Sample size, range of parameters and time-dependent effects on global sensitivity analysis in sugarcane modelling

R. A. A. Pereira[1] ![ORCID], M. S. Vianna[2] and F. R. Marin[1] ![ORCID]

[1]University of São Paulo, 'Luiz de Queiroz' College of Agriculture, Piracicaba, SP 13418-900, Brazil and [2]Institute of Crop Science and Resource Conservation (INRES), University of Bonn, Bonn, Germany

Abstract

Process-based crop models (PBCM) for sugarcane present many genotype parameters compared to other crops, making it harder to calibrate. Global sensitivity analysis (GSA) has thus become an important tool for understanding, calibrating and further developing PBCMs. This work used a recently updated Agronomic Modular Simulator for Sugarcane (SAMUCA) to simulate crop growth and development, conducted with two treatments: with green cane trash blanket (GCTB) and under bare soil (BARE). Using the extended Fourier Amplitude Sensitivity (eFAST) algorithm, GSA was performed on the 24 genotype parameters of SAMUCA. The objective was to determine the sample size (SZ): how many samplings are necessary to quantify the sensitivity indices. Additionally, we aimed to assess the influence of parameter range variation and identify which genotype parameters explain the highest variance in simulations of the SAMUCA model under BARE and GCTB conditions. The results showed that SZ greatly affected the convergence and sensitivity indices, and the SZ required here needed to be >2049 for the analysis to cover all variables. Two sets of parameter ranges were used for analysis (the first used maximum and integer values of each parameter reported in the literature; the second applied 25% perturbation to the previously calibrated values). The results indicated that the parameter range affected the parameters' order of importance. Furthermore, we identified that at different phenological stages of sugarcane development, distinct parameters were responsible for explaining the most variance of the output. However, there was no difference among ratoons or interference in the results of BARE or GCTB.

Introduction

In agriculture, process-based crop models (PBCM) represent the state of the art for simulating crop growth and development (Jones *et al.*, 2017; Marin *et al.*, 2017). When properly calibrated, they are commonly used to simulate crop growth and development under different conditions, thus being able to test hypothetical management, climate and soil scenarios (Faivre *et al.*, 2009). Scientists and decision-makers have used modelling as a tool to address issues related to the sugar and bioenergy sectors, including climate change (Marin *et al.*, 2013; Singels *et al.*, 2013; Jones *et al.*, 2015), plant breeding (Hoffman *et al.*, 2018), risk analysis (Everingham *et al.*, 2002) and crop forecast (Everingham *et al.*, 2016).

Sugarcane is a crucial crop for world bioenergy (Raza *et al.*, 2019), and several authors have studied sugarcane modelling (Keating *et al.*, 1999; Singels and Bezuidenhout, 2002; Inman-Bamber and Smith, 2005; Thorburn *et al.*, 2005; Marin and Jones, 2014; Valade *et al.*, 2014; Jones and Singels, 2018; Vianna *et al.*, 2020). Singels (2013) presented a detailed review of the main sugarcane models in the literature, highlighting sugarcane as one of the crops with highest need to be represented in PBCM given its specific farming systems and logistic requirements. To represent its physiological complexity, sugarcane PBCM have many genotype parameters compared to other crops, such as maize and wheat. For instance, the sugarcane models DSSAT/CASUPRO (Villegas *et al.*, 2005) and DSSAT/CANEGRO (Jones and Singels, 2008) have 42 and 22 cultivar coefficients, respectively, while DSSAT/CERES-MAIZE (Jones *et al.*, 1986) and DSSAT/CERES-WHEAT (Ritchie and Otter, 1985) have only respectively six and seven parameters to be calibrated.

According to Sinclair and Seligman (1996), the development of different PBCM by more research groups allows to improve the understanding of processes. In this sense, Marin and Jones (2014) developed the Agronomic Modular Simulator for Sugarcane (SAMUCA) focusing on the specific characteristics of sugarcane farming systems in Brazil. Recently, the SAMUCA model was improved by reducing the uncertainties around the soil water balance, heat flux and physiological mechanisms such as carbon partition, photosynthesis, tillering and root growth (Vianna *et al.*, 2020).

As any other PBCM, SAMUCA represents a simplification of the real system and requires several parameters whose determination is a problem for practical operational applications (Makowski *et al.*, 2002). Most parameters are acquired through field observations, which are expensive and time-consuming, and the acquisition of certain parameters is difficult.

Furthermore, for reliable simulations, accurate parameter estimation is required (Guérif and Duke, 2000; Wallach *et al.*, 2019a), and so several parameter estimation algorithms have been developed, part of which are based on Bayesian approaches (He *et al.*, 2010, 2009; Sreelash *et al.*, 2012; Marin *et al.*, 2017; Sheng *et al.*, 2019; Zhang *et al.*, 2020). To some extent, these methods solved the problem of difficult-to-acquire parameters, and they are quite efficient but applicable to a small number of parameters only (Varella *et al.*, 2010). Still, the inclusion of many parameters in a PBCM raises a dilemma related to the difficulty in simultaneously estimating all unknown parameters and ensuring at the same time that their biophysical meaning is coherent.

In practice, it is well known that only part of the parameters is usually responsible for most of the model uncertainty, while most of them have only minor influence (Varella *et al.*, 2010; Li *et al.*, 2019; Zhang *et al.*, 2020). The parameter sensitivity analysis (SA) method can identify the most important parameters for a given model output variable, which allows users to focus on the most important model parameters during the calibration process. Furthermore, based on the SA, the balance and robustness of the model can be analysed for future improvement, model development and applications (Fraedrich and Goldberg, 2000; Confalonieri, 2010; Chu-Agor *et al.*, 2011; Hirabayashi *et al.*, 2011).

The SA can be divided into two groups: the local sensitivity analysis (LSA) and the global sensitivity analysis (GSA) (Wallach *et al.*, 2019b). The LSA consists of changing a single parameter at a time, while the other parameters are kept at their reference values; in other words, this method is based on the local derivatives of the model's output concerning the variation of a single parameter, which indicates how strong the output changes around the reference parameter values (Saltelli *et al.*, 1999). The GSA allows you to evaluate the entire uncertainty range of parameters, considering changes in all parameters along with their range, as well as the interactions among parameters (Saltelli *et al.*, 1999).

The GSA methods can also be classified into three groups: screening, regression and variance; all of these follow the Bayesian sampling principle (Wei, 2013). The most used screening method is the Morris method, which permits to define the most important model parameters and it is often considered a qualitative method (Morris, 1991; Dejonge *et al.*, 2012). Regression methods, such as the Partial Rank Correlation Coefficient (PRCC), provide the correlation between the model's output and the selected parameters and they are mandatory when the parameter and the model outputs have a monotonic relationship (Marino *et al.*, 2008; Krishnan and Aggarwal, 2018). Methods based on variance are the most used one, and the three mains are: Sobol (2001), Fourier Amplitude Sensitivity Test (FAST) (Cukier *et al.*, 1978) and the extended Fourier Amplitude Sensitivity Test (eFAST) (Saltelli *et al.*, 1999); they provide what parameters cause the greater variability for the model's output, and they usually demand a high computational cost. To define which method is the most suitable, some properties of model must be known (linearity, prior distribution of parameters and monotonic), furthermore considering the number of parameters to be evaluated and the computational cost (Iooss and Lemaître, 2015). In this sense, Sobol and eFAST are the methods most applicable to any type of PBCM, but while Sobol is very computationally expensive, eFAST integrates the merits of FAST and Sobol's algorithms, representing a method with high efficiency and precision, and the ability to adequately compute interaction effects among parameters (Iooss and Lemaître, 2015).

The GSA has several aspects that can affect sensitivity indices and their uncertainty, regardless of the method adopted. In general, the most important uncertainty sources of GSA are: (i) sample size (SZ), (ii) range of parameters and (iii) complexity of the model (Xu and Gertner, 2011; Gan *et al.*, 2014; Song *et al.*, 2015). To our knowledge, there are no studies in the literature investigating effects in GSA caused by SZ and parameter range on eFAST method in sugarcane models. In the case of parameter SZ, studies that utilized the eFAST method were based on the evidence presented by Wang *et al.* (2013), where the recommended SZ is greater than 1024. However, in studies that did not follow this recommendation, a very large SZ was adopted without clear criteria (Tan *et al.*, 2016). By adopting an SZ like that of Wang *et al.* (2013), the model characteristics are ignored, as the SZ is influenced by various factors such as the number of parameters, the analysed output variable, and the methods used to estimate variables within the model. Additionally, choosing a very large SZ requires excessively long computational time, which can render the analysis infeasible in some cases.

The parameter range is another source of uncertainty in PBCM when using Bayesian approaches for parameter estimation (Makowski *et al.*, 2006), consequently affecting the sensitivity indices in a GSA. Depending on the range of parameters, it is possible to generate calibrations that do not represent the desired genotype (He *et al.*, 2010, 2009; Sexton *et al.*, 2016; Marin *et al.*, 2017) or correctly quantify the uncertainty (Soetaert and Petzoldt, 2010; Dzotsi *et al.*, 2013; Gan *et al.*, 2014; Zhang *et al.*, 2020; Pereira *et al.*, 2021). Wang *et al.* (2013) compared a range of parameters measured for maize and a relative range of 10% in relation to a reference calibration and found important differences in the GSA results. Li *et al.* (2019) evaluated different relative ranges, from 10 to 50%, in relation to a reference value and concluded that the most important parameters for the 10% range diverged from those obtained using the 50% range. Many recent PBCM studies (Vazquez-Cruz *et al.*, 2014; Tan *et al.*, 2016; Jin *et al.*, 2018; Li *et al.*, 2019) have adopted relative parameter ranges to apply a GSA, which can result in serious methodological errors. According to Homma and Saltelli (1996), the GSA principle is to identify the parameters that cause the greatest uncertainty in the model, and this is not possible when all parameters are disturbed to create relative parameters range. Finally, there is still the model complexity, which is a source of considerable uncertainty and a complicated issue to be considered in the GSA (Razavi and Gupta, 2015).

The SAMUCA crop model has been relatively little evaluated compared to well-established models such as DSSAT/CANEGRO and APSIM-Sugar (Thorburn *et al.*, 2005; Marin and Jones, 2014; Marin *et al.*, 2015; Sexton *et al.*, 2017). Pereira *et al.* (2021) used the PRCC to perform a GSA to identify the most important parameters for SAMUCA and then to explore the model uncertainty. However, some issues remained unclear due to the following limitations: (i) the GSA was performed only for the end-of-season output values; (ii) the PRCC method is limited when the parameter relation is not monotonic, and only four out of the 24 genotype parameters are monotonic in

SAMUCA, which means that some parameter responses might be neglected during the GSA; (iii) authors did not consider the effect of the range of parameters on GSA results. The first two limitations may result in the omission of important parameters because they affect certain variables at a different time of simulation (Lamboni *et al.*, 2009; Dejonge *et al.*, 2012), and the third has never been investigated for sugarcane crop models, and even for other crops the range of parameter effect on GSA was little studied (Wang *et al.*, 2013).

Considering the aspects mentioned above, this paper aimed to: (i) determine the optimal SZ for the eFAST method; (ii) investigate whether there is a difference between the ranges of parameters used in GSA; (iii) identify which parameters are responsible for the greatest uncertainty in the SAMUCA model.

Material and methods

SAMUCA crop model

The SAMUCA model is a process-based model that simulates the growth and development of sugarcane. Originating from the argument of Sinclair and Seligman (Sinclair and Seligman, 1996), which emphasizes the importance of research groups developing their own models to incorporate region-specific processes, this approach enhances the understanding of the uncertainties and limitations of the model. Thus, the SAMUCA model was mainly developed by Marin and Jones (2014) and Marin *et al.* (2017), using a large database for different Brazilian production conditions. Subsequently, Vianna *et al.* (2020) improved the model structure by decreasing the uncertainty in the soil water balance and including the effect of straw cover on sugarcane growth and development, modifying routines of soil moisture and the flow of water and heat from soil, compared to the previous versions. The most recent version is also included in the DSSAT platform v.4.8 (https://dssat.net/).

The standalone version of SAMUCA model previously presented by Marin and Jones (2014) simulates fibre and sucrose biomass partitioning at the internode level. The new version keeps the same implementation; however, leaves and internodes are grouped into phytomer units (Vianna *et al.*, 2020). The phytomers, which are associated with tillers that have emerged in the simulation area. The root system is treated as a single set of organs and develops as thermal age progresses. Plant development is calculated separately for roots, tillers, leaves and internodes, considering the accumulation of thermal units. Soil temperature is used for underground processes such as root development and the development of tiller phytomers that have not yet emerged. Air temperature, on the other hand, is used to calculate the development of above-ground phytomers. The phytomer appearance rate for each tiller is simulated based on the *plastochron* interval, which is a measure of thermal time required for the development of each phytomer. During the initial phase of sugarcane development, the tiller appearance rate is calculated based on soil temperature and the emergence interval of new tillers, known as the tiller emergence rate. The senescence phase of tillering is triggered when the amount of light transmitted through the canopy reaches a specific threshold. In this phase, younger tillers gradually senesce over thermal time until the number of tillers reaches the expected final population, determined by the parameter *popmat*. The leaf senescence rate is controlled by the parameter *maxgl*, which defines the maximum number of leaves per tiller. When the number of leaves on a tiller exceeds this value, the

oldest leaf is shed as it has completed its life cycle. Soil temperature is used to calculate thermal age and temperature-dependent processes until the organ in question (such as below-ground internodes, roots and tillering) emerges at the soil surface. After emergence, all new leaves use air temperature to calculate accumulated thermal units and contribute to the formation of the crop canopy, expressed by the leaf area index. As for internodes, the switch from using soil temperature to air temperature for the development of new internodes occurs when the stalk emerges at the soil surface, controlled by the parameter *n_lf_when_stk_emerg*. Internodes are considered 'appeared' at the top of the tiller after a minimum number of leaves have emerged for their formation, while older leaves are attached to already differentiated internodes, enclosed in leaf sheaths. At the apex of the tiller, a set of developing leaves surrounds the apical meristem of the bud, at the centre of the leaf whorl. These leaves will become visible at the tip of the tiller and eventually develop into internodes.

Data and management

The crop growth simulation scenario was based on a field experiment conducted in the College of Agriculture 'Luiz de Queiroz', Piracicaba, São Paulo (Lat: 22°41′55″S, Lon: 47°38′34″W, Alt: 540 m). The sugarcane cultivar was the RB86-7515, a widely used genotype in Brazil (ca. 30% of Brazil's planted area). It was planted on 16 October 2012, with a row spacing of 1.4 m and depth of 0.2 m. The BARE treatment was conducted over four consecutive years, while the GCTB treatment commenced in the first ratoon (October 2013) and was carried out for 3 years. Agricultural practices were adopted to represent high yield farming systems and to ensure the crop was free from pests, diseases and nutritional stress. The climate is characterized by hot and humid summers with dry winters (Cwa-Koppen classification), and the soil classified as Typic Hapludox. The experiment was irrigated by a centre-pivot, based on monitoring the soil moisture by Frequency Domain Reflectometry (FDR) and the evapotranspiration by Bowen Ratio Method (BRM) in both treatments (Table 1).

As the eFAST method requires a high computational time, we divided our study into two steps. In the first step (STp1), we simulated the first ratoon of sugarcane with GCTB and BARE, testing different SZ and two sets of parameter ranges (Table 2). The main objective of STp1 was to define what minimum SZ is needed to obtain a reliable GSA and then apply them to the second step (STp2) of the study. In STp1 we ran the GSA using end-of-season values as reference of the variables: stalk dry mass (SDM), stalk fresh mass (SFM), leaf area index (LAI), sucrose concentration in the fresh matter (POL) and tiller population (TIL). In the STp2, after determining the SZ where the standard deviation of the sensitivity indices tends to zero and there is no risk of altering the parameter importance order, we performed a long simulation considering the different ratoons and ran the GSA in function of the daily values simulated in the whole season of each variable (SDM, SFM, POL, LAI and TIL).

Sensitivity analysis

Extended Fourier Amplitude Sensitivity Test

The eFAST is an algorithm that combines two GSA methods: the FAST and the Sobol (Saltelli *et al.*, 1999, 2010), which in turn, use the model output variance principle. While FAST can scan the entire parameter space and obtain quantitative sensitivity

Table 1. Description of seasons, planting and harvesting dates, duration of the season in days, treatments and variable measurements of the field experiment in Piracicaba, Brazil

Cropping measure	Planting	Harvest	Duration of the season	Measurements	Treatments
Plant cane	10/16/2012	10/15/2013	364	SDM, SFM, TIL, LAI and POL	BARE
1st Ratoon	10/15/2013	07/15/2014	273	SDM, SFM, TIL, LAI and POL	BARE and GCTB
2nd Ratoon	07/15/2014	06/08/2015	328	SDM, SFM, TIL, LAI and POL	BARE and GCTB
3rd Ratoon	06/08/2015	06/08/2016	365	SDM, SFM, TIL, LAI and POL	BARE and GCTB

Green cane trash blanket (GCTB), stalk dry mass (SDM; t/ha) and stalk fresh mass (SFM; t/ha) of leaf area index (LAI; m^2/m^2), sucrose concentration in fresh matter (POL; %) and tiller population (TIL; #/m^2).

Table 2. Description of the processes performed in the first step (STp1) and the second step (STp2); the simulated season, the sample size (SZ) evaluated, soil cover type, number of repetitions (NR) and parameter range set (PRS)

Step	Season	SZ	Treatments	NR	PRS
STp 1	1st Ratoon	65, 129, 257, 513, 1025, 2049, 4097	BARE and GCTB	10	PRS1 PRS2
STp 2	Plant cane to 3rd ratoon	2049	BARE and GCTB	1	PRS1 PRS2

The details about the SZ are presented in the section on sample size, and the details about PRS1 and PRS2 are provided in Table 3.

measures in terms of the main sensitivity index (S_i) of each parameter to output variance, the Sobol calculates the total sensitivity index (ST_i) and provides an indication of the overall effect of a given parameter, considering all possible interactions of that parameter with others (Sobol, 2001). The eFAST is based on the decomposition of the model's output variance, determining which fraction of the variance explained by each parameter. This variation is quantified using the statistical notion of variance (analogous to ANOVA):

$$\sigma^2 = \sum_{i=1}^{N} \frac{(y_i - \bar{y})^2}{(N-1)} \tag{1}$$

where N is number of models runs, y_i is i_{th} model output and $\bar{y}$ sample mean. Partitioning of variance in eFAST works by varying different parameters at different frequencies, encoding the identity of parameters in the frequency of their variation. In recent years, due to these advantageous properties, eFAST has become more popular in hydrological, ecological and agronomy modelling (Varella *et al.*, 2010; Reusser *et al.*, 2011; Xing *et al.*, 2017; Li *et al.*, 2019). We implemented the eFAST method for the SAMUCA model in the sensitivity R-package available at: https://cran.r-project.org/web/packages/sensitivity/index.html.

The main sensitivity index (S_i) of a given parameter (i) is calculated as the variance at a particular parameter's unique frequency (and harmonics of that frequency) divided by total variance (VAR_t). First, variance (VAR_i) is calculated from the Fourier coefficients at the frequency of interest (j):

$$VAR_i = 2(A_j^2 + B_j^2) \tag{2}$$

$$A_j = \int_{-\pi}^{\pi} f(s) \cos(js) \, ds \tag{3}$$

$$B_j = \int_{-\pi}^{\pi} f(s) \sin(js) \, ds \tag{4}$$

$$S_i = \frac{VAR_i}{VAR_t} \tag{5}$$

where s is a scalar variable within the range $-\infty < s < +\infty$; A_j and B_j are the Fourier coefficients (or Fourier amplitude) over the domain of integer frequencies $j \in \{-\infty, \ldots, -1, 0, 1, \ldots, \infty\}$. S_i represents the fraction of the output variance of the model explained by the input variation of a given parameter. ST_i is calculated as the remaining variance after the complementary set contribution is removed. Thus, to estimate ST_i for the given parameter i, the eFAST algorithm first calculates the sensitivity indices except for parameter i using the identification frequencies.

$$ST_i = \frac{VAR_t - VAR_{-i}}{VAR_t} \tag{6}$$

where VAR_{-i} is the sum of all the variance terms that do not include the parameters i.

S_i and ST_i must vary between 0 and 1, and effects are greater the indices reach values closer to 1, whereas values close to 0 indicate negligible effects. ST_i considers both S_i and the interactions between the parameters, such interactions be evaluated by the difference between the ST_i and the S_i. The two sensitivity indices S_i and ST_i are equal if the effect of the parameter i on the model output is independent of the values of the other parameters: in this case, there is no interaction between this parameter and the others, and the model is additive into parameter i. We considered only the parameters that had $S_i > 0.05$ and $ST_i > 0.1$ as significant and relevant in the GSA analysis, as suggested by Dejonge *et al.* (2012) and Xing *et al.* (2017). This allowed us to identify the parameters that individually explain at least 5% of the model's variance through the S_i index. Additionally, using the ST_i index, we could identify the parameter that, when combined with others, explains more than 10% of the variance, indicating that the contribution of the other parameters explains at least an additional 5% of the variance.

Sample size

The SZ refers to the number of samples generated within the range of each parameter. In other words, for a given parameter x, SZ samples are generated while respecting the limits of variation for that parameter. This SZ is a key factor in determining the number of model evaluations (Ne), where the number of evaluations is obtained by multiplying the number of evaluated parameters (Np) by the SZ (Ne = Np × SZ). In some cases, a large number of model evaluations are required, which can restrict the method's use. Therefore, the relationship between SZ and the convergence of the sensitivity measure is of utmost importance. The SZ is determined by the sampling frequency and the exploration curve, as described by Saltelli *et al.* (1999). To investigate the influence of SZ on the convergence of sensitivity indices, an SA was performed with different SZ. Seven SZ cases were used: 65, 129, 257, 513, 1025, 2049 and 4097, as proposed by Wang *et al.* (2013). For each SZ, we conducted ten repetitions considering the 1st ratoon for the BARE and GCTB treatments. We employed two criteria to determine the SZ. The first criterion involved calculating the sum of S_i values and checking if this sum converged to 1. If it did not meet this criterion, the SZ would be discarded. The second criterion involved calculating the mean and standard deviation of S_i values based on the ten repetitions. In this way, we identified the smallest SZ where the standard deviation of S_i was sufficiently small and the sum of S_i approached 1, classifying it as the most suitable for the variable analysis. Additionally, the smallest SZ value that simultaneously met these two criteria for all variables was defined as the most appropriate for the SAMUCA model. In this analysis, we considered two ranges of parameter variation, which will be described in the next item (Parameters range set).

Parameters range set

The GSA is affected by the uncertainty range of the parameters (Wang *et al.*, 2013), so we investigated this factor for the GSA of the SAMUCA model by constructing two ranges of genotype parameters. The first parameter range set (PRS1) was constructed based on the literature, containing the maximum and minimum values of each genotype parameter, regardless of the sugarcane cultivar (Pereira *et al.*, 2021). For the second parameter range set (PRS2), we considered the studies of Wang *et al.* (2013) and Jin *et al.* (2018), and concluded that the order of importance of the parameters converges to above 10% disturbance. Thus, in order not to use an excessively small disturbance that would cause inconsistency in the GSA results, and to avoid an excessively large disturbance to generate parameter values outside their genotype reality, we choose to cause a ±25% perturbation in the values calibrated by Vianna *et al.* (2020). In Table 3, we present the description of the parameters and their respective values for PRS1 and PRS2.

Results

Sample size

Figures 1 and 2 presented the S_i values for the main parameter obtained from the PRS1 intervals (Fig. 1) and PRS2 intervals (Fig. 2), in relation to different SZ in the BARE and GCTB treatments. We observed a dispersion for small SZ. The SZ of 65 and 127 showed high variability both in the sensitivity indices and in the order of the main parameters explaining the model variance. For the lowest SZ evaluated (65 and 127), it was not possible to obtain sensitivity indices and accurately quantify the order of

parameter importance due to the divergence of the sensitivity indices' sum, S_i and ST_i from 1. In these cases, the eFAST method was unable to quantify the sensitivity indices for the PRS1 (Table 4) and PRS2 (Table 5) sets.

We noticed that the SZ varied according to the analysed variable; however, we did not observe any effect of the GCTB or BARE treatments on the SZ (Tables 4 and 5). Tables 4 and 5 present the order of significant parameters for each SZ, considering different parameter sets and treatments. For the SDM variable, the required SZ for the PRS1 set was 1025 (Table 4), which was smaller than the 2049 obtained for the PRS2 set (Table 5). In the case of the SFM, POL and LAI variables, the same SZ were found for both the PRS1 and PRS2 sets: 1025 for SFM and 513 for POL and LAI (Tables 4 and 5). The TIL variable achieved the sensitivity indices with the smallest SZ in both parameter sets. For the PRS1 set, the SZ was 257 (Table 4), while for the PRS2 set, it was 513 (Table 5). We observed a divergence among the variables, and the minimum SZ required to calculate the sensitivity indices using the eFAST method varied depending on the variable. Consequently, the lowest SZ at which it was possible to calculate the sensitivity indices for all variables was 2049 (Tables 4 and 5).

Crop features

We consider two key points to evaluate; (i) the order of importance of the parameters and (ii) the values of S_i and ST_i in both parameter ranges (PRS1 and PRS2). We noticed that GCTB affected the order of the main parameter only in PRS2, while S_i and ST_i values were slightly affected between GCTB and BARE treatments.

In PRS1, the results between BARE and GCTB were similar since the main parameter for all variables was maintained regardless of the presence or absence of GCTB (Fig. 3). For example, the main parameter for SDM and SFM was the *plastochron*, and it was responsible for explaining 29 and 31.6% of the variance in the BARE and GCTB treatments, respectively; for SFM the explained variance was 27.9 and 30.4% for BARE and GCTB, respectively (Tables 6 and 7). The GCTB only influenced the order of importance of the parameters of the variable SDM. In the BARE treatment the parameter *n_lf_when_stk_emerg* was the 6th parameter (explaining 5.4% of the variance) and in the treatment with GCTB it was the 3rd (explaining 12.7% of the variance) (Tables 6 and 7). Furthermore, in the SFM variable, the *plastochron* parameter appears to have greater relevance in the presence of GCTB compared to BARE, as the percentage explained only by it increased, decreasing the values of the parameters *mid_tt_it_growth* and *max_it_dw* (Tables 6 and 7).

In PRS2, the presence of GCTB affected the results regarding the analysis of the variables SDM, SFM and LAI, while there was no influence of GCTB on TIL and POL (Fig. 3). The main parameter for SDM in the BARE treatment was *popmat* (explaining 33.9% of the variance; Table 6), while in the GCTB treatment, the main parameter was *plastochron* (explaining 29.0% of the variance; Table 7). Similarly, in the SFM variable, the main parameter was *popmat* in the BARE treatment (explaining 26.6% of the variance; Table 6) and *plastochron* in the GCTB treatment (explaining 30.5% of the variance; Table 7). In relation to LAI, the presence of GCTB led to the inclusion of the parameter *mla* among the significant parameters, going from three parameters in BARE (*maxgl*, *popmat* and *mid_tt_lf_growth*; Table 6) to four parameters in the GCTB treatment (*maxgl*, *popmat*, *mid_tt_lf_growth* and *mla*; Table 7).

Table 3. Description of parameters, calibrated values for genotype RB867515 (μ) by Vianna *et al.* (2020), first set of parameters (PRS1) extracted from Pereira *et al.* (2021), based on the literature, and second set of parameters (PRS2) with ±25% perturbation relative to μ; CV is the coefficient of variation with respect to μ; Max and Min values are range used for random parameters uniform distribution

Parameters	Description	μ	PRS1 Min	Max	CV (%)	PRS2 Min	Max	CV (%)
amax	Assimilation rate at light saturation point (μmol/m^2/s)	44.9	41.3	60.7	22	33.7	56.1	25
chudec	Heat units for start of tiller abortion (°C/d)	1600	1200	1800	19	1200	2000	25
chumat	Heat units for population establishment (°C/d)	1600	1500	2850	42	1200	2000	25
chupeak	Heat units for population peak (°C/d)	1400	404	1950	55	1050	1750	25
chustk	Heat units for start culm elongation (°C/d)	650	400	1050	50	488	813	25
eff	Carboxylation efficiency (μmol/m^2/s/μmol/m^2/s)	0.069	0.04	0.08	29	0.05	0.09	25
end_tt_it_growth	Thermal time for completion of internode growth (°C/d)	1200	800	1400	25	900	1500	25
end_tt_lf_growth	Thermal time for completion of leaf growth (°C/d)	1300	1100	1500	15	975	1625	25
init_lf_area	Initial leaf area of first appeared leaf (cm^2)	15	10	30	67	11	19	25
max_ini_la	Initial leaf area of leaves appeared after top parts formation (cm^2)	120	80	180	42	90	150	25
max_it_dw	Maximum dry biomass of internodes (g)	28	18	35	30	21	35	25
maxdgl	Maximum number of developed green leaf a tiller can hold (#/tiller)	6	6	12	50	5	8	25
maxgl	Maximum number of green leaf a tiller can hold (#/tiller)	11	10	12	9	8	14	25
mid_tt_it_growth	Thermal time where internodes can achieve half of its maximum biomass (°C/d)	400	380	600	28	300	500	25
mid_tt_lf_growth	Thermal time where leaves can achieve half of its maximum biomass (°C/d)	700	400	800	29	525	875	25
mla	Maximum leaf area (cm^2)	600	450	800	29	450	750	25
n_lf_it_from	Number of leaves appeared before internode formation (#/tiller)	3	2	6	67	2	4	25
n_lf_stk_em	Number of leaves appeared before stalks emerges at soil surface (#/tiller)	4	3	8	63	3	5	25
phyllochron	Phyllochron interval for leaf appearance (°C/d)	132	107	169	23	99	165	25
plastochron	Thermal time required for the appearance of phytomer (°C/d)	132	107	169	23	99	165	25
popmat	Number of tillers on maturation (tiller/m^2)	9.5	8	12	21	7	12	25
poppeak	Maximum number of tillers (tiller/m^2)	27	17	30	24	20	34	25
sla	Specific leaf area (cm^2/g)	120	100	121	9	90	150	25
tillochron	Thermal time required for emergence of new tiller (°C/d)	69	48.1	134.8	63	52	86	25
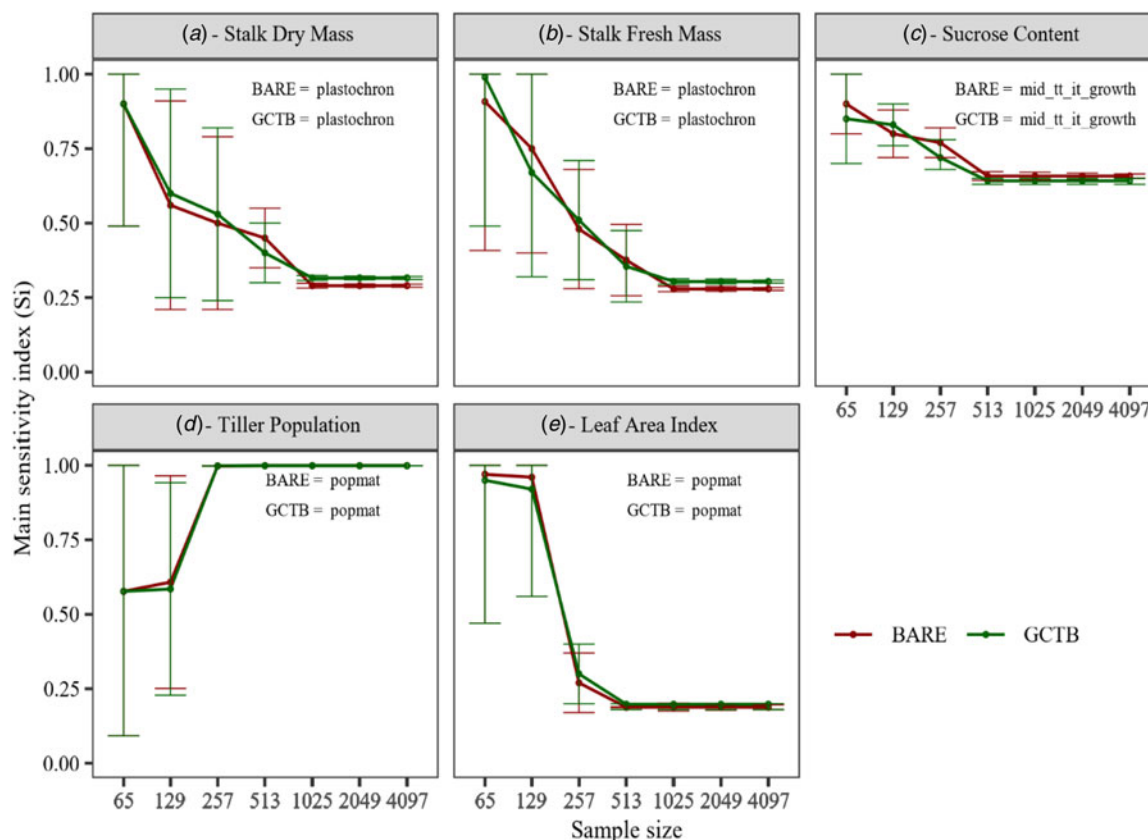
Figure 1. Evolution of sensitivity index of the most important parameter for BARE and GCTB with increasing SZ for variables mass stalk dry (a) and fresh (b), sucrose content (c), tiller population (d) and leaf area index (e), for PRS1; red line is the average of the ten simulations for each SZ and in blue we have the max and min S_i of each SZ.

Effects of parameter range

For SDM, considering the SZ of 2049, we observed a reduction in the number of significant parameters between PRS1 and PRS2. In PRS1 there were six significant parameters (*plastochron*, *max_it_dw*, *n_lf_it_form*, *eff*, *popmat* and *n_lf_when_stk_emerg*), which together were responsible for explaining 81.6% (Table 6) and 82.2% (Table 7) of the variance in the BARE and GCTB treatments, respectively. In the case of PRS2, only four parameters (*popmat*, *plastochron*, *max_it_dw* and *eff*) were significant, responsible for explaining 70% (Table 6) and 70.1% (Table 7) of the variance in the BARE and GCTB treatments, respectively.

Among the output variables analysed, SFM was the one with the highest number of significant parameters, from six to seven parameters. In PRS1 there were six significant parameters in the BARE treatment (*plastochron*, *mid_tt_it_growth*, *max_it_dw*, *end_tt_it_growth*, *eff* and *n_lf_it_form*), which explained 96.8% (Table 6) of the variance, while in the GCTB treatment there were seven parameters (*plastochron*, *max_it_dw*, *mid_tt_it_*, *n_lf_when_stk_emerg*, *end_tt_it_growth*, *eff* and *n_lf_it_form*), which explain 93.3% of the variance (Table 7). In the case of PRS2, there were six significant parameters in both treatments, responsible for explaining around 93.1% (Table 6) and 95% (Table 7) of the variance in the BARE and GCTB treatment, respectively. Two results should be highlighted: (i) in the BARE treatment, *popmat* was not a significant parameter in PRS1, while in PRS2 it was the most important parameter (Table 6); (ii) on the other hand, in the GCTB treatment, *popmat* remained

irrelevant in PRS1 and was the second most important parameter in PRS2; the most important parameter was *plastochron* (Table 7).

In the case of the POL, there was no change in the group of significant parameters or in their order, with the parameters *mid_tt_it_growth* and *end_tt_it_growth* being the only significant ones, regardless of the treatment and the PRS1 and PRS2 set (Tables 6 and 7). However, we noticed that *mid_tt_it_growth* in PRS1 explained 65.8% (Table 6) and 64.2% (Table 7) of the variance, while in PRS2 it was 39.2% (Tables 6 and 7). For the TIL variable, there was no difference between PRS1 and PRS2, being *popmat* responsible for explaining 99.8% of the variance regardless of treatment.

The variable LAI in PRS1 presented only two significant of treatment (BARE or GCTB); the two main parameters were *popmat* and *mla* (Tables 6 and 7). However, in PRS2 the number of significant parameters increased, to 3 (*maxgl*, *popmat* and *mid_tt_lf_growth*; Table 6) in BARE and to 4 (*maxgl*, *popmat*, *mid_tt_lf_growth* and *mla*; Table 7). The biggest discrepancy is in the variance explained by the significant parameters, where in PRS1 they were 33.9% (Table 6) and 34.6% (Table 7), and in PRS2 they were 55.2% (Table 6) and 60% (Table 7).

Time-dependent effects on global sensitivity analysis

Having defined the appropriate SZ (2049), we performed an analysis considering a temporal variation in the different stages of the sugarcane crop season (plant cane to 3rd ratoon). We observed
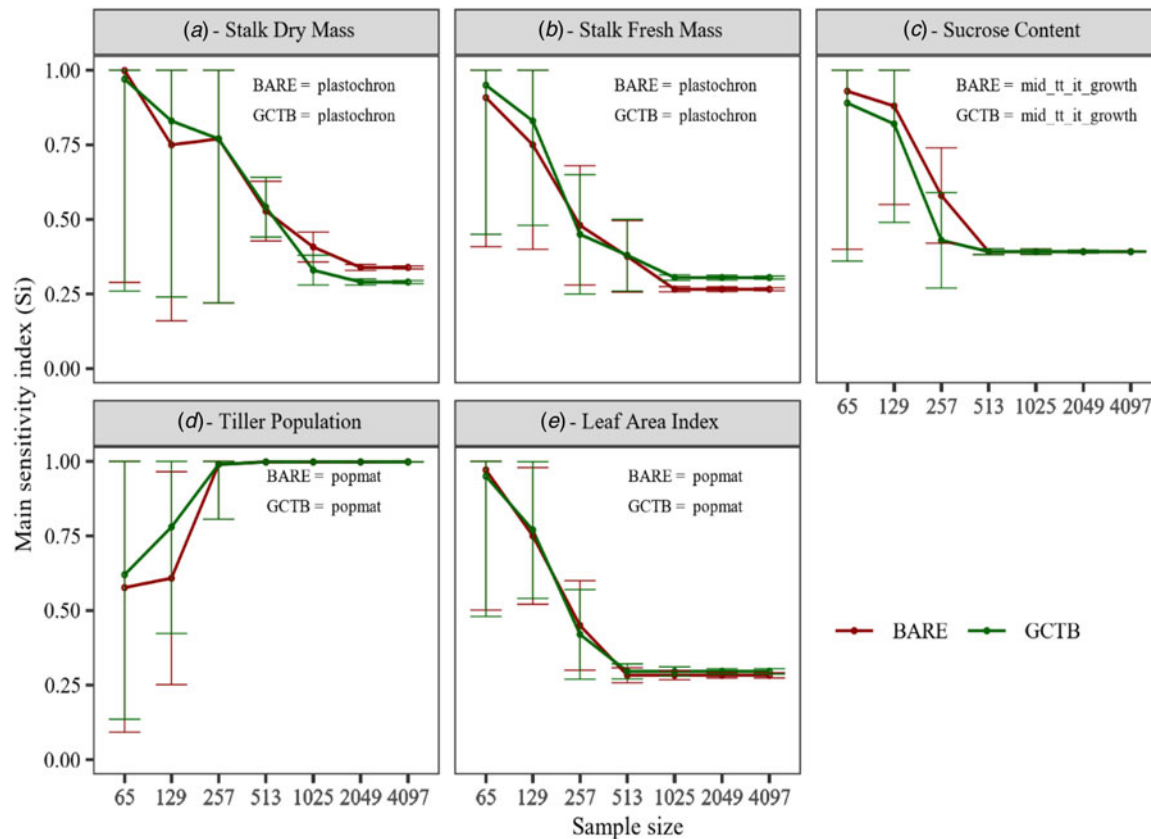
Figure 2. Evolution of sensitivity index of the most important parameter for BARE and GCTB with increasing SZ for variables mass stalk dry (a) and fresh (b), sucrose content (c), tiller population (d) and leaf area index (e), for PRS2; red line is the average of the ten simulations for each SZ and in blue we have the max and min S_i of each SZ.

that the GSA over time provided two important results: (i) the influence of the sugarcane season did not affect the order of importance of parameters. The order of parameters was not changed compared to ratoons; for example, for SDM we noticed that the same parameters have the same degree of importance in all ratoons (Fig. 4). The same pattern was repeated for all variables and their respective parameters and (ii) at different times of sugarcane growth, there were different parameters responsible for explaining the greater variance of the model.

For the variable SDM, during the different growth phases, there was a change in the order of the main parameter. For example, during the crop establishment and development phase, the *plastochron* parameter was responsible for explaining more than 50% of the SDM variance, becoming the main parameter in this period when using the PRS2 set. On the other hand, at the end of the maturation phase, the *plastochron* was the second most important parameter, behind the *popmat* with the PRS2 set (Fig. 4 – PRS2 BARE and GCTB). In PRS1, during the emergence phase until mid-establishment, the main parameter was *n_lf_it_form*, later the *plastochron* became the main parameter until the end of the growth phase (Fig. 4 – PRS1 BARE and GCTB). Furthermore, in short periods and isolated in the growth phase, the parameters *n_lf_it_form*, *end_tt_lf_growth* and *n_lf_when_stk_emerg* were significant using the PRS2 set (Fig. 4 – PRS2 BARE and GCTB); these three parameters were not reported in Tables 6 and 7 as they were not significant at the end of the growth phase. This pattern was not observed in PRS1 (Fig. 4 – PRS1 BARE and GCTB), that is, all significant parameters presented in Tables 6 and 7 appeared in the analysis throughout the growth phases.

For SFM, *plastochron* had a greater impact on the simulation regardless of the parameter set (PRS1 and PRS2) or treatment (Fig. 5 – BARE and GCTB). Among the sets of parameters, PRS1 and PRS2, there was the inclusion of the *popmat* parameter in PRS1 in the BARE treatment; however, its influence was minimal (5%). For PRS2, the *popmat* remained relevant during most parts of the growth phase, but in our view not enough to have more impact on the SFM simulations. The parameter *end_tt_lf_growth* showed significance at specific moments during the tillering phase for the PRS2 set, in both the BARE and GCTB treatments (Fig. 5 – BARE and GCTB), a situation not observed in the PRS1 set. Furthermore, even within the PRS2 set, its relevance was more pronounced in the GCTB treatment, as it was evident in all three ratoons.

For the variable POL, we had the inclusion of several parameters depending on the of the analysed growth phase (Fig. 6). In this case, in PRS1, in addition to *mid_tt_it_growth* and *end_tt_it_growth*, there was the inclusion of *plastochron*, *n_lf_when_stk_emerg* and *n_lf_it_form* (Fig. 6 – PRS1 BARE and GCTB). In PRS2, the parameters *plastochron*, *n_lf_when_stk_emerg*, *n_lf_it_form* and *end_tt_lf_growth* were included (Fig. 6 – PRS2 BARE and GCTB). For PRS1, the parameter *end_tt_it_growth* was not significant in the 1st ratoon (Fig. 6 – PRS1 BARE and GCTB), evidencing a combination of the range of parameters with the weather conditions for that season. Yet, for PRS2, *end_tt_lf_growth* was significant, while in PRS1 the same was not observed.

The TIL variable had two main parameters, from the beginning to the middle of the season, that was *tillochron* and from the middle to the end of the season *popmat* (Fig. 7). In the first half of the season, the *tillochron* influence was around 90%,

Table 4. Order of importance of the genotype parameters for the PRS1 interval, considering the main sensitivity index (S_i) and treatment bare soil (BARE) and green cane trash blanket (GCTB)
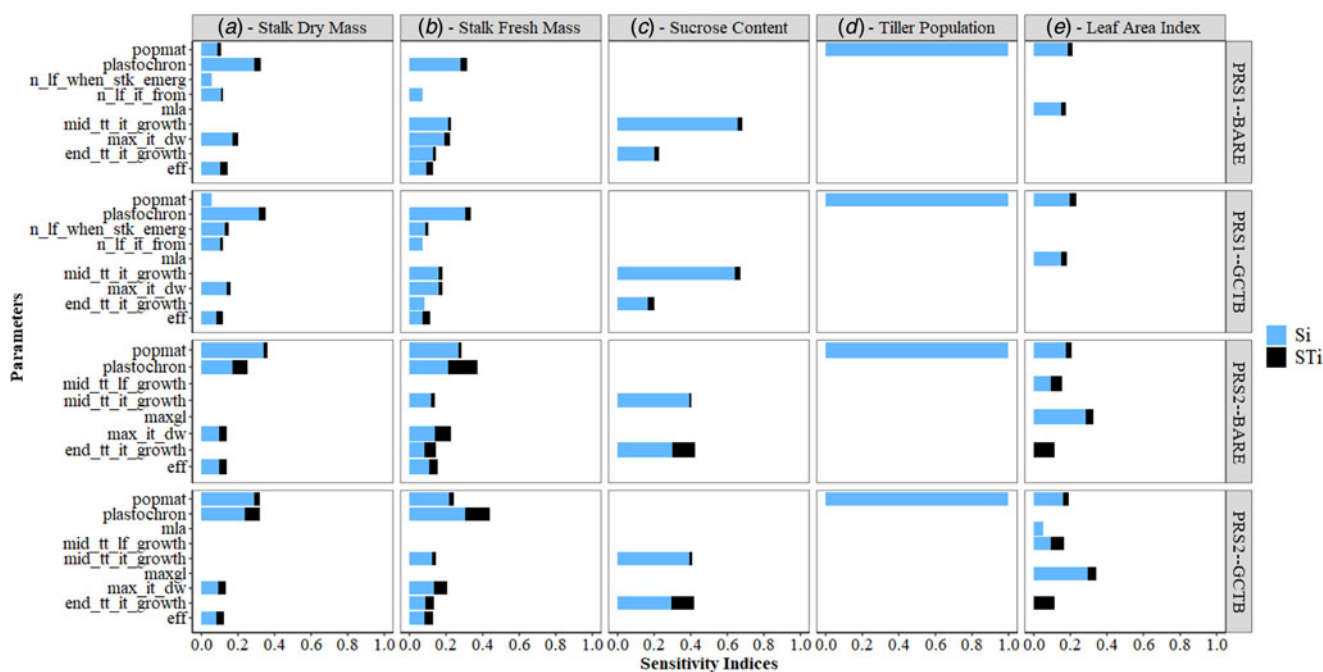
Variable	Rank	SZ 257 BARE	SZ 257 GCTB	SZ 513 BARE	SZ 513 GCTB	SZ 1025 BARE	SZ 1025 GCTB	SZ 2049 BARE	SZ 2049 GCTB	SZ 4097 BARE	SZ 4097 GCTB
SDM (t/ha)	1	–	–	–	–	–	–	plastochron	plastochron	plastochron	plastochron
	2	–	–	–	–	–	–	max_it_dw	max_it_dw	max_it_dw	max_it_dw
	3	–	–	–	–	–	–	n_lf_it_form	n_lf_when_stk_emerg	n_lf_it_form	n_lf_when_stk_emerg
	4	–	–	–	–	–	–	eff	n_lf_it_form	eff	n_lf_it_form
	5	–	–	–	–	–	–	popmat	eff	popmat	eff
	6	–	–	–	–	–	–	n_lf_when_stk_emerg	popmat	n_lf_when_stk_emerg	popmat
SFM (t/ha)	1	–	–	–	–	plastochron	plastochron	plastochron	plastochron	plastochron	plastochron
	2	–	–	–	–	mid_tt_it_growth	max_it_dw	mid_tt_it_growth	max_it_dw	mid_tt_it_growth	max_it_dw
	3	–	–	–	–	max_it_dw	mid_tt_it_growth	max_it_dw	mid_tt_it_growth	max_it_dw	mid_tt_it_growth
	4	–	–	–	–	end_tt_it_growth	n_lf_when_stk_emerg	end_tt_it_growth	n_lf_when_stk_emerg	end_tt_it_growth	n_lf_when_stk_emerg
	5	–	–	–	–	eff	end_tt_it_growth	eff	end_tt_it_growth	eff	end_tt_it_growth
	6	–	–	–	–	n_lf_it_form	eff	n_lf_it_form	eff	n_lf_it_form	eff
	7	–	–	–	–		n_lf_it_form		n_lf_it_form		n_lf_it_form
POL (%[fresh])	1	–	–	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth
	2	–	–	end_tt_it_growth	end_tt_it_growth	end_tt_it_growth	end_tt_it_growth	end_tt_it_growth	end_tt_it_growth	end_tt_it_growth	end_tt_it_growth
TIL (#/m²)	1	popmat	popmat	popmat	popmat	popmat	popmat	popmat	popmat	popmat	popmat
LAI (m²/m²)	1	–	–	popmat	popmat	popmat	popmat	popmat	popmat	popmat	popmat
	2	–	–	mla	mla	mla	mla	mla	mla	mla	mla

It was not possible to define an order of importance for the parameters for the sample size (SZ) values that do not appear in the table; stalk dry mass (SDM), stalk fresh mass (SFM), leaf area index (LAI), sucrose concentration in the fresh matter (POL) and tiller population (TIL).

Table 5. Order of importance of the genotype parameters for the PRS2 interval, considering the main sensitivity index (S_i) and treatment bare soil (BARE) and green cane trash blanket (GCTB)

Variable	Rank	SZ							
		513		1025		2049		4097	
		BARE	GCTB	BARE	GCTB	BARE	GCTB	BARE	GCTB
SDM (t/ha)	1	–	–	popmat	popmat	popmat	popmat	popmat	popmat
	2	–	–	plastochron	plastochron	plastochron	plastochron	plastochron	plastochron
	3	–	–	max_it_dw	max_it_dw	max_it_dw	max_it_dw	max_it_dw	max_it_dw
	4	–	–	eff	eff	eff	eff	eff	eff
SFM (t/ha)	1	–	–	popmat	plastochron	popmat	plastochron	popmat	plastochron
	2	–	–	plastochron	popmat	plastochron	popmat	plastochron	popmat
	3	–	–	max_it_dw	max_it_dw	max_it_dw	max_it_dw	max_it_dw	max_it_dw
	4	–	–	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth
	5	–	–	eff	end_tt_it_growth	eff	end_tt_it_growth	eff	end_tt_it_growth
	6	–	–	end_tt_it_growth	eff	end_tt_it_growth	eff	end_tt_it_growth	eff
POL (% [fresh])	1	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth
	2	end_tt_it_growth	end_tt_it_growth	end_tt_it_growth	end_tt_it_growth	end_tt_it_growth	end_tt_it_growth	end_tt_it_growth	end_tt_it_growth
TIL (#/m²)	1	popmat	popmat	popmat	popmat	popmat	popmat	popmat	popmat
LAI (m²/m²)	1	maxgl	maxgl	maxgl	maxgl	maxgl	maxgl	maxgl	maxgl
	2	popmat	popmat	popmat	popmat	popmat	popmat	popmat	popmat
	3	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth	mid_tt_it_growth
	4		mla		mla		mla		mla

It was not possible to define an order of importance for the parameters for the sample size (SZ) values that do not appear in the table; stalk dry mass (SDM), stalk fresh mass (SFM), leaf area index (LAI), sucrose concentration in the fresh matter (POL) and tiller population (TIL).

Figure 3. Main sensitivity index (S_i) and total sensitivity index (ST_i) of the first parameter range set apply to bare soil (PRS1-BARE) and green cane trash blanket (PRS1-GCTB) treatments, as well as for the second parameter range set applying to bare soil (PRS2-BARE) and green cane trash blanket (PRS2-GCTB) treatments; the sensitivity analysis was calculated for the end-of-season value of each output variable.

Table 6. Relative value of the model output variance (σ^2) explained individually by each parameter, and the variance sum ($\sum^{\sigma^2}$) the parameters; we only considered the parameters that presented $S_i > 0.05$ and the sample size (SZ) of 2049 in treatment bare soil (BARE) to first set of parameters (PRS1) and second set of parameters (PRS2); stalk dry mass (SDM), stalk fresh mass (SFM), leaf area index (LAI), sucrose concentration in the fresh matter (POL) and tiller population (TIL)

Variable	PRS1				PRS2			
	Parameters	σ^2	Rank	$\sum^{\sigma^2}$	Parameters	σ^2	Rank	$\sum^{\sigma^2}$
SDM (t/ha)	plastochron	29.0%	1°	29.0%	popmat	33.9%	1°	33.9%
	max_it_dw	17.2%	2°	46.2%	plastochron	16.7%	2°	50.6%
	n_lf_it_form	10.6%	3°	56.8%	max_it_dw	9.8%	3°	60.4%
	eff	10.5%	4°	67.3%	eff	9.6%	4°	70.0%
	popmat	8.9%	5°	76.2%	–	–	–	–
	n_lf_when_stk_emerg	5.4%	6°	81.6%	–	–	–	–
SFM (t/ha)	plastochron	27.9%	1°	27.9%	popmat	26.6%	1°	26.6%
	mid_tt_it_growth	21.1%	2°	49.0%	plastochron	21.3%	2°	47.9%
	max_it_dw	18.9%	3°	67.9%	max_it_dw	14.2%	3°	62.1%
	end_tt_it_growth	12.7%	4°	80.6%	mid_tt_it_growth	11.9%	4°	74.0%
	eff	9.0%	5°	89.6%	eff	10.6%	5°	84.6%
	n_lf_it_form	7.1%	6°	96.8%	end_tt_it_growth	8.5%	6°	93.1%
POL (%[fresh])	mid_tt_it_growth	65.8%	1°	65.8%	mid_tt_it_growth	39.2%	1°	39.2%
	end_tt_it_growth	20.1%	2°	85.9%	end_tt_it_growth	29.8%	2°	69.1%
TIL (#/m²)	popmat	99.8%	1°	99.8%	popmat	99.8%	1°	99.8%
LAI (m/m²)	popmat	18.8%	1°	18.8%	maxgl	28.3%	1°	28.3%
	mla	15.1%	2°	33.9%	popmat	17.5%	2°	45.8%
	–	–	–	–	mid_tt_lf_growth	9.4%	3°	55.2%
	–	–	–	–	–	–	–	–

Table 7. Relative value of the model output variance (σ^2) explained individually by each parameter, and the variance sum ($\sum^{\sigma^2}$) of the parameters; we only considered the parameters that presented S_i >0.05 and the sample size (SZ) of 2049 in treatment green cane trash blanket (GCTB) to first set of parameters (PRS1) and second set of parameters (PRS2); stalk dry mass (SDM), stalk fresh mass (SFM), leaf area index (LAI), sucrose concentration in the fresh matter (POL) and tiller population (TIL)

| Variable	PRS1				PRS2			
	Parameters	σ^2	Rank	$\sum^{\sigma^2}$	Parameters	σ^2	Rank	$\sum^{\sigma^2}$
SDM (t/ha)	plastochron	31.6%	1°	31.6%	popmat	29.0%	1°	29.0%
	max_it_dw	14.0%	2°	45.5%	plastochron	23.6%	2°	52.6%
	n_lf_when_stk_emerg	12.7%	3°	58.3%	max_it_dw	9.5%	3°	62.1%
	n_lf_it_form	10.4%	4°	68.7%	eff	8.0%	4°	70.1%
	eff	8.0%	5°	76.7%	–	–	–	–
	popmat	5.5%	6°	82.2%	–	–	–	–
SFM (t/ha)	plastochron	30.4%	1°	30.4%	plastochron	30.5%	1°	30.5%
	max_it_dw	15.9%	2°	46.3%	popmat	21.8%	2°	52.3%
	mid_tt_it_growth	15.9%	3°	62.2%	max_it_dw	13.3%	3°	65.6%
	n_lf_when_stk_emerg	8.7%	4°	70.9%	mid_tt_it_growth	12.3%	4°	77.9%
	end_tt_it_growth	8.2%	5°	79.1%	end_tt_it_growth	8.8%	5°	86.7%
	eff	7.2%	6°	86.2%	eff	8.3%	6°	95.0%
	n_lf_it_form	7.1%	7°	93.3%	–	–	–	–
POL (%[fresh])	mid_tt_it_growth	64.2%	1°	64.2%	mid_tt_it_growth	39.2%	1°	39.2%
	end_tt_it_growth	16.7%	2°	80.9%	end_tt_it_growth	29.3%	2°	68.5%
TIL (#/m²)	popmat	99.8%	1°	99.8%	popmat	99.8%	1°	99.8%
LAI (m²/m²)	popmat	19.8%	1°	19.8%	maxgl	29.6%	1°	29.6%
	mla	14.8%	2°	34.6%	popmat	15.8%	2°	45.5%
	–	–	–	–	mid_tt_it_growth	9.5%	3°	55.0%
	–	–	–	–	mla	5.0%	4°	60.0%

while in the second half the *popmat* explained more than 90% of the variance. In the transition period between these two phases of the crop season, that takes place between the end of the establishment and half of the development of the crop, we observed that there was a punctual influence of some other parameters, such as *mla* and *plastochron*. However, during this transition, there were strong indications of the effect of climate on GSA, as these two parameters (*mla* and *plastochron*) were not in all ratoons. To corroborate this result, we did not observe relevant influences of the parameter range (PRS1 or PRS2) and the evaluated treatment (BARE or GCTB).

The LAI variable, contrary to what was seen in sections 'Crop features' and 'Effects of parameter range', with a maximum number of significant parameters of four at the end of growth (Tables 6 and 7), was influenced by 15 parameters at different growth phases (Fig. 8). However, the high number of parameters can be attributed to isolated events. These events occurred at specific moments during the growth phase but did not follow a consistent pattern or a sequential period of influence. In PRS1, we observed that tillochron, *plastochron*, *mla* and *init_leaf_area* parameters had greater relevance, as they explained, individually, more than 40% of the variance of LAI at different moments of emergence phase and of establishment phase (Fig. 8 – PRS1 BARE and GCTB). In the establishment and maturity stages, the variance of the LAI limited to the *mla* and *popmat* parameters with the PRS1 set and, in our view, the *mla* parameter was the most

important parameter as it was significant in all growth phases (Fig. 8 – PRS1 BARE and GCTB). We arrived at this result because the *mla* parameter explains between 20 and 50% of the LAI variance throughout the entire growth phase, while the *popmat* parameter contributes approximately 10–30% of the variance and has influence over a shorter time interval. Thus, even though the variance explained by *popmat* and *mla* may be similar at the end of the cycle, the impact of *mla* is greater and more consistent throughout the entire growth phase. In PRS2, the emergence and establishment phases were dominated by *plastochron* and *mla*, and in the development and maturation phases, the parameters *maxgl*, *popmat* and *mid_tt_lf_growth* were responsible for explaining most of the variance (Fig. 8 – PRS2 BARE and GCTB). In our view, considering the PRS2 set, the most important parameter was *maxgl* because it explained more than 30% of the variance in almost all growth phases. On the other hand, *popmat* and mid_tt_lf_growth explained 10–20% of the variance, and over a shorter period compared to *maxgl*. This result showed the uncertainty present in the range of the chosen parameters, since *maxgl* was not significant at growth phases in the PRS1 set (Fig. 8 – PRS1 BARE and GCTB).

Discussion

The SZ can vary depending on the variable of interest. For example, the variable POL required a minimum size of 513 for
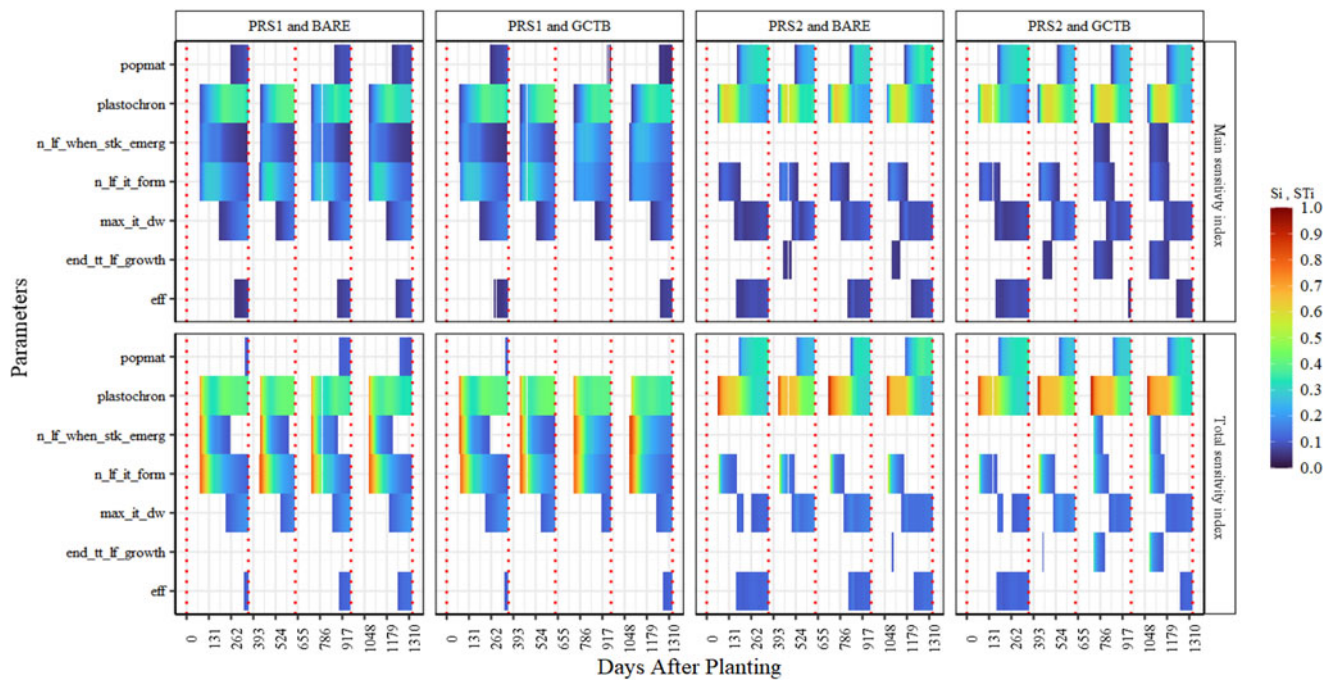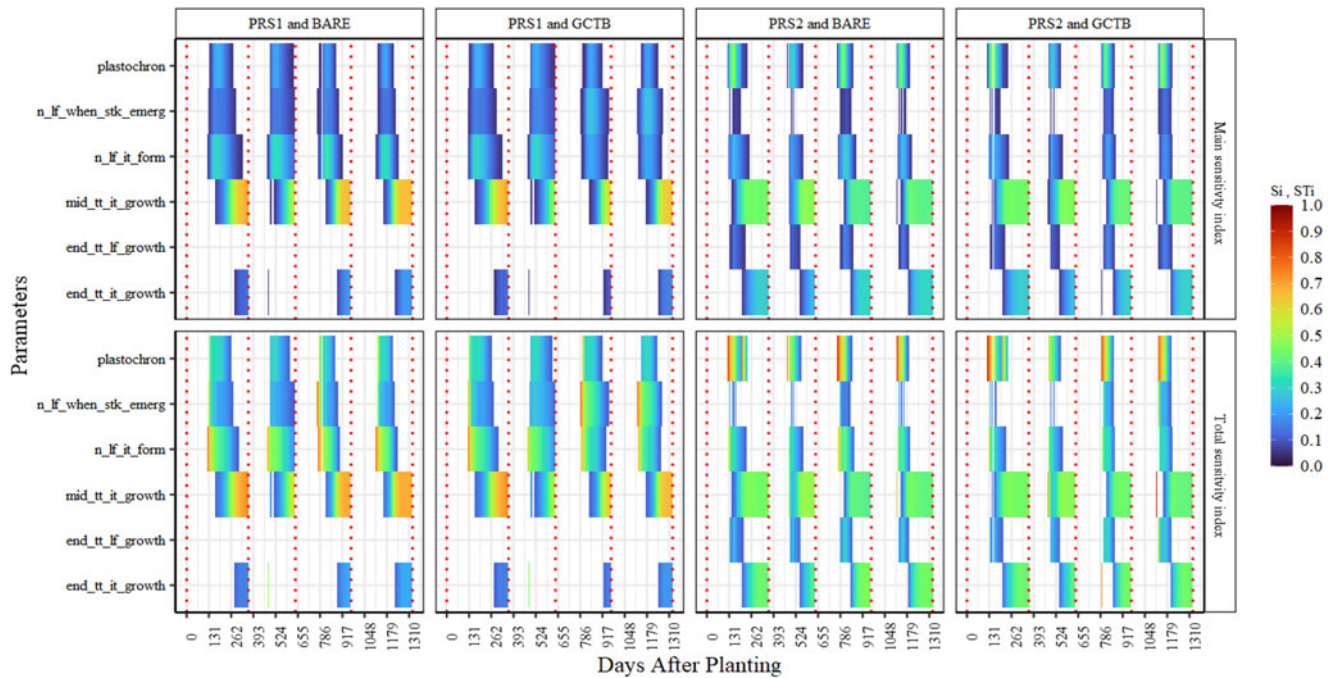
Figure 4. Main sensitivity index (S_i) and total sensitivity index (ST_i) calculated for first parameter set (PRS1) and second parameter set (PRS2) for bare soil (BARE) and green cane trash blanket (GCTB) treatments across plant cane, 1st, 2nd and 3rd ratoons for stalk dry mass (SDM). The red dashed line separates the stages from plant cane to the 3rd ratoons; 'days after planting' represents the time in days from plant cane to the end of 3rd ratoons.
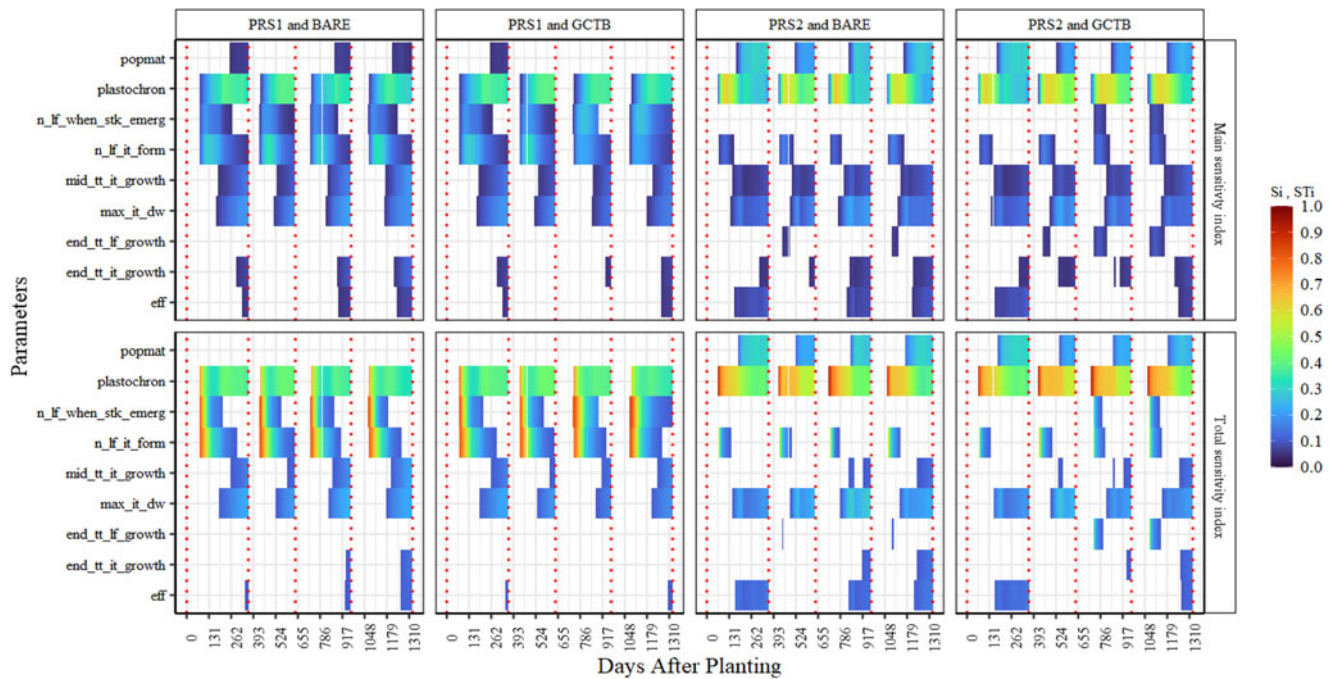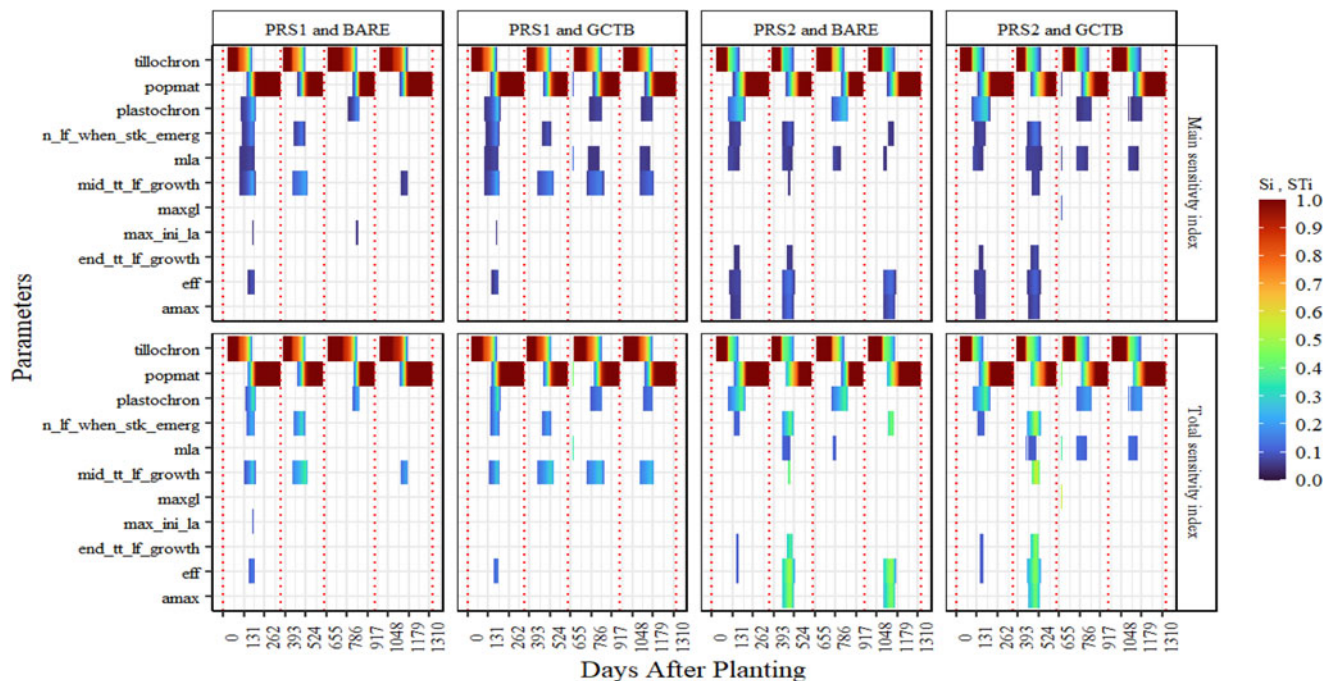

Figure 5. Main sensitivity index (S_i) and total sensitivity index (ST_i) calculated for first parameter set (PRS1) and second parameter set (PRS2) for bare soil (BARE) and green cane trash blanket (GCTB) treatments across plant cane, 1st, 2nd and 3rd ratoons for stalk fresh mass (SFM). The red dashed line separates the stages from plant cane to the 3rd ratoons; 'days after planting' represents the time in days from plant cane to the end of 3rd ratoons.

convergence (Tables 4 and 5), while the variables SDM and SFM required an SZ of at least 2049 (Tables 4 and 5). Additionally, we observed that for the TIL variable, the SZ was 257 for PRS1 and 513 for PRS2 (Tables 4 and 5). This difference implies that computational time can be reduced depending on the output variable.

Therefore, the high computational cost, which is one of the major limitations of GSA (Jeuffroy *et al.*, 2006; Marino *et al.*, 2008; Gilardelli *et al.*, 2018), can be optimized through previous studies exploring the SZ. Thus, the most suitable SZ for conducting SA in the SAMUCA model was 2049, as it encompasses all the analysed

Figure 6. The main sensitivity index (*S*ᵢ) and total sensitivity index (*ST*ᵢ) calculated for first parameter set (PRS1) and second parameter set (PRS2) for bare soil (BARE) and green cane trash blanket (GCTB) treatments across plant cane, 1st, 2nd and 3rd ratoons for sucrose content (POL). The red dashed line separates the stages from plant cane to the 3rd ratoons; 'days after planting' represents the time in days from plant cane to the end of 3ʳᵈ ratoons.
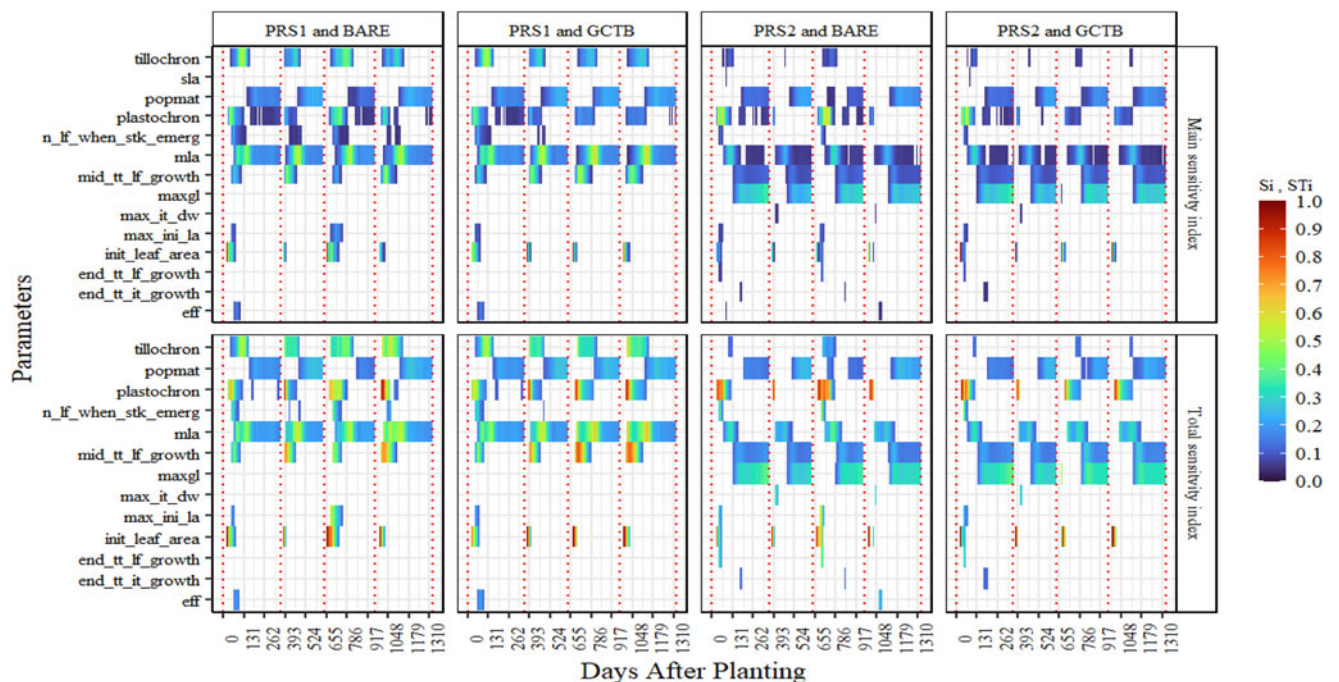

Figure 7. The main sensitivity index (*S*ᵢ) and total sensitivity index (*ST*ᵢ) calculated for first parameter set (PRS1) and second parameter set (PRS2) for bare soil (BARE) and green cane trash blanket (GCTB) treatments across plant cane, 1st, 2nd and 3rd ratoons for tiller population (TIL). The red dashed line separates the stages from plant cane to the 3rd ratoons; 'days after planting' represents the time in days from plant cane to the end of 3ʳᵈ ratoons.

variables. This result is in line with the findings reported by Wang *et al.* (2013), where an SZ of 2049 yielded the most stable sensitivity indices.

As far as we know, Wang *et al.* (2013) is the main study that used the eFAST method in a PBCM and evaluated different SZ,

being considered the main reference source for other GSA studies. Another study, conducted by Ma *et al.* (2017), also determined the most suitable SZ for their model and found lower values compared to those observed by Wang *et al.* (2013) and obtained in our study. This highlights the uniqueness of GSA, where the

Figure 8. The main sensitivity index (S_i) and total sensitivity index (ST_i) calculated for first parameter set (PRS1) and second parameter set (PRS2) for bare soil (BARE) and green cane trash blanket (GCTB) treatments across plant cane, 1st, 2nd and 3rd ratoons for leaf area index (LAI). The red dashed line separates the stages from plant cane to the 3rd ratoons; 'days after planting' represents the time in days from plant cane to the end of 3rd ratoons.

adopted boundary conditions and the analysed model have an impact on the convergence of sensitivity indices. Our study has two distinct characteristics compared to the works of Ma *et al.* (2017) and Wang *et al.* (2013): we evaluated the impact of SZ on different variables and considered two sets of parameter ranges. We observed that, in addition to the output variable, the SZ differed depending on the applied parameter sets (PRS1 and PRS2; Tables 4 and 5). Although there are several studies (Xu and Gertner, 2011; Gan *et al.*, 2014; Song *et al.*, 2015) that consider the effect of SZ and parameter range on sensitivity indices, there are currently no available studies on PBCM that consider the relationship between SZ in different analysis variables and the parameter range. Therefore, we can state that the SZ in GSA may vary depending on (i) the analysed output variable and (ii) the parameter ranges.

As in many studies on the influence of the environment on sensitivity indices (Dejonge *et al.*, 2012; Sexton *et al.*, 2017; Zhang *et al.*, 2020), there was an effect of GCTB on SAMUCA sensitivity indices, but it was not sufficient to change the order of importance of the parameters and it did not influence the parameter range (PRS1 and PRS2) (Figs 4–8). For example, the SFM had the *plastochron* parameter as the main parameter in PRS1 in both treatments (BARE and GCTB), and in PRS2, the main parameter was *plastochorn* in BARE treatment and *popmat* in GCTB treatment. To consider that a GCTB was sufficient to change the order of importance of the parameters, an alteration between the parameters should be observed in both sets of parameters (PRS1 and PRS2; Tables 6 and 7). In the case of GCTB, it has a direct influence on the soil heat flux and on the soil water dynamics (Pereira *et al.*, 2021; Vianna *et al.*, 2020), which would hypothetically refer to the lower effect of GCTB on GSA for the variables LAI, POL, SDM, SFM and TIL; if we would evaluate soil temperature or evapotranspiration, we would possibly have observed a greater impact of

GCTB, as it is directly related to these variables. We assume this based on the results of Dejonge *et al.* (2012), who identified the radiation use efficiency as the most important parameter for yield, both for irrigated and rainfed treatments. However, for crop evapotranspiration, in the irrigated environment, the main parameters were related to the crop, while in the rainfed environment the main parameters were related to the soil (Dejonge *et al.*, 2012). Thus, the effect of management on the SA is dependent on the variable of interest and did not affect all simulation variables in our study.

The range of parameters was the main source of uncertainty for the SA, changing the order of parameters for the variables SDM, SFM and LAI. Li *et al.* (2019) have already identified that parameter range affected the order and variance explained by each parameter. Many studies have applied parameter ranges with relative values from GSA studies in different PBCM (Vazquez-Cruz *et al.*, 2014; Tan *et al.*, 2016; Jin *et al.*, 2018; Li *et al.*, 2019), which, in our view, may not be the most appropriate approach, based on the influence of the parameter range. According to Santelli *et al.* (1999), the principle of the GSA is to quantify the model uncertainty based on perturbations on the environmental variables. In this sense, the use of relative values in the parameter ranges without experimental foundations may not represent well the influence of the parameters of a PBCM. For example, for the variable LAI, *mla* and *maxgl* were observed as the main parameters in PRS1 and PRS2, respectively. The CV of these parameters was 29% (*mla*) and 9% (*maxgl*), in PRS1, and 25% for both parameters in PRS2; for this configuration of PRS1, *maxgl* was not significant, but it was the most important parameter in PRS2. Yet, this can be related to the correlation that exists between the parameters, neglected in many Bayesians and PBCM applications (Marin *et al.*, 2017; Pereira *et al.*, 2021), which demonstrated that any uncertainty analysis using relative values seems to be inadequate.

In SA, it is necessary to use a time interval for one evaluation, as changes in the order of importance of main parameters occurred during the growth of the crop. For example, considering our final SFM yield value, the parameters *popmat* and *plastochron* accounted for between 21.8 and 30.4% of the variation, respectively (Table 7). However, the *popmat* is less relevant than the *plastochron*, as the *plastochron* influenced a longer period of the season, being responsible for explaining up to 60% of the variance in the first 150 days of each ratoon in the PRS2 set (Fig. 5, BARE and GCTB). For the TIL variable, this was even more evident, having distinct parameters and explaining over 90% of the variation in different times. In addition, the occurrence of significant different parameters between ratoon years indicated a possible influence of climate variation (Fig. 7). Several GSA studies have demonstrated the impact of seasonality on sensitivity indices, regardless of the method employed (Vazquez-Cruz *et al.*, 2014; Tan *et al.*, 2016; Sexton *et al.*, 2017; Xing *et al.*, 2017; Li *et al.*, 2019; Attia *et al.*, 2021). In our study, we conducted GSA using a daily time interval from plant cane to the 3rd ratoon and observed that the ranking of parameter importance between ratoons remained consistent (Figs 4–8). Over a short span of 4 years, we did not detect any impact of climate variability on the sensitivity indices. However, there is evidence that climate variability influenced the sensitivity index, as noted by Anderson *et al.*, 2014; Attia *et al.*, 2021; Gilardelli *et al.*, 2018; Sexton *et al.*, 2017. Therefore, for sugarcane models, it is crucial to incorporate time intervals in GSA, and future studies should consider the influence of climate in GSA.

Conclusion

The results showed that the SZ and parameter range were important for GSA, and that an SZ of at least 2049 was required for the sensitivity indices to converge regardless of the variable. However, some variables needed smaller sizes, such as the case of TIL that predicted an SZ of 257. The range of parameters must be carefully investigated, and we demonstrate that the use of relative values without biophysical basis to determine the parameter ranges is inappropriate for the uncertainty analysis, and measured thresholds should always be used, even if from different genotypes, to determine the model's response across the full range of parameters. We did not identify the influence of GCTB and the different ratoons on the order of importance of the parameters, they only slightly affected the values (S_i and ST_i) of the sensitivity indices.

References

Anderson B, Borgonovo E, Galeotti M and Roson R (2014) Uncertainty in climate change modeling: can global sensitivity analysis be of help? *Risk Analysis* **34**, 271–293.

Attia A, El-Hendawy S, Al-Suhaibani N, Tahir MU, Mubushar M, Vianna MdS, Ullah H, Mansour E and Datta A (2021) Sensitivity of the DSSAT model in simulating maize yield and soil carbon dynamics in arid Mediterranean climate: effect of soil, genotype, and crop management. *Field Crops Research* **260**, 107981.

Chu-Agor ML, Muñoz-Carpena R, Kiker G, Emanuelsson A and Linkov I (2011) Exploring vulnerability of coastal habitats to sea level rise through global sensitivity and uncertainty analyses. *Environmental Modelling and Software* **26**, 593–604.

Confalonieri R (2010) Monte Carlo-based sensitivity analysis of two crop simulators and considerations on model balance. *European Journal of Agronomy* **33**, 89–93.

Cukier RI, Levine HB and Shuler KE (1978) Nonlinear sensitivity analysis of multiparameter model systems. *Journal of Computational Physics* **26**, 1–42.

Dejonge KC, Ascough Ii JC, Ahmadi M, Andales AA and Arabi M (2012) Global sensitivity and uncertainty analysis of a dynamic agroecosystem model under different irrigation treatments. *Ecological Modelling* **231**, 113–125.

Dzotsi KA, Basso B and Jones JW (2013) Development, uncertainty, and sensitivity analysis of the simple SALUS crop model in DSSAT. *Ecological Modelling* **260**, 62–76.

Everingham YL, Muchow RC, Stone RC, Inman-Bamber NG, Singels A and Bezuidenhout CN (2002) Enhanced risk management and decision-making capability across the sugarcane industry value chain based on seasonal climate forecasts. *Agricultural Systems* **74**, 459–477.

Everingham Y, Sexton J, Skocaj D and Inman-Bamber G (2016) Accurate prediction of sugarcane yield using a random forest algorithm. *Agronomy for Sustainable Development* **36**, 27.

Faivre R, Leenhardt D, Voltz M, Benoît M, Papy F, Dedieu G and Wallach D (2009) Spatialising crop models BT. In Lichtfouse E, Navarrete M, Debaeke P, Véronique S and Alberola C (eds), *Sustainable Agriculture*. Dordrecht: Springer Netherlands, pp. 687–705.

Fraedrich D and Goldberg A (2000) Methodological framework for the validation of predictive simulations. *European Journal of Operational Research* **124**, 55–62.

Gan Y, Duan Q, Gong W, Tong C, Sun Y, Chu W, Ye A, Miao C and Di Z (2014) A comprehensive evaluation of various sensitivity analysis methods: a case study with a hydrological model. *Environmental Modelling and Software* **51**, 269–285.

Gilardelli C, Confalonieri R, Cappelli GA and Bellocchi G (2018) Sensitivity of WOFOST-based modelling solutions to crop parameters under climate change. *Ecological Modelling* **368**, 1–14.

Guérif M and Duke CL (2000) Adjustment procedures of a crop model to the site-specific characteristics of soil and crop using remote sensing data assimilation. *Agriculture, Ecosystems and Environment* **81**, 57–69.

He J, Dukes MD, Jones JW and Judge J (2009) Applying GLUE for estimating CERES-maize genetic and soil parameters for sweet corn production. *Transactions of the ASABE* **52**, 1907–1921.

He J, Jones JW, Graham WD and Dukes MD (2010) Influence of likelihood function choice for estimating crop model parameters using the generalized likelihood uncertainty estimation method. *Agricultural Systems* **103**, 256–264.

Hirabayashi S, Kroll CN and Nowak DJ (2011) Component-based development and sensitivity analyses of an air pollutant dry deposition model. *Environmental Modelling and Software* **26**, 804–816.

Hoffman N, Singels A, Patton A and Ramburan S (2018) Predicting genotypic differences in irrigated sugarcane yield using the Canegro model and independent trait parameter estimates. *European Journal of Agronomy* **96**, 13–21.

Homma T and Saltelli A (1996) Importance measures in global sensitivity analysis of nonlinear models. *Reliability Engineering and System Safety* **52**, 17.

Inman-Bamber NG and Smith DM (2005) Water relations in sugarcane and response to water deficits. *Field Crops Research* **92**, 185–202.

Iooss B and Lemaître P (2015) A review on global sensitivity analysis methods. *Operations Research/ Computer Science Interfaces Series* **59**, 101–122.

Jeuffroy MH, Barbottin A, Jones JW and Leoeur J (2006) Crop models with genotype parameters. In Wallach D, Makowski D and Jones JW (eds), *Working with Crop Models—Evaluation, Analysis, Parameterization and Applications*. Amsterdam, The Netherlands: Elsevier, pp. 281–308.

Jin X, Li Z, Nie C, Xu X, Feng H, Guo W and Wang J (2018) Parameter sensitivity analysis of the AquaCrop model based on extended Fourier amplitude sensitivity under different agro-meteorological conditions and application. *Field Crops Research* **226**, 1–15.

Jones M and Singels A (2008) *DSSAT v4.5-Canegro Sugarcane Plant Module User Documentation*. Mount Edgecombe, South Africa: South African Sugarcane Research Institute, 56 pp. International Consortium for Sugarcane Modelling (ICSM). http://sasri.sasa.org.za/misc/icsm.html.

Jones MR and Singels A (2018) Refining the Canegro model for improved simulation of climate change impacts on sugarcane. *European Journal of Agronomy* **100**, 76–86.

Jones CA, Kiniry JR and Dyke PT (1986) CERES-Maize: a simulation model of maize growth and development.

Jones MR, Singels A and Ruane AC (2015) Simulated impacts of climate change on water use and yield of irrigated sugarcane in South Africa. *Agricultural Systems* **139**, 260–270.

Jones JW, Antle JM, Basso B, Boote KJ, Conant RT, Foster I, Godfray HCJ, Herrero M, Howitt RE, Janssen S, Keating BA, Munoz-Carpena R, Porter CH, Rosenzweig C and Wheeler TR (2017) Toward a new generation of agricultural system data, models, and knowledge products: state of agricultural systems science. *Agricultural Systems* **155**, 269–288.

Keating BA, Robertson MJ, Muchow RC and Huth NI (1999) Modelling sugarcane production systems I. Development and performance of the sugarcane module. *Field Crops Research* **61**, 253–271.

Krishnan P and Aggarwal P (2018) Global sensitivity and uncertainty analyses of a web-based crop simulation model (Web InfoCrop Wheat) for soil parameters. *Plant and Soil* **423**, 443–463.

Lamboni M, Makowski D, Lehuger S, Gabrielle B and Monod H (2009) Multivariate global sensitivity analysis for dynamic crop models. *Field Crops Research* **113**, 312–320.

Li ZH, Jin XL, Liu HL, Xu XG and Wang JH (2019) Global sensitivity analysis of wheat grain yield and quality and the related process variables from the DSSAT-CERES model based on the extended Fourier amplitude sensitivity test method. *Journal of Integrative Agriculture* **18**, 1547–1561. https://doi.org/10.1016/S2095-3119(18)62046-5.

Ma C, Li X, Wang J, Wang C, Duan Q and Wang W (2017) A comprehensive evaluation of microwave emissivity and brightness temperature sensitivities to soil parameters using qualitative and quantitative sensitivity analyses. *IEEE Transactions on Geoscience and Remote Sensing* **55**, 1025–1038.

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

Makowski D, Naud C, Jeuffroy MH, Barbottin A and Monod H (2006) Global sensitivity analysis for calculating the contribution of genetic parameters to the variance of crop model prediction. *Reliability Engineering & System Safety* **91**, 1142–1147.

Marin FR and Jones JW (2014) Process-based simple model for simulating sugarcane growth and production. *Scientia Agricola* **71**, 1–16.

Marin FR, Jones JW, Singels A, Royce F, Assad ED, Pellegrino GQ and Justino F (2013) Climate change impacts on sugarcane attainable yield in southern Brazil. *Climatic Change* **117**, 227–239.

Marin FR, Thorburn PJ, Nassif DSP and Costa LG (2015) Sugarcane model intercomparison: structural differences and uncertainties under current and potential future climates. *Environmental Modelling and Software* **72**, 372–386.

Marin FR, Jones JW and Boote KJ (2017) A stochastic method for crop models: including uncertainty in a sugarcane model. *Agronomy Journal* **109**, 483–495.

Marino S, Hogue IB, Ray CJ and Kirschner DE (2008) A methodology for performing global uncertainty and sensitivity analysis in systems biology. *Journal of Theoretical Biology* **254**, 178–196.

Morris MD (1991) Factorial sampling plans for preliminary computational experiments. *Technometrics* **33**, 161.

Pereira RAA, Vianna MS, Nassif DSP, Carvalho KS and Marin FR (2021) Global sensitivity and uncertainty analysis of a sugarcane model considering the trash blanket effect. *European Journal of Agronomy* **130**, 126371. https://doi.org/10.1016/j.eja.2021.126371.

Raza G, Ali K, Hassan MA, Ashraf M, Khan MT and Khan IA (2019) Sugarcane as a bioenergy source. *Sugarcane Biofuels*, 3–19. https://doi.org/10.1007/978-3-030-18597-8_1.

Razavi S and Gupta HV (2015) What do we mean by sensitivity analysis? The need for comprehensive characterization of 'global' sensitivity in earth and environmental systems models. *Water Resources Research* **51**, 3070–3092.

Reusser DE, Buytaert W and Zehe E (2011) Temporal dynamics of model parameter sensitivity for computationally expensive models with the Fourier amplitude sensitivity test. *Water Resources Research* **47**, W07551.

Ritchie JT and Otter S. (1985) Description and performance of CERES-Wheat: A user-oriented wheat yield model. In *ARS Wheat Yield Project*. ARS-38. Springfield, VA: National Technical Information Service (NTIS), pp. 159–175.

Saltelli A, Tarantola S and Chan KPS (1999) A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics* **41**, 39–56.

Saltelli A, Annoni P, Azzini I, Campolongo F, Ratto M and Tarantola S (2010) Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer Physics Communications* **181**, 259–270.

Sexton J, Everingham Y and Inman-Bamber G (2016) A theoretical and real-world evaluation of two Bayesian techniques for the calibration of variety parameters in a sugarcane crop model. *Environmental Modelling and Software* **83**, 126–142.

Sexton J, Everingham YL and Inman-Bamber G (2017) A global sensitivity analysis of cultivar trait parameters in a sugarcane growth model for contrasting production environments in Queensland, Australia. *European Journal of Agronomy* **88**, 96–105.

Sheng M, Liu J, Zhu AX, Rossiter DG, Liu H, Liu Z and Zhu L (2019) Comparison of GLUE and DREAM for the estimation of cultivar parameters in the APSIM-maize model. *Agricultural and Forest Meteorology* **278**, 107659.

Sinclair TR and Seligman NG (1996) Crop modeling: from infancy to maturity. *Agronomy Journal* **88**, 698–704.

Singels A (2013) Crop models. In Morre PH and Botha FC (eds), *Sugarcane: Physiology, Biochemistry, and Functional Biology*, John Wiley & Sons, Inc., pp. 541–577. https://doi.org/10.1002/9781118771280.ch20.

Singels A and Bezuidenhout CN (2002) A new method of simulating dry matter partitioning in the Canegro sugarcane model. *Field Crops Research* **78**, 151–164.

Singels A, Jones M, Marin F, Ruane A and Thorburn P (2013) Predicting climate change impacts on sugarcane production at sites in Australia, Brazil and South Africa using the Canegro model. *Sugar Tech: An International Journal of Sugar Crops & Related Industries* **16**, 347–355.

Sobol IM (2001) Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates. *Mathematics and Computers in Simulation* **55**, 271–280.

Soetaert K and Petzoldt T (2010) Inverse modelling, sensitivity, and Monte Carlo analysis in R using package FME. *Journal of Statistical Software* **33**, 1–28.

Song X, Zhang J, Zhan C, Xuan Y, Ye M and Xu C (2015) Global sensitivity analysis in hydrological modeling: review of concepts, methods, theoretical framework, and applications. *Journal of Hydrology* **523**, 739–757.

Sreelash K, Sekhar M, Ruiz L, Tomer SK, Guérif M, Buis S, Durand P and Gascuel-Odoux C (2012) Parameter estimation of a two-horizon soil profile by combining crop canopy and surface soil moisture observations using GLUE. *Journal of Hydrology* **456**, 57–67.

Tan J, Cui Y and Luo Y (2016) Global sensitivity analysis of outputs over rice-growth process in ORYZA model. *Environmental Modelling and Software* **83**, 36–46. https://doi.org/10.1016/j.envsoft.2016.05.001

Thorburn PJ, Meier EA and Probert ME (2005) Modelling nitrogen dynamics in sugarcane systems: recent advances and applications. *Field Crops Research* **92**, 337–251.

Valade A, Ciais P, Vuichard N, Viovy N, Caubel A, Huth N, Marin F and Martiné J-F (2014) Modeling sugarcane yield with a process-based model from site to continental scale: uncertainties arising from model structure and parameter values. *Geoscientific Model Development* **7**, 1225–1245.

Varella H, Guérif M and Buis S (2010) Global sensitivity analysis measures the quality of parameter estimation: the case of soil parameters and a crop model. *Environmental Modelling and Software* **25**, 310–319.

Vazquez-Cruz MA, Guzman-Cruz R, Lopez-Cruz IL, Cornejo-Perez O, Torres-Pacheco I and Guevara-Gonzalez RG (2014) Global sensitivity analysis by means of EFAST and Sobol' methods and calibration of reduced state-variable TOMGRO model using genetic algorithms. *Computers and Electronics in Agriculture* **100**, 1–12.

Vianna MdS, Nassif DSP, dos Santos Carvalho K and Marin FR (2020) Modelling the trash blanket effect on sugarcane growth and water use. *Computers and Electronics in Agriculture* **172**, 105361.

Villegas FD, Daza OH, Jones JW and Royce FS (2005) CASUPRO: an industry-driven sugarcane model. In 2005 ASAE Annual Meeting, p. 1. American Society of Agricultural and Biological Engineers.

Wallach D, Makowski D, Jones JW and Brun F (2019a) Data assimilation for dynamic models. *Working with Dynamic Crop Models*, 487–518. https://doi.org/10.1016/B978-0-12-811756-9.00014-9.

Wallach D, Makowski D, Jones JW and Brun F (2019b) Uncertainty and sensitivity analysis. *Working with Dynamic Crop Models*, 209–250. https://doi.org/10.1016/B978-0-12-811756-9.00006-X.

Wang J, Li X, Lu L and Fang F (2013) Parameter sensitivity analysis of crop growth models based on the extended Fourier amplitude sensitivity test method. *Environmental Modelling and Software* **48**, 171–182.

Wei T (2013) A review of sensitivity analysis methods in building energy analysis. *Renewable and Sustainable Energy Reviews* **20**, 411–419.

Xing Hm, Xu Xg, Li Zh, Chen Yj, Feng Hk, Yang Gj and Chen Zx (2017) Global sensitivity analysis of the AquaCrop model for winter wheat under different water treatments based on the extended Fourier amplitude sensitivity test. *Journal of Integrative Agriculture* **16**, 2444–2458.

Xu C and Gertner G (2011) Understanding and comparisons of different sampling approaches for the Fourier amplitudes sensitivity test (FAST). *Computational Statistics and Data Analysis* **55**, 184–198.

Zhang Y, Arabi M and Paustian K (2020) Analysis of parameter uncertainty in model simulations of irrigated and rainfed agroecosystems. *Environmental Modelling and Software* **126**, 104642.