A Gaussian-Process-Based Global Sensitivity Analysis of Cultivar Trait Parameters in APSIM-Sugar Model: Special Reference to Environmental and Management Conditions in Thailand

W. B. M. A. C. Bandara 1,2,*, Kazuhito Sakai 1,3,*, Tamotsu Nakandakari 1,3, Preecha Kapetch 4 and R. H. K. Rathnappriya 1

1 Faculty of Agriculture, University of the Ryukyus, 1 Senbaru, Nishihara-cho, Okinawa 903-0213, Japan; zhunai@agr.u-ryukyu.ac.jp (T.N.); himashakithminir@gmail.com (R.H.K.R.)
2 Department of Agricultural Engineering, Faculty of Agriculture, University of Ruhuna, Kamburupitiya 81100, Sri Lanka
3 United Graduate School of Agricultural Sciences, Kagoshima University, 1-21-24 Korimoto, Kagoshima-shi, Kagoshima 890-0065, Japan
4 Nakhon Sawan Agricultural Research and Development Center, Moo 2, Udomthanya, Takfa 60190, Thailand; p.kapetch@gmail.com
* Correspondence: chathu.anushk@gmail.com or cbandara@ageng.ruh.ac.lk (W.B.M.A.C.B.);
ksakai@agr.u-ryukyu.ac.jp (K.S.)

Received: 22 June 2020; Accepted: 7 July 2020; Published: 9 July 2020

Abstract: Process-based crop models are advantageous for the identification of management strategies to cope with both temporal and spatial variability of sugarcane yield. However, global optimization of such models is often computationally expensive. Therefore, we performed global sensitivity analysis based on Gaussian process emulation to evaluate the sensitivity of cane dry weight to trait parameters implemented in the Agricultural Productions System Simulator (APSIM)-Sugar model under selected environmental and management conditions in Khon Kaen (KK), Thailand. Emulators modeled 30 years, three soil types and irrigated or rainfed conditions, and emulator performance was investigated. rue, green_leaf_no, transp_eff_cf, tt_emerg_to_begcane and cane_fraction were identified as the most influential parameters and together they explained more than 90% of total variance on the simulator output. Moreover, results indicate that the sensitivity of sugarcane yield to the most influential parameters is affected by water stress conditions and nitrogen stress. Our findings can be used to improve the efficiency and accuracy of modeling and to identify appropriate management strategies to address temporal and spatial variability of sugarcane yield in KK.

Keywords: APSIM; Gaussian process emulation; global sensitivity analysis; sugarcane

1. Introduction

Sugarcane plays a critical role in Thailand’s economy and has become one of the most important agricultural crops of the country [1]. Being the major sugarcane production region of Thailand, the Northeast is responsible for 43.2% of the total produced sugarcane and 44.2% of the total sugarcane harvesting area [2]. Recently, paddy fields that produce lower net value per hectare in the Khon Kaen (KK) area of the Northeast have been converted into sugarcane fields [3]. Increasing evidence indicates that global climate change could reduce sugarcane production. According to Preecha et al. [4], climate change is the most obvious factor responsible for spatial and temporal yield variability in the Northeast of Thailand. Thus, identification of suitable management strategies to cope with both temporal and
spatial variability is of a paramount importance. For instance, sugar mills require forecasting and estimation of cane yield to manage their strategies.

In this respect, it is advantageous to study how different cultivars perform under different environmental and management conditions. Process-based crop models that can simulate cultivar differences are used by researchers to simulate how the cultivars perform under various production environments and to identify advantageous traits in defined environments [5]. However, recent advances in crop models for cultivar–environment interaction studies have created a requirement for quantifying and analyzing uncertainty in crop models. For instance, Ojeda et al. [6] has quantified the input uncertainty for their study on assessing effect of data aggregation in regional scale crop modeling. Sensitivity analysis (SA) is useful in studying how the uncertainty of the model input affects the uncertainty of the model output and to what extent model outputs are sensitive to model parameters [7].

Song et al. [8] suggested a way of dividing SA into local and global SA. Local one-at-a-time sensitivity indices are efficient if linear output responses are produced by all the factors in a model. In general, as explained by Ewert et al. [9] variations in input factors generate non-linear model output responses. Therefore, an alternative global SA (GSA) approach is required, in which the whole model parameter space is analyzed for all input factors at once [10]. In comparison with local SA, GSA can provide a better understanding of how cultivar parameters influence the simulated output [11], because GSA ranks parameters according to their importance, and generate information about main and interaction effects of individual parameters on output [12].

Various GSA methods have been used for process-based crop models (e.g., Fourier amplitude sensitivity test (FAST) [13], random-based-design FAST and extended FAST [14], Sobol method [15–17]), which all operate by separating the variance of the model output into different groups according to sources of input variation. However, because process-based crop models are often computationally expensive, carrying out the required number of simulations may not be feasible and SA may be extremely time consuming [18,19]. A widely used solution is the statistical approximation of a simulator by generating a meta-model [20,21], which is called an emulator [22]. Running the emulator is computationally less expensive because it is simplified relative to the actual simulator. The original simulator can be substituted by an emulator of sufficient accuracy (cross-validated root-mean-squared standardized error (RMSSE) close to 1.0), and SAs can be based on the emulator [20,23].

Emulators are usually implemented as Gaussian process (GP) regression models that use a finite set of design points to approximate the simulator mapping [24]. GP emulators are a category of surrogate models, and a detailed discussion of the theory and implementation of GP emulation can be found in Kennedy and O’Hagan [7] and Rasmussen and Williams [25]. Sexton et al. [11,26] and Gunarathna et al. [27] have used GP for GSA of trait parameters used in the Agricultural Productions System Simulator (APSIM)-Sugar model. These studies have emphasized the need to study the influence of sugarcane cultivar parameters under various environmental and management conditions.

Here, we assessed the sensitivity of the model output (cane dry weight, CDW) to trait parameters used in the APSIM-Sugar model under different environmental and management conditions in KK (three soil types, and irrigated (Ir) or rainfed (Rf) conditions) using emulator-based GSA. As suggested by Sexton et al. [11], we considered the effect of soil and climate interactions on trait parameters.

2. Materials and Methods

2.1. Study Field

KK, northeast Thailand (16.48° N 102.82° E; 181 m elevation), was selected for the study. Climate in KK is classified as Aw (tropical wet-dry climate) by the Köppen-Geiger system [28]. Study was conducted based on crop performance of sugarcane under different environmental and management conditions in KK in years between 1980 and 2010. Figure 1 shows average, mean monthly rainfall, mean daily maximum and minimum temperatures and mean daily solar radiation of each month between
1980–2010 in KK. We observed similar values of mean daily maximum and minimum temperatures and mean daily solar radiation among the years. However, mean monthly rainfall values were highly varied among years. Textural classes and physical and chemical properties of each selected soil type are shown in Table 1. Available water content varies as; S1 > S46 > S44 (Table 1).

![Graph showing average monthly climatic data of Khon Kaen (KK) between 1980–2010](image)

**Figure 1.** Average monthly climatic data of Khon Kaen (KK) between 1980–2010; Rain: mean monthly rainfall (mm); Radn: mean daily solar radiation (MJ/m²); Maxt: mean daily maximum temperature (°C); Mint: mean daily minimum temperature (°C).

**Table 1.** Physical and chemical properties of selected soil types of KK [4].

| Soil Group | Soil Depth (cm) | Texture Class | Wilting Point (mm/mm) | Field Capacity (mm/mm) | Hydraulic Conductivity (cm/h) | Bulk Density (g/cm³) | Clay % | Silt % | Sand % | pH |
|------------|-----------------|---------------|-----------------------|------------------------|-------------------------------|----------------------|-------|-------|--------|----|
| S1         | 0–100           | Clay soil     | 0.328                 | 0.461                  | 0.06                          | 1.44                 | 68.0  | 29.0  | 3      | 5.4 |
| S44        | 0–100           | Sandy soil    | 0.038                 | 0.120                  | 13.34                         | 1.7                  | 1     | 9.5   | 89.5   | 5.6 |
| S46        | 0–100           | Clay loam     | 0.133                 | 0.231                  | 0.36                          | 1.52                 | 29.2  | 29    | 41.8   | 5.1 |

* Soil texture classes according to the USDA Soil textural triangle [29].

2.2. APSIM Simulation

APSIM [30] is a modeling platform that can be used to simulate the performance of a single crop or a cropping system under different soil and climatic conditions and permits evaluation of management interventions via tillage, fertilization, irrigation and selection, timing and crop sequencing (in fixed or flexible rotations) [31]. For instance, Ojeda et al. [32] has used APSIM for forage crop by considering crop sequences.

The APSIM 7.10 Sugar model was used for the simulations. APSIM-Sugar uses radiation use efficiency (rue) to simulate CDW accumulation by converting intercepted radiation into biomass [11]. On the basis of the crop phenological stage, biomass is partitioned among different plant components (sucrose, leaf, structural stem, cabbage and roots). It uses six different crop phenological stages to define sugarcane growth, i.e., “sowing (from sowing to sprouting), sprouting (from sprouting to emergence), emergence (from emergence to the beginning of cane growth), begin cane (from the beginning of cane growth to flowering), flowering (from flowering to the end of the crop) and the end of the crop (crop is not currently in the simulated system)” [30]. The model is designed for the simulation of a uniform cane field using daily time-steps, and predicts cane yield, crop biomass, sucrose yield, commercial sucrose concentration, water use and crop nitrogen uptake on an area basis [33]. Cultivar-specificity,
CDW of sugarcane plant-crop at harvest was simulated for 30 years (1980–2010) for three selected soil types under Ir or Rf conditions to study soil and climate interactions on trait parameters. Management criteria used for the APSIM simulation setup are indicated in Table 2. Planting date was selected as 28 November of each year in accordance with the previous simulation study [4] to represent realistic management practices in the region. Soil data (Table 1) and daily weather data for KK from 1980 to 2010 collected by Preecha et al. [4] were used for the simulations.

| Criteria                        | Value                                      |
|---------------------------------|--------------------------------------------|
| Planting date                   | November 28                                |
| Crop duration                   | 360 days                                   |
| Stalk density                   | 6.8 stalks/m²                              |
| Planting depth                  | 100 mm                                     |
| Fertilizer application          |                                            |
| Fertilization at planting       | Urea_N—46.75 kg/ha                        |
| Fertilization at 100 days after planting | Urea_N—46.75 kg/ha                        |
| Water supply by irrigation      |                                            |
| Rainfed condition (Rf)          | 24 mm of irrigation at 7, 14, 21 and 28 days after planting a |
| Irrigation condition (Ir)       | 24 mm of irrigation with 7 days’ time interval from planting to end date of the simulation b |

* The irrigation schedule was manually induced; a to ensuring the crop establishment, b based on the actual management practices. The amount of irrigation (24 mm) was assumed with the purpose of avoiding water stress and to confirm the difference between Ir and Rf conditions in the simulation. Therefore, Irrigation efficiency is considered as one.

2.3. Sensitivity Analysis

2.3.1. Preparation of Training Design Points

Parameters which control the underlying biophysical process of sugarcane growth in APSIM-Sugar are categorized into cultivar specific parameters, plant and ratoon class parameters and soil and climate parameters [11]. Table 3 shows the cultivar-specific parameters implemented in APSIM-Sugar which we used in this study. Leaf development (leaf size, green_leaf_no, tiller leaf size), phenological development based on thermal time (tt_emerge_to_begcane, tt_begcane_to_flowering, tt_flowering_to_crop_end) and partitioning of assimilates (cane_fraction, sucrose_fraction_stalk, stress_factor_stalk, sucrose_delay, min_sstem_sucrose, min_sstem_sucrose_redn) are controlled by cultivar specific parameters. Parameters such as green_leaf_no are directly related to express the cultivar traits [11]. Some of the traits are expressed via a parameter combination. For instance, parameters such as leaf_size and leaf_size_no (position of the leaf along stalk) together control canopy development [35]. Although, parameters which are related to later phenological development stages such as tt_begcane_to_flowering and tt_flowering_to_crop_end included in APSIM-sugar, they remain deactivated until a better physiological basis for prediction is available [33,35]. Both rue and transp_eff Cf are not considered as cultivar specific parameters in APSIM-Sugar [11]. However, rue and transp_eff Cf were included in the analysis. In APSIM, dry matter assimilation is governed by radiation interception and rue in the conditions which soil water availability is not limited. If the soil water supply is not enough to meet the transpiration demand, dry matter assimilation is governed by water supply, transp_eff Cf and the vapor pressure deficit. Moreover, results of SA studies conducted by Sexton et al. [11,35], Gunaratnaya et al. [27] and Sexton and Everingham [26] have indicated that both rue and transp_eff Cf may improve simulations based on cultivar differences.
Table 3. Description of the trait parameters and parameter space [36].

| Parameter Name | Description | Level | Code | Units | Range |
|----------------|-------------|-------|------|-------|-------|
| leaf_size      | Leaf area of the respective leaf | leaf_size_no = 1 | LS1 | mm² | 500–2000 |
|                |             | leaf_size_no = 14 and 20 | LS2&3 | mm² | 25,000–70,000 |
| cane_fraction  | Fraction of accumulated biomass partitioned to cane | CF | g/g | 0.65–0.80 |
| sucrose_fraction_stalk | Fraction of accumulated biomass partitioned to sucrose | SF1 | g/g | 0.50–0.70 |
| stress_factor_stalk | Stress factor for sucrose accumulation | SF2 | n/a | 0.2–1.0 |
| sucrose_delay  | Sucrose accumulation delay | SD | g/m² | 0–600 |
| min_sstem_sucrose | Minimum stem biomass before partitioning to sucrose commences | MSS | g/m² | 450–1500 |
| min_sstem_sucrose_redn | Reduction to minimum stem sucrose under stress | MSSR | g/m² | 0–20 |
| tt_emerg_to_begcane | Accumulated thermal time from emergence to beginning of cane | EB | °C day | 1200–1900 |
| tt_begcane_to_flowering | Accumulated thermal time from beginning of cane to flowering | BF | °C day | 5500–6500 |
| tt_flowering_to_crop_end | Accumulated thermal time from flowering to end of the crop | FC | °C day | 1750–2250 |
| green_leaf_no  | Maximum number of fully expanded green leaves | GLN | No. | 9–14 |
| tillerf_leaf_size | Tilling factors according to the leaf numbers | Tiller_leaf_size_no = 1 | TLS1 | mm²/mm² | 1–6 |
|                |             | Tiller_leaf_size_no = 4 | TLS2 | mm²/mm² | 1–6 |
|                |             | Tiller_leaf_size_no = 10 | TLS3 | mm²/mm² | 1–6 |
|                |             | Tiller_leaf_size_no = 16 | TLS4 | mm²/mm² | 1–6 |
|                |             | Tiller_leaf_size_no = 26 | TLS5 | mm²/mm² | 1–6 |
| transp_eff Cf  | Transpiration efficiency coefficient | TEC | kg kPa/kg | 0.008–0.014 |
| rue            | Radiation use efficiency | RUE | g/MJ | 1.2–2.5 |

Initially, 500 parameter combinations were generated based on the parameter ranges indicated in Table 3 using APSIM package [37] of R software [38]. These training design points were generated as uniform random numbers distributed between the minimum and maximum values of selected cultivar trait input parameters (listed in Table 3). Minimum and maximum values of parameters were selected based on available literature on previous research and APSIM-Sugar documentation (Table 3).

The ranges for leaf_size, cane_fraction, sucrose_fraction_stalk, stress_factor_stalk, sucrose_delay, min_sstem_sucrose, min_sstem_sucrose_redn, tt_emerg_to_begcane, tt_begcane_to_flowering, tt_flowering_to_crop_end, green_leaf_no and tillerf_leaf_size were selected based on APSIM-Sugar documentation [36], Sexton and Everingham [26] and Sexton et al. [11,35].

According to Sinclair [39], transp_eff Cf ranges between 0.009–0.010 kg kPa/kg for C4 plants like sugarcane. A recent research conducted by Jackson et al. [40] has reported that under water stressed conditions, higher transpiration efficiency could be identified for sugarcane cultivars. Therefore, in order to represent the response of sugarcane to water stressed conditions, the present study has used the range of transp_eff Cf as 0.008–0.014 kg kPa/kg following Gunarathna et al. [27].
In the APSIM-Sugar model, \( r_{ue} \) parameter is fixed as 1.8 g/MJ for plant crops and 1.65 g/MJ for ratoon crops [36]. It is reported that a considerable variation in \( r_{ue} \) could be occurred due to temperature variations, soil water deficit or excess [33], crop class and age, lodging, soil fertility (Nitrogen deficit) and culm death [41]. Considering the intercepted photosynthetically active radiation, Ferreira et al. [42] have found that \( r_{ue} \) of sugarcane in single and combined spacing as 2.73 (±0.09) and 2.78 (±0.25) g/MJ, respectively. Further, Olivier et al. [43] have reported \( r_{ue} \) value of 1.75 g/MJ for N19 sugarcane variety. Meki et al. [44] have obtained \( r_{ue} \) value of 2.06 g/MJ for their study on modeling of specific crop parameter attributes of two-year sugarcane growth cycle. Hence, by considering values of previous studies, we have used 1.2–2.5 g/MJ as the range for \( r_{ue} \).

Above mentioned parameter combinations were then simulated in APSIM-Sugar for 30 years under three selected soil types and Ir or Rf conditions (described in Section 2.2). APSIM output values (CDW) corresponding to each parameter combination and environmental and management condition (180 APSIM output files and each file including 500 outputs) were obtained from the simulations. Both parameter combination files and corresponding APSIM output files were used as training design points for emulator generation and validation during the GSA.

2.3.2. Gaussian-Process-Based Global Sensitivity Analysis

To conduct SA for complex simulation models, an increasing number of studies have focused on model emulation. According to Villa-Vialaneix et al. [45], it is a common approach to use GP when generating emulators [20,46], even though other options are available as well. GP can be defined as a distribution for a function. According to O’Hagan et al. [46], each value of a function has a normal distribution, and a set of function values has a multivariate normal distribution. Therefore, GPs and the normal distribution both have equal mathematically convenient properties. During emulator building, the original model is described by assigning a GP prior, and then the prior is updated using a series of model runs by applying the Bayes theorem. The emulator is the resulting posterior distribution [47].

SA was conducted by using GP-based emulation implemented in the Gaussian emulation machine for sensitivity analysis (GEM-SA) software package [47]. A more detailed description of the mathematics underlying GEM-SA can be found in Kennedy and O’Hagan [7] and Oakley and O’Hagan [22]. The GEM-SA package calculates two variance-based sensitivity indices, the main-effect index (\( S_i \)) and the total-effect index (\( ST_i \)).

Equation (1) defines the main-effect index as:

\[
S_i = \frac{\text{Var}[E(f(X)|x_i)]}{\text{Var}[f(X)]}
\]

where “Var\( [f(X)] \) is the total variance in the output given variations of all parameters and Var\( [E(f(X)|x_i)] \) is the variance in the expected output of \( f(X) \) given \( x_i \). Therefore, \( S_i \) represents the expected reduction in output variance if parameter \( x_i \) is known” [22].

Equation (2) defines the total-effect index as:

\[
TS_i = 1 - \frac{\text{Var}[E(f(X)|x_i)]}{\text{Var}[f(X)]}
\]

where “Var\( [E(f(X)|x_i)] \) is the variance in the expected output of \( f(X) \) if all parameters except \( x_i \) are known. Therefore, \( TS_i \) represents the total effect of the parameter \( x_i \) and all its interactions. If parameters are linearly additive (i.e., no strong interactions), \( S_i \) and \( ST_i \) should be equal” [22].

Prepared training design points for the outputs of the APSIM-Sugar and the parameter combinations (described in Section 2.3.1) were run in GEM-SA. 300 training points of APSIM outputs and corresponding parameter combinations were used to generate emulators. 180 emulators (30 × 3 × 2) for combinations of each year (30), each soil type (3) and Ir or Rf condition were generated. While developing the emulators, remaining 200 parameter combinations were used in GEM-SA to obtain
emulator predictions (emulator predicted CDW). These predictions were graphed with remaining APSIM-output training design points and coefficient of determination ($R^2$) were calculated to evaluate how well emulator can predict the APSIM simulator outputs. $R^2$ range between 1 to 0 and emulators with higher accuracy can be identified with $R^2$ values close to one.

When running the GEM-SA, linear term for each input was set as the prior mean option for the outputs, as it allows us to observe the output trend in response to input changes. Leave-one-out cross-validation was selected to evaluate the accuracy of emulators built by GEM-SA. GEM-SA calculated $S_i$ and $TS_i$ values for each parameter corresponding to each emulator were recorded. Further, GEM-SA calculates leave-one-out cross-validated RMSSE (Equation (3)) and sigma-squared value ($\sigma^2$) for each emulator [48]; these values were used to represent the performance of model emulators with reference to environmental and management conditions in addition to $R^2$.

Equation (3) defines cross-validated RMSSE as:

$$CVRMSE = \sqrt{\frac{\sum_{i=1}^{n}((y_i - \hat{y})/s_i)^2}{n}}$$

where, “$y_i$ is the true output for the $i$th training run, $\hat{y}$ is the corresponding emulator approximation, $s_i$ is the standard deviation calculated with the $i$th training point removed and $n$ is the number of runs” [48].

The cross-validated RMSSE is close to 1 if the actual error variance is accurately estimated by the emulator variance [49], while lower and higher values indicate overestimation and underestimation, respectively [50].

The $\sigma^2$ value is an effective measure that indicates an emulator’s non-linearity by expressing emulator variance after output standardization [51]. The values of $\sigma^2$ ranges near 0 for a linear model and has greater values (without a defined cutoff value) for moderately to highly nonlinear models [49].

The sensitivity of the model outputs to cultivar parameters was explored using stacked column bar chart with $S_i$ indices. Parameters with the strongest effects on simulated CDW were identified from $S_i$ values and were further examined with reference to environmental and management conditions in KK by using main effect plots.

3. Results and Discussion

3.1. Emulator Accuracy

$R^2$ calculated by using the APSIM simulated CDW and emulator predicted CDW, and GEM-SA internally calculated $\sigma^2$ and cross-validated RMSSE values were used to evaluate the performances of generated emulators. Scatter plots of Figure 2 indicate the linear relationship between the APSIM simulated CDW and emulator predicted CDW. As it is complicated to show all the graphs, here we present only some examples to represent all the conditions. However, calculated $R^2$ for all conditions were ranged between 0.85 to 0.99 and closer to one indicating that all emulators can successfully approximate the APSIM simulators.

The calculated $\sigma^2$ values of all emulators ranged between 0.08 and 0.89 (Figure 3a). Petropoulos et al. [51] obtained $\sigma^2$ values ranging from 0.13 to 1.6 for their emulators and concluded that their parameters deviated only moderately from linearity. Gunarathna et al. [27] obtained $\sigma^2$ values ranging from 0.10 to 1.43 and concluded that their models showed good to moderate linearity. Hence, we can conclude that our emulators showed good linearity in each environmental and management condition.
Graphical representation of the relationships between APSIM simulated sugarcane dry weight (CDW) and emulator predicted CDW for different soil types (S1, S44, and S46) and irrigation conditions (Ir and Rf) for years 10 and 20. The graphs show high correlation coefficients ($R^2$) ranging from 0.96 to 0.99, indicating a strong linear relationship between simulated and emulator predicted values.

**Figure 2.** Relationship between APSIM simulated cane dry weight (CDW) and emulator predicted CDW values. The plots illustrate the strong linear fits for all conditions, with $R^2$ values ranging from 0.96 to 0.99, indicating excellent model performance.

**Figure 3.** Box plots of (a) $\sigma^2$ and (b) cross-validated root-mean-squared standardized error (RMSSE) values for the emulator build across different soil types (S1, S44, and S46) and irrigation conditions (Ir and Rf). The plots show the distribution of $\sigma^2$ and RMSSE values, with whiskers indicating 1.5 times the interquartile range (IQR) and black points representing outliers.

Computed cross-validated RMSSE values of emulators ranged between 0.82 and 1.21 (Figure 3b). These values were lower than the values reported by Kennedy et al. [47] and Petropoulos et al. [51], indicating improved model performance.
and were close to one in all the SA experiments, suggesting that the true model can be well represented by the generated emulators.

3.2. Determination of Parameter Sensitivity

Studying the sensitivity of model outputs to cultivar parameters under different environmental and management conditions would help to improve the calibration efficiency of the model. Moreover, when determining the appropriate management practices for sugarcane cultivation, it is important to consider parameters that strongly affect sugarcane yield. Therefore, to determine parameter sensitivity across environmental and management conditions, we examined $S_i$ and $ST_i$ computed by GEM-SA. However, we disregard the $ST_i$ values because of the observation of less difference among $S_i$ and $ST_i$ values and a greater fraction of variability being explained by $S_i$. The $S_i$ values of each parameter for all simulated conditions are shown in Figure 4. In order to explain the differences of $S_i$ among each condition APSIM simulated CDW was included (Figure 4).

Figure 4. Parameter $S_i$ values and CDW of soil type S1, S44 and S46, under Ir and Rf conditions across 30 simulated years from 1980–2010.
Based on the $S_i$ values of 30 years of each soil type under Ir or Rf conditions, $\text{rue}$ (RUE), $\text{green leaf no}$ (GLN), $\text{transp eff cf}$ (TEC), $tt_{\text{emerg to becane}}$ (EB) and $\text{cane fraction}$ (CF) were identified as the most influential parameters on CDW (these parameters together explained >90% of the variability of CDW) while, TLS5, LS2 and 3, TLS4, SF1, BF, MSSR, MSS, FC, SD, TLS2, TLS1, TLS3, LS1, SF2 were identified as the insensitive parameters (each parameter explained <1% of the variability of CDW). (Parameters were defined in Table 3). However, our results are evidence that $S_i$ values of influential parameters are widely differed across simulated years under selected environmental and management conditions (Figure 4).

For CDW, RUE was the most influential parameter under each environmental and management condition in KK except in few years (Figure 4). Ojeda et al. [52] has also found similar results for SA of APSIM-Sugar. In APSIM-Sugar, CDW accumulation is largely governed by intercepted radiation via RUE. Higher sensitivity of RUE is not surprising because as a tropical country, Thailand receives high solar radiation throughout the year (Figure 1).

Nitrogen stress arise due to higher rainfall affects to reduce the sensitivity of RUE and cause for the reduction of CDW. Under Ir conditions, we could observe lower $S_i$ of RUE and lower CDW only for year 1982 than the other years of Ir (Figure 4). For this condition under S1, RUE became the lowest and CF and EB became more sensitive than RUE. In APSIM-Sugar RUE is affected from temperature extremes, excess or shortage of available water or nitrogen deficit for photosynthesis [11]. Therefore, we analyzed these factors using APSIM and could find that nitrogen deficit for photosynthesis ($nfact_{\text{photo}}$) in the year 1982 between April to November. Figure 5 indicates APSIM simulated $nfact_{\text{photo}}$ under S1 and Ir as this condition represents the most severe nitrogen stress during the period (Y2 in Figure 5). Though we simulated similar urea application among all conditions, the nitrogen stress occurred due to higher rainfall prevailing during these months.

![Figure 5. APSIM simulated nitrogen stress factor (1 = no stress and 0 = full stress) for photosynthesis ($nfact_{\text{photo}}$) under S1 soil type under Ir conditions across 30 simulated years from 1980–2010. Each Y no. indicates 360 days of crop duration from planting date (November 28) to the harvesting date.](image)

Moreover, results prove that severity of nitrogen stress arise due to higher rainfall, varies with the soil type. When compared to S1 under Ir of year 1982, $S_i$ of RUE became higher under S46 and the highest under S44 (Figure 4). This is because soil water content in selected soil types are varied according to textural properties; S1 (clay 68%, silt 29%, sand 3%) > S46 (clay 29.2%, silt 29%, sand 41.8%) > S44 (clay 1%, silt 9.5%, sand 89.5%) (Table 1). As indicated in Figure 4 this will largely reduce
the CDW and therefore it is crucial to manage nitrogen application when molding higher rainfall periods to reduce the nitrogen stress specially in S1 soil type.

It is observed that sensitivity of RUE reduced with available water content. Sensitivity of RUE became lower in Rf condition than in Ir condition (Figure 4). Under Rf, sensitivity of RUE was the highest in S1 and weakened in S46 and S44 soils. This is because the available soil water content in selected soil types are varied; S1 > S46 > S44 (Table 1). This was more evident in years which represent lower annual rainfall (year: 1981, 1984, 1991, 1992, 1993 and 2006 in Figure 4) than the other years during the study period.

For CDW, TEC was the second most influential parameter under each environmental and management condition in KK based on average S_i values across study period. However, the sensitivity of TEC is higher in Rf than Ir (Figure 4) indicating that TEC is highly sensitive to water stressed conditions. Sexton and Everingham [26] has also found similar results for their study. This is because in APSIM, dry matter assimilation is governed by radiation interception and RUE in the conditions which soil water availability is not limited. However, in case the soil water supply is not enough to meet the transpiration demand, dry matter assimilation is governed by water supply, TEC and the vapor pressure deficit.

GLN is highly influential under water stressed conditions. GLN indicated higher S_i under Rf than Ir. Although GLN was the third most influential parameter based on the average S_i values, it became the second most influential one in the year of 1981, 1984, 1991, 1992, 1993 and 2006 (Figure 4). These years indicated lower rainfall compared to other years and under Rf condition water stress becomes more severe. Higher water stresses may create leaf emergence rate reduction and leaf senescence rate increment, causing significant reduction in GLN and reduce CDW [53]. This was more evident in our results of year 1991, 1992 and 1993 (the years with lowest rainfall) (Figure 4) under S44 (soil type with lowest water availability) and Rf.

For CDW, CF was the fourth most influential parameter and EB was the fifth most influential parameter. Both indicated higher S_i for Rf than Ir indicating high sensitivity for water stresses (Figure 4). In addition, they indicated high sensitivity for year 1982 under S1 and Ir which we previously identified as nitrogen stressed condition. This is not surprisingly because in APSIM, water deficit and nitrogen deficit both cause for limiting the biomass partitioning in the stem (CF) and phenological development based on thermal time (EB).

Computed S_i indicated that sensitivity of RUE, GLN, TEC, EB and CF explains more than 90% of total variance for most of the simulator outputs across all simulated years, while other parameters had much weaker effects (Figure 4). Similar studies on cultivar-by-environment interactions conducted by Sexton et al. [11] and Gunarathna et al. [27] also found these parameters among highly influential parameters under their selected environmental and management conditions. Therefore, when modeling CDW using APSIM-Sugar those influential parameters can be used to calibrate the model. When such calibrations are streamlined, non-influential (low S_i) parameters could be fixed to default values. Parameters such as RUE and TEC are ideal for statistical calibration of APSIM-Sugar as they are difficult in measuring. By measuring comparatively simple-to-obtain parameters like GLN, it can be reduced the number of parameters used for calibration.

3.3. Sensitivity of Highly Influential Parameters

Knowledge of sensitive parameters is needed for the improvement of simulations of sugarcane growth under various environmental and management conditions. Therefore, we further analyzed the response of outputs to selected environmental and management conditions. The response of CDW to the highly influential parameters (CF, EB, GLN, TEC and RUE) was visualized by plotting the mean of the emulator’s main effects from 6000 randomly selected iterations (Figure 6).
Statistical calibration of RUE and TEC parameters would improve the simulation of cultivar differences in CDW. When parameter value of RUE increased from 1.2 g/MJ to 2.5 g/MJ we could observe high increment in CDW (Figure 6). This relationship was stronger in Ir condition when compared with Rf condition and the strongest in soil types with the highest available water content (S1, S46) and weakened with the lowest available water content (S1 > S46 > S44). These results confirm that CDW is highly sensitive to RUE, which is directly connected with the availability of moisture for plants. In addition, when increasing the TEC parameter value from 0.008–0.014 kg kPa/kg, CDW tends to increase slightly in all conditions, however this was more evident in Rf conditions than in Ir conditions (Figure 6). In APSIM, both RUE and TEC do not differ by default [36,54]. Therefore, statistical calibration of highly influential RUE and TEC parameters are crucial to achieve higher accuracy when modeling the CDW in KK.

Cultivars with higher GLN and lower CF values will be more beneficial when modeling CDW under any of the environmental and management conditions in KK. This is because under all simulated conditions, the influence of GLN on CDW shows an increasing trend while CF indicting the declining trend when increasing the parameter values from 9 to 14 and 0.65 to 0.8 g/g, respectively (Figure 6). However, it seems that increasing the parameter value from 1200 to 1900 °C day of EB may cause lower increment in CDW compared to other parameters (Figure 6). These results are very important when parameterizing the crop model for KK because they are useful in reducing the number of parameters to be calibrated and avoiding over-parameterization.

Figure 6. Parameter main effect of highly influential parameters under soil types (S1, S44 and S46), and Ir (blue) and Rf (orange) conditions for CDW.
We could study the sensitivity of model outputs to cultivar parameters under different environmental conditions of tropical sugarcane production in KK, Thailand. Determination of variability in the influence of model input parameters on model output could have a considerable impact on studies of cultivar-by-environment interactions. Such studies would improve the efficiency and accuracy of crop modeling, which is computationally expensive, and will be ultimately important for identification of appropriate management strategies to cope with both temporal and spatial variability of crop yield. Therefore, we encourage future research focused on a range of soil types, climate interactions and different water regimes.

4. Conclusions

Our study focused on the use of GP-based emulators to analyze parameter sensitivity in the APSIM-Sugar model under different environmental and management conditions in KK, Thailand. The emulators we obtained, which corresponded to each environmental and management condition across simulated years showed satisfactory results, as evidenced by $R^2$, $\sigma^2$ and cross-validated RMSSE values, indicate that these emulators can successfully replace the simulators. $\text{rue}$ (RUE), $\text{green\_leaf\_no}$ (GLN), $\text{transp\_eff\_cf}$ (TEC), $\text{tt\_emerg\_to\_begcane}$ (EB) and $\text{cane\_fraction}$ (CF) were the most influential parameters regardless of soil type, Ir or Rf conditions. Other analyzed parameters had little influence on the simulator output. Outcomes of our study are beneficial in enhancing the efficiency and accuracy of crop modeling. Further, findings can be used to identify appropriate management strategies to address temporal and spatial variability of sugarcane yield in KK.

Author Contributions: Conceptualization, methodology and formal analysis, W.B.M.A.C.B and K.S.; investigation and writing—Original draft preparation, W.B.M.A.C.B; writing—Review and editing, W.B.M.A.C.B and R.H.K.R.; supervision, K.S., P.K. and T.N. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Manivong, P.; Bourgois, E. White Paper: Thai Sugarcane Sector and Sustainability; FairAgora Asia Co. Ltd.: Bangkok, Thailand, 2017.
2. Hongthong, P.; Patanonthai, A. Variations in Sugarcane Yield among Farmers’ Fields and Their Causal Factors in Northeast Thailand. Int. J. Plant Prod. 2017, 11, 533–548. [CrossRef]
3. Rambo, A.T. The Agrarian Transformation in Northeastern Thailand: A Review of Recent Research. Southeast Asian Stud. 2017, 6, 211–245. [CrossRef]
4. Preecha, K.; Sakai, K.; Pisanjaroen, K.; Sansayawichai, T.; Cho, T.; Nakamura, S.; Nakandakari, T. Calibration and Validation of Two Crop Models for Estimating Sugarcane Yield in Northeast Thailand. Trop. Agric. Dev. 2016, 60, 31–39. [CrossRef]
5. Jeuffroy, M.H.; Barbottin, A.; Jones, J.W.; Lecoeur, J. Crop Models with Genotype Parameters. In Working with Crop Models, 1st ed.; Wallach, D., Makowski, D., Jones, J.W., Eds.; Elsevier: Amsterdam, The Netherlands, 2006; pp. 281–308.
6. Ojeda, J.J.; Rezaei, E.E.; Remenyi, T.A.; Webb, M.A.; Webber, H.A.; Kamali, B.; Harris, R.M.B.; Brown, J.N.; Kidd, D.B.; Mohammed, C.L.; et al. Effects of Soil and Climate Data Aggregation on Simulated Potato Yield and Irrigation Water Requirement. Sci. Total Environ. 2020, 710, 135589. [CrossRef]
7. Kennedy, M.C.; O’Hagan, A. Bayesian Calibration of Computer Models. J. R. Stat. Soc. Ser. B (Stat. Methodol.) 2001, 63, 425–464. [CrossRef]
8. Song, X.; Zhan, C.; Kong, F.; Xia, J. Advances in the Study of Uncertainty Quantification of Large-Scale Hydrological Modeling System. J. Geogr. Sci. 2011, 21, 801–819. [CrossRef]
9. Ewert, F.; van Ittersum, M.K.; Heckelesi, T.; Therond, O.; Bezlepkina, I.; Andersen, E. Scale Changes and Model Linking Methods for Integrated Assessment of Agri-Environmental Systems. Agric. Ecosyst. Environ. 2011, 142, 6–17. [CrossRef]
10. Song, X.M.; Kong, F.Z.; Zhan, C.S.; Han, J.W.; Zhang, X.H. Parameter Identification and Global Sensitivity Analysis of Xin’anjiang Model Using Meta-Modeling Approach. *Water Sci. Eng.* 2013, 6, 1–17. [CrossRef]

11. 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. *Eur. J. Agron.* 2017, 88, 96–105. [CrossRef]

12. Muñoz-Carpena, R.; Zajac, Z.; Kuo, Y.M. Global Sensitivity and Uncertainty Analyses of the Water Quality Model VFSMOD-W. *Trans. ASABE* 2007, 50, 1719–1732. [CrossRef]

13. Cukier, R.I.; Fortuin, C.M.; Shuler, K.E.; Petschek, A.G.; Schaibly, J.H. Study of the Sensitivity of Coupled Reaction Systems to Uncertainties in Rate Coefficients. I Theory. *J. Chem. Phys.* 1973, 59, 3873–3878. [CrossRef]

14. Mara, T.A.; Tarantola, S. Application of Global Sensitivity Analysis of Model Output to Building Thermal Simulations. *Build. Simul.* 2008, 1, 290–302. [CrossRef]

15. Saltelli, A. Making Best Use of Model Evaluations to Compute Sensitivity Indices. *Comput. Phys. Commun.* 2002, 145, 280–297. [CrossRef]

16. Homma, T.; Saltelli, A. Importance Measures in Global Sensitivity Analysis of Nonlinear Models. *Reliab. Eng. Syst. Saf.* 1996, 52, 1–17. [CrossRef]

17. Sobol’, I.M. On Sensitivity Estimation for Nonlinear Mathematical Models. *Matem. Mod.* 1990, 2, 112–118.

18. Saltelli, A.; Chan, K.; Scott, M. *Sensitivity Analysis*; Probability and Statistics Series; John Wiley Sons: Chichester, UK, 2000.

19. Specka, X.; Nendel, C.; Wieland, R. Temporal Sensitivity Analysis of the MONICA Model: Application of Two Global Approaches to Analyze the Dynamics of Parameter Sensitivity. *Agriculture* 2019, 9, 37. [CrossRef]

20. O’Hagan, A. Bayesian Analysis of Computer Code Outputs: A Tutorial. *Reliab. Eng. Syst. Saf.* 2006, 91, 1290–1300. [CrossRef]

21. Sacks, J.; Welch, W.J.; Mitchell, T.J.; Wynn, H.P. Design and Analysis of Computer Experiments. *Stat. Sci.* 1989, 4, 409–423. [CrossRef]

22. Oakley, J.E.; O’Hagan, A. Probabilistic Sensitivity Analysis of Complex Models: A Bayesian Approach. *J. R. Stat. Soc. Ser. B Stat. Methodol.* 2004, 66, 751–769. [CrossRef]

23. Uusitalo, L.; Lehikoinen, A.; Helle, I.; Myrberg, K. An Overview of Methods to Evaluate Uncertainty of Deterministic Models in Decision Support. *Environ. Model. Softw.* 2015, 63, 24–31. [CrossRef]

24. Boukouvalas, A.; Cornford, D.; Maniyar, D.; Singer, A. Gaussian Process Emulation of Stochastic Models: Developments and Application to Rabies Modelling. In Proceedings of the RSS 2008 Conference, Nottingham, UK, 1–5 September 2008.

25. Rasmussen, C.E.; Williams, C.K.I. *Gaussian Processes for Machine Learning* Cambridge; MIT Press: Cambridge, MA, USA, 2006.

26. Sexton, J.; Everingham, Y. Global Sensitivity Analysis of Key Parameters in a Process-Based Sugarcane Growth Model—A Bayesian Approach. In Proceedings of the 7th International Congress on Environmental Modelling and Software, San Diego, CA, USA, 15–19 June 2014.

27. Gunarathna, M.H.J.P.; Sakai, K.; Nakandakari, T.; Momii, K.; Kumari, M.K.N. Sensitivity Analysis of Plant and Cultivar-Specific Parameters of APSIM-Sugar Model: Variation between Climates and Management Conditions. *Agronomy* 2019, 9, 242. [CrossRef]

28. Khon Kaen Climate. Available online: https://en.climate-data.org/asia/thailand/khon-kaen-province/khon-kaen-4291/ (accessed on 19 November 2019).

29. USDA. Soil Texture Calculator. Available online: https://www.nrcs.usda.gov (accessed on 15 November 2019).

30. Holzworth, D.P.; Huth, N.I.; de Voil, P.G.; Zurcher, E.J.; Herrmann, N.I.; McLean, G.; Chenu, K.; van Oosterom, E.J.; Snow, V.; Murphy, C.; et al. APSIM—evolution towards a new generation of agricultural systems simulation. *Environ. Model. Softw.* 2014, 62, 327–350. [CrossRef]

31. Wang, E.; Robertson, M.J.; Hammer, G.I.; Carberry, P.S.; Holzworth, D.; Meinke, H.; Chapman, S.C.; Hargreaves, J.N.G.; Huth, N.I.; McLean, G. Development of a Generic Crop Model Template in the Cropping System Model APSIM. *Eur. J. Agron.* 2002, 18, 121–140. [CrossRef]

32. Ojeda, J.J.; Pembleton, K.G.; Caviglia, O.P.; Islam, M.R.; Agnusdei, M.G.; Garcia, S.C. Modelling Forage Yield and Water Productivity of Continuous Crop Sequences in the Argentinian Pampas. *Eur. J. Agron.* 2018, 92, 84–96. [CrossRef]
33. Keating, B.A.; Robertson, M.J.; Muchow, R.C.; Huth, N.I. Modelling sugarcane production systems. I. Description and validation of the sugarcane module. *F. Crop. Res.* **1999**, *61*, 253–271. [CrossRef]

34. Dias, H.B.; Inman-Bamber, G.; Bermejo, R.; Sentelhas, P.C.; Christodoulou, D. New APSIM-Sugar Features and Parameters Required to Account for High Sugarcane Yields in Tropical Environments. *F. Crop. Res.* **2019**, *235*, 38–53. [CrossRef]

35. Sexton, J.; Everingham, Y.; Inman-Bamber, G. A Theoretical and Real-World Evaluation of Two Bayesian Techniques for the Calibration of Variety Parameters in a Sugarcane Crop Model. *Environ. Model. Softw.* **2016**, *83*, 126–142. [CrossRef]

36. Keating, B. The APSIM Sugar Model. Available online: [http://apsrunet.apsim.info/svn/development/trunk/apsim/sugar/docs/sugar_pseudo.html#sugar_dm_partition_pot](http://apsrunet.apsim.info/svn/development/trunk/apsim/sugar/docs/sugar_pseudo.html#sugar_dm_partition_pot) (accessed on 23 November 2019).

37. Stanfill, B. Apsimr: Edit, Run and Evaluate APSIM Simulations Easily Using R. Available online: [https://cran.r-project.org/web/packages/apsimr/index.html](https://cran.r-project.org/web/packages/apsimr/index.html) (accessed on 23 November 2019).

38. R Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria. Available online: [https://www.R-project.org/](https://www.R-project.org/) (accessed on 25 November 2019).

39. Sinclair, T.R. Is Transpiration Efficiency a Viable Plant Trait in Breeding for Crop Improvement? *Funct. Plant Biol.* **2012**, *39*, 359–365. [CrossRef]

40. Jackson, P.A.; Basnayake, J.; Inman-Bamber, G.; Lakshmanan, P. Selecting Sugarcane Varieties with Higher Transpiration Efficiency. In *The Australian Society of Sugar Cane Technologists, Broadbeach, Australia, 28 April–1 May 2014*; Volume 36.

41. Park, S.E.; Robertson, M.; Inman-Bamber, N.G. Decline in the Growth of a Sugarcane Crop with Age under High Input Conditions. *F. Crop. Res.* **2005**, *92*, 305–320. [CrossRef]

42. Ferreira, R.A.; de Souza, J.L.; Lyra, G.B.; Escobedo, J.F.; Santos, M.V.C. Energy Conversion Efficiency in Sugarcane under Two Row Spacings in Northeast of Brazil. *Rev. Bras. Eng. Agríc. e Ambient.* **2015**, *19*, 741–747. [CrossRef]

43. Olivier, F.C.; Singels, A.; Eksteen, A.B. Water and Radiation Use Efficiency of Sugarcane for Bioethanol Production in South Africa, Benchmarked against Other Selected Crops. *S. Afr. J. Plant Soil* **2016**, *33*, 1–11. [CrossRef]

44. Meki, M.N.; Kiniry, J.R.; Youkhana, A.H.; Crow, S.E.; Ogoshi, R.M.; Nakahata, M.H.; Tirado-Corbalá, R.; Anderson, R.G.; Osorio, J.; Jeong, J. Two-Year Growth Cycle Sugarcane Crop Parameter Attributes and Their Application in Modeling. *Agron. J.* **2015**, *107*, 1310–1320. [CrossRef]

45. Villa-Vialaneix, N.; Follador, M.; Ratto, M.; Leip, A. A Comparison of Eight Metamodeling Techniques for the Simulation of N₂O Fluxes and N Leaching from Corn Crops. *Environ. Model. Softw.* **2012**, *34*, 51–66. [CrossRef]

46. O’Hagan, A. Probabilistic Uncertainty Specification: Overview, Elaboration Techniques and Their Application to a Mechanistic Model of Carbon Flux. *Environ. Model. Softw.* **2012**, *36*, 35–48. [CrossRef]

47. Kennedy, M.C.; Anderson, C.W.; Conti, S.; O’Hagan, A. Case Studies in Gaussian Process Modelling of Computer Codes. *Reliab. Eng. Syst. Saf.* **2006**, *91*, 1301–1309. [CrossRef]

48. Kennedy, M.C.; Petropoulos, G.P. GEM-SA: The Gaussian Emulation Machine for Sensitivity Analysis. In *Sensitivity Analysis in Earth Observation Modelling*; George, P.P., Prashant, K.S., Eds.; Elsevier: Amsterdam, The Netherlands, 2017; pp. 341–361. ISBN 978-0-12-803011-0.

49. Qin, X.; Wang, H.; Li, Y.; Li, Y.; McConkey, B.; Lemke, R.; Li, C.; Brandt, K.; Gao, Q.; Wan, Y.; et al. A Long-Term Sensitivity Analysis of the Denitrification and Decomposition Model. *Environ. Model. Softw.* **2013**, *43*, 26–36. [CrossRef]

50. Sexton, J. Bayesian Statistical Calibration of Variety Parameters in Asugarcan Crop Model. Master’s Thesis, James Cook University, Townsville, Australia, April 2015.

51. Petropoulos, G.; Wooster, M.J.; Carlson, T.N.; Kennedy, M.C.; Scholze, M. A Global Bayesian Sensitivity Analysis of the 1d SimSphere Soil-Vegetation-Atmospheric Transfer (SVAT) Model Using Gaussian Model Emulation. *Ecol. Modell.* **2009**, *220*, 2427–2440. [CrossRef]

52. Ojeda, J.J.; Volene, J.J.; Brouder, S.M.; Caviglia, O.P.; Agnusdei, M.G. Evaluation of Agricultural Production Systems Simulator as Yield Predictor of Panicum Virgatum and Miscanthus x Giganteus in Several US Environments. *GCB Bioenergy* **2017**, *9*, 796–816. [CrossRef]
53. Smit, M.A.; Singels, A.; van Antwerpen, A. Differences in Canopy Development of Two Sugarcane Cultivars under Conditions of Water Stress: Preliminary Results. *Proc. S. Afr. Sugar Technol. Assoc.* **2004**, *78*, 149–152.

54. Ojeda, J.J.; Pembleton, K.G.; Islam, M.R.; Agnusdei, M.G.; Garcia, S.C. Evaluation of the Agricultural Production Systems Simulator Simulating Lucerne and Annual Ryegrass Dry Matter Yield in the Argentine Pampas and South-Eastern Australia. *Agric. Syst.* **2016**, *143*, 61–75. [CrossRef]

© 2020 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (http://creativecommons.org/licenses/by/4.0/).