models (Mo and Beven, 2004; Pathak et al., 2012), in this study we follow previous studies that have not included such a reduction (Makowski et al., 2002; Marin et al., 2011; Wang et al., 2005).

The foundation of the GLUE methodology is the idea of equifinality (Beven and Freer 2001; Beven, 2006). That is, that there may be many models or parameter sets that describe the observed data equally well, but exist in different regions of the parameter space. GLUE is often referred to as a pseudo-Bayesian approach as the likelihood function $L(\boldsymbol{\theta}_i|\mathbf{Y})$ can be defined subjectively by the modeller (Beven and Binley, 1992). This allows for a likelihood function that does not necessarily reflect the parameter structure or model error variance (Stedinger et al., 2008) but does attempt to reflect all sources of error in the modelling process (Vrugt et al., 2009). The choice of likelihood function and how likelihood functions are combined can affect the accuracy of the estimated

parameter values and their posterior distribution (He et al., 2010). Similarly, choice of likelihood and likelihood combinations can result in output prediction credible intervals that do not include the specified proportion of observations (i.e. a 95% credible interval may not contain 95% of observations as would be expected of a more formal statistical approach) (Montanari, 2009; Stedinger et al., 2008).

2.2.2. Markov Chain Monte Carlo

The MCMC approach approximates the posterior distribution by simulating a random walk which converges to the posterior probability distribution $P(\boldsymbol{\theta}|\mathbf{Y})$ that describes our updated beliefs of probable parameter values. A Markov chain is formed as each sequential approximate realization from the distribution (from iteration t = 1, …, N) is based on the previous draw. The MCMC process can be summarized as:

1. Define the Markov Chain as $\boldsymbol{\Theta} = [\boldsymbol{\theta}_0, \ldots, \boldsymbol{\theta}_{t-1}, \boldsymbol{\theta}_t, \ldots, \boldsymbol{\theta}_N]$ for a chain of length N and select a starting parameter set $\boldsymbol{\theta}_0$.
2. Generate a candidate parameter set $\boldsymbol{\theta}_*$ based on a symmetric transition kernel $P(\boldsymbol{\theta}_*|\, \boldsymbol{\theta}_{t-1})$ such that: $P(\boldsymbol{\theta}_*|\, \boldsymbol{\theta}_{t-1}) = P(\boldsymbol{\theta}_{t-1}|\, \boldsymbol{\theta}_*)$.
3. Calculate the acceptance criteria (r) based on the ratio of densities (Metropolis et al., 1953) as:
$$r = \frac{P(\boldsymbol{\theta}_*|\mathbf{Y})}{P(\boldsymbol{\theta}_{t-1}|\mathbf{Y})} = \frac{P(\boldsymbol{\theta}_*)P(\mathbf{Y}|\, \boldsymbol{\theta}_*)}{P(\boldsymbol{\theta}_{t-1})P(\mathbf{Y}|\, \boldsymbol{\theta}_{t-1})} \quad (6)$$
4. If $\begin{cases} r > U[0,1] & \boldsymbol{\theta}_t = \boldsymbol{\theta}_* \\ r < U[0,1] & \boldsymbol{\theta}_t = \boldsymbol{\theta}_{t-1} \end{cases}$ where U[0,1] is a random draw from the uniform distribution between 0 and 1.

The Metropolis algorithm requires a symmetric transition kernel but can be generalized to the Metropolis-Hastings algorithm to allow for asymmetric transition kernels (Gelman et al., 1997; Hastings, 1970). In the above process, 2 to 4 are repeated for N iterations. A period of M iterations (referred to as the burn-in period) is required for the chain to stabilize. Therefore the remaining $n = N - M$ iterations represent draws from the posterior distribution. The total number of iterations required (N), the length of the burn-in period (M) and the transition kernel must be defined. The number of iterations must be long enough for the chain to converge. Convergence can be monitored by running several (J) consecutive chains from disparate starting parameter sets and computing the potential scale reduction statistic ($\sqrt{\hat{R}}$) (Gelman et al., 1997) as:

$$\sqrt{\hat{R}} = \sqrt{\frac{\frac{n-1}{n}W + \frac{1}{n}B}{W}} \quad (7)$$

for each parameter of interest. For a single parameter (θ), B is the between chain variance defined as:

$$B = \frac{n}{J-1}\sum_{j=1}^{J}(\overline{\theta}_{\cdot j} - \overline{\theta}_{\cdot\cdot})^2 \tag{7.1}$$

where

$$\overline{\theta}_{\cdot j} = \frac{1}{n}\sum_{i=1}^{n}\theta_{ij}$$

$$\overline{\theta}_{\cdot\cdot} = \frac{1}{J}\sum_{j=1}^{J}\overline{\theta}_{\cdot j}$$

and W is the within chain variance defined as:

$$W = \frac{1}{J}\sum_{j=1}^{J}s_j^2 \tag{7.2}$$

where

$$s_j^2 = \frac{1}{n-1}\sum_{i=1}^{n}(\theta_{ij} - \overline{\theta}_{\cdot j})^2.$$

Chains approach convergence as the variance between chains (Eq. (7.1)) approaches the variance within chains (Eq. (7.2)) based on the final n iterations from each of J chains. Gelman et al. (1997) recommend $\sqrt{\hat{R}}$ values of less than 1.2 are acceptable for most applications. For multi-dimensional problems convergence should be reached for all parameters. The posterior parameter distributions for each parameter can then be described using the posterior mean:

$$\hat{\mu}_{post} = \frac{1}{J \times n}\sum_{j=1}^{J}\sum_{i=1}^{n}\theta_{ij} \tag{8}$$

and variance:

$$\hat{\sigma}_{post}^2 = \frac{1}{(J \times n) - 1}\sum_{j=1}^{J}\sum_{i=1}^{n}(\theta_{ij} - \hat{\mu}_{post})^2 \tag{9}$$

of the n-by-J draws.

In contrast to the GLUE approach, the MCMC approach to parameter estimation requires a formal likelihood function and has a formal statistical foundation (Makowski et al., 2002). The advantage of such a formal approach is that the assumption of model errors is stated explicitly and can be verified (Dumont et al., 2014). The disadvantage of this is that these assumptions can be violated in crop models (Dumont et al., 2014; Wallach et al., 2011). Previous studies concluded that more formal MCMC approaches can be more efficient than the GLUE approach (Jones et al., 2011; Vrugt et al., 2009).

3. Materials and methods

Two approaches were taken to evaluate the use of GLUE and MCMC as calibration tools for variety parameters in APSIM-Sugar. A theoretical and real world evaluation were performed. In the theoretical evaluation variety parameters were pre-defined for two varieties (V001 and V002). Glue and MCMC were then used to calibrate APSIM-Sugar using modified simulation outputs of green biomass and sucrose yield as 'observed' yields. This evaluated the ability of GLUE and MCMC calibrations to approach known parameter values. In the real world evaluation GLUE and MCMC were used to calibrate APSIM-Sugar for Australian variety Q117 based on observed green biomass and sucrose yield. The GLUE and MCMC estimated parameter values were validated using independent field trial data. This evaluated the ability of GLUE and MCMC calibrations to simulate observed yields. Uncertainty in model outputs due to uncertainty in the estimated parameter values was also analysed. Finally, GLUE and MCMC evaluation results are compared quantitatively and qualitatively. A flow diagram of the methods used in this analysis is provided in Fig. 2.

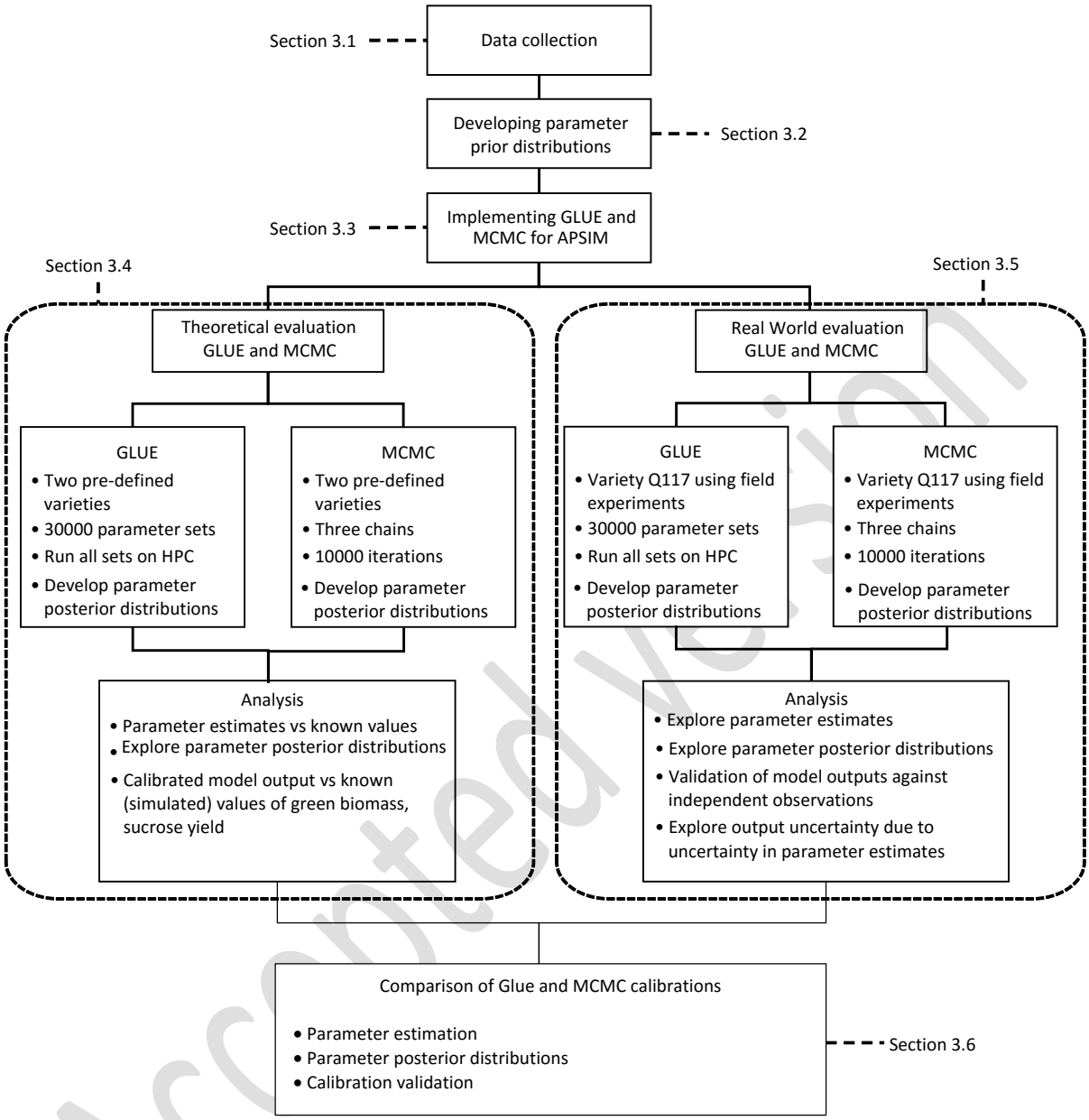

Fig. 2. Diagrammatic representation of the materials and methods described in this paper. Labels identify the relevant sections in the paper. Section 3.1 to 3.3 describe general data collection and implementation of GLUE and MCMC for APSIM. Sections 3.4 and 3.5 detail the two approaches used to evaluate GLUE and MCMC. A brief comparison of GLUE and MCMC calibrations is performed.

3.1. Data

Calibrations were performed using data from three field trials for variety Q117 run in Australia (Table 2). Each of the field experiments used for calibration were used in the development of APSIM-Sugar (Keating et al., 1999) and had the soil, climate and management data (e.g. irrigation and fertilizer applications) required to run APSIM simulations. Although other data sets exist, only APSIM-Sugar data sets for Q117 were considered in this study as the calibration for Q117 for APSIM was previously validated (Keating et al., 1999) providing a good comparison for the results of the two statistical calibration techniques used in this study.

The three field experiments were conducted at Harwood, New South Wales and Ingham, Queensland. The "Harwood(1993-94)" (Hughes et al., 1995) and "Ingham(1992-93)" (Robertson et al., 1996) experiments were single 12 month plant crops well irrigated and fertilized to avoid stress. The "Ingham(1992-94)" experiment (Muchow et al., 1996) had three nitrogen treatments. Crops in each treatment were harvested in 1993 and then ratooned (allowed to regrow) before being harvested again in 1994. The Harwood field experiment was characterised by cooler temperatures and lower levels of rainfall compared to the warmer and wetter Ingham field experiments (Table 2). Green biomass (g m^{-2}) and sucrose yield (g m^{-2}) were sampled throughout the growing cycle for each treatment in the field experiments. From the three field experiments (Harwood(1993-94), Ingham(1992-93), Ingham(1992-94)); eight yield samples were used as part of the calibration procedure (Harwood(1993-94), Ingham(1992-93), Ingham(1992-94) [low nitrogen; plant crop], Ingham(1992-94) [low nitrogen; ratoon crop], Ingham(1992-94) [med nitrogen; plant crop], Ingham(1992-94) [med nitrogen; ratoon crop], Ingham(1992-94) [high nitrogen; plant crop], Ingham(1992-94) [high nitrogen; ratoon crop]). The sample yields used for calibration were considered representative of harvest yields and the dates of these samples are recorded as 'Harvest Date' in Table 2.

Table 2. Details of field experiments used in both GLUE and MCMC calibrations of APSIM-Sugar. Average climate data were calculated as the average daily value for the period from planting to harvest. Total rainfall was calculated from plant to harvest or ratoon to harvest for ratooned crops.

	Harwood(1993-94)	Ingham(1992-93)	Ingham(1992-94)
Reference	(Hughes et al., 1995) [Dataset 1, Keating et al. (1999)]	(Robertson et al., 1996) [Dataset 2, Keating et al. (1999)]	(Muchow et al., 1996) [Dataset 16, Keating et al. (1999)]
Soil	PAWC = 180.0 mm	PAWC = 216.0 mm	PAWC = 216.0 mm
Avg. Min Temp.	18.11 ^{O}C	18.73 ^{O}C	18.80 ^{O}C
Avg. Max Temp.	24.67 ^{O}C	28.18 ^{O}C	28.39 ^{O}C
Avg. Daily Rainfall	1.86 mm	3.86 mm	4.44 mm
Total Rainfall	834.75 mm	1762.2 mm	Plant = 1631.4 mm; Ratoon = 2025.0 mm
Plant Date	24-Sep-1993	23-Jul-1992	23-Jul-1992
Harvest Date	15-Dec-1994	21-Oct-1993	23-Oct-1994 (ratoon 18-Aug-1993)
Experimental Design	Plant crop; Fumigated	Plant crop; Fumigated	Plant crop and First ratoon; Fumigated Treatments: low, med and high Nitrogen

Two additional field experiments were used as independent validations of GLUE and MCMC calibrations in the real world evaluation. These field experiments were conducted at Grafton, New South Wales ("Grafton(1994-95)") and Ayr, Queensland ("Ayr(1992-94)"). These two field experiments were also used in the design of APSIM-Sugar (Keating et al., 1999). The Grafton(1994-95) field experiment was characterised by cooler conditions and higher rainfall than the Ayr(1992-94) experiment (Table 3). Both field experiments used for validation were well irrigated and fertilized to remove water and nutrient stresses. Green biomass, sucrose yield and LAI measurements for the validation experiments were taken 7 times during the Grafton(1994-95) experiment and 24 times over the plant and first ratoon for the Ayr(1992-94) experiment.

Table 3. Field experiments used for validation of GLUE and MCMC calibrations of APSIM-Sugar. Average climate data is calculated as the daily average for the period from planting to harvest. Total rainfall was calculated from plant to harvest or ratoon to harvest for ratooned crops.

	Grafton(1994-95)	Ayr(1992-94)
Reference	[Dataset 11, Keating et al. (1999)]	[Dataset 6, Keating et al. (1999)]
Soil	PAWC = 180.0 mm	PAWC = 197.0 mm
Avg. Min Temp.	17.80 ^{O}C	21.05 ^{O}C
Avg. Max Temp.	25.96 ^{O}C	29.13 ^{O}C
Avg. Daily Rainfall	2.45 mm	1.01 mm
Total Rainfall	707.30	Plant = 435.8 mm Ratoon = 341.1 mm
Plant Date	28-Sep-1994	31-Jul-1992
Harvest Date	25-Jul-1995	06-Sep-1994 (ratoon 29-Jul-1993)
Experimental Design	Plant crop; Fumigated	Plant crop and First ratoon; Non-fumigated

3.2. Parameter prior distributions

Sexton and Everingham (2014) performed a global sensitivity analysis of the APSIM-Sugar model in order to identify variety parameters to which simulated biomass and sugar yields were sensitive. In their study, Sexton and Everingham (2014) tested 14 APSIM-Sugar parameters that represent biophysical traits that may help distinguish between varieties (Table 1). Of the 14 parameters tested, Sexton and Everingham found that four had little to no influence on simulated biomass or sugar yields under well irrigated or water stressed conditions. These included two phenology parameters (*tt_begcane_to_flowering* and *tt_flowering_to_crop_end*); a sucrose partitioning parameter (*min_sstem_sucrose_redn*) and the leaf area profile parameter (*tillerf_leaf_size*). The remaining 10 parameters from the results of Sexton and Everingham (2014), were chosen for this calibration study as parameters that may help distinguish between varieties and have been shown to influence simulated biomass and sugar yields. Table 4 lists the 10 parameters and their assumed prior distributions.

Parameter prior distributions generally followed those described by Sexton and Everingham (2014) and were based on APSIM-Sugar documentation, available literature and previous experimental data. The ranges of uniform prior distributions for parameters *cane_fraction*, *sucrose_fraction_stalk*, *sucrose_delay*, *min_sstem_sucrose* and *tt_emerg_to_begcane*, were based on varieties defined in APSIM-Sugar documentation as little literature was available on the genetic variability of these parameters.

The range for the *leaf_size* parameter was selected to capture previously reported leaf areas for a range of commercial varieties (Inman-Bamber, 2013). For simplicity and consistency among varieties, the *leaf_size* parameter was modified relative to the default value for variety Q117. For example APSIM-Sugar values for Q117 *leaf_size* are 1500 mm^2 for *leaf_size_no* 1 and 55000 mm^2 for *leaf_size_no* 14 and 20 (Table 2). The minimum value of *leaf_size* (maximum) reported in Table 2 (20000 mm^2) represents a *leaf_size* value of 20000 for *leaf_size_no* 14 and 20 and a *leaf_size* value of 545.45 mm^2 for *leaf_size_no* 1 (1500 mm^2/55000 mm^2 *20000 mm^2).

Transpiration efficiency (*transp_eff_cf*) for C_4 crops such as sugarcane is generally accepted to be around 0.009 g kPa g^{-1} to 0.010 g kPa g^{-1} (Sinclair, 2012). In this study the range of values for *transp_eff_cf* was extended lower and higher to represent genetic variability and response to water stressed conditions identified in recent research for sugarcane (Jackson et al., 2014). In a greenhouse study of 20 sugarcane varieties, Jackson et al. (2014) found that transpiration efficiency for sugarcane varieties could be higher under water stressed conditions. Data from Jackson et al. (2014) were used

to identify a maximum and minimum range in *transp_eff_cf* under both well irrigated and water stressed conditions relative to variety Q117. Values based on water stressed conditions were included in order to capture the possibility of higher *transp_eff_cf* values under brief periods of water stress that can occur even in well irrigated trials.

The range for parameters *green_leaf_no*, *rue* and root conductance (k_L) were based on data reported for a breeding field trial conducted at Home Hill, Australia (Basnayake et al., 2012). The mean and standard deviation of observed numbers of green leaves across 89 varieties grown under irrigated conditions, was used to describe a prior distribution for *green_leaf_no*. Parameters *rue* and k_L were adjusted to represent observed variation of conductance under well irrigated conditions (*rue*) and the difference between well irrigated and stressed conditions (k_L). Adjustment factors for each of 89 varieties were calculated relative to variety Q117 (Sexton, 2015). The distributions of *rue* and k_L adjustment factors calculated for 89 varieties were found to be approximately normally distributed.

Table 4. Prior distributions assumed for statistical calibration of variety parameters. Uniform Distributions were described as maximum and minimum allowed values. Normal distributions were described using mean and standard deviation (SD).

Parameter	Units	Distribution	Mean	SD	min	max
leaf_size (maximum)	mm^2	Uniform			20000	70000
cane_fraction	$g\ g^{-1}$	Uniform			0.65	0.80
sucrose_fraction_stalk	$g\ g^{-1}$	Uniform			0.4	0.7
sucrose_delay	$g\ m^2$	Uniform			0	600
min_sstem_sucrose	$g\ m^2$	Uniform			400	1500
tt_emerg_to_begcane	^oC day	Uniform			1200	2000
green_leaf_no	leaves	Normal	11.73	1.58	-	-
transp_eff_cf	$g\ kPa\ g^{-1}$	Uniform			0.0060	0.0140
rue adjustment factor[1]	nil	Normal	0.90	0.084	-	-
k_L adjustment factor[2]	nil	Normal	1.248	0.425	-	-

[1]an *rue* adjustment factor of 1 relates to the APSIM default *rue* of 1.8 g MJ^{-1} (plant) and 1.65 g MJ^{-1} (ratoon)
[2]a k_L adjustment factor of 1 relates to the default k_L based on APSIM Q117 and soil data.

3.3. Implementing GLUE and MCMC

Here we describe the computational process as well as the formulation of the transition kernel, likelihood and acceptance criteria. GLUE and MCMC were implemented within the R statistical program (R Core Team, 2013). Likelihoods and prior distributions were generated using the basic statistical package in R. Multivariate normal distributions were generated using the MASS package (Venables and Ripley, 2002). The APSIMBatch (Zheng, 2012) Package 'generateSim' function was modified to generate the simulation files needed to run APSIM-Sugar. The functions written in the R environment to run GLUE and MCMC were compiled into an R package called Bayes4APSIM.

Both GLUE and MCMC calibrations were based on the field experiments described in Table 2. Model errors were assumed normally distributed and a normal likelihood function based on (Eq. (2)) was used. The Likelihood value for each parameter set was calculated based on observed green biomass and sucrose yield at harvest. Error variance was estimated for harvest green biomass ($\sigma^2_{biomass}$) and sucrose yield ($\sigma^2_{sucrose}$). These variances were assumed known with no uncertainty. To avoid problems with computer precision within R likelihoods were calculated as log-likelihoods.

3.3.1. Implementing GLUE

Following earlier GLUE calibration studies such as He et al. (2010), prior distributions were used to generate 30000 (Q) parameter sets. Previous GLUE calibration studies such as Marin et al. (2011) and Jones et al. (2011) have used Q values of as low as 6000 as a trade-off between statistical power and time efficiency. However, a larger Q can help improve parameter posterior distributions (He et al., 2010). Parameter sets were then run simultaneously on a high performance cluster. Harvest green biomass and sucrose yield were simulated to match the scenarios of the calibration field experiments (Table 2). The equation for GLUE posterior probabilities (Eq. (3)) was modified to use log-likelihoods and posterior probabilities were calculated as

$$p(\boldsymbol{\theta}_i) = \frac{L(\boldsymbol{\theta}_i|\mathbf{Y})}{\sum_{i=1}^{Q} L(\boldsymbol{\theta}_i|\mathbf{Y})} = \frac{\exp(\ln[L(\boldsymbol{\theta}_i|\mathbf{Y})])}{\sum_{i=1}^{Q} \exp(\ln[L(\boldsymbol{\theta}_i|\mathbf{Y})])}. \quad (10)$$

The posterior distribution for each parameter was defined by the posterior mean (Eq. (4)) and variance (Eq. (5)). A diagrammatic representation of the GLUE process is presented in Fig. 3.

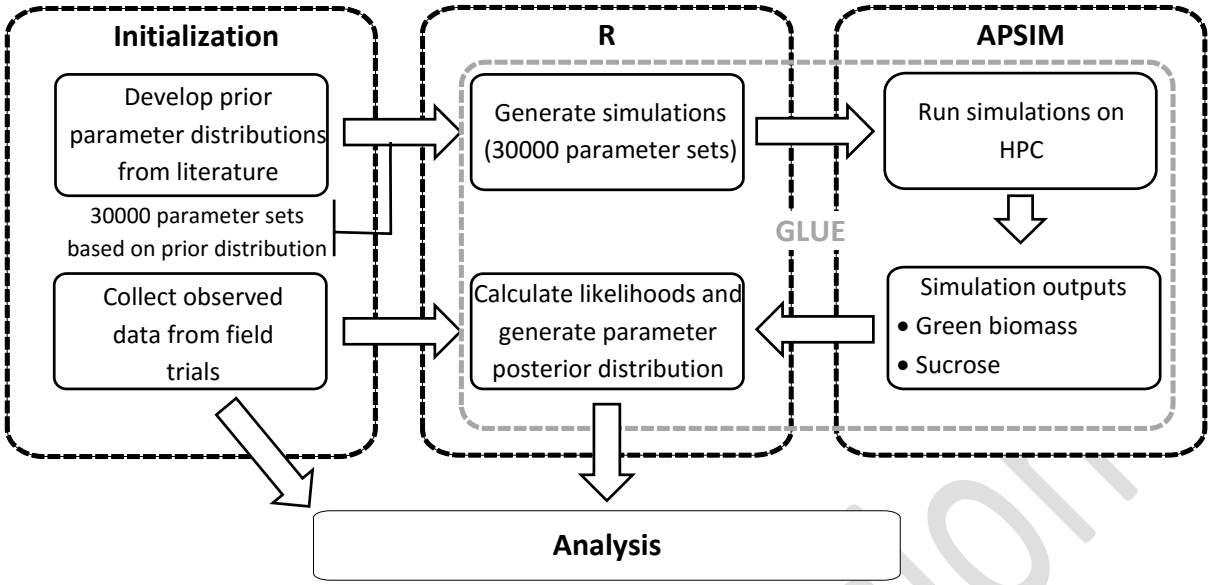

Fig. 3. Diagrammatic representation of the application of GLUE to APSIM used in this paper. Prior parameter distributions were used to generate 30000 parameter sets. These were run on a HPC cluster for efficiency. Likelihood values were calculated for all 30000 sets and final posterior parameter distributions were developed for each parameter.

3.3.2. Implementing MCMC

For implementation of MCMC it is simplest to consider a single chain. For each chain, a candidate parameter set ($\boldsymbol{\theta}_*$) was generated from the multivariate normal distribution ($\boldsymbol{\theta}_*|\boldsymbol{\theta}_{t-1} \sim N(\boldsymbol{\theta}_{t-1}, k\boldsymbol{\Sigma})$). Parameters were assumed independent such that the covariance matrix ($\boldsymbol{\Sigma}$) was diagonal with variances derived from the prior distributions. The covariance matrix was modified by k = 0.01 as this was found – using a grid based optimization search – to produce an acceptance rate between 30 and 40 percent (Gelman et al., 1997). As log-likelihoods were calculated, the acceptance criterion (Eq. (6)) was calculated using log-likelihoods as

$$r = \frac{P(\boldsymbol{\theta}_*)L(\boldsymbol{\theta}_*|\mathbf{Y})}{P(\boldsymbol{\theta}_{t-1})L(\boldsymbol{\theta}_{t-1}|\mathbf{Y})} = \exp((\ln[P(\boldsymbol{\theta}_*)] + \ln[L(\boldsymbol{\theta}_*|\mathbf{Y})]) - (\ln[P(\boldsymbol{\theta}_{t-1})] - \ln[L(\boldsymbol{\theta}_{t-1}|\mathbf{Y})])) . \quad (11)$$

Chains were set to a length of 10000 iterations (N). Three simultaneous chains (J = 3) were run to monitor convergence of the posterior distribution giving a total number of simulations equal to that of the GLUE implementation (30000). Following Eq. (8) and Eq. (9), posterior means and variances were calculated as the final 5000 (M) iterations of each chain. Posterior distributions were generated on samples of 15000 (n) parameter sets. A diagrammatic representation of the MCMC process is presented in Fig. 4.

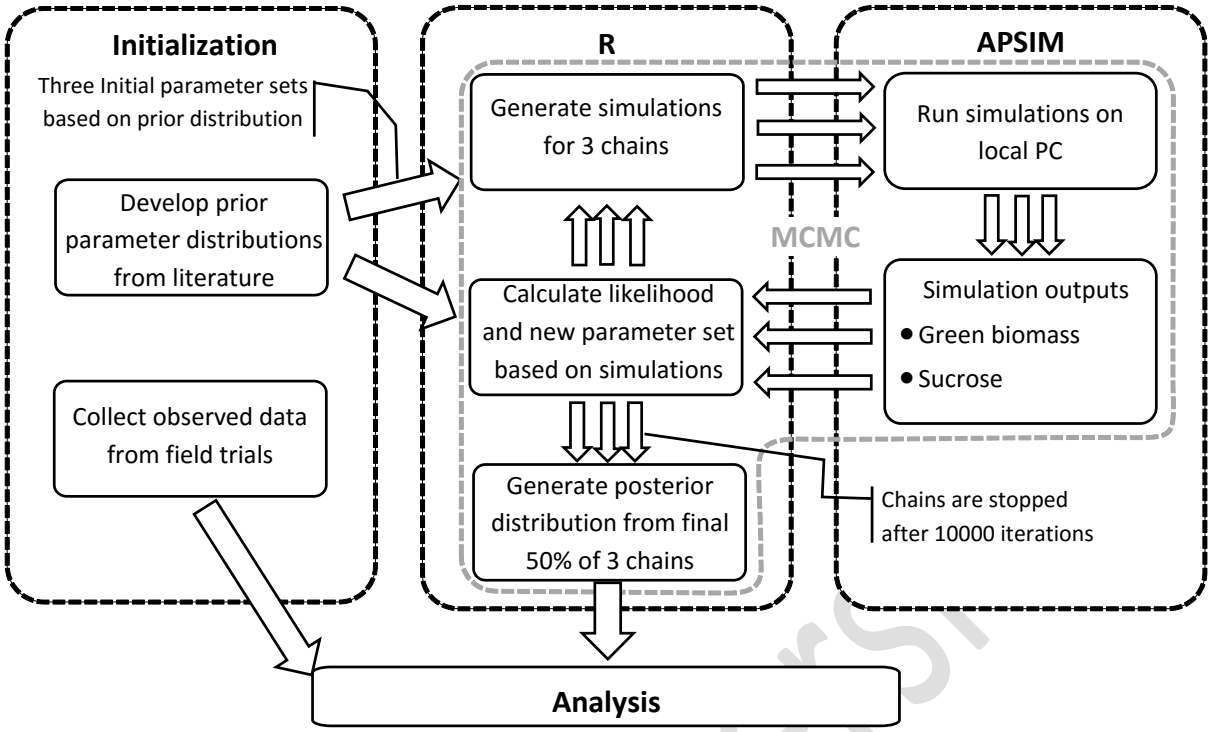

Fig. 4. Diagrammatic representation of the application of MCMC to APSIM used in this paper. Prior parameter distributions were used to generate three initial parameter sets. These initial parameter sets were used to generate three chains from 10000 iterations. Convergence of the three chains to a single posterior distribution was monitored and the final posterior parameter distribution was developed from the final 50% of all chains.

3.4. Theoretical evaluation of GLUE and MCMC

Two varieties referred to as V001 and V002 were pre-defined using 10 parameters (Table 5). The leaf area parameter (*leaf_size*) for V001 was made larger than typically seen in Australian field experiments while parameters *cane_fraction*, *min_sstem_sucrose* and *sucrose_delay* were set to lower values to increase sugar and cane yields. Sucrose fraction was raised higher to increase sucrose yields. Variety V002 represented a smaller plant with a lower sucrose yield. Parameters representing the transpiration efficiency coefficient (*transp_eff_cf*), RUE (*rue*) and root conductance (k_L) were higher in V002 than V001. An increased *transp_eff_cf* and increased k_L can improve performance in water stressed conditions (Inman-Bamber et al., 2012).

Table 5. Parameter values for two pre-defined varieties V001 and V002. The parameter *leaf_size* is expressed as leaf size for each leaf number modelled on variety Q117. Parameters *rue* and k_L were modified relative to the default simulation values used in Keating et al. (1999) for variety Q117 and are represented here as a unit-less fraction of the original value. All other parameters appear as the values required by APSIM-Sugar.

Parameter	Units	V001			V002		
leaf_size	mm^2	1718.182	63000	63000	954.546	35000	35000
leaf_size_no	leaf	1	14	20	1	14	20
cane_fraction	$g\ g^{-1}$			0.66			0.74
sucrose_fraction_stalk	$g\ g^{-1}$			0.65			0.46
sucrose_delay	$g\ m^{-2}$			100			400
min_sstem_sucrose	$g\ m^{-2}$			550			1000
tt_emerg_to_begcane	°C day			1500			1860
green_leaf_no	leaves			14			11
transp_eff_cf	kPa			0.0078			0.0110
rue adjustment factor	nil			0.89[1]			1.02[1]
k_L adjustment factor	nil			0.60[2]			1.20[2]

[1] adjustment factors resulted in an actual *rue* of 1.602 (plant) and 1.469 (ratoon) for V001 and 1.836 (plant) and 1.02 (ratoon) for V002

[2] adjustment factors resulted in an actual k_L of 0.096 and 0.192 for V001 and V002 respectively

APSIM was used to produce simulated yields for the pre-defined varieties V001 and V002 for calibration field experiments (Table 2). To represent realistic uncertainty in data collection, a normal error was added to simulated yields to generate 'observed' yields for varieties V001 and V002. Error was added as a random number drawn from a normal distribution with zero mean and standard deviation (SD) of 5% of simulated yield. This represented the increasing uncertainty in field measurements as values increased and allowed variance to be smaller early in the season or where limiting factors may have reduced yields and yield variances. For the theoretical evaluation of both GLUE and MCMC calibrations error variances for biomass and sucrose yield were calculated for each observation based on a standard deviation of 5% of the 'observed' yields. The likelihood for the simulation studies was then calculated as

$$L(\boldsymbol{\theta}_i|\mathbf{Y}) = \prod_{o\,=\,1\mathrm{s}\,=\,1}^{OS} \frac{1}{\sqrt{2\pi\sigma_{os}{}^2}} \exp\left(\frac{[y_{os} - \hat{y}_{os}(\boldsymbol{\theta}_i)]^2}{2\sigma_{os}{}^2}\right). \quad (12)$$

Here S is the number of productivity measures considered (S = 2), specifically green biomass and sucrose yield and O is the number of observations of each measure (O = 8), specifically harvest yields from experiments Harwood(1993-94), Ingham(1992-93), Ingham(1992-94) (low, med and high nitrogen; plant and first ratoon).

Both GLUE and MCMC were used to estimate the 10 variety parameters for V001 and V002. Posterior distributions for each parameter were recorded. The absolute relative error

$$ARE = \frac{|\overline{\theta}_{post} - \theta_o|}{\theta_o} \times 100 \quad (13)$$

was calculated between the posterior mean and known value for each parameter of the defined varieties (Table 5). To further explore the accuracy of the posterior distribution, the percentile of the known value was estimated from the empirical posterior distribution.

The fit between calibrated simulation outputs and observed data was assessed using normalized root mean square error (NRMSE), Willmott's agreement index (D) (Marin et al., 2011; Willmott, 1982) and regression r^2. The slope and intercept of the linear regression between calibrated simulation outputs and observations were also recorded and scatter plots of observed and simulated results were produced. GLUE and MCMC were applied with the goal of simulating green biomass (g m^{-2}) and sucrose yield (g m^{-2}). Root mean square error was normalized and expressed as a percentage by dividing by the range of observed values. This allowed comparison between model outputs.

Bennet et al. (2013) highlight some of the key elements of evaluating model performance including visual analysis to gain an overview of model performance and selection of basic performance criteria. The validation statistics used in this paper have been widely used in the assessment of crop models and other process based models such as environmental and hydrological models (Bennett et al., 2013).Visually plotting observed data against simulation outputs allows for a quick qualitative analysis. The slope and intercept of the linear regression can further be used to identify model bias. The r^2 statistic is commonly reported as a quantitative measure of model efficiency and is included to aid communication of results to a wider audience, while Willmott's agreement index can be used as an alternative to r^2 designed to account for differences in the distributions of model results and observations. The root mean square error statistics provide a quantitative assessment of the magnitude of bias in the model. By providing a range of performance criteria it is possible to be more confident in our results as weaknesses in one criteria may be compensated by strengths in another (Bennet et al., 2013).

3.5. Real world evaluation of GLUE and MCMC for Q117

Observed harvest yields for Q117 for the calibration experiments (Table 2) were used to estimate the 10 variety parameters for Q117. As with the theoretical evaluation, posterior distributions of the parameter values were generated. Error variance for biomass and sucrose yields were estimated from a well calibrated example using default parameter values (Pathak et al., 2012; Wang et al., 2005). The error variance was estimated from simulations of the available field experiments such that

$$L(\boldsymbol{\theta}_i|\mathbf{Y}) = \prod_{o=1s=1}^{OS} \frac{1}{\sqrt{2\pi\sigma_s^2}} \exp\left(\frac{[y_{os} - \hat{y}_{os}(\boldsymbol{\theta}_i)]^2}{2\sigma_s^2}\right). \quad (14)$$

Here S is the number of productivity measures considered (S = 2), specifically green biomass and sucrose yield and O is the number of observations of each measure (O = 8), specifically harvest yields from experiments Harwood(1993-94), Ingham(1992-93), Ingham(1992-94) (low, med and high nitrogen; plant and first ratoon).

GLUE and MCMC parameter estimates were used to simulate sugarcane yields for two independent validation field trials (Table 3). As several measurements were made throughout the validation trials, the observed growth curves were simulated. Regression statistics (slope, intercept and r^2), D and NRMSE were calculated for each experiment using all available data and data at harvest only. Uncertainty in model outputs for green biomass, sucrose yield and LAI were investigated by producing 95% credible intervals from 1000 random draws of the parameter posterior distributions.

3.6. Comparison of GLUE and MCMC calibrations

The quantitative results of the GLUE and MCMC evaluations were compared and contrasted to explore the advantages and disadvantages of each calibration technique. The quantitative comparisons were put in perspective of the qualitative differences in philosophy between the two approaches.

4. Results and discussion

4.1. Theoretical evaluation of GLUE and MCMC

Both GLUE and MCMC were able to accurately estimate known variety parameters for two pre-defined varieties based on ARE (Table 6). Standard deviations for each parameter posterior distributions were recorded to identify uncertainty in the parameter estimates. For both pre-defined varieties 6/10 GLUE parameter estimates and 7/10 MCMC parameter estimates were within 10% of the defined values. This is a somewhat better rate than a previous study by Makowski et al. (2002). Of the 20 parameters calibrated in Makowski et al. (2002) using 5 data samples, 6/20 GLUE parameter estimates were within 10% of defined values while 7/20 MCMC parameter estimates were within 10% of defined values. In contrast He et al. (2010) were able to use GLUE to estimate all 9 parameters within 10% of defined values using a specific likelihood combination method.

Importantly estimated parameter values differed accurately between variety V001 and V002. For example *transp_eff_cf*, *rue* and k_L were particularly well estimated for V001 and V002 using the MCMC calibration resulting in realistic differences in these parameters between the two varieties. These parameters in particular may play an important role in improving differences in response to water stress between sugarcane varieties (Inman-Bamber et al., 2012) but are difficult to measure directly. Therefore it is a critical result that these parameters can be accurately estimated using only harvest data.

Based on ARE, the parameter *sucrose_delay* was particularly poorly estimated for V001 using both GLUE and MCMC. This may have been a consequence of weak influence of the parameter on sucrose yield. Sexton and Everingham (2014) showed that *sucrose_delay* was weakly influential for sucrose yield under both irrigated and water stressed conditions. Particularly *sucrose_delay* had less influence on sucrose yields at low values (Sexton and Everingham (2014)). This may explain why *sucrose_delay* estimates for V002 were closer than for V001. Similarly *leaf_size* and *transp_eff_cf* estimates were closer when known values were low and high respectively, reflecting regions where these parameter values had a greater influence on yields.

Table 6. GLUE and MCMC evaluation of estimated parameter values for two pre-defined varieties (V001 and V002). Mean and standard deviation of the empirical posterior distributions are recorded. Posterior percentile is the estimated percentile of the known parameter value in the posterior distribution. Posterior percentile values closer to the 50th percentile are preferred. The ARE measures the absolute difference between the posterior mean and the defined value as a percentage. For MCMC all chains were assumed to converge for all parameters as $\sqrt{\hat{R}}$ was less than 1.2 for all parameters (see supplementary material 2 Fig. S2.1 to Fig. S2.4).

Variety	Parameter	Value	GLUE			MCMC		
			Mean (SD)	Posterior Percentile	ARE (%)	Mean (SD)	Posterior Percentile	ARE (%)
V001	*leaf_size*	63000	56065.03(3181.31)	91[3]	11.0	56630.50(7956.9)	75	10.1
	cane_fraction	0.66	0.665(0.021)	65	0.8	0.671(0.0159)	30	1.7
	sucrose_fraction_stalk	0.65	0.666(0.014)	16[3]	2.5	0.621(0.0413)	75	4.5
	sucrose_delay	100	373.96(94.91)	2[2]	274.0	305.56(154.10)	11[3]	205.6
	min_sstem_sucrose	550	661.98(100.54)	7[3]	20.4	700.13(161.25)	20[3]	27.3
	tt_emerg_to_begcane	1500	1461.19(95.37)	54	2.6	1445.03(136.57)	69	3.7
	green_leaf_no	14	12.37(0.441)	-[1]	11.7	13.04(0.801)	89[3]	6.8
	transp_eff_cf	0.0078	0.0077(0.0003)	70	1.9	0.0078(0.0004)	55	0.4
	rue adjustment factor	0.89	0.942(0.012)	<1[2]	5.8	0.934(0.035)	10[3]	4.9
	k_L adjustment factor	0.60	0.548(0.039)	84[3]	8.6	0.577(0.085)	67	3.9
				Average ARE (%)	33.9		**Average ARE (%)**	26.9
V002	*leaf_size*	35000	27952.67(5592.94)	91[3]	20.1	34289.87(4928.9)	56	2.0
	cane_fraction	0.74	0.780(0.031)	18[3]	5.4	0.710(0.029)	82[3]	4.0
	sucrose_fraction_stalk	0.46	0.473(0.015)	15[3]	2.8	0.469(0.029)	38	2.1
	sucrose_delay	400	416.02(63.78)	55	4.0	507.99(71.02)	10[3]	27.0
	min_sstem_sucrose	1000	995.04(94.85)	70	0.5	799.52(232.20)	79[3]	20.1
	tt_emerg_to_begcane	1860	1876.21(73.87)	19[3]	0.9	1834.42(101.85)	55	1.4
	green_leaf_no	11	10.38(0.765)	83[3]	5.6	11.25(0.712)	38	2.2
	transp_eff_cf	0.011	0.0126(0.00083)	11[3]	15.0	0.0124(0.00102)	14[3]	12.6
	rue adjustment factor	1.02	1.14(0.072)	14[3]	12.3	1.014(0.046)	58	0.5
	k_L adjustment factor	1.20	1.32 (0.244)	16[3]	10.2	1.138(0.35)	53	5.2
				Average ARE (%)	7.7		**Average ARE (%)**	7.7

[1]The empirical posterior distribution did not capture the known value.

[2] The known value was not within a 95% credible interval of the posterior mean.

[3] The known value was not within a 50% credible interval of the posterior mean.

A closer consideration of the empirical posterior distribution for parameter values showed that MCMC posterior distributions were more likely to capture the defined value than GLUE posterior distributions (Table 6). The defined values for 4 parameter of V001 (*sucrose_delay*, *min_sstem_sucrose*, *green_leaf_no* and *rue*) were not within a 95% credible interval of their GLUE empirical posterior distribution. While all MCMC posterior distributions captured the defined values, several parameters were located in the tail of the distributions. For example, for variety V001 the defined value for *rue* was in the lower tail of the posterior distribution and that of *transp_eff_cf* was in the centre of the distribution, while for variety V002 the reverse was true. Plots of the parameter empirical posterior distributions highlighted the differences between GLUE and MCMC empirical posterior distributions (see supplementary material 1, Fig. S1.1 to Fig. S1.4).

Fig. 5 shows the linear regression between the APSIM-Sugar generated and calibrated APSIM-Sugar simulated green biomass and sucrose yield for V001 and V002. As would be expected from the accurate estimation of parameter values, both GLUE and MCMC calibrations accurately reproduced the APSIM-Sugar generated biomass and sucrose yield used in the model calibration. For both pre-defined varieties, both GLUE and MCMC calibrated simulations of green biomass were closer to the calibration data than sucrose yield based on NRMSE. This was likely a consequence of the poorer estimation of sucrose parameters. The real world evaluation provided a better indication of GLUE and MCMC to replicate observed yields using independent validation datasets.

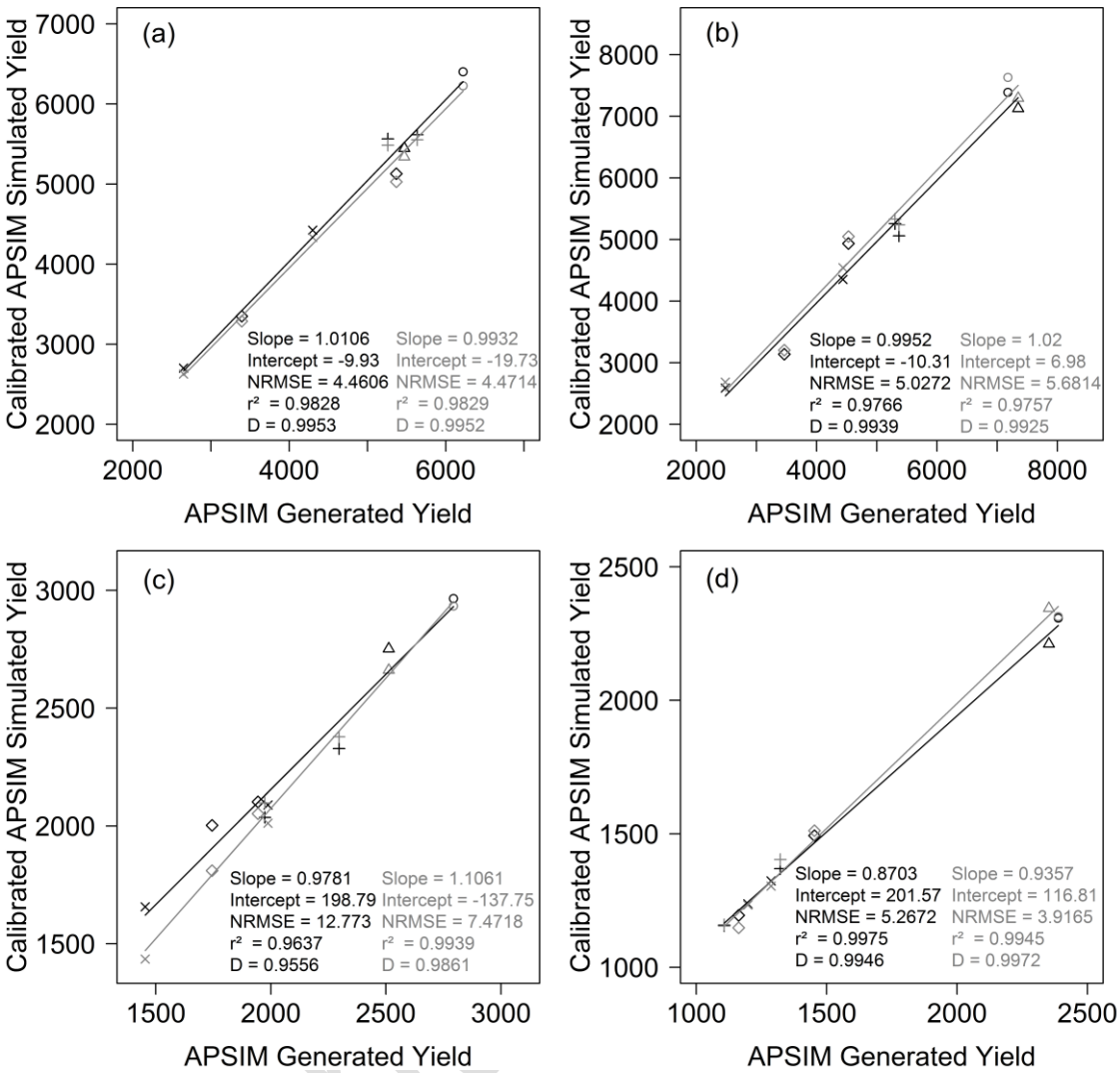

Fig. 5. Linear regression between APSIM generated yields used in the calibration and calibrated APSIM simulated values for (a, b) green biomass (g m^{-2}) and (c, d) and sucrose yield (g m^{-2}) for pre-defined varieties V001 (a, c) and V002 (b, d). Results from MCMC (black) and GLUE (grey) are plotted. Data represents simulations of V001 and V002 for Harwood(1993-94) (○), Ingham(1993-94) (△), Ingham(1992-94) low Nitrogen (×), Ingham(1992-94) med Nitrogen (◇) and Ingham(1992-94) high Nitrogen (+) trials.

4.2. Real world evaluation of GLUE and MCMC for variety Q117

Estimated parameter values using GLUE were closer to APSIM-Sugar default values for variety Q117 than using MCMC (Table 7). Compared to APSIM-Sugar default values GLUE estimated a higher *leaf_size* and low *min_sstem_sucrose* while MCMC estimated a lower *leaf_size* and *tt_emerg_to_begcane* and a higher *green_leaf_no*. Both GLUE and MCMC estimated values for *sucrose_delay* differed greatly from default values. As with the theoretical evaluation this was likely a consequence of the low influence of *sucrose_delay* on simulated yields. Similarly GLUE and MCMC estimates of k_L varied greatly with the expected default value. It is possible that a weak influence on yields complicated the estimation techniques. Both k_L and *sucrose_delay* had total effect indices of < 5% under irrigated or stressed conditions (Sexton and Everingham, 2014).

Table 7. GLUE and MCMC evaluation of estimated parameter values for Q117. Mean and standard deviation of the empirical posterior distributions are recorded. For MCMC all chains converge for all parameters ($\sqrt{\hat{R}}$ > 1.2) except *sucrose_delay* ($\sqrt{\hat{R}}$ > 1.2) (see supplementary material 2 Fig. S2.5 and Fig. S2.6).

Parameter	Q117	GLUE Mean (SD)	MCMC Mean (SD)
leaf_size	55000	61971.11(6704.96)	39921.81(3665.07)
cane_fraction	0.70	0.730(0.010)	0.670(0.0216)
sucrose_fraction_stalk	0.55	0.581(0.025)	0.553(0.0596)
sucrose_delay	0	133.82(84.36)	253.20(101.94)
min_sstem_sucrose	800	589.56(283.28)	1253.56(165.95)
tt_emerg_to_begcane	1900	1853.37(199.81)	1396.49(131.89)
green_leaf_no	13	13.97(0.343)	16.38(0.916)
transp_eff_cf	0.0080	0.0081(0.00051)	0.00890(0.00055)
rue adjustment factor	1	1.030(0.019)	1.033(0.0314)
k_L adjustment factor	1	0.367 (0.026)	0.266(0.0349)

In contrast to k_L, *rue* and *transp_eff_cf* parameters agreed closely with expected (default) values. Radiation use efficiency is a difficult value to measure and the *rue* parameter in APSIM greatly influences model outputs (Sexton and Everingham, 2014; Sexton et al., 2015). The agreement between *transp_eff_cf* and default values is especially encouraging given the wide prior distribution. The default value for *transp_eff_cf* was based on literature for other C_4 crops such as sorghum which have a value of 0.009 g kPa g^{-1} (Keating et al., 1999; Sinclair, 2012). Inman-Bamber and McGlinchey (2003) derived a *transp_eff_cf* of 0.0087 g kPa g^{-1} for variety Q138. The MCMC and GLUE estimates while based on only a small sample of sites, closely aligns with this recorded value and provides a measure of uncertainty not available with the original calibration.

The lower *leaf_size* value of the MCMC calibration may more closely represent field measured values than the higher GLUE estimate. While maximum leaf area has been measured as high as 70000

mm^2 under high input glasshouse studies (Inman-Bamber, 2013), under field conditions in Australia values of 38000 mm^2 were more common (Robertson et al., 1998). This would seem a more realistic calibrated value based on field measures of yield.

To further explore uncertainty in parameter posterior uncertainty, posterior probability distributions were analysed. Parameter uncertainty was lower for GLUE parameter posterior distributions based on standard deviations for all parameters except *leaf_size*, *min_sstem_sucrose* and *tt_emerg_to_begcane* (Table 7). However, the empirical posterior distribution did not always follow a normal distribution. Fig. 6 shows, as an example, the posterior distribution of *tt_emerge_to_begcane* for GLUE and MCMC calibrations. The GLUE empirical distribution of *tt_emerg_to_begcane* had a distinct major and minor mode while the MCMC empirical distribution of *tt_emerge_to_begcane* was biased towards lower values.

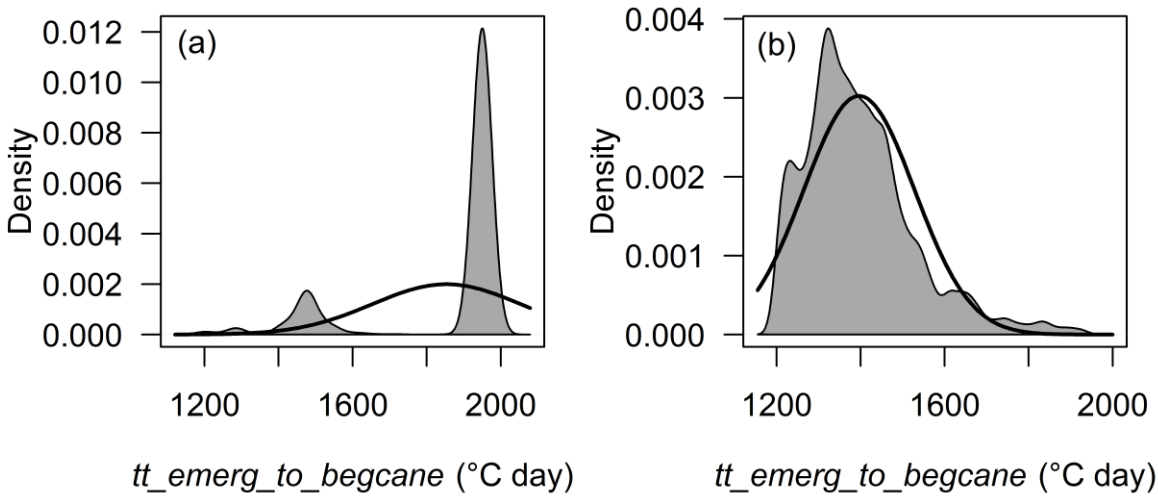

Fig. 6. Empirical posterior probability density functions (grey) for parameter *tt_emerge_to_begcane* and approximate normal distribution (—) based on the mean and standard deviations reported in Table 7 for (a) GLUE and (b) MCMC. The empirical distributions for all other parameters are available as supplementary material (see supplementary material 1, Fig. S1.5 and Fig. S1.6).

Based on all available samples from two validation experiments both GLUE and MCMC calibrations produced lower NRMSE and higher r^2 values for green biomass and sucrose yield model outputs compared to default APSIM values for Q117 (Table 8). Green biomass was most accurately simulated in all cases while simulated LAI had the highest NRMSE and lowest r^2 and D values of the three outputs analysed. This agrees with results from the original model validation (Keating et al., 1999). LAI was not used in the calibration process as it is not routinely measured in breeding trials or commercial production and is therefore not generally available. By excluding LAI from the calibration process, the calibration performance was more representative of what could be achieved in an operational setting. Although LAI was not used in the calibration process, both GLUE and MCMC

calibrations produced similar NRMSE, r^2 and D statistics to the default APSIM-Sugar settings. Based on NRMSE and D index, green biomass, sucrose yield and LAI simulations were slightly better using the MCMC calibration compared to the GLUE calibration.

Table 8. Comparison of simulated to observed green biomass, sucrose yield and LAI using APSIM default parameter values for Q117, GLUE and MCMC calibration posterior means. NRMSE was calculated as the root mean square error divided by the output range expressed as a percentage. The r^2 was calculated from the linear regression between observed and simulated values. Willmott's agreement index (D) is a non-parametric goodness of fit measure. Similarly to r^2 a value close to one is desired.

Output	Experiment	N	APSIM Q117 Default			MCMC			GLUE		
			NRMSE	r^2	D	NRMSE	r^2	D	NRMSE	r^2	D
Green Biomass	Grafton(1994-95)	7	7.61	0.99	0.99	7.13	0.98	0.99	4.83	0.98	1.00
	Ayr(1992-94)	24	16.11	0.80	0.95	9.24	0.93	0.98	12.37	0.88	0.97
	Harvest	3	22.58	0.85	0.94	8.37	0.99	0.99	15.65	0.95	0.97
	All Data	31	14.28	0.83	0.95	8.30	0.94	0.98	10.94	0.89	0.97
Sucrose yield	Grafton(1994-95)	4	20.11	0.99	0.94	19.37	0.98	0.95	21.84	0.99	0.94
	Ayr(1992-94)	17	16.64	0.77	0.93	13.97	0.84	0.98	16.87	0.87	0.94
	Harvest	3	16.88	0.99	0.97	17.76	0.99	0.96	23.93	1.00	0.94
	All Data	21	15.14	0.81	0.95	12.76	0.88	0.96	15.38	0.89	0.95
LAI	Grafton(1994-95)	7	11.03	0.94	0.98	16.95	0.86	0.95	15.13	0.92	0.97
	Ayr(1992-94)	24	17.42	0.70	0.90	15.75	0.85	0.91	16.47	0.70	0.91
	Harvest	3	13.23	0.99	0.97	11.03	1.00	0.98	10.92	0.98	0.99
	All Data	31	14.84	0.78	0.93	14.92	0.85	0.93	14.97	0.78	0.94

Despite both calibrations having high r^2 and D values, simulated outputs could differ noticeably (Fig. 7). The high simulated LAI using GLUE parameter estimates is likely a result of the high value of *leaf_size*. Fig. 8 shows the time course of observations and simulated values using both GLUE and MCMC calibrations (solid lines) as well as a 95% credible interval based on 1000 random draws from the normal approximation to the posterior distributions (dashed lines). The solid grey line of Fig. 8 (c) shows that this high *leaf_size* helped simulate peak values of LAI in the Grafton(1994-95) experiment but likely also lead to overestimating lower values of LAI. Conversely, simulated LAI using MCMC parameter estimates underestimated peak LAI in the Grafton experiment but accurately estimated lower value (Fig. 8 (c) solid black line).

The different parameter estimates of GLUE and MCMC lead to similar simulated green biomass and sucrose yield accumulation curves. However, the difference in parameter uncertainty lead to very different uncertainty in simulated sucrose yield (Fig. 8 (b)). The wide credible interval for sucrose yield based on the GLUE posterior distribution is likely a result of the higher uncertainty in *min_sstem_sucrose* and *tt_emerge_to_begcane* parameters which can affect when sucrose accumulation starts. The large difference in output uncertainty due to parameter uncertainty

between the GLUE and MCMC calibrations highlights the importance of reporting the uncertainty along with estimated parameter values when using statistical calibration techniques.

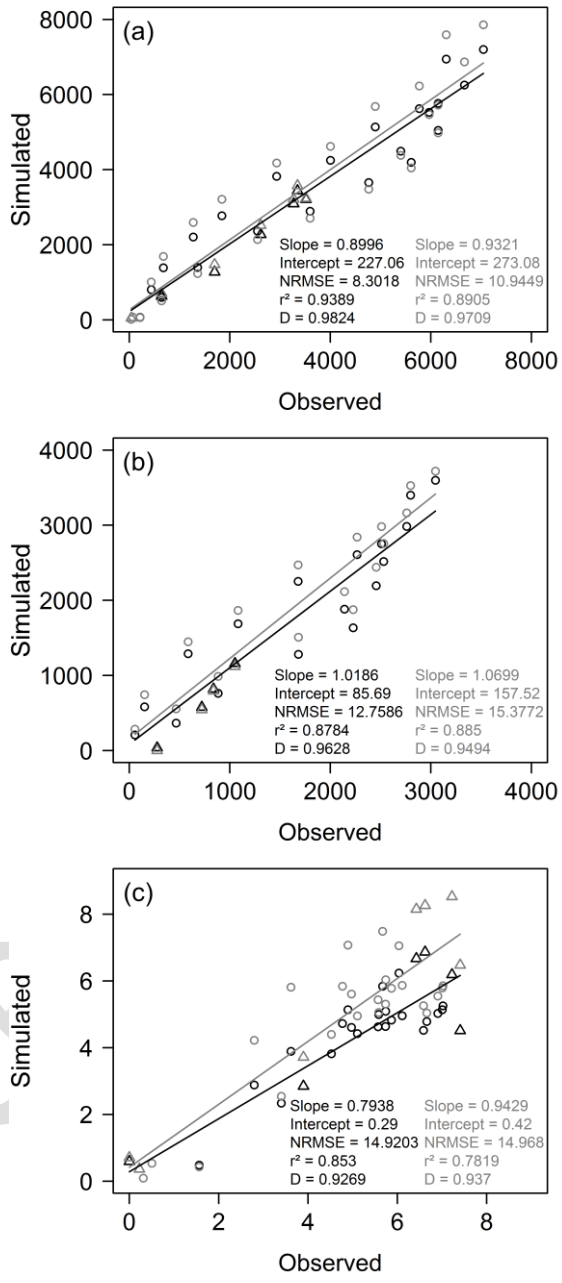

Fig. 7. Linear regression between observed and simulated (a) green biomass (g m^{-2}), (b) sucrose yield (g m^{-2}) and (c) leaf area index for Q117. Results from GLUE (—) and MCMC (—) are plotted across all samples taken from Grafton(1994-95) (△) and Ayr(1992-94) (○) validation experiments.

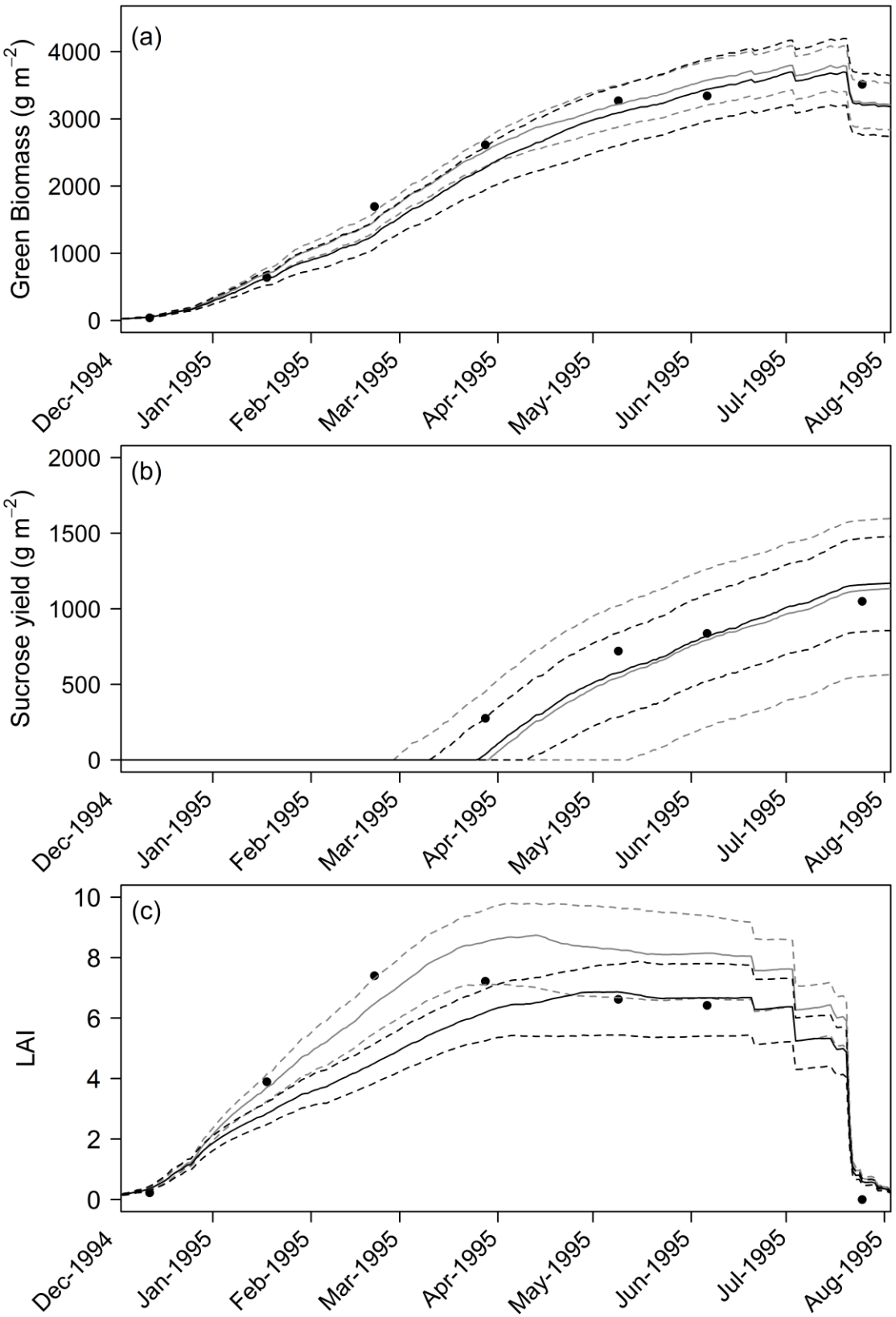

Fig. 8. Time series comparison of simulated and observed values (•) for Q117 in Grafton(1994-95). (a) green biomass (b) sucrose yield and (c) leaf area index (LAI). Solid lines represent simulated values using parameter posterior means for GLUE (—) and MCMC (—). Dashed lines represent a 95% credible interval based on 1000 random draws from the parameter posterior distributions.

4.3. Comparisons of GLUE and MCMC calibrations

Both GLUE and MCMC calibrations were able to accurately simulate green biomass and sucrose yield in both a theoretical and real world evaluation. In fact based on NRMSE, r^2 and D statistics, there was little difference in the accuracy of GLUE and MCMC. This matches earlier studies which have found that GLUE and MCMC can perform similarly under theoretical conditions (Makowski et al., 2002; Vrugt et al., 2009). Based on the absolute relative error between the estimate parameter values (posterior mean) and pre-defined values, both GLUE and MCMC were able to estimate the differences in parameter values between two pre-defined varieties. This kind of analysis has not been performed before for a sugarcane model and strongly supports the use of statistical calibration techniques in calibrating APSIM-Sugar for different varieties.

Although there was little difference in the skill of GLUE and MCMC in either the theoretical or real world evaluation, the authors recommend modellers consider the use of MCMC in calibrating variety parameters in sugarcane models. MCMC produced slightly lower ARE between estimated and pre-defined parameter values than GLUE in the theoretical evaluation. Furthermore, the MCMC posterior parameter distributions captured the pre-defined parameter values more often than did the GLUE approach in this study. The authors agree with previous comparison studies such as Makowski et al. (2002) that formal techniques such as the MCMC offer advantages in having a well-documented statistical background. Makowski et al. (2002) concluded that by discretising the parameter space the GLUE process may inaccurately represent the posterior parameter distribution. Similarly studies such as Montanari (2005) have shown that GLUE can underestimate the posterior output uncertainty.

Although crop models are highly complex, they are still simplifications of reality. In this sense the GLUE philosophy seems an intuitive fit. Given the simplifications made it seems likely that there should be parameter sets throughout the parameter space capable of performing equally well (equifinality). However, in practice crop models are largely used deterministically such that the calibrated parameter values are likely the only values used. In this case using a calibration technique aimed at finding the optimum solution such as MCMC may be better aligned to the crop modellers' objective.

4.4. Limitations

There are several opportunities to improve either estimation technique. The main limitation of this research was the use of simplistic GLUE and MCMC algorithms. Future research should consider more advanced algorithms such as including the removal of non-behavioural parameter sets form

GLUE (Pathak et al., 2012) or the use of adaptive MCMC algorithms (Dumont et al., 2014). Algorithms could also be improved by including measurements made throughout the season and better reflecting the structure of the error variance by use of an appropriate transformation of the likelihood function. For example, Dumont et al. (2014) replaced the error variance with a coefficient of variation calculated as the ratio of the standard deviation and the value of the observation. This allowed for the inclusion of observed data throughout the growing season. Better defining the structure of error variance within the likelihood function may have resulted in improved calibration performance in both the theoretical and real-world analysis of this study. Calibration efficiency could also be improved by including covariance between parameters in the prior distribution. In this study, the prior distributions of parameters were assumed to be independent as no data on the relationship between parameters were available. One option would be to use preliminary MCMC runs to estimate the structure of parameter covariance.

A further limitation identified in both the theoretical and real world evaluation was the inclusion of parameters with relatively weak influence on the outputs used in calibration. The *sucrose_delay* parameter was poorly estimated in the theoretical evaluation while the estimated value of k_L in the real world evaluation could not be explained physiologically. Future research could avoid calibrating these and other weakly influential parameters. While Sexton and Everingham (2014) identified k_L as a weakly influential parameter, later work (Sexton et al., 2015) suggested k_L could be highly influential under certain growing conditions and may be influenced by soil characteristics. The difficulties that arose in estimating the *sucrose_delay* and k_L parameter values highlights the need for future calibration methods to more formally consider the different sources of model error as described by Kennedy and O'Hagan (2001). Unusual or unexpected parameter values from calibrations may reflect some form of model limitation and may be used to identify a source of error in the specific simulation or highlight an area in the model where further research in the underlying physiology could be explored. The performance of GLUE and MCMC algorithms in the real world analysis may have been improved by including data from more field trials with replicated measurements or by incorporating the available data on LAI into the calibration process.

5. Conclusions

In this paper we evaluated the use of GLUE and MCMC as techniques for calibrating APSIM-Sugar for different sugarcane varieties. Both GLUE and MCMC calibrations were able to accurately simulate green biomass and sucrose yield in both a theoretical and real world evaluation. Parameters that are difficult to measure can be estimated using a limited amount of data such as biomass and

sucrose yields which are routinely collected in breeding programs. Although there was little difference in the skill of GLUE and MCMC in either the theoretical or real world application, the authors recommend modellers consider the use of MCMC in calibrating variety parameters in sugarcane models as the MCMC has a well-documented statistical background. In this study, the MCMC posterior parameter distributions were also able to capture pre-defined parameter values.

The MCMC and GLUE techniques provided a reproducible and easily documented approach to model calibration. Furthermore, quantifying parameter uncertainty through the posterior distribution allowed the associated uncertainty in model outputs to be explored. The ability to implement these statistical techniques in freely available programs such as R means that calibration toolboxes could readily be developed for other crop models. Statistical calibration techniques such as MCMC should be used as a systematic approach to updating models such as APSIM-Sugar as varieties are developed. With appropriate variety definitions, APSIM-Sugar could be used for early risk assessment of adopting new varieties in different growing environments.

Acknowledgements

The authors would like to thank Peter deVoil and Bangyou Zheng for assistance with setting up and running APSIM simulations on a cluster. The authors would also like to thank Michael Robertson for help with accessing field experiment data. This work was funded by Sugar Research Australia and James Cook University as part of a Master's Thesis project (STU076). The authors would like to thank Abraham Singles and Brian Kennedy for reviewing an earlier version of this paper and three anonymous reviewers for providing constructive feedback.

References

Aggarwal, P.K., Kropff, M.J., Cassman, K.G., ten Berge, H.F.M., 1997. Simulating genotypic strategies for increasing rice yield potential in irrigated, tropical environments. Field Crop. Res. 51(1–2) 5-17. DOI: http://dx.doi.org/10.1016/S0378-4290(96)01044-1.

Archontoulis, S.V., Miguez, F.E., Moore, K.J., 2014. A methodology and an optimization tool to calibrate phenology of short-day species included in the APSIM PLANT model: Application to soybean. Environ. Model. Softw. 62 465-477. DOI: http://dx.doi.org/10.1016/j.envsoft.2014.04.009.

Basnayake, J., Jackson, P.A., Inman-Bamber, N.G., Lakshmanan, P., 2012. Sugarcane for water-limited environments. Genetic variation in cane yield and sugar content in response to water stress. J. Exp. Bot. 63(16) 6023-6033. DOI: http://dx.doi.org/10.1093/jxb/ers251.

Bennett, N.D., Croke, B.F.W., Guariso, G., Guillaume, J.H.A., Hamilton, S.H., Jakeman, A.J., Marsili-Libelli, S., Newham, L.T.H., Norton, J.P., Perrin, C., Pierce, S.A., Robson, B., Seppelt, R., Voinov, A.A., Fath, B.D., Andreassian, V., 2013. Characterising performance of environmental models. Environ. Model. Softw. 40 1-20. DOI: http://dx.doi.org/10.1016/j.envsoft.2012.09.011.

Beven, K., Binley, A., 1992. The future of distributed models: Model calibration and uncertainty prediction. Hydrol. Process. 6(3) 279-298. DOI: http://dx.doi.org/10.1002/hyp.3360060305.

Beven, K., Freer, J., 2001. Equifinality, data assimilation, and uncertainty estimation in mechanistic modelling of complex environmental systems using the GLUE methodology. J. Hydrol. 249(1–4) 11-29. DOI: http://dx.doi.org/10.1016/S0022-1694(01)00421-8.

Bezuidenhout, C., Singels, A., Hellmann, D., 2002. Whole farm harvesting strategy optimisation using the Canegro model: a case study for irrigated and rainfed sugarcane, Proceedings of the South African Sugarcane Technologists' Association: Kwa-Shukela, Mount Edgecombe, pp. 250-259.

Biggs, J.S., Thorburn, P.J., Crimp, S., Masters, B., Attard, S.J., 2013. Interactions between climate change and sugarcane management systems for improving water quality leaving farms in the Mackay Whitsunday region, Australia. Agric., Ecosyst. Environ. 180 79-89. DOI: http://dx.doi.org/10.1016/j.agee.2011.11.005.

Boote, K.J., Kropff, M.J., Bindraban, P.S., 2001. Physiology and modelling of traits in crop plants: implications for genetic improvement. Agric. Syst. 70(2–3) 395-420. DOI: http://dx.doi.org/10.1016/S0308-521X(01)00053-1.

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P., Bussière, F., Cabidoche, Y.M., Cellier, P., Debaeke, P., Gaudillère, J.P., Hénault, C.,

Maraux, F., Seguin, B., Sinoquet, H., 2003. An overview of the crop model STICS. Eur. J. Agron. 18(3–4) 309-332. DOI: http://dx.doi.org/10.1016/S1161-0301(02)00110-7.

Chapman, S., 2008. Use of crop models to understand genotype by environment interactions for drought in real-world and simulated plant breeding trials. Euphytica 161(1-2) 195-208. DOI: http://dx.doi.org/10.1007/s10681-007-9623-z.

Dumont, B., Leemans, V., Mansouri, M., Bodson, B., Destain, J.P., Destain, M.F., 2014. Parameter identification of the STICS crop model, using an accelerated formal MCMC approach. Environ. Model. Softw. 52 121-135. DOI: http://dx.doi.org/10.1016/j.envsoft.2013.10.022.

Everingham, Y.L., Muchow, R.C., Stone, R.C., Inman-Bamber, N.G., Singels, A., Bezuidenhout, C.N., 2002. Enhanced risk management and decision-making capability across the sugarcane industry value chain based on seasonal climate forecasts. Agric. Syst. 74 459-477. DOI: http://dx.doi.org/10.1016/S0308-521X(02)00050-1.

Gelman, A., Carlin, J.B., Stern, H.S., Rubin, D.B., 1997. Markov chain simulation, In: Chatfield, C., Zidek, J. (Eds.), Bayesian data analysis, 1 ed. Chapman & Hall: London.

Hastings, W.K., 1970. Monte Carlo sampling methods using Markov chains and their applications. Biometrika 57(1) 97-109. DOI: http://dx.doi.org/10.1093/biomet/57.1.97.

He, J., Jones, J.W., Graham, W.D., Dukes, M.D., 2010. Influence of likelihood function choice for estimating crop model parameters using the generalized likelihood uncertainty estimation method. Agric. Syst. 103(5) 256-264. DOI: http://dx.doi.org/10.1016/j.agsy.2010.01.006.

Holzworth, D.P., Huth, N.I., deVoil, P.G., Zurcher, E.J., Herrmann, N.I., McLean, G., Chenu, K., van Oosterom, E.J., Snow, V., Murphy, C., Moore, A.D., Brown, H., Whish, J.P.M., Verrall, S., Fainges, J., Bell, L.W., Peake, A.S., Poulton, P.L., Hochman, Z., Thorburn, P.J., Gaydon, D.S., Dalgliesh, N.P., Rodriguez, D., Cox, H., Chapman, S., Doherty, A., Teixeira, E., Sharp, J., Cichota, R., Vogeler, I., Li, F.Y., Wang, E., Hammer, G.L., Robertson, M.J., Dimes, J.P., Whitbread, A.M., Hunt, J., van Rees, H., McClelland, T., Carberry, P.S., Hargreaves, J.N.G., MacLeod, N., McDonald, C., Harsdorf, J., Wedgwood, S., Keating, B.A., 2014. APSIM – Evolution towards a new generation of agricultural systems simulation. Environ. Model. Softw. 62 327-350. DOI: http://dx.doi.org/10.1016/j.envsoft.2014.07.009.

Holzworth, D.P., Snow, V., Janssen, S., Athanasiadis, I.N., Donatelli, M., Hoogenboom, G., White, J.W., Thorburn, P., 2015. Agricultural production systems modelling and software: Current status and future prospects. Environ. Model. Softw. 72 276-286. DOI: http://dx.doi.org/10.1016/j.envsoft.2014.12.013.

Hughes, R., Robertson, M., Muchow, R., Wood, A., 1995. A comparison of 12-month sugarcane crop production between North Queensland and Northern New South Wales, Proceedings of the Australian Society of Sugar Cane Technologists, pp. 149-154.

Iizumi, T., Yokozawa, M., Nishimori, M., 2009. Parameter estimation and uncertainty analysis of a large-scale crop model for paddy rice: Application of a Bayesian approach. Agric. For. Meteorol. 149(2) 333-348. DOI: http://dx.doi.org/10.1016/j.agrformet.2008.08.015.

Iizumi, T., Yokozawa, M., Nishimori, M., 2011. Probabilistic evaluation of climate change impacts on paddy rice productivity in Japan. Clim. Change 107(3-4) 391-415. DOI: http://dx.doi.org/10.1007/s10584-010-9990-7.

Inman-Bamber, G., 2013. Sugarcane Yields and Yield-Limiting Processes, In: Moore, P.H., Botha, F.C. (Eds.), Sugarcane Physiology Biochemistry and Functional Biology. Wiley-Blackwell, pp. 579-600. DOI: http://dx.doi.org/10.1002/9781118771280.ch21.

Inman-Bamber, N.G., Lakshmanan, P., Park, S., 2012. Sugarcane for water-limited environments: Theoretical assessment of suitable traits. Field Crop. Res. 134(0) 95-104. DOI: http://dx.doi.org/10.1016/j.fcr.2012.05.004.

Inman-Bamber, N.G., McGlinchey, M.G., 2003. Crop coefficients and water-use estimates for sugarcane based on long-term Bowen ratio energy balance measurements. Field Crop. Res. 83(2) 125-138. DOI: http://dx.doi.org/10.1016/S0378-4290(03)00069-8.

Jackson, P.A., Basnayake, J., Inman-Bamber, G., Lakshmanan, P., 2014. Selecting sugarcane varieties with higher transpiration efficiency, Proceedings of the Australian Society of Sugar Cane Technologists: Gold Coast, pp. 82-92.

Jones, C.A., Wegener, M.K., Russell, J.S., McLeod, I.M., Williams, J.R., 1989. AUSCANE -- Simulation of Australian sugarcane with EPIC. CSIRO, Canberra, Australia.

Jones, J.W., He, J., Boote, K.J., Wilkens, P., Porter, C.H., Hu, Z., 2011. Estimating DSSAT Cropping System Cultivar-Specific Parameters Using Bayesian Techniques, In: Ahuja, L.R., Ma, L. (Eds.), Advances in agricultural systems modeling 2. Methods of Introducing System Models into Agricultural Research. Book and Multimedia Publishing Committee, pp. 365-394. DOI: http://dx.doi.org/10.2134/advagricsystmodel2.c13.

Jones, M., Singels, A., Ruane, A., 2014. Simulated impacts of climate change on water use and yield of irrigated sugarcane in South Africa, 86th Annual Congress of the South African Sugar Technologists' Association (SASTA 2013), Durban, South Africa, 6-8 August 2013. South African Sugar Technologists' Association, pp. 184-189.

Keating, B.A., Robertson, M.J., Muchow, R.C., Huth, N.I., 1999. Modelling sugarcane production systems I. Development and performance of the sugarcane module. Field Crop. Res. 61(3) 253-271. DOI: http://dx.doi.org/10.1016/S0378-4290(98)00167-1.

Kennedy, M.C., O'Hagan, A., 2001. Bayesian calibration of computer models. Journal of the Royal Statistical Society: Series B (Statistical Methodology) 63(3) 425-464. DOI: http://dx.doi.org/10.1111/1467-9868.00294.

Knox, J.W., Rodríguez Díaz, J.A., Nixon, D.J., Mkhwanazi, M., 2010. A preliminary assessment of climate change impacts on sugarcane in Swaziland. Agric. Syst. 103(2) 63-72. DOI: http://dx.doi.org/10.1016/j.agsy.2009.09.002.

Laurila, H., Mäkelä, P., Kleemola, J., Peltonen, J., 2012. A comparative ideotype, yield component and cultivation value analysis for spring wheat adaptation in Finland. Agric. Food Sci. 21(4) 384-408.

Lethbridge, P., Cox, M., 2010. Managing variety improvement program operations using an integrated database and handheld computers to improve efficiency and reduce errors, Proceedings of the Australian Society of Sugar Cane Technologists: Bundaberg, Queensland, Australia, pp. 251-261.

Lisson, S.N., Inman-Bamber, N.G., Robertson, M.J., Keating, B.A., 2005. The historical and future contribution of crop physiology and modelling research to sugarcane production systems. Field Crop. Res. 92(2–3) 321-335. DOI: http://dx.doi.org/10.1016/j.fcr.2005.01.010.

Liu, D.L., Bull, T.A., 2001. Simulation of biomass and sugar accumulation in sugarcane using a process-based model. Ecol. Model. 144(2–3) 181-211. DOI: http://dx.doi.org/10.1016/S0304-3800(01)00372-6.

Magarey, R., Croft, B., Willcox, T., 2001. An epidemic of orange rust on sugarcane in Australia, Proceedings of the International Society of Sugar Cane Technologists, pp. 410-416.

Makowski, D., Hillier, J., Wallach, D., Andrieu, B., Jeuffroy, M., 2006. Parameter estimation for crop models, In: Wallach, D., Makowski, D., Jones, J. (Eds.), Working with dynamic crop models. Evaluation, analysis, parameterization and applications. Elsevier: Amsterdam, pp. 101-150.

Makowski, D., Wallach, D., 2002. It pays to base parameter estimation on a realistic description of model errors. Agronomie-Sciences des Productions Vegetales et de l'Environnement 22(2) 179-190. DOI: http://dx.doi.org/10.1051/agro:2002002.

Makowski, D., Wallach, D., Tremblay, M., 2002. Using a Bayesian approach to parameter estimation; comparison of the GLUE and MCMC methods. Agronomie 22(2) 191-203. DOI: http://dx.doi.org/10.1051/agro:2002007.

Marin, F., Ribeiro, R., Marchiori, P.R., 2014. How can crop modeling and plant physiology help to understand the plant responses to climate change? A case study with sugarcane. Theor. Exp. Plant Physiol. 26(1) 49-63. DOI: http://dx.doi.org/10.1007/s40626-014-0006-2.

Marin, F.R., Jones, J.W., Royce, F., Suguitani, C., Donzeli, J.L., Filho, W.J.P., Nassif, D.S.P., 2011. Parameterization and Evaluation of Predictions of DSSAT/CANEGRO for Brazilian Sugarcane. Agron. J. 103(2) 304-315. DOI: http://dx.doi.org/10.2134/agronj2010.0302.

McGlinchey, M., Inman-Bamber, N., 1996. Effect of irrigation scheduling on water use efficiency and yield, Proceedings of the South African Sugar Technologists' Association, pp. 55-56.

Metropolis, N., Rosenbluth, A., Rosenbluth, M., Teller, A., Teller, E., 1953. Equation of state calculations by fast computing machines. J. Chem. Phys. 21 1087-1092. DOI: http://dx.doi.org/10.1063/1.1699114.

Mitchell, M., 1999. An introduction to genetic algorithms. The MIT Press, Cambridge, Massachusetts.

Mo, X., Beven, K., 2004. Multi-objective parameter conditioning of a three-source wheat canopy model. Agric. For. Meteorol. 122(1–2) 39-63. DOI: http://dx.doi.org/10.1016/j.agrformet.2003.09.009.

Montanari, A., 2005. Large sample behaviours of the generalized likelihood uncertainty estimation (GLUE) in assessing the uncertainty of rainfall-runoff simulations. Water Resources Research 41(8) p. 13. DOI: http://dx.doi.org/10.1029/2004WR003826.

Muchow, R.C., Robertson, M., Wood, A., Keating, B.A., 1996. Effect of nitrogen on the time-course of sucrose accumulation in sugarcane. Field Crop. Res. 47 143-153. DOI: http://dx.doi.org/10.1016/0378-4290(96)00022-6.

Nott, D.J., Marshall, L., Brown, J., 2012. Generalized likelihood uncertainty estimation (GLUE) and approximate Bayesian computation: What's the connection? Water Resources Research 48(12) p. 7. DOI: http://dx.doi.org/10.1029/2011WR011128.

Pathak, T.B., Jones, J.W., Fraisse, C.W., Wright, D., Hoogenboom, G., 2012. Uncertainty Analysis and Parameter Estimation for the CSM-CROPGRO-Cotton Model. Agron. J. 104(5) 1363-1373. DOI: http://dx.doi.org/10.2134/agronj2011.0349.

Putto, C., Patanothai, A., Jogloy, S., Boote, K.J., Hoogenboom, G., 2013. Determination of plant traits that affect genotype × location (G×L) interaction in peanut using the CSM-CROPGRO-peanut model. Int. J. Plant. Prod. 7(3) 537-568.

R Core Team, 2013. R: A language and environment for statistical computing., 3.0.0 ed. R Foundation for Statistical Computing: Vienna, Austria.

Robertson, M., Bonnett, G., Hughes, R.M., Muchow, R.C., Campbell, J.A., 1998. Temperature and leaf area expansion of sugarcane: integration of controlled-environment, field and model studies. Aust. J. Plant Physiol. 25 819-828. DOI: http://dx.doi.org/10.1071/PP98042.

Robertson, M.J., Wood, A.W., Muchow, R.C., 1996. Growth of sugarcane under high input conditions in tropical Australia. I. Radiation use, biomass accumulation and partitioning. Field Crop. Res. 48, 11-25. DOI: http://dx.doi.org/10.1016/0378-4290(96)00041-X.

Sexton, J., 2015. Bayesian statistical calibration of variety parameters in a sugarcane crop model. College of Science, Technology and Engineering, James Cook University: Townsville, Australia, p. 111.

Sexton, J., Everingham, Y., 2014. Global sensitivity analysis of key parameters in a process-based sugarcane growth model - A Bayesian approach, In: Ames, D.P., Quinn, N.W.T., Rizzoli, A.E. (Eds.), Proceedings of the 7th International Congress on Environmental Modelling and Software: San Diego, California, USA, p. 8.

Sexton, J., Everingham, Y.L., Inman-Bamber, G., A global sensitivity analysis of cultivar trait parameters in a sugarcane growth model for contrasting production environments in Queensland, Australia. European Journal of Agronomy. DOI: http://dx.doi.org/10.1016/j.eja.2015.11.009.

Sinclair, T.R., 2012. Is transpiration efficiency a viable plant trait in breeding for crop improvement? Funct. Plant Biol. 39(5) 359-365. DOI: http://dx.doi.org/10.1071/FP11198.

Singels, A., 2013. Crop Models, In: Moore, P.H., Botha, F.C. (Eds.), Sugarcane Physiology Biochemistry and Functional Biology. Wiley-Blackwell, pp. 541-577. DOI: http://dx.doi.org/10.1002/9781118771280.ch20.

Singels, A., Jones, M., Marin, F., Ruane, A., Thorburn, P., 2013. Predicting Climate Change Impacts on Sugarcane Production at Sites in Australia, Brazil and South Africa Using the Canegro Model. Sugar Tech 1-9.

Singels, A., Jones, M., van den Berg, M., 2008. DSSAT v4. 5-CANEGRO Sugarcane Plant Module: Scientific documentation. South African Sugarcane Research Inst. Mount Edgecombe, South Africa.

Skocaj, D., Everingham, Y., Schroeder, B., 2013. Nitrogen Management Guidelines for Sugarcane Production in Australia: Can These Be Modified for Wet Tropical Conditions Using Seasonal Climate Forecasting? Springer Science Reviews 1(1-2) 51-71. DOI: http://dx.doi.org/10.1007/s40362-013-0004-9.

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(12) W00B06. DOI: http://dx.doi.org/10.1029/2008WR006822.

Stewart, L.K., Charlesworth, P.B., Bristow, K.L., Thorburn, P.J., 2006. Estimating deep drainage and nitrate leaching from the root zone under sugarcane using APSIM-SWIM. Agric. Water Manage. 81(3) 315-334. DOI: http://dx.doi.org/10.1016/j.agwat.2005.05.002.

Tao, F., Yokozawa, M., Zhang, Z., 2009. Modelling the impacts of weather and climate variability on crop productivity over a large area: A new process-based model development, optimization, and uncertainties analysis. Agric. For. Meteorol. 149(5) 831-850. DOI: http://dx.doi.org/10.1016/j.agrformet.2008.11.004.

Tao, F., Zhang, Z., 2013. Climate change, wheat productivity and water use in the North China Plain: A new super-ensemble-based probabilistic projection. Agric. For. Meteorol. 170(0) 146-165. DOI: http://dx.doi.org/10.1016/j.agrformet.2011.10.003.

Thorburn, P.J., Biggs, J.S., Collins, K., Probert, M.E., 2010. Using the APSIM model to estimate nitrous oxide emissions from diverse Australian sugarcane production systems. Agric., Ecosyst. Environ. 136(3–4) 343-350. DOI: http://dx.doi.org/10.1016/j.agee.2009.12.014.

van der Laan, M., Miles, N., Annandale, J.G., du Preez, C.C., 2011. Identification of opportunities for improved nitrogen management in sugarcane cropping systems using the newly developed Canegro-N model. Nutr. Cycl. Agroecosyst. 90(3) 391-404. DOI: http://dx.doi.org/10.1007/s10705-011-9440-6.

Venables, W.N., Ripley, B.D., 2002. Modern Applied Statistics with S, Fourth ed. Springer, New York. DOI: http://dx.doi.org/10.1007/978-0-387-21706-2.

Vrugt, J., ter Braak, C.F., Gupta, H., Robinson, B., 2009. Equifinality of formal (DREAM) and informal (GLUE) Bayesian approaches in hydrologic modeling? Stoch. Environ. Res. Risk. Assess. 23(7) 1011-1026. DOI: http://dx.doi.org/10.1007/s00477-008-0274-y.

Wallach, D., Buis, S., Lecharpentier, P., Bourges, J., Clastre, P., Launay, M., Bergez, J.E., Guerif, M., Soudais, J., Justes, E., 2011. A package of parameter estimation methods and implementation for the STICS crop-soil model. Environ. Model. Softw. 26(4) 386-394. DOI: http://dx.doi.org/10.1016/j.envsoft.2010.09.004.

Wang, X., He, X., Williams, J., Izaurralde, R., Atwood, J., 2005. Sensitivity and uncertainty analyses of crop yields and soil organic carbon simulated with EPIC. Trans. ASAE 48(3) 1041-1054. DOI: http://dx.doi.org/10.13031/2013.18515.

Webster, A.J., Thorburn, P.J., Roebeling, P.C., Horan, H.L., Biggs, J.S., 2009. The expected impact of climate change on nitrogen losses from wet tropical sugarcane production in the Great Barrier Reef region. Mar. Freshwat. Res. 60(11) 1159-1164. DOI: http://dx.doi.org/10.1071/MF08348.

Willmott, C.J., 1982. Some Comments on the Evaluation of model performance. Bull. Am. Meteorol. Soc. 1309-1313. DOI: http://dx.doi.org/10.1175/1520-0477(1982)063<1309:SCOTEO>2.0.CO;2.

Zheng, B., 2012. APSIMBatch: Analysis the output of APSIM software, R package version 0.1.0.2347 ed.