Agricultural Water Management 301 (2024) 108938

Contents lists available at ScienceDirect

Agricultural Water Management

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

The importance of model structure and soil data detail on the simulations of crop growth and water use: A case study for sugarcane

Murilo dos Santos Vianna[a,b,1,*], Klaas Metselaar[d], Quirijn de Jong van Lier[c], Thomas Gaiser[a], Fábio Ricardo Marin[b]

[a] Institute of Crop Science and Resource Conservation (INRES), University of Bonn, Katzenburgweg 5, Bonn 53115, Germany
[b] College of Agriculture "Luiz de Queiroz" (ESALQ), University of Sao Paulo (USP), Av. Paduas Dias 235, Piracicaba 13418-900, Brazil
[c] Center for Nuclear Energy in Agriculture (CENA), University of Sao Paulo (USP), Av. Centenário 303, Piracicaba 13400-970, Brazil
[d] Wageningen University, Chairgroup Soil Physics and Land Management, Droevendaalsesteeg 3, Wageningen 6708 PB, the Netherlands

ARTICLE INFO

Handling Editor - Dr Z Xiying

Keywords:
Process-based crop models
Model Structure
Soil Data Detail
Sugarcane
Irrigation Scheduling

ABSTRACT

Process-based crop models have faced rapid development over the last years, and many modelling platforms are now available and can be used in a wide range of conditions. Whilst the selection of a model should be suited to the purpose of its application, very few studies focused on the impact of choosing different model structures and data details on the simulation outputs. One important aspect is the soil water dynamics, which can be simulated at different levels of details in terms of data and approaches. In this study, we investigated the impact of model structure and data detail on simulations of sugarcane growth and irrigation scheduling. Three different soil water routines (Standalone, Tipping-Bucket, SWAP) were coupled with the SAMUCA model and calibrated with a comprehensive field experiment dataset. We also tested the influence of using simplified homogeneous (SL) and detailed (DL) soil profile information in model performance. The model framework was evaluated against independent field experiments across Brazil and used to simulate long-term sugarcane growth and irrigation scheduling. After calibration, the SWAP-DL showed the highest accuracy in soil moisture predictions, with a 6 % error (RRMSE), but the difference from TippingBucket-DL was small (8 %). While the performance of stalk dry mass, LAI and water-use efficiency simulations were within the range found in literature, comprehensive field experiments showing significant impacts of drought on sugarcane growth are still lacking for a more rigorous evaluation. Both SWAP and tipping-bucket approaches showed higher robustness to soil data detail as compared to the Standalone method, which should be avoided when soil water is critical for sugarcane growth. The use of tipping-bucket method may still be preferred when the research goal is focused on crop growth and soil parameters are limited. SWAP-SAMUCA may provide an extended ability to represent agrohydrological processes in sugarcane plantations and process understanding.

1. Introduction

Process-based crop models have become instrumental in agricultural systems analysis as they enable the explicit formulation and testing of how physiological and soil processes interact and explain emerging system behaviour (Jones et al., 2017). Due to the large multi-dimensional nature of agroecosystems, a unified model is improbable and not existing. Therefore, models are typically employed using the "fit-for-purpose" concept, where the contextualisation of the scientific question and data availability are central aspects in the selection and definition of the model boundaries, methods and processes (Hamilton et al., 2022). With the rapid evolution and access to computer-based modelling, a large number of process-based simulation models have become available that could be arguably used in a wide range of applications and spatiotemporal scales (Rosenzweig et al., 2013; Antle et al., 2017).

To simulate dynamic crop growth, modellers often need to incorporate key processes governing the soil-plant-atmosphere system into their simulation frameworks. Although many modules share similar concepts that evolved from pioneering work from the 1960s (de Wit,

* Corresponding author at: Institute of Crop Science and Resource Conservation (INRES), University of Bonn, Katzenburgweg 5, Bonn 53115, Germany.
E-mail address: murilodsv@gmail.com (M. dos Santos Vianna).
[1] ORCID: http://orcid.org/0000-0003-1139-4589

https://doi.org/10.1016/j.agwat.2024.108938
Received 5 February 2024; Received in revised form 31 May 2024; Accepted 25 June 2024
Available online 5 July 2024

1958; van Ittersum et al., 2003; Jones et al., 2017), the uncertainty on crop model simulation is still considerably large (Martre et al., 2015; Tao et al., 2018). In addition, very few modelling frameworks have a collection of modules allowing modellers to easily link and test different model structures according to the end-user needs (Enders et al., 2023). Therefore, the process of developing new modules and their incorporation into an existing simulation platform is still not easy to achieve.

The representation of soil water dynamics and belowground process is a typical example of which process-based crop models may differ in their approach and detail. One model may use physically-based principles (e.g., mass and energy flow controlled by gradients) and detailed soil information to simulate water flow and use (Kroes et al., 2017), whereas another approach could use a simplified method (e.g., cascade flow) that mimics the real system behaviour (van Ittersum et al., 2003). Recently, the wide use of the latter has been brought into debate under the hypothesis that an exacerbated unnecessary empiricism also limits the use of crop models for process understanding, while the physical descriptions of many water flow processes in the soil-crop system are now available (Jarvis et al., 2022). While in one hand, many large-scale applications of crop models can only be achieved with a more simplified approach due to the lack of detailed and accurate soil information. On the other, land surface models such as JULES and CLM simulate large-scale carbon (including crop growth) and water cycles using Richard's approach since a few years (Best et al., 2011; Osborne et al., 2015; Lawrence et al., 2019).

One of the aims of our study is to investigate the impact of choosing different model approaches and soil data details on the simulation of crop growth and water use. We chose the sugarcane crop as it is the main source of sugar globally and has emerged as the second-largest source of biofuel (Goldemberg et al., 2014). On a global scale, more than 70 % of sugarcane biomass is produced in Brazil, India, China, Thailand and Pakistan (FAO, 2022). Brazil alone corresponds to ca. 40 % of the global production, where the crop is largely cultivated under rainfed conditions.

In the last years, sugarcane areas strongly expanded towards the central-western regions of the country production and faces difficulties due to higher water deficits and poor soil conditions, making supplementary irrigation essential for crop production (Vianna et al., 2016; Dias and Sentelhas, 2018b; Marin et al., 2020). Whilst fully-irrigated sugarcane accounts for only 6.7 % of all Brazilian fields, fertirrigation using sewage and vinasse is widely adopted, corresponding to 29 % of all planted sugarcane in Brazil (ANA, 2019). Irrigation is also a key strategy for maintaining stable sugarcane yields in other countries like South Africa (Singels et al., 2019). In this context, soil-water and hydrological studies have become increasingly important for decision-makers in the sugarcane industry (Stenzel et al., 2021).

Although a handful of sugarcane models were developed and described in the literature (Dias and Inman-Bamber, 2020; Singels, 2013), few of them are modular and freely available for end users: DSSAT/CANEGRO (Jones and Singels, 2018), APSIM/Sugar (Keating et al., 1999) and the Agronomic Modular Simulator for Sugarcane (SAMUCA) (Marin and Jones, 2014). The latter was recently updated by Vianna et al. (2020), operating with the widely employed "tipping-bucket" soil water balance method as in the DSSAT platform (Ritchie, 1998; Jones et al., 2003). On its first version, SAMUCA was implemented with a simple standalone soil water routine, which employs Darcy-law principle at a limited 4-layer compartment soil profile (Marin and Jones, 2014; Teh, 2006). Due to its modular structure, the SAMUCA model can be coupled to other soil water balance routines and simulation platforms.

The SWAP platform is currently one of the most robust hydrological models available for end users that evolved over the last 50 years (Kroes et al., 2017; Heinen et al., 2024). The model employs an implicit numerical solution scheme to solve Richard's equation, simulating soil water movement in saturated or unsaturated soils. Soil physics modules are included to simulate solute transport, macropore flow, water repellency, soil heat flow and lateral drainage flow at the field level. In SWAP, crop growth can be simulated either with static crop parameters or dynamically with the generic crop model WOFOST (World Food Studies).

In this study, we aimed to compare simulations of sugarcane growth and water consumption using three water balance routines with increasing complexity (standalone, tipping-bucket, SWAP). Therefore, we coupled the most recent version of the SAMUCA crop module into the SWAP modelling platform. The resulting model (SWAP-SAMUCA) was intercompared with the SAMUCA model which can still be executed with two soil water balance routines: the original standalone (Marin and Jones, 2014), and the tipping-bucket approach (Vianna et al., 2020). We also investigated the model performance when using detailed soil profile data or homogenous soil properties, which is often necessary when the vertical description of the soil profile is not available. The methods are calibrated and evaluated against field observations in different regions of Brazil. Finally, we tested the implications of choosing each of the models and detailed soil data on the simulations of sugarcane growth and irrigation scheduling.

2. Material and Methods

2.1. Modelling framework description and coupling

All the crop growth simulations in this study were executed with the most recent version of the SAMUCA crop module (Vianna et al., 2020). By default, the model is executed with the tipping-bucket empirical method, but users can still select the early soil water balance routine described by Marin and Jones (2014). The preliminary standalone version of SAMUCA utilizes a simple 4-layered soil water routine based on the Darcy Law to simulate soil water available to the crop over the growing season (Teh, 2006), termed here as the "Standalone" water balance method. The crop module code of SAMUCA is currently maintained in the open-source DSSAT platform as of version 4.8 (https://github.com/DSSAT/dssat-csm-os).

The SWAP model platform simulates soil water, heat and solute fluxes following physically based approaches (Kroes et al., 2017). Three modular routines to simulate crop growth and development are included in SWAP: (i) a simple module considering static growth; (ii) the WOFOST model, a generic dynamic model that simulates the growth of many kinds of crops (de Wit et al., 2019); (iii) and a modification of the WOFOST model for perennial grass simulations (Kroes et al., 2017).

To investigate the effect of using a physically-based approach for sugarcane simulations we coupled the corresponding crop-modules of the SAMUCA model into SWAP (Fig. 1). The implementation and variables' linkage were performed directly in the Fortran source code of SWAP v4.0.1 (www.swap.alterra.nl) using Visual Studio 2015 IDE. In this way, SAMUCA could be utilized as the 4th crop module option of SWAP alongside WOFOST. The source code and executable version for the SWAP-SAMUCA model can be found in: https://github.com/Murilodsv/SWAP-SAMUCA. All the data pre-post processing and analysis were executed with the R language.

As our goal was to isolate the effect of soil water routines, we attempted to use similar approaches for processes that are mainly governed by the atmosphere or crop physiology. For example, all three model configurations utilized the same approach for potential evapotranspiration (Penman-Monteith) and partitioning into transpiration and soil evaporation components (Beer law). We also switched off any effect of soil temperature on crop development and considered that water intercepted by the canopy is neglectable. Furthermore, our simulations considered that there was no effect of soil salinity, deep groundwater (e.g., free drainage at the bottom layer), mulch cover or even surface barriers that could cause water ponding. Table 1 summarizes the main configuration used for each of the three soil water balance routines and their differences.

Except for potential evapotranspiration, all the routines differ on the

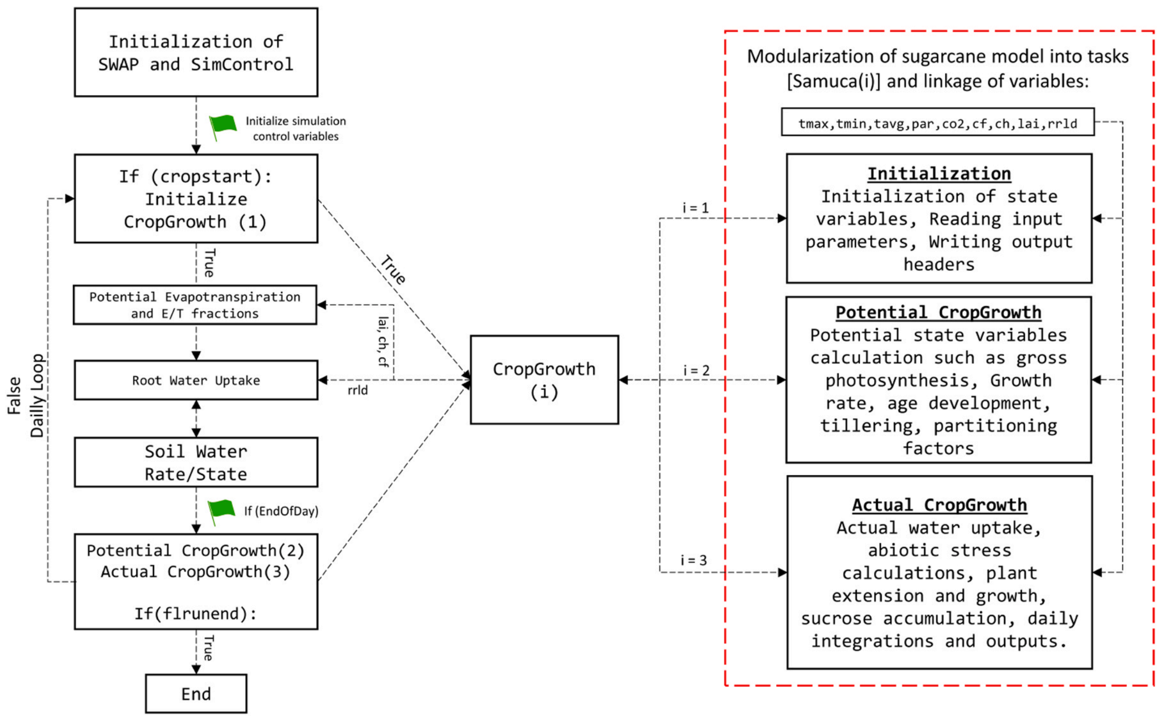

Fig. 1. Simplified diagram of SWAP platform structure and its CropGrowth module, including the coupled SAMUCA crop module components (in red dotted box). Daily maximum, minimum and average air temperatures are expressed as tmax, tmin and tavg; photosynthetically active radiation and atmospheric CO2 concentration as par and co2, respectively; crop variables linked with SWAP are the crop coefficient (cf), canopy height (ch), leaf area index (lai), and relative root length density (rrld).

Table 1
List of main agrohydrological processes and their definitions that are incorporated in each modelling approach coupled with the SAMUCA model: Standalone, Tipping-Bucket and SWAP.

Process	Modelling structure		
	Standalone	Tipping-Bucket	SWAP
Infiltration	All water infiltrates	Incoming water after runoff	Infiltration capacity based on Richards' equation
Runoff	Not considered	CN curve number	Infiltration capacity + Ponding threshold
Reference Evapotranspiration (ETo)	Penman-Monteith	Penman-Monteith	Penman-Monteith
Potential Evapotranspiration (ETp)	ETo x Crop coefficient	ETo x Crop coefficient	ETo x Crop coefficient
Potential soil evaporation (Ep) and crop transpiration (Tp)	Beer Law (LAI and light extinction)	Beer Law (LAI and light extinction)	Beer Law (LAI and light extinction)
Actual soil evaporation (Ea)	Reduction of Ep based on relative water content of topsoil (Keulen and Seligman, 1988)	Reduction of Ep with the soil diffusion method of Suleiman–Ritchie	Reduction of Ep to maximum Darcy flux
Potential root water uptake (PRWU) *	Consider root length density, soil moisture and empirical coefficients (Ritchie, 1998)	Consider root length density, soil moisture and empirical coefficients (Ritchie, 1998)	Not needed
Actual root water uptake (ARWU)	Minimum between the total PRWU and Tp	Minimum between the total PRWU and Tp	Based on soil matric potential (Feddes method)
Actual crop transpiration (Ta)	Sum of ARWU for each soil layer	Sum of ARWU for each soil layer	Weighted integral of ARWU with root length density
Downward soil water flux	Based on the estimated unsaturated hydraulic conductivity	Cascade based on water holding capacity	Gradient of soil matric potentials (Richards' equation)
Upward soil water flux	Not considered	Estimate of soil water diffusivity and volumetric soil water gradient	Gradient of soil matric potentials (Richards' equation)
Bottom flux	Free drainage as a function of hydraulic conductivity of bottom layer	Free drainage of excessive volumetric water limited to saturated hydraulic conductivity	Free drainage as a function of hydraulic conductivity of bottom layer
Vertical soil discretization	Fixed 4-layered profile	Any number of soil layers	Any number of soil layers and sub-compartments
Timestep	Daily	Daily	Sub-daily (dynamic)

representation of soil water infiltration and runoff, vertical movement and the spatiotemporal discretization scheme. The standalone and tipping-bucket routines share similar approaches for the simulation of root water uptake, which is calculated based on fixed empirical coefficients, the current status of soil moisture and root length density, and must not exceed the potential crop transpiration (Tp) (Ritchie, 1998). For SWAP, we used the default empirical method of Feddes et al. (1978) due to its simplicity and because our calibration dataset does not allow the precise determination of root water uptake and soil evaporation. More mechanistic approaches that can simulate water uptake under oxygen stress or detailed drought are also available in SWAP (Bartholomeus et al., 2008; de Jong van Lier et al., 2008), but may require extra parametrization and a more comprehensive dataset.

The standalone routine uses a fixed 4-layer approach, whereas the tipping bucket and SWAP allow for multiple soil compartments. To minimize the effect of vertical discretization, we harmonized the soil inputs so that all simulations considered the same soil depth and had similar sizes of compartments. Both standalone and tipping-bucket operate at daily timesteps, whereas SWAP solves water transport at sub-daily dynamic timesteps controlled by user-defined parameters and mass balance. In our runs, the minimum and maximum possible timesteps ranges from 1 s to 5 h, respectively.

2.2. Modelling dataset

The modelling dataset of our study can be separated into three parts: model calibration; evaluation across different sites; and simulations for long-term sugarcane growth and irrigation scheduling (Table 2). We calibrated the three configurations of the SAMUCA water balance routines with the comprehensive field experiment conducted at the College of Agriculture "Luiz de Queiroz" (ESALQ/USP) in Piracicaba, Brazil (Lat: 22°41'55"S Lon: 47°38'34"W Alt: 540 m). The experimental dataset included 4-season observations of fully-irrigated and well-fertilized sugarcane growth and development, soil moisture and evapotranspiration under two treatments: bare soil and mulch cover. Daily evapotranspiration rates (ET) were determined by the integration of 15-minute latent heat flux measurements taken by the Bowen Ratio method (Perez et al., 1999). Soil moisture (SM) was measured every 3 days or at one day after rainfall/irrigation event with a Frequency Domain Reflectometry (FDR) probe (Vianna et al., 2020). Stalk dry mass (SDM) was measured by regular destructive sample within experimental plots. Leaf area index (LAI) was measured by LAI-2000 sensor (LICOR) at different crop stages. For this simulation study, we used only the data from the bare soil treatment. The site's climate is characterized by a hot and humid summer with a dry winter (Cwa- Köppen classification), and the soil is classified as Typic Hapludox. The experimental setup is fully described by Gonçalves et al. (2023) and Vianna et al. (2020).

Six independent 1-season field trials across Brazil were used to evaluate the calibrated models with respect to crop growth simulations under contrasting conditions. For all of these sites, sugarcane SDM and LAI were measured, except Uniao, which had no LAI measurements. Finally, we utilized 30-year meteorological data (1980–2010) at four distinct sugarcane-producing regions to simulate the effect of choosing each version of SAMUCA and soil data detail on crop growth and irrigation scheduling. This dataset was also used in previous studies with SAMUCA and other sugarcane models for benchmarking models' performance (Marin et al., 2015; Vianna et al., 2020, 2022).

Weather data for all sites were obtained from the nearest meteorological station and consisted of daily maximum and minimum air temperature, solar radiation, precipitation, wind speed and relative air humidity. The cultivar in all sites was the RB867515, still one of Brazil's most widely planted cultivars. For the calibration and evaluation sites, the soil water retention curves were determined from undisturbed soil samples at different depths (Vianna et al., 2020). For the long-term simulations, the predominant soil type characteristic was used (Tab. A1) (Vianna et al., 2016; Dias and Sentelhas, 2018a).

2.3. Model calibration and evaluation runs

The soil and crop parameters of the three combinations of SAMUCA and soil water routines were calibrated using the concept of inverse modelling, i.e. including soil and crop parameters (Ines and Droogers, 2002). This was done because the Feddes method of SWAP generally requires calibration of both soil and crop properties simultaneously (de Melo and de Jong van Lier, 2021). Furthermore, we assume that applying the same calibration procedure to all three models would minimise any bias effect from a previous calibration in our analysis (Vianna et al., 2020). Thus, we tested the suitability of global optimization methods contained in the R package "nloptr" v2.0.3 (Ypma et al., 2022) for minimizing our objective function (Fig. A1). The method that provided better and more stable results was the global optimization method based on controlled random search with local mutation (CRS2) from Kaelo and Ali (2006).

The objective function for the calibration procedure was formulated as the average of relative root mean squared error (RRMSE) between simulated and observed SDM, LAI, ET and SM. The crop and soil parameters selected to minimize the objective function are shown in Tab. A2, where its initial values and plausible ranges were taken either from field measurements or literature (de Melo and de Jong van Lier, 2021; Pereira et al., 2021). This calibration setup was applied to both types of soil data (detailed profile - DL; and homogenous profile - SL) resulting in six sets of parameter values, one per each combination of SAMUCA water routine and soil data type.

While in the calibration runs we used the calibrated soil properties for SL or DL conditions, in the evaluation sites where the SWC and ET are not known the soil input parameters were replaced with the observed soil profile information, which we used directly for different soil layers (DL) or the mean value for the whole profile to represent the SL conditions (Tab. A1). The number of layers for both tipping-bucket and SWAP approaches used the same number of soil layers as specified in the soil profile, while the soil data was aggregated for standalone approach to conform with the 4-layered simulation profile. To minimize an effect of spatial resolution, we preserved the same vertical discretization for each model in both DL and SL conditions, i.e. SL used the same number of layers as in DL but with fixed soil properties for all layers (Tab. A1).

To evaluate model performance after calibration and across different sites, we used the statistical indexes of performance described by Wallach et al. (2018): Relative root mean squared error (RRMSE), precision (R2), accuracy (d), and Nash–Sutcliffe efficiency (EF). In the calibration site, the performance was assessed with SDM, LAI, ET and SM data,

Table 2
Summary of sugarcane field experiments datasets across Brazil used for model calibration (c), evaluation (e), and long-term simulations (s). All long-term simulations (s) considered a 1-year growing season, totaling 30 seasons. Except for Piracicaba (c), which have 4 seasons, all the other experiments were carried out for a 1 season planted cane.

Site	Lat	Lon	Alt	Soil	T (°C)	P (mm)	Koppen	Water Regime	Start Date	End Date
Piracicaba (c)	22°41'S	47°38'W	540	Typic Hapludox	22,0	1331	Cwa	Irrigated	2012–10–16	2016–06–08
Uniao-Irrig (e)	4°51'S	42°52'W	68	Oxisol	27,0	1500	Aw	Irrigated	2007–09–09	2008–06–16
Uniao-Rain (e)	4°51'S	42°52'W	68	Oxisol	27,0	1500	Aw	Rainfed	2007–09–09	2008–06–16
Coruripe (e)	10°07'S	36°10'W	16	Fragiudult	21,6	1401	As	Rainfed	2005–08–16	2006–09–15
Ap. Taboado (e)	20°05 S	51°18'W	335	Typic Hapludox	23,5	1560	Aw	Rainfed	2006–07–01	2007–09–08
Colina (e)	20°25'S	48°19'W	590	Typic Hapludox	22,8	1363	Cwa	Rainfed	2004–02–10	2005–12–01
Olimpia (e)	20°26'S	48°32'W	500	Typic Hapludox	23,3	1349	Cwa	Rainfed	2004–02–10	2005–12–01
Piracicaba (s)	22°41'S	47°38'W	540	Typic Hapludox	22,0	1331	Cwa	Irrigated	1980–07–01	2010–07–01
Jatai (s)	17°53'S	51°43'W	663	Typic Hapludox	23,7	1657	Aw	Irrigated	1980–07–01	2010–07–01
Petrolina (s)	9°22'S	40°28'W	370	Cambisol	27,1	485	BSh	Irrigated	1980–07–01	2010–07–01
Recife (s)	8°3'S	34°57'W	7	Typic Hapludox	25,8	2267	Am	Irrigated	1980–02–01	2010–02–01

whereas on evaluation sites only SDM and LAI data was available for the analysis.

2.4. Long-term simulations of crop growth and irrigation scheduling

After the calibration procedure, we applied the three versions of SAMUCA-water balance routines and soil data type to simulate sugarcane growth and irrigation scheduling. We employed the same approach as in the evaluation runs where the observed soil properties were provided for different layers or homogenized to represent the DL and SL conditions, respectively. Simulations were configured to represent a 1-year sugarcane season, continuously replicated for the years 1980–2010. In each location we selected the most predominant cutting date: February for Recife; and July for Jatai, Petrolina and Piracicaba (Vianna et al., 2022).

An automatic irrigation routine was implemented where irrigation applications are triggered when the relative available water in the rooting zone drops below 80 %. To mimic the drying-off period necessary for sucrose accumulation, the irrigation applications were halted 1 month before harvesting (Inman-Bamber, 2004; Dias and Sentelhas, 2018b). Although farmers may choose different irrigation strategies depending on their availability to water resources and irrigation systems, we adopt this irrigation scheduling to represent intensive irrigation practices without an irrigation deficit in Brazil (Nassif et al., 2019; Marin et al., 2020). The simulated SDM and total water application (T_{irr}) were then statistically compared separately against each model configuration and soil type at each location by the ANOVA test with post-hoc Tukey HSD (Honestly Significant Difference). The simulated water use efficiency in terms of irrigated water ($IWUE=SDM/T_{irr}$) and evapotranspired during the crop season ($EWUE=SDM/T_{et}$) were also used to assess the model configurations.

3. Results

3.1. Impact on model performance and parameter values at the calibration site of Piracicaba

Results showed similar RRMSE values between models after the optimization procedure (Fig. 2). The highest impact was in LAI for SL simulations for which RRMSE was 39 % for the tipping-bucket and 31 % for SWAP, whereas in DL, it was 38 % (tipping-bucket) and 34 % (standalone). For all the other variables, RRMSE differences between modelling approaches were lower than 4 %. Although small differences were found, none of the model approaches showed the best fit for all variables simultaneously. For example, SWC simulations had the best agreement indexes for SWAP-SAMUCA when different soil layers (RRMSE=6 %) were considered, but this approach did not show the best agreement for ET and LAI simulations.

The performance of simulations with different soil layers showed a relatively lower RRMSE of 25 % (mean across all variables) as compared to simulations assuming a homogeneous soil profile (28 %). This is more evident when comparing the difference in modelling efficiency (EF) and precision (R2) between simulations considering multiple layers and single-layer for SWC and ET; while the impact was more pronounced in SWAP-SAMUCA (difference of −0,70 EF and −0,65 R2) as compared to the other soil water routines (Fig. 2). In an overview between variables, the simulations of ET and LAI presented higher values of RRMSE of about 36 % (±3 %), followed by SDM with 27 % (±2 %), and SWC with

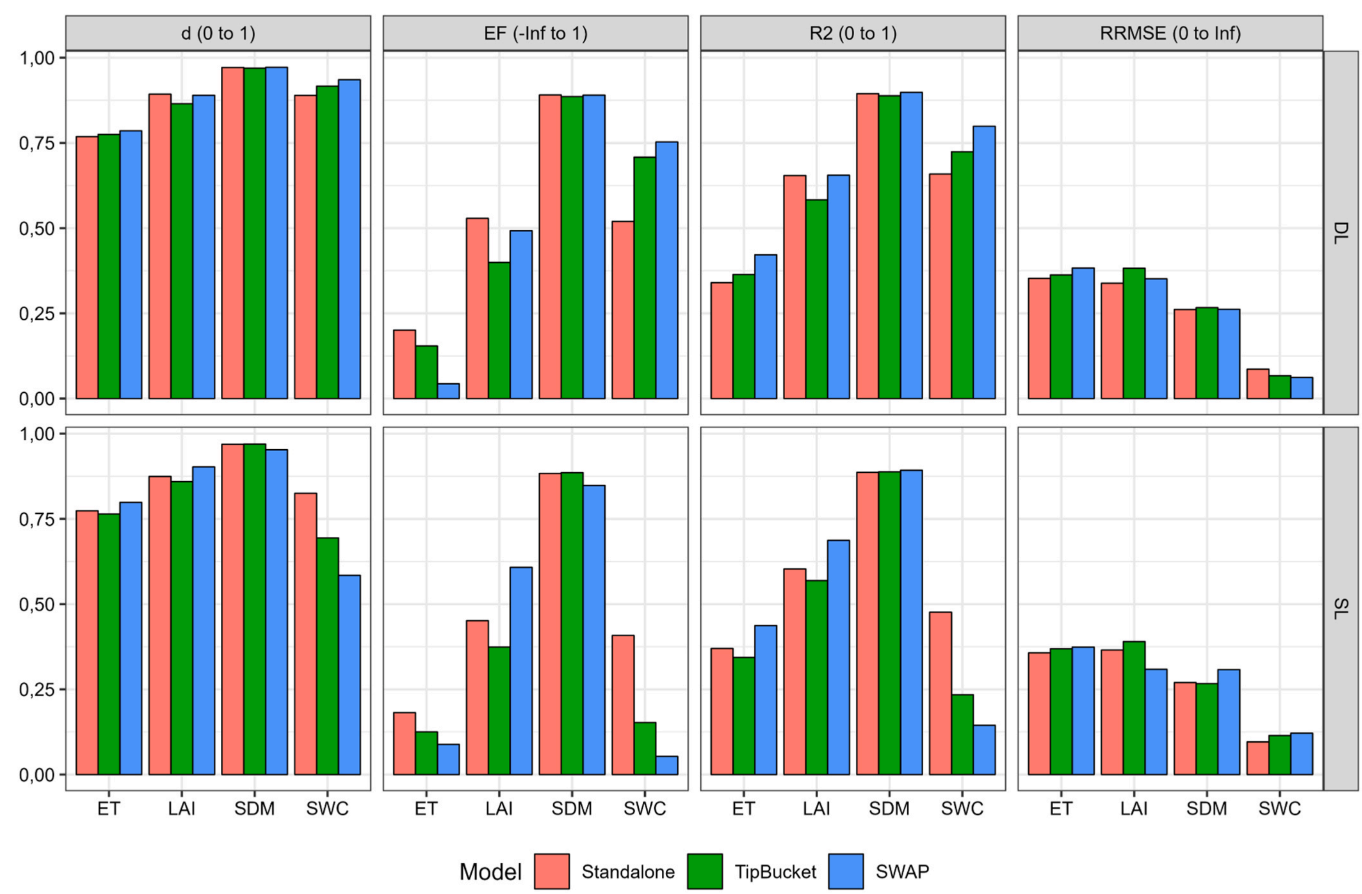

Fig. 2. Statistical indices of performance of each modelling approach (SWAP, TipBucket, Standalone), using different soil data types (SL and DL) in simulating the evapotranspiration (ET), leaf area index (LAI), stalk dry mass (SDM) and soil water content (SWC) at the calibration experiment of Piracicaba. RRMSE is the Relative root mean squared error (fraction, or conversely, %), R2 and d are respectively the model precision and accuracy, and EF the Nash–Sutcliffe efficiency.

9 % (±2 %). The simulations of SDM had the highest values of model accuracy (d=0,96), precision (R2=0,87) and efficiency (EF=0,86) as compared to other variables.

The temporal dynamics of soil moisture content simulated by SWAP and tipping-bucket showed a similar pattern (Fig. 3). Our results showed that the correlation indexes (r) between these two approaches were relatively high, ranging from 0,73 to 0,83 (Fig. A3). The standalone soil water routine presented a more unstable performance, with higher oscillation in soil moisture simulations as compared to the SWAP and tipping-bucket approaches. Simulations were in general more accurate under detailed soil (DL) than considering a homogenous profile (SL) (Fig. A4 and Fig. 2). In SL simulations of both the SWAP and tipping bucket, we observed a systematic overprediction for the topsoil (10 cm) and underprediction at deeper soil (60 cm) which indicates a compensatory effect during the calibration procedure to cope with the lack of soil data (Fig. 3).

Daily ET simulations showed the lowest accuracy among the studied variables (Fig. 2). However, the seasonal pattern of observed ET could be replicated by all the models, with higher values reaching around 6 mm day^{-1} during summer and lower than 2 mm day^{-1} in winter (Fig. 4). We observed an overestimation of the cumulative ET for the standalone model, notably under the SL conditions, which may be associated with excessive ET rates in the early season.

The differences in crop-related variables between model approaches were not large, but more pronounced in SL than DL simulations. For the DL case, only the second and third-season simulations of LAI presented noticeable differences between models (Fig. 5). On the other hand, the SL simulations showed clear differences between models in all seasons for both LAI and SDM variables, where SWAP-SAMUCA simulations were systematically lower than the other two model approaches.

The impact of utilizing either DL or SL on crop and soil parameters is shown in Fig. 6. Optimized parameter controlling photosynthesis (<amax>, <eff>) and crop transpiration (<kc_min>, <eoratio>) presented very similar values for both DL and SL in tipping bucket and standalone versions. This was also the case for SWAP-SAMUCA, except for <amax> and <eff> parameters. Although the shape parameters controlling the rooting density profile presented some differences between SL and DL (<y_ini>, <tm>, <delta>), the pattern of the root density profile for all the models was similar (Fig A2). We also found differences between soil properties parameters when employing either SL or DL. For SWAP-SAMUCA, a gradient was observed for the soil parameters <osat>, <ksat> and <lexp> between top-to-bottom soil layers (Fig. 6). A similar pattern was observed for parameters <fcp>, <wpp> and <ksat> for the tipping bucket approach and to some extent for the standalone version. In some cases, the soil parameter values of SL were within the range of variation of its corresponding DL parameter, but it was not observed for some parameters such as <npar> and <lexp> (SWAP-SAMUCA).

3.2. Model evaluation across different Brazilian regions

The performance SWAP-SAMUCA and the tipping bucket approach were superior to the standalone approach in the evaluation dataset (Fig. 7, Table 2). This was evident for LAI simulations where the standalone version presented an error of about 47 % (RRMSE) while SWAP and tipping bucket approaches showed errors between 25 % and 28 % (Fig. 7). Further, the standalone approach was the only one presenting negative EF index values (-0,52) for LAI simulations in both DL and SL conditions. We observed that LAI simulations were consistently underpredicted at the CLER site (Fig. 8), which is mainly owed to the exacerbated water stress representation by this approach. As a result, SDM simulations were also consistently lower than the observations at this site for the standalone approach. On the other hand, the SWAP simulations of SDM at this site showed systematically higher values under DL conditions but presented good agreement for the SL simulations (Fig. 8).

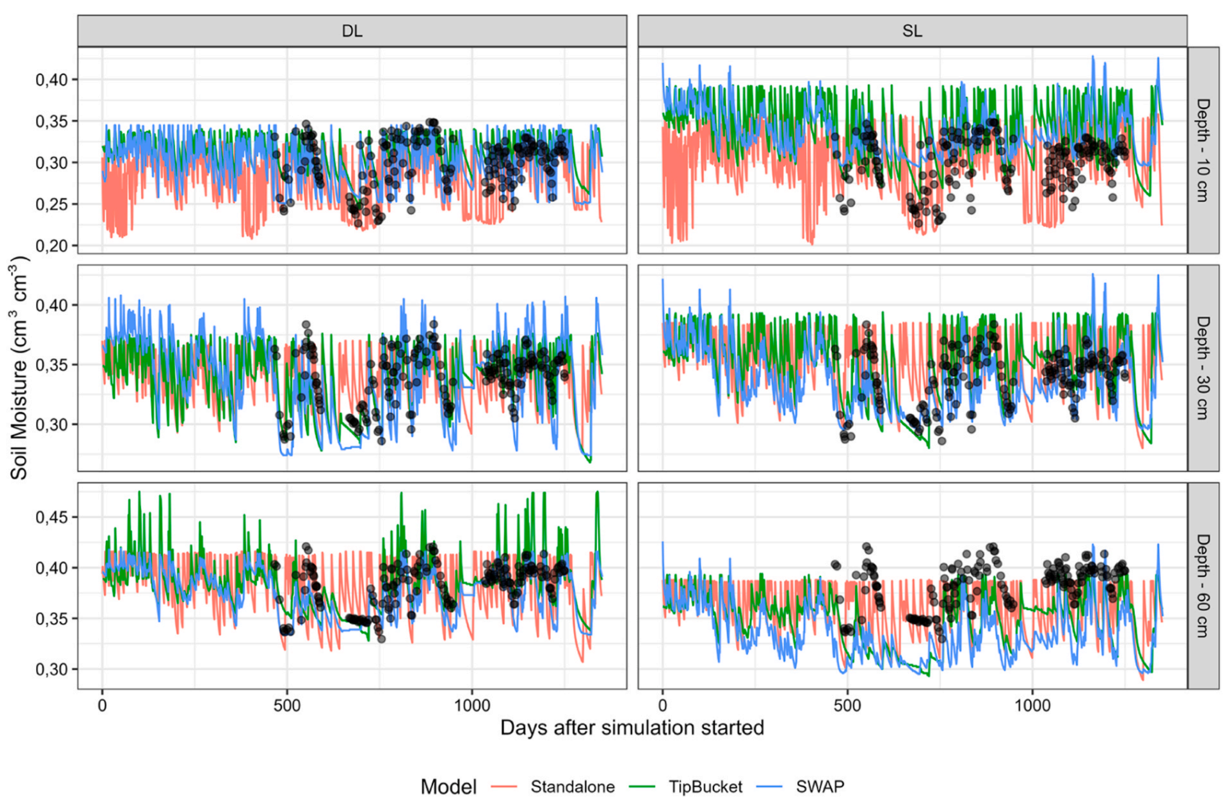

Fig. 3. Comparison between observed (dark points) and simulated (coloured lines) soil water content by each modelling approach (SWAP, TipBucket, Standalone), using different soil data types (SL and DL) at 10 cm, 30 cm and 60 cm soil depths at the calibration experiment of Piracicaba.

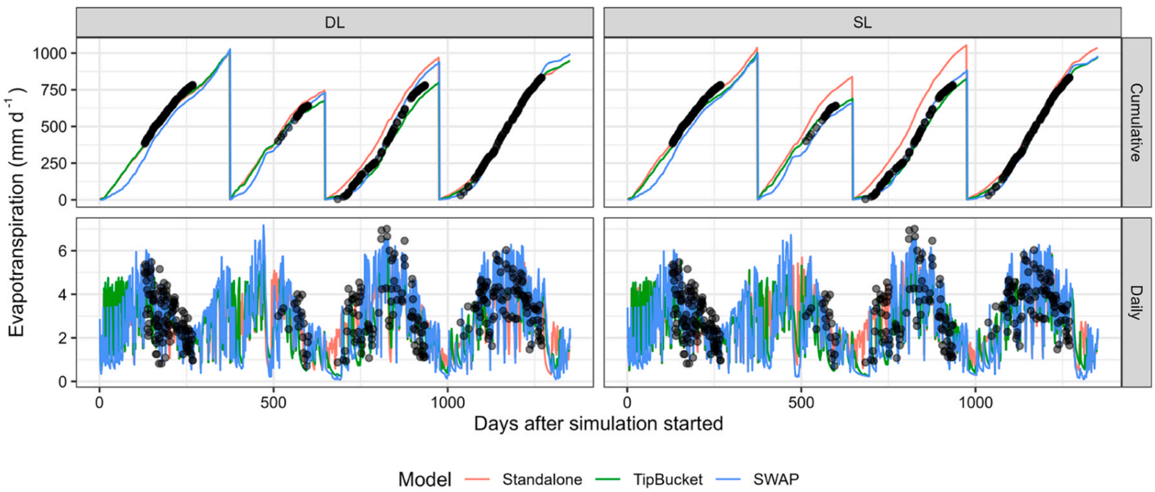

Fig. 4. Comparison between observed (dark points) and simulated (coloured lines) cumulative and daily evapotranspiration by each modelling approach (SWAP, TipBucket, Standalone), using different soil data types (SL and DL) at the calibration experiment of Piracicaba.

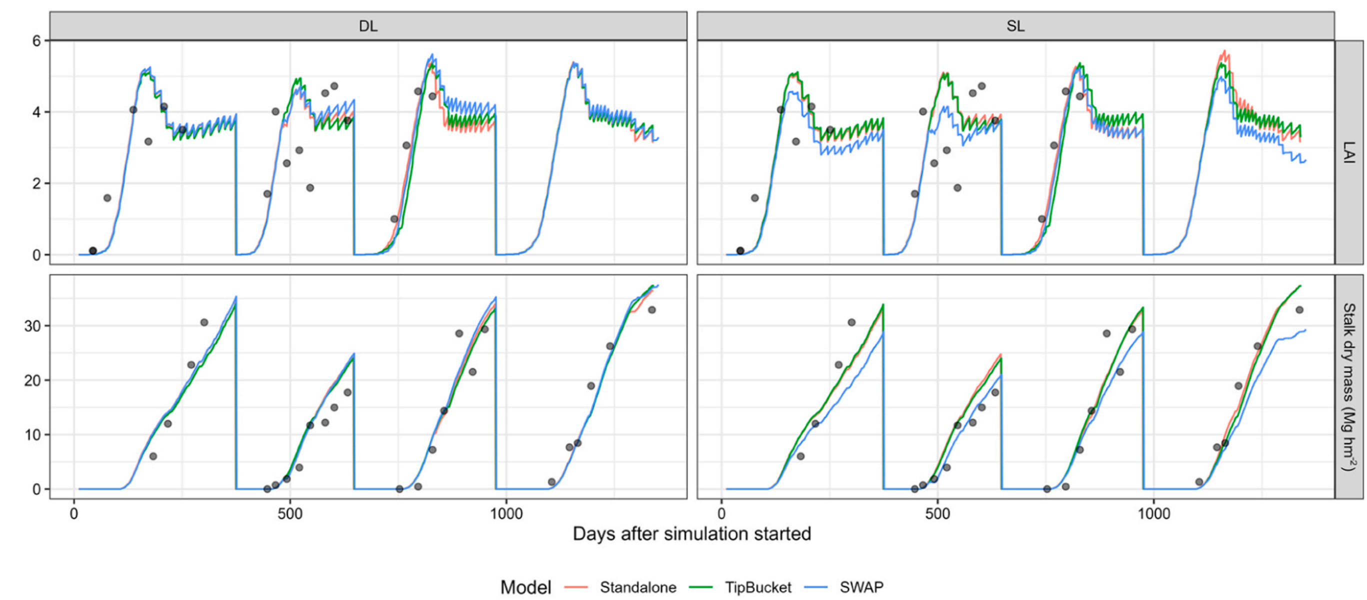

Fig. 5. Comparison between observed (dark points) and simulated (coloured lines) leaf area index (LAI) and stalk dry mass (SDM) by each modelling approach (SWAP, TipBucket, Standalone), using different soil data types (SL and DL) at the calibration experiment of Piracicaba.

The simulations for all other sites reproduced the observed pattern for LAI and SDM (Fig. 8). Comparing DL and SL results, SL showed 6 % and 11 % lower EF for SDM simulations compared to the DL with the standalone and SWAP approaches, respectively. For LAI, the DL simulations presented 23 % higher EF as compared to the SL condition for the tipping bucket approach (Fig. 7). All the other variables and indexes were generally improved under DL conditions, but differences were lower than 6 %.

3.3. The effect of model structure and soil detail on the simulations of sugarcane growth and water use in the long-term dataset

Using different model structures and soil data types significantly impacted the simulations of SDM and total irrigation in all sites of the long-term dataset (Fig. 9, Table 2). Simulations of SDM showed statistical difference between the SWAP and tipping-bucket approaches in all sites, with an average 7 % higher SDM simulations using SWAP. In most of DL cases, the SDM simulations between the standalone and SWAP method did not show statistical difference (except in Jatai). However, under SL data, the simulations of SDM for the standalone method were consistently lower than both SWAP (-10 %) and tipping-bucket (-3 %) approaches. When comparing soil data types (boxplots of same colour in Fig. 9), only the standalone approach showed statistical difference in SDM, where DL simulations were in average 9 % higher than SL simulations in across all sites.

The irrigation simulations were statistically different between modelling approaches in most of the sites. The exceptions are in DL simulations between SWAP and tipping-bucket in Piracicaba, and the standalone and tipping-bucket in Recife (Fig. 9). Although statistically different, the magnitude of differences between irrigation simulations of SWAP and tipping-bucket were smaller as compared to the standalone approach. For example, under SL conditions, the irrigation simulations of the standalone approach were between 21 % and 421 % times higher than SWAP and tipping-bucket approaches. Soil data type also substantially impacted the irrigation simulations for the standalone method. In all sites, the total irrigation simulated by the standalone

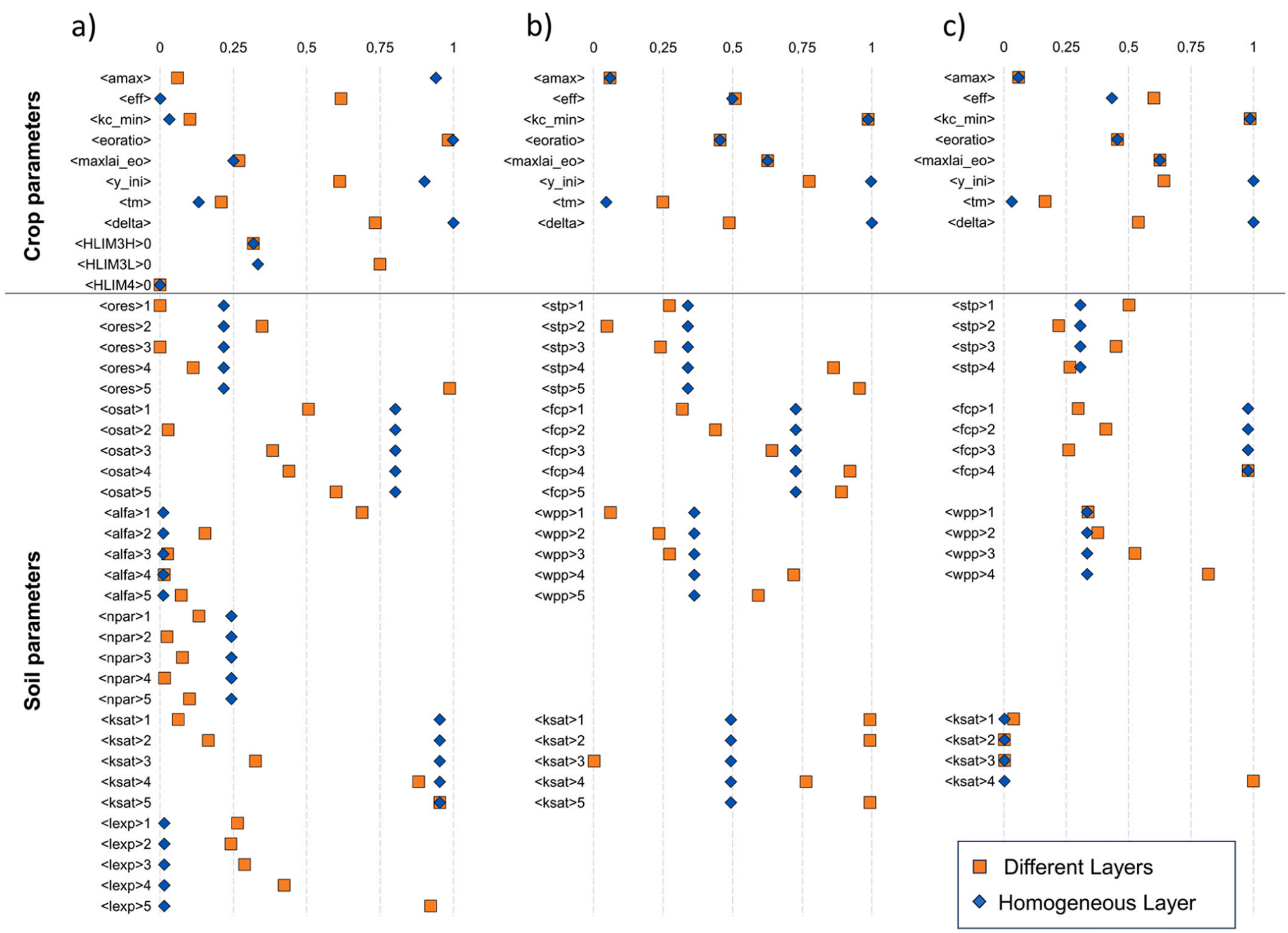

Fig. 6. Normalized parameter values after the optimization procedure for SWAP (a), Tipping-Bucket (b), and Standalone(c) approaches and different soil data types (Different Layers-DL, and Homogenous Layers-SL) utilising the calibration dataset of Piracicaba. Parameters were normalized within their minimum and maximum limits, which are defined in Tab. A2.

method was in average 47 % higher in SL conditions than in DL conditions. SWAP and tipping-bucket simulations between DL and SL soil data types were also statistically different in Petrolina and Recife, but the differences were smaller. In average, the irrigation simulations with SWAP-DL were 7 % lower than SWAP-SL, whereas tipping-bucket-DL were 20 % smaller than tipping-bucket-SL across all sites.

These differences were translated to the water use efficiency based on irrigated applications (IWUE) (Table 3). Therefore, IWUE values were in average 109 % higher in standalone-DL than standalone-SL simulations, whereas the relative differences between DL and SL simulations were +26 % for the tipping-bucket, and +6 % for the SWAP methods. The highest IWUE in dry mass were obtained using SWAP (16,5 g L^{-1}), followed by the tipping-bucket (12,5 g L^{-1}) and standalone (7,4 g L^{-1}) approaches. For EWUE, the differences between soil data type were rather small but the pattern was similar to IWUE, with higher relative differences between DL and SL for standalone approach (+7 %), followed by tipping-bucket (+4 %) and SWAP (+1 %). The EWUE simulations in Petrolina were in average (2,6 g L^{-1}) lower than other locations (3,5 g L^{-1}). Except for EWUE, we observed statistically significant interaction between model approaches and soil data types in all sites and other variables (Tabellen A4, A5).

Overall, the impact of choosing different soil data types was smaller for SWAP and tipping-bucket approaches but had a stronger effect on the standalone approach where all of the SDM and irrigation simulations were statistically different in all locations. The spatial pattern of higher-to-lower yields simulations were preserved in all combinations of model approaches and soil data types, with higher yields in Recife (47–52 Mg hm^{-2}), followed by Petrolina and Jatai (39–48 Mg hm^{-2}), and Piracicaba (34–43 Mg hm^{-2}). The simulations of total irrigation showed the highest values in the semi-arid site Petrolina (840–2262 mm $year^{-1}$), whereas in other sites it ranged from 48 to 1134 mm $year^{-1}$.

4. Discussion

4.1. Model calibration and evaluation: how good is enough?

While process-based crop models are powerful tools to formulate and test hypothesis about the functioning of agroecosystems, comprehensive field experiments with measured properties of the system are difficult to be conducted and are, therefore, scarce. As a result, process-based crop models are often calibrated and evaluated with limited information, which should expose the model to a range of conditions over which the key processes can be assessed. In our study, the field experiment of Piracicaba provides measured crop growth information and water fluxes both in soil and atmosphere that is used to calibrate and assess the models. Although the analysis in Section 3.1 showcases the model features and the impact of different model approaches and soil data detail on model performance, it can be prone to model over-fitting. In Section 3.2 we evaluated the models in a more rigorous manner, by using independent field trials across different pedoclimatic conditions, though soil moisture and evapotranspiration information were not available.

Nevertheless, the models showed similar performance in the

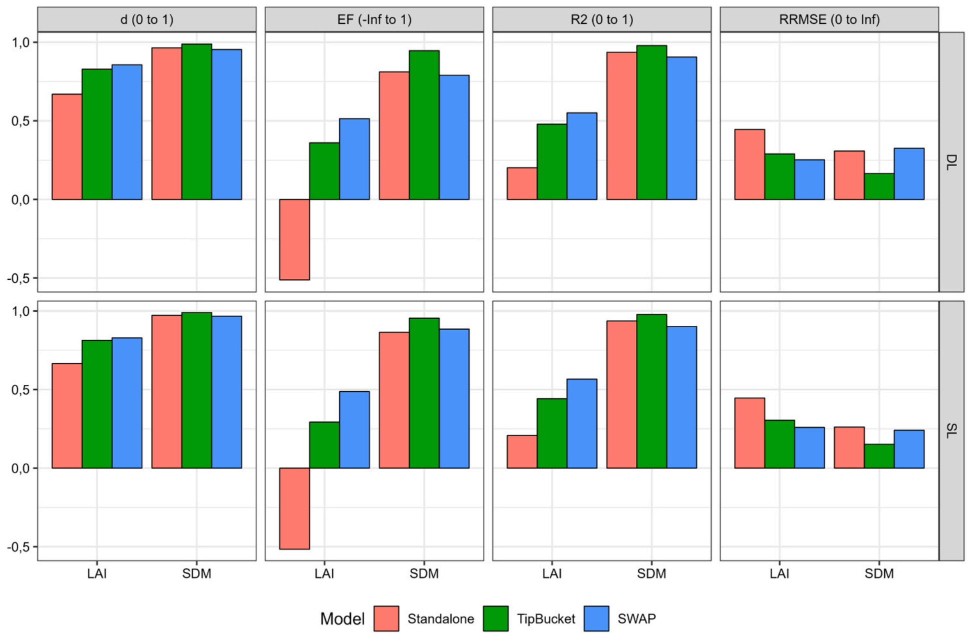

Fig. 7. Statistical indices of performance of each modelling approach (SWAP, TipBucket, Standalone), using different soil data types (SL and DL) in simulating the leaf area index (LAI) and stalk dry mass (SDM) at the evaluation experiments across Brazil. RRMSE is the Relative root mean squared error, R2 and d are respectively the model precision and accuracy, and EF the Nash–Sutcliffe efficiency.

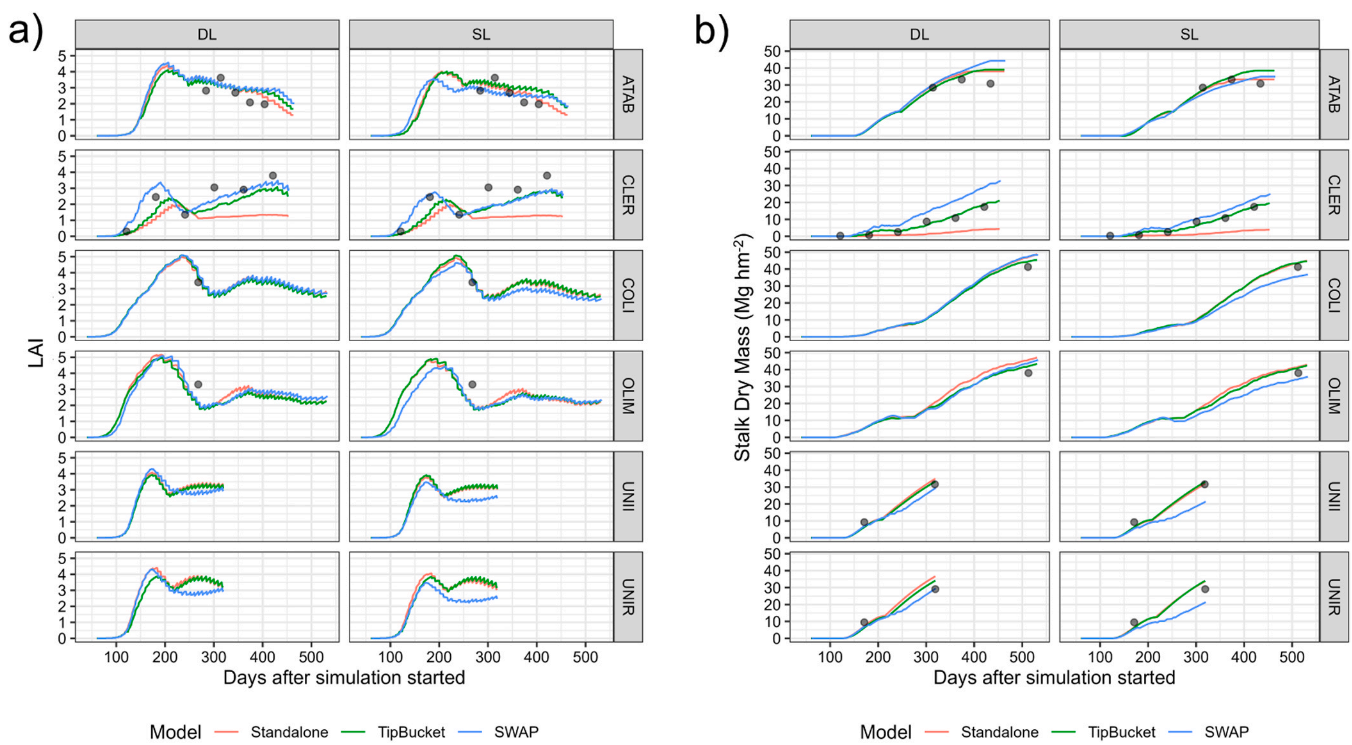

Fig. 8. Comparison between observed (dark points) and simulated (coloured lines) leaf area index (LAI) (a) and stalk dry mass (SDM) (b) by each modelling approach (SWAP, TipBucket, Standalone), using different soil data types (SL and DL) at the evaluation experiments across Brazil. ATAB: Ap. Taboado; CLER: Coruripe; COLI: Colina; OLIM: Olimpia; UNII: Uniao-Irrig; UNIR: Uniao-Rain.

calibration dataset and performed relatively well across the different regions of the evaluation set. The reference for our comparison is based on previous studies which assessed the SAMUCA (tipping-bucket) model using the same dataset, and found a modelling efficiency (EF) for soil moisture of 0,61 (Vianna et al., 2020), while the re-calibrated versions of the model using SWAP and the tipping-bucket approaches in our study showed higher EF values of 0,75 and 0,70, respectively. The other variables showed only small improvement, for example, SDM with the best EF of 0,89 for SWAP as compared to 0,87 by Vianna et al. (2020), or a reduction from 0,55 to 0,50 for LAI when using SWAP-SAMUCA. These performance metrics were also at the same level of accuracy as other sugarcane models, such as DSSAT-CANEGRO, APSIM-Sugar, MOSICAS

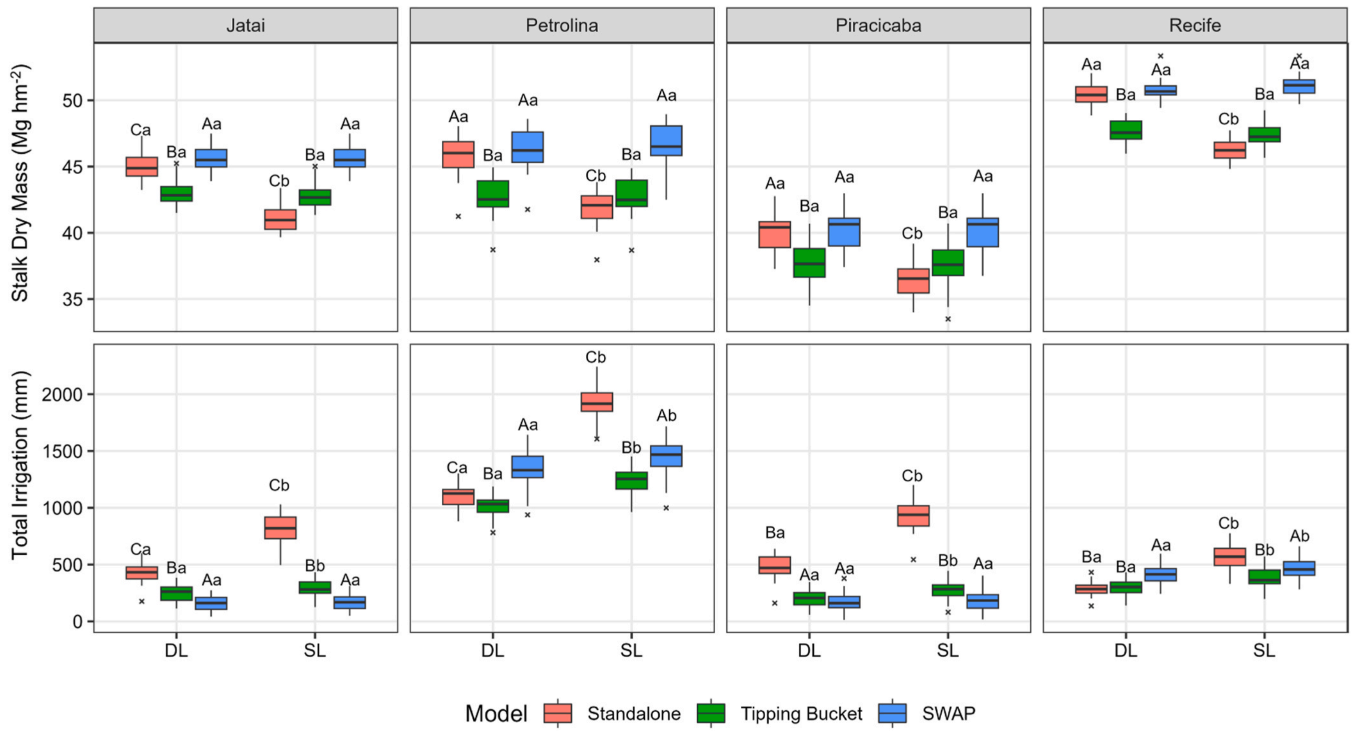

Fig. 9. Boxplots of final stalk dry mass yields and total irrigation applications simulated for each year from 1980–2010 by each modelling approach (Standalone, SWAP and Tipping Bucket) and using different soil data types (SL and DL) across four producing regions in Brazil (Jatai, Petrolina, Piracicaba and Recife). Capitalized letters compare the model approaches in each location for a given soil type [DL or SL] (i.e. compare between coloured boxplots), whereas the lowercase letters compare the soil data types in each location for a given model approach [standalone or tipping-bucket or SWAP] (i.e. compare between boxplots of the same colour).

Table 3
Simulated mean and standard deviation (brackets) of water use efficiencies in a dry mass (g L^{-1}) based on applied irrigation (IWUE) and evapotranspiration (EWUE) by each combination of modelling approach (SWAP, TipBucket, Standalone) and soil data types (SL and DL) at the four long-term simulations sites (1980–2010). Capitalized letters compare the model approaches in each location for a given soil type [DL or SL], whereas the lowercase letters compare the soil data types in each location for a given model approach [standalone or tipping-bucket or SWAP].

WUE in dry mass (g L-1)	Site	Standalone		Tipping Bucket		SWAP	
		DL	SL	DL	SL	DL	SL
IWUE	Jatai	10.4 (±1.6)^{Ca}	5 (±0.7)^{Cb}	17.2 (±1)^{Ba}	14.5 (±1.3)^{Bb}	28.7 (±2.9)^{Aa}	27.4 (±2.7)^{Ab}
	Petrolina	4.2 (±0.5)^{Ba}	2.2 (±0.2)^{Bb}	4.2 (±0.6)^{Ba}	3.5 (±0.5)^{Ab}	3.5 (±0.4)^{Aa}	3.3 (±0.3)^{Aa}
	Piracicaba	8.4 (±0.6)^{Ca}	3.9 (±0.5)^{Cb}	18.6 (±2)^{Ba}	13.7 (±1.5)^{Bb}	23.3 (±2.1)^{Aa}	22.4 (±1.1)^{Ab}
	Recife	17.4 (±1.9)^{Ca}	8.2 (±0.8)^{Cb}	15.7 (±1.6)^{Ba}	12.4 (±1.6)^{Bb}	12.2 (±1.5)^{Aa}	11 (±0.7)^{Ab}
EWUE	Jatai	3.7 (±0.4)^{Aa}	3.4 (±0.3)^{Aa}	3.8 (±0.6)^{Aa}	3.7 (±0.5)^{Aa}	3.8 (±0.6)^{Aa}	3.7 (±0.4)^{Aa}
	Petrolina	2.4 (±0.1)^{Aa}	2.4 (±0.2)^{Aa}	3.2 (±0.2)^{Ba}	2.8 (±0.3)^{Aa}	2.4 (±0.1)^{Aa}	2.4 (±0.1)^{Aa}
	Piracicaba	3.5 (±0.2)^{Aa}	3.2 (±0.4)^{Aa}	3.5 (±0.4)^{Aa}	3.5 (±0.3)^{Aa}	3.6 (±0.4)^{Aa}	3.6 (±0.4)^{Aa}
	Recife	3.6 (±0.5)^{Aa}	3.3 (±0.4)^{Aa}	3.7 (±0.2)^{Aa}	3.6 (±0.4)^{Aa}	3.3 (±0.2)^{Aa}	3.3 (±0.2)^{Aa}
IWUE	Mean	10.1 (±4.8)^{Ca}	4.8 (±2.2)^{Cb}	13.9 (±5.7)^{Ba}	11 (±4.4)^{Bb}	16.9 (±9.8)^{Aa}	16 (±9.4)^{Ab}
EWUE	Mean	3.3 (±0.5)^{Aa}	3.1 (±0.4)^{Aa}	3.6 (±0.2)^{Aa}	3.4 (±0.4)^{Aa}	3.3 (±0.5)^{Aa}	3.3 (±0.5)^{Aa}

and JULES, with RMSE values ranging between 4,3 to 12,8 Mg hm^{-2} for SDM and 1,1 to 1,6 for LAI (Marin et al., 2015; Hoffman et al., 2018; Jones and Singels, 2018; Dias et al., 2019; Christina et al., 2021; Vianna et al., 2022) (Tab. A6).

Although performance indices can be used as an indicator to monitor the progress of model improvement, it should not be solely the criteria for model suitability and selection. Model calibration typically requires optimisation of a given objective function to a calibration dataset and respective target variables. If not carefully analyzed, it could lead to misleading interpretation. For example, the best agreement for SWC simulations (SWAP-DL, Fig. 2) did not lead to the absolute best agreement for the LAI or ET variables (Fig. 2). In this case, the role of the modeller, who may identify an overfitting and avoid getting the "right answer for the wrong reason" is crucial (Keating, 2020). In addition, the human decision factor during the calibration steps has shown to be significant (Albanito *et al.*, 2022), whereas the definition of a consensus about a general calibration protocol for process-based crop models is still a work in progress (Wallach et al., 2021, 2023).

To assess how good is enough, the process-based modeler should take into account not only a referential level of performance, but also carry out a heuristic analysis on the spatiotemporal dynamics of the key variables and processes in question. We observed that the SWC dynamics simulated by the standalone approach showed a substantially different temporal dynamics as compared to SWAP and tipping-bucket approach (Fig. 3 and Fig. A3). This departure was not translated to other variables (e.g., LAI and SDM) for most cases, except in the Coruripe site where rooting depth is restricted to 40 cm and SWC becomes more important for crop growth. The drop in performance results between the calibration and evaluation simulations for the standalone approach also indicates that this approach has more limitations as compared to the SWAP and tipping-bucket methods. Yet, our analysis would be improved if more comprehensive datasets were available to evaluate other important processes concerning soil evaporation and root growth and water uptake dynamics (Singels et al., 2010).

4.2. How did we improve SAMUCA, limitations and opportunities for further research

Each modelling framework inherits unique features that may suit better for the given application or target system it was initially developed for, i.e. the “fit-for-purpose” concept. One model may require a simplified representation of soil processes whereas others would incorporate complex interactions mimicking the real system. Passioura (1996) hypothesises that this practice could lead to the categorisation of models for “scientific” or “engineering” applications, and the misuse of them could be detrimental to education or decision-making (comparing it to “snake oil” salesmen). Keating (2020) have recently revisited Passioura’s work and reinforced the warning of models in the context of “snake oil salesmen”, which could be mitigated with a rigorous scientific approach to model development, parameterisation and application.

In the context of soil water dynamics, Jarvis et al. (2022) postulate that the use of overly simplified soil water routines is unnecessary as physics-based flow models are at least as parsimonious and not difficult to parameterise. They also acknowledge the challenge on the parametrization of physically-based approaches to represent, for example, the complex role of soil macropores on the saturated hydraulic conductivity. With all that, modelers must face the challenge to investigate when enough detail is included parsimoniously depending on the research goal, scale and data availability (George et al., 2024). While tipping-bucket methods have been proved to work well for mimicking soil water fluxes for crop growth simulations (van Ittersum et al., 2003), the amount of “work-around fixes” needed to incorporate new features (e.g., capillary rise, leaching, dynamic watertable) may very well increase the number of empirical parameters needed, which are harder to measure or estimate as compared to physically-based properties. On the other hand, gradient-based models connect well with plant hydraulics theory and atmospheric sciences and provide a solid foundation for process understanding (Jarvis et al., 2022; George et al., 2024).

In this study, we equipped SAMUCA with two widely used soil water balance techniques for process-based crop models (tipping-bucket and Richards’ equation). An intercomparison of its performance under different locations and input soil information was assessed, including the simpler standalone water balance routine. Although the improvements in performance indices of using SWAP over tipping-bucket were small, having SAMUCA incorporated into SWAP opens new opportunities to explore and improve understanding of new aspects, for example, solute flow (vinasse application) (Christofoletti et al., 2013a, 2013b), heat-flow (mulch-cover) (Vianna et al., 2020), intercropping (Pinto et al., 2019), waterlogging (Nóia Júnior et al., 2023) and root water and nutrients uptake (de Melo and de Jong van Lier, 2021). This rather qualitative improvement should be acknowledged and used in favour of future applications of the SAMUCA model.

For our simulations, we did not consider soil temperature or mulch cover effect and also calibrated the soil hydraulic parameters. We opted to re-calibrate the three model approaches using the same optimization scheme to avoid any bias effect from a previous calibration procedure (Wallach et al., 2023), thereby isolating the impact of model structure and soil data quality. As a result, we observed that for each combination of model approach and data type, the crop and soil parameter values slightly changed within their limit range that was set according to the literature (de Melo and de Jong van Lier, 2021; Pereira et al., 2021). The reasons for that may very well lie in the differences in model structure that the parameter values would compensate for during the optimization procedure. Additionally, even within the expected plausible range of variation, multiple combinations of parameter values can lead to the same or very similar model performance. This makes it difficult to compare parameter values between modelling approaches. However, physically-based approaches such as SWAP may provide good approximations of the real system, i.e., via inverse modelling concept (Ines and Droogers, 2002; de Jong van Lier et al., 2015).

Another overlooked aspect is the natural variability of observations which are generally not taken into account and difficult to infer for such comprehensive datasets. Stochastic techniques can be used to explore this aspect (Marin et al., 2017; Pereira et al., 2021), but the definition of a plausible range of parameter values and their interdependencies is still very challenging. The use of simplified input information also has a significant effect on the model results, as observed in our study. This is especially important when the goal of the work is directly related to the input information in question (e.g., soil water dynamic). In the same direction, model ensembles have proved their role in increasing prediction accuracy and quantification of uncertainty (Martre et al., 2015; Wallach et al., 2018), but also at the cost of the heuristic feature of the models, limiting the understanding of emerging properties of the system (Yin et al., 2021).

Although we did not modify crop-related algorithms, our results provide insights into the indirect effect of below-ground processes and water dynamics on crop growth and irrigation simulations. The standalone water routine neglects important components of water balance, such as runoff and capillary rise. But most importantly, the highly oscillating pattern in SWC simulations with this routine may be directly associated with the fact that the original method presented by Teh (2006) was not developed and tested for daily time step and the limited number of soil layers. This is more evident especially in dry periods (e.g., between DAS 500–750 in Fig. 3), where the mass balance using this approach not always match (Fig. A5).

Moreover, the standalone approach showed the higher differences in long-term simulations when the model was employed under DL or SL conditions (Fig. 9 and Table 3). In contrast, the SWAP and tipping-bucket approaches showed the highest robustness to soil data detail as the difference between long-term simulations under DL and SL were small for both methods (Fig. 9 and Table 3). However, it must be noted that our SL soil properties were generated with observed average DL data, which does not imply that gridded and/or coarser resolution datasets always agree with local DL information. Furthermore, there’s a growing number of methods that can be used to estimate vertical soil properties, and some are freely available for end-users (Hengl et al., 2017). Future studies could systematically assess the impact of choosing different techniques for acquiring soil properties on process-based crop simulations.

In a cross-site analysis, our results showed higher differences for IWUE than EWUE simulations, with EWUE values consistently lower in Petrolina site than in other sites which agrees with the higher soil evaporation in this semi-arid location. Our simulated EWUE values were also close to the range observed by Olivier and Singels (2015) of between 2,4 to 3,1 dry mass (g L^{-1}) (assuming a typical 70 % moisture in fresh stalks). A comprehensive field experiment with clear water stress impact on sugarcane is still missing. As observed in our results, only the CLER site presents a significant water stress condition, and this is mainly imposed by rooting depth restrictions as pointed out in previous studies (Marin et al., 2015; Vianna et al., 2022). Sugarcane roots can surpass 4 m in depth giving more access to the water supply in the soil, which is reported as the main cause for small or non-significant differences of SDM in field trials testing irrigated versus rainfed conditions (Suguitani, 2006; Laclau and Laclau, 2009). This is also evident at the Uniao site with irrigated and rainfed conditions, where less than a 5 % difference in SDM was observed between both treatments. The conjunction of this factor with the long-lasting characteristics of the sugarcane growing season and subsequent regrows prompts the development and assessment of algorithms that better represent root growth dynamics and water and nutrient uptake (de Jong van Lier et al., 2008; de Melo and de Jong van Lier, 2021).

5. Conclusions

This study examined the influence of different soil water balance approaches and soil data detail on sugarcane growth and irrigation scheduling simulations. The use of Richard's equation approach (SWAP) showed slightly superior performance than the tipping-bucket method for soil moisture simulations, while both were superior than the standalone method. Other variables, such as LAI and SDM presented similar performance using these both approaches, but this may be due to lack of comprehensive dataset on sugarcane growth under extreme drought or wet conditions. In terms of soil data detail, both methods (SWAP and tipping-bucket) showed higher robustness in long-term simulations as compared to the standalone approach which was more sensitive to soil detail conditions. In addition, the original standalone approach showed inconsistent spatiotemporal pattern as compared to the other routines, and therefore, should be avoided with the SAMUCA model. The differences between SWAP and tipping-bucket simulations were small, so modelers may select the appropriate approach based on the research question and data availability. While the tipping-bucket approach may be, at current date, easier to upscale and be employed at regional scale, the SWAP method opens new opportunities for process-understanding as it approximates to the physical system.

CRediT authorship contribution statement

Murilo dos Santos Vianna: Writing – review & editing, Writing – original draft, Validation, Software, Methodology, Investigation, Formal analysis, Data curation, Conceptualization. **Fabio Ricardo Marin:** Writing – review & editing, Supervision, Project administration, Funding acquisition, Conceptualization. **Thomas Gaiser:** Writing – review & editing, Supervision. **Quirijn de Jong van Lier:** Writing – review & editing, Supervision, Conceptualization. **Klaas Metselaar:** Writing – review & editing, Supervision, Software, Project administration, Conceptualization.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Data will be made available on request.

Acknowledgements

Primary funding sources of this research include the São Paulo Research Foundation (FAPESP grants 2014/05887-6, 2017/20925-0, 2017/50445-0, 2015/16214-5 and 2021/00720,0) and the Brazilian National Council for Scientific and Technological Development (CNPq grants 301424/2015-2, 401662/2016-0, 425174/2018-2, 302597/2021-2). We also would like to express our gratitude to the staff of the Group of Experimentation and Research on Agriculture and Modelling (Esalq-Clima) for the support given in field data collection and maintenance, and to the College of Agriculture "Luiz de Queiroz" (ESALQ/USP) for providing the field experiment area. The authors gratefully acknowledge the granted access to the Bonna cluster hosted by the University of Bonn along with the support provided by its High-Performance Computing & Analytics Lab.

Appendix A

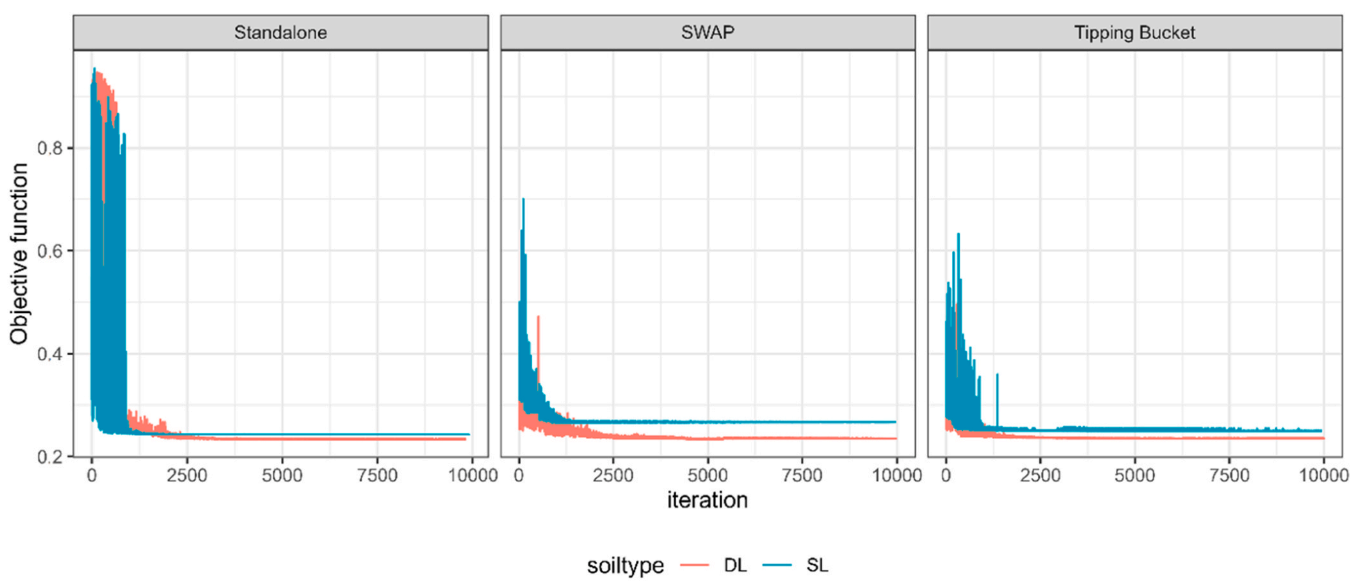

Fig A1. Objective function values as a function of iteration steps in the optimization process for each combination of the SAMUCA model and soil water balance routine (Standalone, SWAP, Tipping Bucket) using detailed soil profile data (DL) or a homogenous single layer (SL). Iterations were set to a maximum of 10000 and the test was repeated ten times to check the robustness of results.

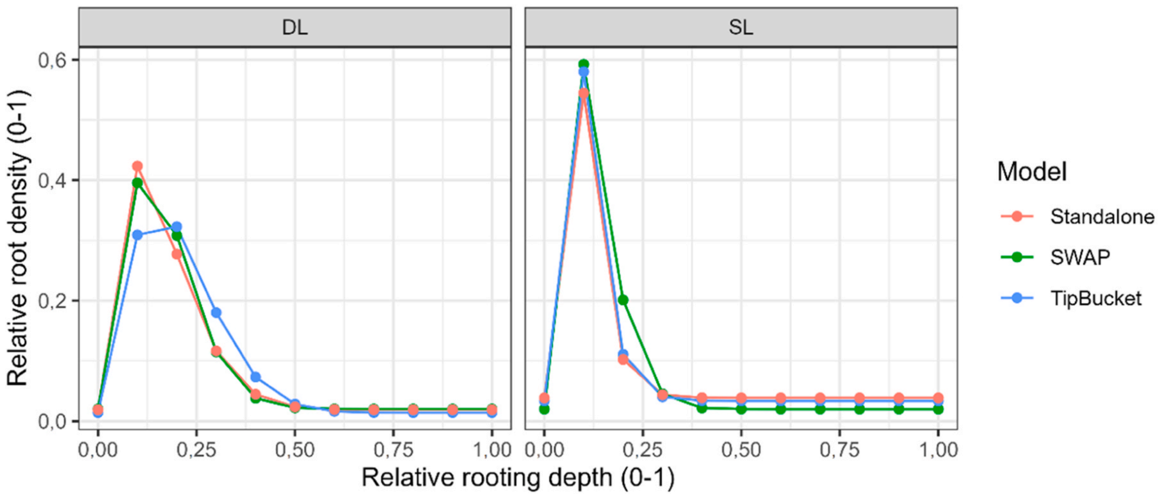

Fig A2. Relative root profile as a function of the relative rooting depth after the optimization procedure for the SAMUCA model and soil water balance routine (Standalone, SWAP, Tipping Bucket) using detailed soil profile data (DL) or a homogenous single layer (SL). Curves were adjusted with a sigmoid curve controlled by parameters <y_ini>, <tm> and <delta> (Tab. A2).

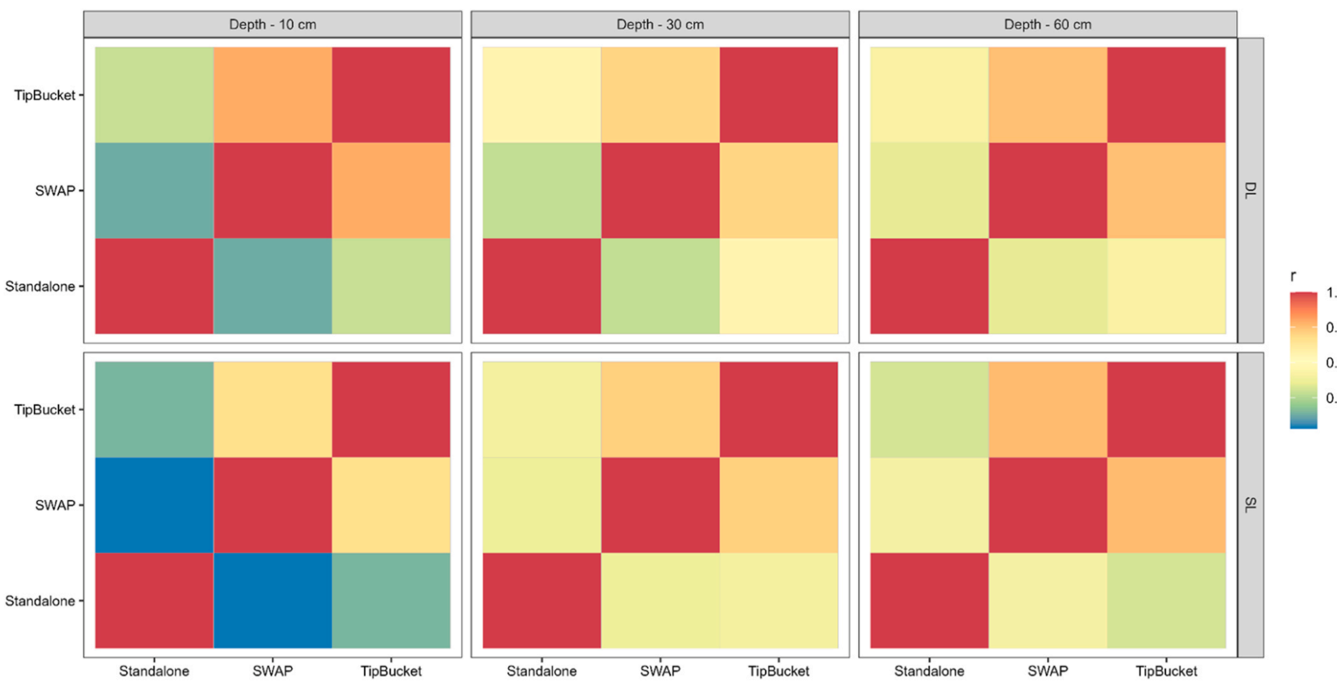

Fig A3. Pearson correlation indices (r) between simulated soil moisture content by each modelling approach (SWAP, Tipping-Bucket, Standalone), using different soil data types (SL and DL), and for the 10 cm, 30 cm and 60 cm depth in the calibration site Piracicaba, Brazil. The color scale represents the Pearson correlation indices (r) between each modelling approach. When the values of Pearson correlation indices (r) are closer to 1 (red) it indicates high correlation between the corresponding modelling approaches.

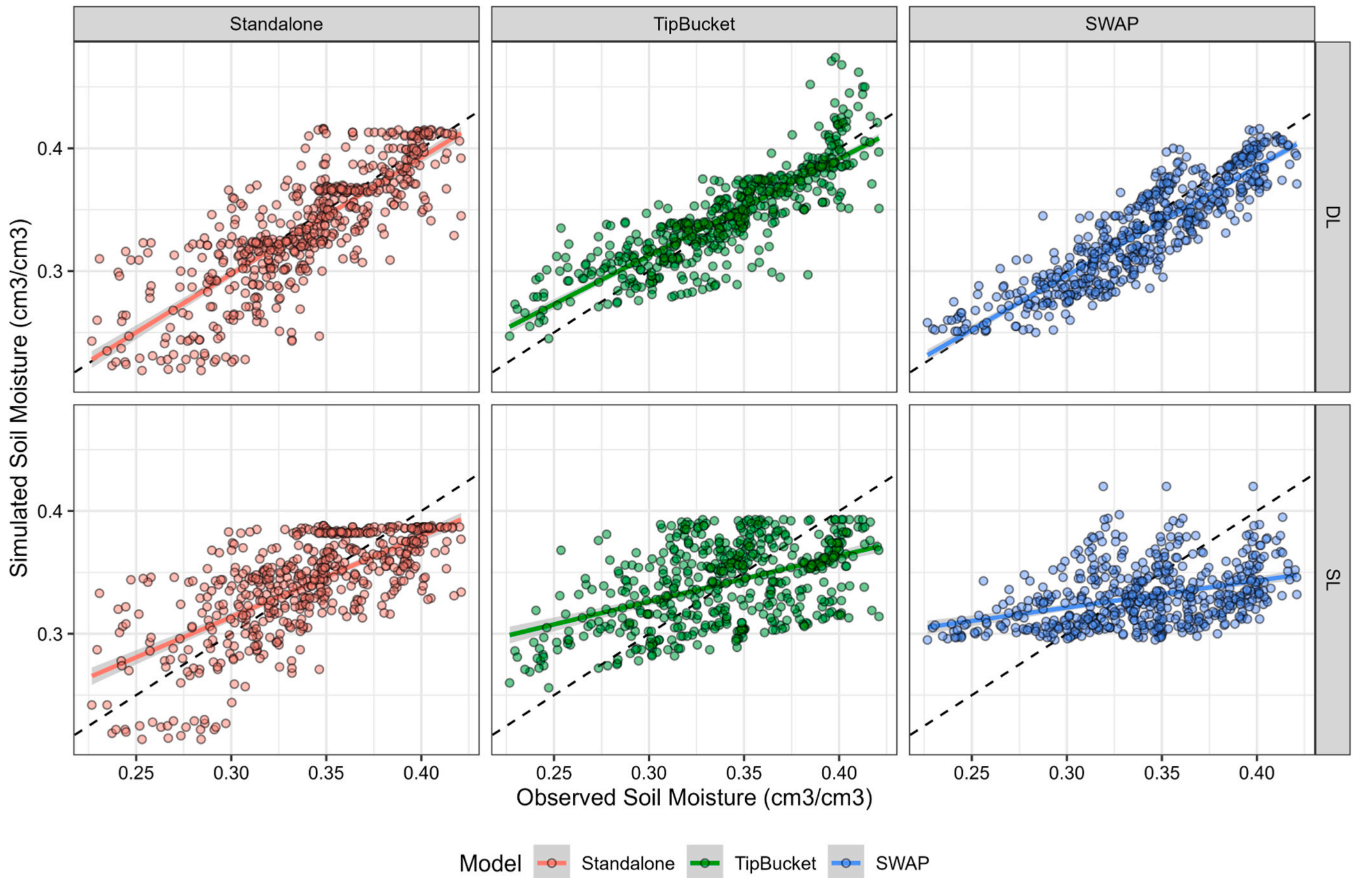

Fig A4. Scatter plot of observed versus simulated soil moisture content by each modelling approach (SWAP, Tipping-Bucket, Standalone), using different soil data types (SL and DL), for the calibration site in Piracicaba, Brazil.

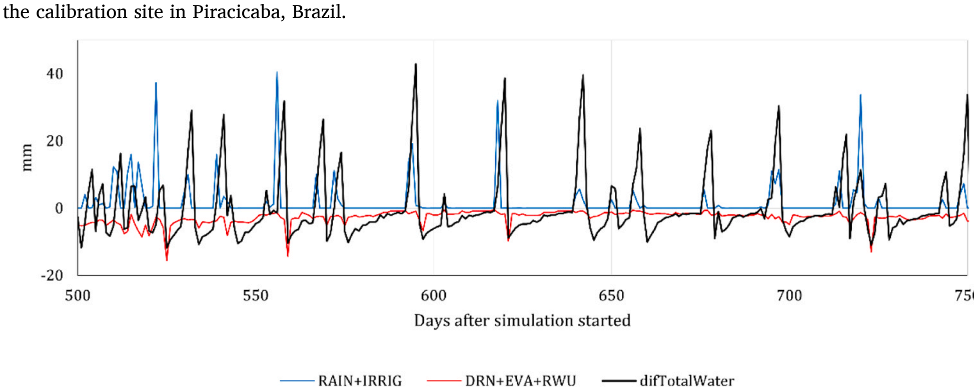

Fig A5. Time course simulations of water balance components by the Standalone version (Marin and Jones, 2014) at the Piracaba site between 500 and 750 days after simulation started (Fig. 3). RAIN+IRRIG is the incoming daily amount of water considered by the model as infiltration; DRN+EVA+RWU is outcoming (negative) daily amount of water considered by the model as deep drainage, soil evaporation and root water uptake; difTotalWater is daily difference in total water volume stored in the soil column. Days when difTotalWater > RAIN+IRRIG or difTotalWater < DRN+EVA+RWU indicate mass balance mismatch.

Table. A1. Soil hydraulic characteristics for different soil (DL) layers and single layer (SL). Where WP, FC and SP are the soil water contents at the wilting point, field capacity and saturation, in cm^3 cm^{-3}; ThetaRes and ThetaSat are the residual and saturation water content; alpha and n are the Mualen van Genuchten parameters; KSat is the soil saturated hydraulic conductivity, in cm day^{-1}; Size Sub-Comp is the corresponding sub-compartment size for each layer in SWAP, in cm; and the standalone layers defined for each of the 1–4 layers (average used for same layer number). SL data was calculated as the weighted mean of layers' depth provided in DL, and used the same vertical discretization as DL

Soil Type	Site	Depth	WP	FC	SP	alpha	n	ThetaRes	ThetaSat	KSat	Size Sub-Comp	Standalone Layer
DL	Piracicaba	5	0210	0280	0379	0153	1076	0000	0379	40,8	2,50	1
		15	0240	0300	0361	0069	1058	0000	0361	24,0	2,50	2
		30	0250	0320	0365	0020	1066	0000	0365	24,0	5,00	3
		60	0320	0390	0440	0029	1052	0000	0440	14,4	5,00	3
		100	0230	0399	0475	0005	1165	0000	0475	4,8	10,00	4
	Ap, Taboado	15	0200	0360	0430	0007	1163	0000	0430	57,6	2,50	1

(continued on next page)

(*continued*)

Soil Type	Site	Depth	WP	FC	SP	alpha	n	ThetaRes	ThetaSat	KSat	Size Sub-Comp	Standalone Layer
		60	0210	0340	0470	0038	1126	0000	0470	57,6	5,00	2
		90	0220	0330	0480	0100	1106	0000	0480	57,6	7,50	3
		120	0250	0330	0490	0323	1081	0000	0490	57,6	10,00	4
	Colina	15	0110	0210	0410	0137	1173	0000	0410	69,6	2,50	1
		50	0120	0270	0350	0008	1222	0000	0350	45,6	5,00	2
		90	0120	0290	0400	0010	1238	0000	0400	40,8	10,00	3
	Olimpia	21	0150	0360	0490	0009	1236	0000	0490	43,2	3,00	1
		58	0160	0370	0480	0007	1230	0000	0480	36,0	7,40	2
		112	0160	0370	0480	0007	1230	0000	0480	19,2	9,00	3
		130	0160	0370	0480	0007	1230	0000	0480	16,8	9,00	4
	Coruripe	5	0044	0076	0287	0254	1273	0000	0287	76,8	2,50	1
		15	0050	0100	0241	0196	1204	0000	0241	76,8	2,50	2
		40	0112	0164	0232	0095	1100	0000	0232	76,8	5,00	3
	Uniao	5	0106	0201	0344	0074	1167	0000	0344	28,8	2,50	1
		15	0090	0220	0332	0016	1236	0000	0332	24,0	2,50	2
		75	0077	0201	0310	0015	1254	0000	0310	14,4	6,00	3
		120	0077	0201	0300	0013	1255	0000	0300	14,4	9,00	4
	Jatai	15	0126	0208	0451	0282	1160	0000	0451	43,2	3,00	1
		30	0115	0189	0430	0296	1167	0000	0430	26,4	5,00	2
		70	0129	0207	0422	0286	1148	0000	0422	19,2	5,00	3
		90	0140	0220	0402	0252	1131	0000	0402	4,8	10,00	4
	Petrolina	45	0216	0321	0436	0057	1103	0000	0436	4,8	3,00	1
		90	0208	0313	0434	0063	1106	0000	0434	36,0	7,50	2
		120	0221	0324	0439	0061	1100	0000	0439	36,0	10,00	3
	Recife	15	0089	0245	0461	0032	1264	0000	0461	36,0	2,50	1
		30	0274	0436	0491	0005	1133	0000	0491	38,4	5,00	2
		60	0230	0350	0564	0185	1114	0000	0564	36,0	5,00	3
		90	0201	0302	0474	0174	1109	0000	0474	36,0	10,00	4
	Piracicaba	5	0216	0285	0380	0147	1073	0000	0380	33,6	2,50	1
		15	0240	0303	0352	0033	1061	0000	0352	31,2	2,50	2
		45	0278	0347	0390	0019	1059	0000	0390	31,2	5,00	3
		90	0307	0394	0428	0007	1070	0000	0428	26,4	9,00	3
		120	0253	0393	0456	0008	1122	0000	0456	33,6	10,00	4
SL	Piracicaba	100	0260	0373	0432	0011	1098	0000	0432	40,8	Same as DL	
	Ap, Taboado	120	0221	0338	0473	0062	1110	0000	0473	26,4		
	Colina	90	0118	0269	0382	0013	1218	0000	0382	21,6		
	Olimpia	130	0158	0368	0482	0008	1231	0000	0482	4,8		
	Coruripe	40	0088	0137	0241	0236	1125	0000	0241	4,8		
	Uniao	120	0079	0203	0310	0015	1248	0000	0310	14,4		
	Jatai	90	0129	0207	0424	0284	1149	0000	0424	57,6		
	Petrolina	120	0214	0319	0436	0060	1103	0000	0436	48,0		
	Recife	90	0204	0331	0505	0084	1126	0000	0505	26,4		
	Piracicaba	120	0277	0370	0417	0012	1079	0000	0417	76,8		

c: data used in the calibration runs; s: data used in the long-term runs (Vianna and Sentelhas, 2016)

Table. A2. Calibrated parameter values for each model approach (SWAP, Tipping-Bucket, and Standalone) and soil data type (SL, DL), with their respective minimum and maximum ranges (defined from literature and field data). The relative results of this table are shown in Fig. 6, whereas the definition of each parameter is given in Tab. A3. The resulting roots distribution are shown in Fig. A2.

SWAP					Tipping-Bucket					Standalone				
Parameter	min	max	DL	SL	Parameter	min	max	DL	SL	Parameter	min	max	DL	SL
<amax>	40	62	41,30	60,68	<amax>	40	62	41,30	41,30	<amax>	40	62	41,30	41,30
<eff>	0,04	0,08	0,06	0,04	<eff>	0,04	0,08	0,06	0,06	<eff>	0,04	0,08	0,06	0,06
<kc_min>	0,05	0,8	0,13	0,07	<kc_min>	0,05	0,8	0,79	0,79	<kc_min>	0,05	0,8	0,79	0,79
<eoratio>	0,75	1,3	1,29	1,30	<eoratio>	0,75	1,3	1,00	1,00	<eoratio>	0,75	1,3	1,00	1,00
<maxlai_eo>	2	6	3,07	3,00	<maxlai_eo>	2	6	4,50	4,50	<maxlai_eo>	2	6	4,50	4,50
<y_ini>	0,5	20	12,43	18,08	<y_ini>	0,5	20	15,60	19,94	<y_ini>	0,5	20	13,00	20,00
<tm>	0001	0,6	0,1260	0,0800	<tm>	0001	0,6	0,1500	0,0280	<tm>	0001	0,6	0,1000	0,0200
<delta>	0,5	20	14,8	20,0	<delta>	0,5	20	10,0	20,0	<delta>	0,5	20	11,0	20,0
<HLIM3H>0	-1600	-400	-1219	-1218										
<HLIM3L>0	-3326	-1600	-2033	-2750										
<HLIM4>0	-16000	-10000	-16000	-16000										
<ores>1	0	0,23	0,00	0,05	<stp>1	0,34	0,54	0,39	0,41	<stp>1	0,34	0,54	0,44	0,40
<ores>2	0	0,23	0,08	0,05	<stp>2	0,34	0,54	0,35	0,41	<stp>2	0,34	0,54	0,38	0,40
<ores>3	0	0,23	0,00	0,05	<stp>3	0,34	0,54	0,39	0,41	<stp>3	0,34	0,54	0,43	0,40
<ores>4	0	0,23	0,03	0,05	<stp>4	0,34	0,54	0,51	0,41	<stp>4	0,34	0,54	0,39	0,40
<ores>5	0	0,23	0,23	0,05	<stp>5	0,34	0,54	0,53	0,41					
<osat>1	0,34	0,52	0,43	0,48	<fcp>1	0,26	0,4	0,30	0,36	<fcp>1	0,26	0,4	0,30	0,41
<osat>2	0,34	0,52	0,35	0,48	<fcp>2	0,26	0,4	0,32	0,36	<fcp>2	0,26	0,4	0,32	0,41
<osat>3	0,34	0,52	0,41	0,48	<fcp>3	0,26	0,4	0,35	0,36	<fcp>3	0,26	0,4	0,30	0,41
<osat>4	0,34	0,52	0,42	0,48	<fcp>4	0,26	0,4	0,39	0,36	<fcp>4	0,26	0,4	0,41	0,41
<osat>5	0,34	0,52	0,45	0,48	<fcp>5	0,26	0,4	0,38	0,36					
<alfa>1	0	0,98	0675	0011	<wpp>1	0,19	0,35	0200	0248	<wpp>1	0,19	0,35	0244	0243

(*continued on next page*)

(*continued*)

SWAP					Tipping-Bucket					Standalone				
Parameter	min	max	DL	SL	Parameter	min	max	DL	SL	Parameter	min	max	DL	SL
<alfa>2	0	0,98	0150	0011	<wpp>2	0,19	0,35	0228	0248	<wpp>2	0,19	0,35	0250	0243
<alfa>3	0	0,98	0025	0011	<wpp>3	0,19	0,35	0234	0248	<wpp>3	0,19	0,35	0274	0243
<alfa>4	0	0,98	0014	0011	<wpp>4	0,19	0,35	0305	0248	<wpp>4	0,19	0,35	0321	0243
<alfa>5	0	0,98	0071	0011	<wpp>5	0,19	0,35	0285	0248					
<npar>1	1,04	1,5	1101	1152										
<npar>2	1,04	1,5	1051	1152										
<npar>3	1,04	1,5	1075	1152										
<npar>4	1,04	1,5	1047	1152										
<npar>5	1,04	1,5	1086	1152										
<ksat>1	2,64	1000	63,9	953,2	<ksat>1	0,11	42	39,7	19,8	<ksat>1	0,11	42	1,7	0,2
<ksat>2	2,64	1000	166,0	953,2	<ksat>2	0,11	42	39,7	19,8	<ksat>2	0,11	42	0,2	0,2
<ksat>3	2,64	1000	326,6	953,2	<ksat>3	0,11	42	0,2	19,8	<ksat>3	0,11	42	0,2	0,2
<ksat>4	2,64	1000	881,3	953,2	<ksat>4	0,11	42	30,6	19,8	<ksat>4	0,11	42	42,0	0,2
<ksat>5	2,64	1000	953,3	953,2	<ksat>5	0,11	42	39,7	19,8					
<lexp>1	-5,3	-0,1	-3,93	-5,23										
<lexp>2	-5,3	-0,1	-4,04	-5,23										
<lexp>3	-5,3	-0,1	-3,80	-5,23										
<lexp>4	-5,3	-0,1	-3,10	-5,23										
<lexp>5	-5,3	-0,1	-0,51	-5,23										

Table. A3

Model parameter names, definitions and units

Parameter	Definition	Unit
<amax>	Assimilation rate at light saturation point	μmol/m2/s
<eff>	Carboxylation efficiency	μmol/m2/s (μmol/m2/s)-1
<kc_min>	Minimum crop coefficient (Kcb_ini)	dimensionless
<eoratio>	Ratio between maximum LAI and crop coefficient at maximum canopy formation	dimensionless
<maxlai_eo>	Leaf area index when maximum evapotranspiration occurs	m2/m2
<y_ini>	Empirical coefficient defining vertical root profile	dimensionless
<tm>	Empirical coefficient defining vertical root profile (mid-point)	dimensionless
<delta>	Empirical coefficient defining vertical root profile (shape)	dimensionless
<HLIM3H>	Feddes root water uptake pressure head limit 3High (SWAP)	cm
<HLIM3L>	Feddes root water uptake pressure head limit 3Low (SWAP)	cm
<HLIM4>	Feddes root water uptake pressure head limit 4 (SWAP)	cm
<ores>	Residual soil water content (SWAP)	cm3/cm3
<osat>	Saturated soil water content (SWAP)	cm3/cm3
<alfa>	Parameter alfa of main drying curve (SWAP)	1/cm
<npar>	Parameter n of MvG model (SWAP)	dimensionless
<ksat>	Soil hydraulic conductivity at saturation	cm/d, cm/h
<lexp>	Exponent in hydraulic conductivity function (SWAP)	dimensionless
<stp>	Saturated water content (Tipping-bucket/Standalone)	cm3/cm3
<fcp>	Field capacity point (Tipping-bucket/Standalone)	cm3/cm3
<wpp>	Wilting point (Tipping-bucket/Standalone)	cm3/cm3

Table. A4

Two-way ANOVA results comparing 30-year simulations of Total irrigation (Tirr) and Stalk Dry Mass (SDM) with different Models (Standalone, Tipping-Bucket, SWAP) and Soils (DL and SL). P-values lower than 0.01 represent statistically significance of 1 %

Variable	Site	Sources of variation	Degrees of Freedom	Sum of Squares	Mean of Squares	F-value	P-value
Tirr	Recife	Model	2	3.35E+05	1.67E+05	20.2	1.2E-08
		Soil	1	8.50E+05	8.50E+05	102.7	2.2E-19
		Model:Soil	2	4.59E+05	2.29E+05	27.7	3.3E-11
		Residuals	180	1.49E+06	8.28E+03		
	Piracicaba	Model	2	1.04E+07	5.22E+06	538.7	1.0E-76
		Soil	1	1.51E+06	1.51E+06	156.2	3.3E-26
		Model:Soil	2	1.86E+06	9.30E+05	96.0	4.2E-29
		Residuals	180	1.74E+06	9.68E+03		
	Jatai	Model	2	7.19E+06	3.60E+06	489.2	1.7E-73
		Soil	1	9.88E+05	9.88E+05	134.4	1.4E-23
		Model:Soil	2	1.34E+06	6.72E+05	91.4	4.1E-28
		Residuals	180	1.32E+06	7.35E+03		
	Petrolina	Model	2	4.96E+06	2.48E+06	140.4	1.8E-37
		Soil	1	6.59E+06	6.59E+06	373.4	9.2E-46
		Model:Soil	2	4.53E+06	2.27E+06	128.3	2.3E-35
		Residuals	180	3.18E+06	1.77E+04		
SDM	Recife	Model	2	383.7	191.8	313.6	2.2E-59

(*continued on next page*)

Table. A4 (*continued*)

Variable	Site	Sources of variation	Degrees of Freedom	Sum of Squares	Mean of Squares	F-value	P-value
		Soil	1	88.2	88.2	144.2	9.0E-25
		Model:Soil	2	187.5	93.7	153.2	1.4E-39
		Residuals	180	110.1	0.6		
	Piracicaba	Model	2	215.9	108.0	45.7	8.7E-17
		Soil	1	64.9	64.9	27.5	4.4E-07
		Model:Soil	2	118.9	59.5	25.2	2.3E-10
		Residuals	180	424.9	2.4		
	Jatai	Model	2	307.4	153.7	171.5	2.1E-42
		Soil	1	87.3	87.3	97.4	1.2E-18
		Model:Soil	2	150.9	75.5	84.2	1.6E-26
		Residuals	180	161.4	0.9		
	Petrolina	Model	2	469.5	234.8	119.3	1.0E-33
		Soil	1	61.1	61.1	31.0	9.1E-08
		Model:Soil	2	181.5	90.7	46.1	6.8E-17
		Residuals	180	354.2	2.0		

Table. A5
Two-way ANOVA results comparing 30-year simulations of IWUE and EWUE with different Models (Standalone, Tipping-Bucket, SWAP) and Soils (DL and SL). P-values lower than 0.01 represent statistically significance of 1 %

Variable	Site	Sources of variation	Degrees of Freedom	Sum of Squares	Mean of Squares	F-value	P-value
IWUE	Recife	Model	2	186,1	93,1	103,5	1,2E-30
		Soil	1	969,7	969,7	1078,2	6,4E-78
		Model:Soil	2	533,3	266,7	296,5	1,1E-57
		Residuals	180	161,9	0,9		
	Piracicaba	Model	2	8758,1	4379,1	4868,7	2,0E-157
		Soil	1	548,1	548,1	609,4	1,1E-59
		Model:Soil	2	150,5	75,2	83,6	2,1E-26
		Residuals	180	161,9	0,9		
	Jatai	Model	2	13007,3	6503,6	7230,8	1,2E-172
		Soil	1	456,5	456,5	507,6	2,9E-54
		Model:Soil	2	134,6	67,3	74,8	2,2E-24
		Residuals	180	161,9	0,9		
	Petrolina	Model	2	13,7	6,9	7,6	0001
		Soil	1	43,5	43,5	48,3	6,5E-11
		Model:Soil	2	26,8	13,4	14,9	1,0E-06
		Residuals	180	161,9	0,9		
EWUE	Recife	Model	2	3,8	1,9	2,1	0122
		Soil	1	0,8	0,8	0,9	0339
		Model:Soil	2	0,7	0,4	0,4	0670
		Residuals	180	161,9	0,9		
	Piracicaba	Model	2	2,0	1,0	1,1	0338
		Soil	1	0,5	0,5	0,5	0473
		Model:Soil	2	0,9	0,5	0,5	0597
		Residuals	180	161,9	0,9		
	Jatai	Model	2	1,7	0,8	0,9	0401
		Soil	1	1,3	1,3	1,4	0232
		Model:Soil	2	0,4	0,2	0,2	0795
		Residuals	180	161,9	0,9		
	Petrolina	Model	2	14,9	7,4	8,3	3,7E-04
		Soil	1	0,8	0,8	0,9	0339
		Model:Soil	2	1,7	0,8	0,9	0401
		Residuals	180	161,9	0,9		

Table. A6
Statistical indices of performance for stalk dry mass (SDM) and leaf area index (LAI) for each modelling approach (SWAP, Tipping-Bucket, Standalone), using different soil data types (SL and DL). RRMSE is the relative root mean square error (%), EF is the model efficiency (dimensionless), R2 is the determination index of precision (dimensionless), d is the accuracy index of Wilmot (dimensionless), and RMSE is the root mean squared error (Mg hm^{-2} for SDM and unitless for LAI)

Model	SoilType	Variable	RRMSE	EF	R2	d	RMSE
Standalone	DL	SDM	0,31	0,81	0,94	0,96	6,00
		LAI	0,45	-0,51	0,20	0,67	1,16
	SL	SDM	0,26	0,86	0,94	0,97	5,09
		LAI	0,45	-0,52	0,21	0,67	1,16
TipBucket	DL	SDM	0,16	0,95	0,98	0,99	3,21
		LAI	0,29	0,36	0,48	0,83	0,75
	SL	SDM	0,15	0,95	0,98	0,99	2,95
		LAI	0,30	0,29	0,44	0,81	0,79

(*continued on next page*)

Table. A6 (*continued*)

Model	SoilType	Variable	RRMSE	EF	R2	d	RMSE
SWAP	DL	SDM	0,33	0,79	0,91	0,95	6,33
		LAI	0,25	0,51	0,55	0,86	0,66
	SL	SDM	0,24	0,88	0,90	0,97	4,69
		LAI	0,26	0,49	0,57	0,83	0,67

References

ANA, 2019. Levantamento da Cana-de-Açúcar Irrigada e Fertirrigada no Brasil. Ag. ência Nac. De. Águas e Saneam. BáSci.

Antle, J.M., Basso, B., Conant, R.T., Godfray, H.C.J., Jones, J.W., Herrero, M., Wheeler, T.R., 2017. Towards a new generation of agricultural system data, models and knowledge products: Design and improvement. Agric. Syst. 155, 255–268.

Bartholomeus, R.P., Witte, J.-P.M., van Bodegom, P.M., van Dam, J.C., Aerts, R., 2008. Critical soil conditions for oxygen stress to plant roots: Substituting the Feddes-function by a process-based model. J. Hydrol. 360 (1), 147–165.

Best, M., Pryor, M., Clark, D., Rooney, G., Essery, R., Menard, C.B., Harding, R., 2011. The Joint UK Land Environment Simulator (JULES), model description – Part 1: Energy and water fluxes. Geosci. Model Dev. 4, 677–699.

Christina, M., Jones, M.R., Versini, A., Mézino, M., Le Mezo, L., Auzoux, S., Gérardeaux, E., 2021. Impact of climate variability and extreme rainfall events on sugarcane yield gap in a tropical Island. Field Crops Res. 274, 108326 https://doi.org/10.1016/j.fcr.2021.108326.

Christofoletti, C.A., Escher, J.P., Correia, J.E., Marinho, J.F.U., Fontanetti, C.S., 2013b. Sugarcane vinasse: environmental implications of its use. Waste Manag. 33 (12), 2752–2761.

Christofoletti, C.A., Escher, J.P., Correia, J.E., Marinho, J.F.U., Fontanetti, C.S., 2013a. Sugarcane vinasse: environmental implications of its use. Waste Manag. 33 (12), 2752–2761.

Dias, H.B., Inman-Bamber, G., 2020. Sugarcane: Contribution of Process-Based Models for Understanding and Mitigating Impacts of Climate Variability and Change on Production. In: Ahmed, M. (Ed.), Systems Modeling. Springer Singapore, Singapore, pp. 217–260.

Dias, H.B., Sentelhas, P.C., 2018b. Dimensioning the Impact of Irrigation on Sugarcane Yield in Brazil. Sugar Tech. https://doi.org/10.1007/s12355-018-0619-x.

Dias, H.B., Sentelhas, P.C., 2018a. Sugarcane yield gap analysis in Brazil - A multi-model approach for determining magnitudes and causes. Sci. Total Environ. 637-638, 1127–1136.

Dias, H.B., Inman-Bamber, G., Bermejo, R., Sentelhas, P.C., Christodoulou, D., 2019. New APSIM-Sugar features and parameters required to account for high sugarcane yields in tropical environments. Field Crops Res. 235, 38–53.

Enders, A., Vianna, M., Gaiser, T., Krauss, G., Webber, H., Srivastava, A.K., Ewert, F., 2023. SIMPLACE - A versatile modelling and simulation framework for sustainable crops and agroecosystems. silico Plants, diad006.

FAO, 2022. Food and Agriculture Organization Corporate Statistical Database. FAOSTAT.

Feddes, R.A., Kowalik, P.J., Zaradny, H., 1978. Simulation of Field Water Use and Crop Yield, 188. Wiley.

George, T.S., Bulgarelli, D., Carminati, A., Chen, Y., Jones, D., Kuzyakov, Y., Roose, T., 2024. Bottom-up perspective – The role of roots and rhizosphere in climate change adaptation and mitigation in agroecosystems. Plant Soil. https://doi.org/10.1007/s11104-024-06626-6.

Goldemberg, J., Mello, F.F.C., Cerri, C.E.P., Davies, C.A., Cerri, C.C., 2014. Meeting the global demand for biofuels in 2021 through sustainable land use change policy. Energy Policy 69, 14–18.

Gonçalves, I.Z., Vianna, M.S., Nassif, D.S.P., Carvalho, K., Marin, F.R., 2023. Effects of Residue from Harvested Green Cane on Evapotranspiration, Growth, and Development of Irrigated Sugarcane in Southern Brazil. Sugar Tech. 25 (6), 1445–1455.

Hamilton, S.H., Pollino, C.A., Stratford, D.S., Fu, B., Jakeman, A.J., 2022. Fit-for-purpose environmental modeling: Targeting the intersection of usability, reliability and feasibility. Environ. Model. Softw. 148, 105278.

Heinen, M., Mulder, M., van Dam, J., Bartholomeus, R., van Lier, Q. de J., de Wit, J., Broeke, M.H.-T., 2024. SWAP 50 year: Advances in modelling soil-water-atmosphere-plant interactions. Copernic. Meet.

Hengl, T., Mendes de Jesus, J., Heuvelink, G.B.M., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., Kempen, B., 2017. SoilGrids250m: Global gridded soil information based on machine learning. PloS One 12 (2), e0169748.

Hoffman, N., Singels, A., Patton, A., Ramburan, S., 2018. Predicting genotypic differences in irrigated sugarcane yield using the Canegro model and independent trait parameter estimates. Eur. J. Agron.: J. Eur. Soc. Agron. 96, 13–21.

Ines, A.V.M., Droogers, P., 2002. Inverse modelling in estimating soil hydraulic functions: a Genetic Algorithm approach. Hydrol. Earth Syst. Sci. 6 (1), 49–66.

van Ittersum, M.K., Leffelaar, P.A., van Keulen, H., Kropff, M.J., Bastiaans, L., Goudriaan, J., 2003. On approaches and applications of the Wageningen crop models. Eur. J. Agron.: J. Eur. Soc. Agron. 18 (3), 201–234.

Jarvis, N., Larsbo, M., Lewan, E., Garré, S., 2022. Improved descriptions of soil hydrology in crop models: The elephant in the room? Agric. Syst. 202, 103477.

Jones, J.W., Hoogenboom, G., Porter, C.H., Boote, K.J., Batchelor, W.D., Hunt, L.A., Ritchie, J.T., 2003. The DSSAT cropping system model. Eur. J. Agron.: J. Eur. Soc. Agron. 18 (3), 235–265.

Jones, J.W., Antle, J.M., Basso, B., Boote, K.J., Conant, R.T., Foster, I., Wheeler, T.R., 2017. Brief history of agricultural systems modeling. Agric. Syst. 155, 240–254.

Jones, M.R., Singels, A., 2018. Refining the Canegro model for improved simulation of climate change impacts on sugarcane. Eur. J. Agron.: J. Eur. Soc. Agron. 100, 76–86.

de Jong van Lier, Q., van Dam, J.C., Metselaar, K., de Jong, R., Duijnisveld, W.H.M., 2008. Macroscopic root water uptake distribution using a matric flux potential approach. Vadose zone J.: VZJ 7 (3), 1065–1078.

de Jong van Lier, Q., Wendroth, O., van Dam, J.C., 2015. Prediction of winter wheat yield with the SWAP model using pedotransfer functions: An evaluation of sensitivity, parameterization and prediction accuracy. Agric. Water Manag. 154, 29–42.

Kaelo, P., Ali, M.M., 2006. Some Variants of the Controlled Random Search Algorithm for Global Optimization. J. Optim. Theory Appl. 130 (2), 253–264.

Keating, B.A., 2020. Crop, soil and farm systems models – science, engineering or snake oil revisited. Agric. Syst. 184 (102903), 102903.

Keating, B.A., Robertson, M.J., Muchow, R.C., Huth, N.I., 1999. Modelling sugarcane production systems I. Development and performance of the sugarcane module. Field Crops Res. 61 (3), 253–271.

Keulen V.H. & Seligman N.G. (1988). Simulation of Water Use, Nitrogen Nutrition and Growth of a Spring Wheat Crop. 310 pages. Wageningen: Pudoc. 1987. Price Df1 100.00 (hard covers). ISBN 90 220 0905 X. The Journal of agricultural science, 110 (2), 428–428.

Kroes J.G., Van Dam J.C. & Bartholomeus R.P. (2017). SWAP version 4.

Laclau, P.B., Laclau, J.P., 2009. Growth of the whole root system for a plant crop of sugarcane under rainfed and irrigated environments in Brazil. Field Crops Res.

Lawrence, D.M., Fisher, R.A., Koven, C.D., Oleson, K.W., Swenson, S.C., Bonan, G., Zeng, X., 2019. The Community Land Model Version 5: Description of New Features, Benchmarking, and Impact of Forcing Uncertainty. J. Adv. Model. Earth Syst. 11 (12), 4245–4287.

Marin, F., Jones, J.W., Boote, K.J., 2017. A Stochastic Method for Crop Models: Including Uncertainty in a Sugarcane Model. Agron. J. 109, 483–495.

Marin, F.R., Jones, J.W., 2014. Process-based simple model for simulating sugarcane growth and production. Sci. Agric. 71 (1), 1–16.

Marin, F.R., Thorburn, P.J., Nassif, D.S.P., Costa, L.G., 2015. Sugarcane model intercomparison: Structural differences and uncertainties under current and potential future climates. Environ. Model. Softw. 72, 372–386.

Marin, F.R., Inman-Bamber, G., Silva, T.G.F., Vianna, M.S., Nassif, D.S.P., Carvalho, K.S., 2020. Sugarcane evapotranspiration and irrigation requirements in tropical climates. Theor. Appl. Climatol. 140 (3-4), 1349–1357.

Martre, P., Wallach, D., Asseng, S., Ewert, F., Jones, J.W., Rötter, R.P., Wolf, J., 2015. Multimodel ensembles of wheat growth: many models are better than one. Glob. Change Biol. 21 (2), 911–925.

de Melo, M.L.A., de Jong van Lier, Q., 2021. Revisiting the Feddes reduction function for modeling root water uptake and crop transpiration. J. Hydrol. 603, 126952.

Nassif, D.S.P., da Costa, L.G., dos Santos Vianna, M., dos Santos Carvalho, K., Marin, F.R., 2019. The role of decoupling factor on sugarcane crop water use under tropical conditions. Exp. Agric. 1–11.

Nóia Júnior, R. de S., Asseng, S., García-Vila, M., Liu, K., Stocca, V., dos Santos Vianna, M., Harrison, M.T., 2023. A call to action for global research on the implications of waterlogging for wheat growth and yield. Agric. Water Manag. 284, 108334.

Olivier, F.C., Singels, A., 2015. Increasing water use efficiency of irrigated sugarcane production in South Africa through better agronomic practices. Field Crops Res. 176, 87–98.

Osborne, T., Gornall, J.L., Hooker, J., Williams, K., Wiltshire, A., Betts, R.A., Wheeler, T., 2015. JULES-crop: a parametrisation of crops in the Joint UK Land Environment Simulator. Geosci. Model Dev. 8 (4), 1139–1155.

Passioura, J.B., 1996. Simulation models: Science, snake oil, education, or engineering? Agron. J. 88 (5), 690–694.

Pereira, R.A., de, A., Vianna, M., dos, S., Nassif, D.S.P., Carvalho, K., dos, S., Marin, F.R., 2021. Global sensitivity and uncertainty analysis of a sugarcane model considering the trash blanket effect. Eur. J. Agron.: J. Eur. Soc. Agron. 130, 126371.

Perez, P.J., Castellvi, F., Ibañez, M., Rosell, J.I., 1999. Assessment of reliability of Bowen ratio method for partitioning fluxes. Agric. For. Meteorol. 97 (3), 141–150.

Pinto, V.M., van Dam, J.C., de Jong van Lier, Q., Reichardt, K., 2019. Intercropping Simulation Using the SWAP Model: Development of a 2×1D Algorithm. Collect. FAO: Agric. 9 (6), 126.

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

Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P., Winter, J.M., 2013. The Agricultural Model Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies. Agric. For. Meteorol. 170, 166–182.

Singels, A., 2013. Crop Models. Sugarcane: Physiology, Biochemistry, and Functional Biology. John Wiley & Sons Ltd, Chichester, UK, pp. 541–577.

Singels, A., van den Berg, M., Smit, M.A., Jones, M.R., van Antwerpen, R., 2010. Modelling water uptake, growth and sucrose accumulation of sugarcane subjected to water stress. Field Crops Res. 117 (1), 59–69.

Singels, A., Paraskevopoulos, A.L., Mashabela, M.L., 2019. Farm level decision support for sugarcane irrigation management during drought. Agric. Water Manag. 222, 274–285.

Stenzel, F., Greve, P., Lucht, W., Tramberend, S., Wada, Y., Gerten, D., 2021. Irrigation of biomass plantations may globally increase water stress more than climate change. Nat. Commun. 12 (1), 1512.

Suguitani C. (2006). Entendendo o crescimento e produção da cana de açúcar: avaliação do modelo Mosicas. Universidade de São Paulo.

Tao, F., Rötter, R.P., Palosuo, T., Gregorio Hernández Díaz-Ambrona, C., Mínguez, M.I., Semenov, M.A., Schulman, A.H., 2018. Contribution of crop model structure, parameters and climate projections to uncertainty in climate change impact assessments. Glob. Change Biol. 24 (3), 1291–1307.

Teh, C., 2006. Introduction to Mathematical Modeling of Crop Growth. Universal Publishers, p. 280.

Vianna, M., dos, S., Sentelhas, P.C., 2016. Performance of DSSAT CSM-CANEGRO Under Operational Conditions and its Use in Determining the "Saving Irrigation" Impact on Sugarcane Crop. Sugar Tech. 18 (1), 75–86.

Vianna, M., dos, S., Nassif, D.S.P., dos Santos Carvalho, K., Marin, F.R., 2020. Modelling the trash blanket effect on sugarcane growth and water use. Comput. Electron. Agric. 172, 105361.

Vianna, M.S., Williams, K.W., Littleton, E.W., Cabral, O., Cerri, C.E.P., De Jong van Lier, Q., Galdos, M.V., 2022. Improving the representation of sugarcane crop in the Joint UK Land Environment Simulator (JULES) model for climate impact assessment. Glob. Change Biol. Bioenergy 14 (10), 1097–1116.

Wallach, D., Martre, P., Liu, B., Asseng, S., Ewert, F., Thorburn, P.J., Zhang, Z., 2018. Multimodel ensembles improve predictions of crop–environment–management interactions. Glob. Change Biol. 24 (11), 5072–5083.

Wallach, D., Palosuo, T., Thorburn, P., Hochman, Z., Gourdain, E., Andrianasolo, F., Seidel, S.J., 2021. The chaos in calibrating crop models: Lessons learned from a multi-model calibration exercise. Environ. Model. Softw. 145, 105206.

Wallach D., Buis S., Seserman D.-M., Palosuo T., Thorburn P., Mielenz H., … Seidel S.J. (2023). A calibration protocol for soil-crop models aimed at reducing prediction error and inter-model variability. bioRxiv, 2023.10.26.564162.

de Wit, A., Boogaard, H., Fumagalli, D., Janssen, S., Knapen, R., van Kraalingen, D., van Diepen, K., 2019. 25 years of the WOFOST cropping systems model. Agric. Syst. 168, 154–167.

de Wit C.T. (1958). Transpiration and crop yields.

Yin, X., Struik, P.C., Goudriaan, J., 2021. On the needs for combining physiological principles and mathematics to improve crop models. Field Crops Res. 271, 108254.

Ypma J., Borchers H.W. & Eddelbuettel D. (2022). nloptr: R Interface to NLopt (v2.0.3). R package version.