

Sensitivity analysis using Morris: Just screening or an effective ranking method?

Livia Paleari^a, Ermes Movedi^a, Michele Zoli^b, Andrea Burato^b, Irene Cecconi^b, Jabir Errahouly^b, Eleonora Pecollo^b, Carla Sorvillo^b, Roberto Confalonieri^{a,*}

^a University of Milan, ESP, Cassandra lab, via Celoria 2, 20133 Milano, Italy

^b University of Milan, Cropping Systems MS course, via Celoria 2, 20133 Milano, Italy

ARTICLE INFO

Keywords:

Crop model

CropSyst

E-FAST

Fourier amplitude sensitivity test

STICS

WOFOST

ABSTRACT

Sensitivity analysis (SA) is a fundamental practice for analyzing model behavior under different conditions of application. A number of SA techniques were proposed, ranging from simple screening methods to computationally expensive variance-based ones. In this study, we compared the Morris and E-FAST methods by applying them to three widely used generic crop models largely differing for complexity and for the approaches used to formalize knowledge on crop physiology, i.e., STICS, CropSyst and WOFOST. SA experiments were carried out at sub-model level on rice crops grown under different environmental conditions. Results highlighted the lack of linearity between the total-order sensitivity estimates provided by E-FAST and Morris, although the concordance (TDCC) between the parameter rankings obtained with the two methods was always significant at the 0.05 level for parameters involved with crop growth and for those involved with phenological development for STICS, whereas it was significant at the 0.10 level for the phenology parameters of CropSyst and WOFOST. Given Morris required less than 3% of the model executions needed by E-FAST, our results allow considering Morris as a suitable alternative to more demanding SA methods when ranking parameters or discriminating between influential and non-influential model factors are the SA goals, especially in computationally expensive SA studies.

1. Introduction

Sensitivity analysis (SA) – i.e., the investigation of how changes in input factors influence numerical model outputs – is increasingly used in different research areas (Saltelli, 2002; Pianosi et al., 2015) for a variety of purposes, including the support to model analysis (Tarantola and Saltelli, 2003), reduction (Degenring et al., 2004; Hsieh et al., 2018) and calibration (Makowski et al., 2006; Mathieu et al., 2018), the analysis of system drivers (Pastres et al., 1999; Confalonieri, 2012), and the support to decision making (Anderson et al., 2014).

There is a number of methods available to run SA (e.g., Borgonovo and Plischke, 2016; Pianosi et al., 2016) and their performance were compared for a variety of modeling domains (e.g., Hamby, 1995; Confalonieri et al., 2010; Nguyen and Reiter, 2015) in light of different criteria like, e.g., their computational efficiency and the capability of managing interactions among factors as well as model non-linearities. In other cases, the evaluation of approaches for SA was carried out with the

aim of identifying the most appropriate SA method for a given type of model (Ligmann-Zielinska et al., 2020). Moreover, some studies analysed the robustness of some of these methods with respect to changes in the values of their own parameters (e.g., Confalonieri et al., 2010; Herman et al., 2013; Vanrolleghem et al., 2015).

Different classifications were proposed for SA methods, each deriving from the classification criteria that were adopted. In case, e.g., the criterion is parsimony (number of model executions needed to derive the sensitivity metrics), SA methods can be ranked from screening methods (highest parsimony) to variance-based ones (lowest parsimony). Regardless of the type of model, Morris (Morris, 1991) and E-FAST (Saltelli et al., 1999) play a key role among the most popular SA methods. They have been widely used alone (e.g., Campolongo et al., 2007; Xu and Gertner, 2011) or in an integrated fashion (e.g., Martre et al., 2015), with Morris used to screen for non-influential parameters in case of computationally-expensive simulation experiments before the application of the more demanding E-FAST method on the remaining

* Corresponding author.

E-mail address: roberto.confalonieri@unimi.it (R. Confalonieri).

<https://doi.org/10.1016/j.ecolmodel.2021.109648>

Received 25 January 2021; Received in revised form 12 May 2021; Accepted 26 May 2021

Available online 9 June 2021

0304-3800/© 2021 Elsevier B.V. All rights reserved.

ones. Although they are classified in two distinct categories of SA methods – i.e., screening in the case of Morris and variance-based for E-FAST – and the theoretical principles behind their algorithms are drastically different, their popularity is likely due to the same peculiarity: the efficiency in sampling the hyperspace of parameters within their category of SA techniques. Indeed, Morris is considered as one of the most effective screening method, whereas E-FAST is less computationally-expensive than other methods based on the decomposition of output variance, like Sobol' (1993) (Confalonieri et al., 2010).

However, in spite of the continuous increase in the power of computers, of the already mentioned possibility to combine screening and variance-based methods and of a certain tendency to instability of Morris in case of strong model non-linearities (Confalonieri et al., 2010; Hsieh et al., 2018), the extreme efficiency of Morris makes this method somehow seductive. This is particularly true when SA techniques are applied to spatially-distributed models, for which variance-based methods sometimes require a prohibitive number of model executions (Herman et al., 2013). Moreover, although variance-based methods are considered as benchmark approaches for quantifying total-order effects (e.g., Saltelli and Sobol', 1995), SA is often carried out *just* to discriminate between influential and non-influential parameters for practical purposes, like, e.g., in cases when target parameters for calibration need to be identified (Mathieu et al., 2018). For this purpose, ranking parameters according to their relative importance is what is needed, and modelers in this case aim at reaching a trade-off between accuracy and computational cost of SA (Herman et al., 2013).

Nevertheless, findings of studies where Morris was evaluated against variance-based methods – also for ranking purposes – are not unambiguous, with good Morris performances highlighted in some of these studies (e.g., Herman et al., 2013) and poor reliability of the method discussed in others (e.g., Vanrolleghem et al., 2015). Despite the reasons for this are not clear, some authors suggested that the differences in Morris performance could deal with intrinsic factors of models from different disciplines or with the fact that SA is not always properly performed (Saltelli et al., 2019). In some cases, when poor performances were achieved, they were explained with a tendency of Morris to be prone to Type II error (e.g., Campolongo et al., 2007) – i.e., to underestimate the relevance of input factors with a large effect on model outputs – especially in case of strong model non-linearities (Confalonieri et al., 2010; Hsieh et al., 2018). Other authors underlined that disagreement between Morris and variance-based results could arise when non-uniform distributions of input factors are used (e.g., Feng et al., 2018).

This study aimed at comparing Morris and E-FAST methods using three generic crop models that greatly differ for complexity and for the approaches used to formalize knowledge on crop growth and development by running SA experiments at sub-model level. Crop models can be considered as particularly suitable for evaluating and comparing SA methods, given they are strongly non-linear and they often present correlations among parameters. Moreover, SA should be considered as mandatory before using crop models, since their calibration is often performed under conditions characterized by unfavorable number of parameters to number of observations ratio. Therefore, running SA to identify target parameters for calibration for the specific conditions of application (SA is *situational*; Stearns, 1992) reduces the risk of losing robustness because of equifinality and overfitting (Whittaker et al., 2010; Her and Chaubey, 2015). Simulations were carried out for rice while exploring different environmental conditions, by paying particular attention to parameter distributions and related uncertainty, this being something whose impact on SA results, if quantified (Paleari and Confalonieri, 2016), can be considerable, especially when results from Morris and variance-based methods are compared using non-uniform distributions for model parameters (Feng et al., 2018).

2. Materials and methods

2.1. Sensitivity analysis methods

2.1.1. The Morris method

Assuming $X = (x_1, \dots, x_N)$ as the vector of model parameters, each x_i is forced to assume values in the set $\left\{0, \frac{1}{p-1}, \frac{2}{p-1}, \frac{3}{p-1}, \dots, \frac{p-2}{p-1}, 1\right\}$, with p being the number of levels (i.e., values over which parameters can be sampled) in the N -dimensional, p -level unit hypercube parameter space Ω . After defining $\Delta = \frac{1}{(p-1)}$, an elementary effect R_i for the i th parameter is calculated as:

$$R_i(X, \Delta) = \frac{y(x_1, \dots, x_{i-1}, x_i + \Delta, x_{i+1}, \dots, x_N) - y(X)}{\Delta} \quad (1)$$

where $y(X)$ is the model output. The Ω hyperspace is sampled by selecting r trajectories each composed by $N + 1$ points, with Δ being applied at each point to a single parameter. Compared to traditional one-factor-at-a-time SA design, model evaluations are carried out while varying single parameters along trajectories through the Ω hyperspace. The Morris sampling method prevents parameters from assuming any of the previous values when varied, and it allows limiting the number of model evaluations to $r(N + 1)$. After the sampling, parameters are scaled back from the unit hypercube to their original ranges according to the specified parameter distributions (based on quantiles). Mean (μ) and standard deviation (σ) are calculated for each distribution of R_i that represent the sensitivity measures for the parameter x_i , μ (strength) quantifying the overall influence of x_i on $y(X)$ (total-order effect) and σ (spread) identifying parameters with a non-linear effect on $y(X)$ or that are involved in interactions with other parameters (Saltelli et al., 2004). In this study, we used the version of the Morris method proposed by Campolongo et al. (2007), where μ is derived by averaging $|R_i|$ instead of R_i . Further details on the Morris method are available, e.g., in Morris (1991), Campolongo et al. (2007) and King and Perera (2013).

2.1.2. The e-fast method

Variance-based methods use variance ratios to quantify the importance of parameters. The total variance of model output $V(y)$ is decomposed into terms of increasing dimension (partial variances) that represent the contribution of single parameters (1st-order effects), and of pairs of parameters (2nd-order effects), triplets (3rd-order effects), etc. to the overall variability of y :

$$V(y) = \sum_{i=1}^N D_i + \sum_{i < j \leq N} D_{ij} + \dots + \sum_{i < j < k \leq N} D_{ijk} \quad (2)$$

where $D_i = V[E(y|x_i)]$ is the first order effect of x_i , $D_{ij} = V[E(y|x_i, x_j)] - (D_i + D_j)$ is the 2nd-order effect (interaction between x_i and x_j), and terms up to $D_{i...N}$ represent higher-order effects. The 1st-order effect sensitivity index for x_i is calculated as:

$$S_i = \frac{V[E(y|x_i)]}{V(y)} \quad (3)$$

Given the high computational cost required to estimate the sensitivity indices of higher-order interactions, Homma and Saltelli (1996) proposed the concept of total sensitivity index, which allows markedly reducing the computational effort. At the cost of a reasonable number of model runs, the total sensitivity index (St_i , Eq. (4)) provides a measure of the overall effect of a parameter considering the interactions with all the others.

$$St_i = S_i + \sum_{j \neq i} S_{ij} + \dots + S_{i...N} \quad (4)$$

Compared to other variance-based methods (e.g., Sobol', 1993), E-FAST explores the N -dimensional parameter hyperspace by using a

search curve defined by a set of parametric equations:

$$x_i(s) = 0.5 + \frac{1}{\pi}(\omega_i s + \varphi_i) \quad (5)$$

where ω_i ($i = 1, \dots, N$) is the angular frequency associated to the factor x_i , s is a scalar variable ranging from $-\infty$ to $+\infty$, and φ_i ($0 - 2\pi$) is a random factor (random phase-shift) indicating the starting point of the search curve. The sampling procedure driven by the transformation presented in Eq. (5) is repeated different times by selecting different values of φ_i , and final sensitivity metrics are obtain as the arithmetic means of the metrics obtained for the different φ_i . This transformation allows E-FAST exploring very efficiently the hyperspace, reducing the number of model executions compared to other variance-based methods and it is one of the reasons for its great popularity. Further details on the E-FAST method are available, e.g., in Saltelli et al. (1999) and Ryan et al. (2018).

2.2. The crop models

The three crop models used in this study – CropSyst (Stöckle et al., 1994), STICS (Brisson et al., 1998) and WOFOST (van Keulen and Wolf, 1986) – are generic (able to reproduce growth and development of different species with different sets of parameters) and widely used within the modellers' community. They all simulate phenological development based on thermal time accumulation, although they use different functions to reproduce crop response to mean daily temperature, with options to account for photoperiod sensitivity. The main differences among the three models are in the algorithms used to simulate crop growth, with CropSyst and STICS that are based on monolayer canopy representations (Beer's law analogies for intercepting radiation) and net photosynthesis, estimated using radiation use efficiency (RUE) in the case of STICS and both RUE and vapor pressure deficit (VPD)-corrected transpiration use efficiency (Tanner and Sinclair, 1983) for CropSyst. On the contrary, WOFOST uses a gross photosynthesis approach based on leaf light response curve, with instantaneous gross CO_2 assimilation estimated at three moments in the day and at three canopy depths. For this model, daily biomass accumulation is estimated after growth and maintenance respiration are considered. The main driver for estimating leaf area index (LAI) in STICS is crop development, whereas WOFOST derives the same variable based on a development-dependent specific leaf area (SLA) and the amount of assimilates daily partitioned to leaves. For CropSyst, LAI is estimated according to the early-stage SLA, the daily rate and state of aboveground biomass (AGB), and an empiric parameter named stem/leaf partition coefficient (SLP). Leaf senescence is estimated by the three models as a function of the thermal time accumulated by daily-emitted LAI units, with WOFOST considering also the effect of self-shading. Yield is derived based on the concept of harvest index for STICS and CropSyst, and with a dynamic partitioning of assimilates for WOFOST. Concerning WOFOST, we used the version proposed by Stella et al. (2014), based on the substitution of AFGEN tables (Arbitrary Function GENerator) with functions driven by few parameters. Further details on the models algorithms are available in the seminal literature and in the models' documentation, whereas their suitability for simulating rice growth and development has been demonstrated by Confalonieri and Bocchi (2005) (CropSyst), Stella et al. (2014) (WOFOST) and Li et al. (2015) (STICS), who evaluated the performances of the three models for rice under a variety of environmental conditions.

The three models and related documentations are available at:

- http://modeling.bsyse.wsu.edu/CS_Suite/cropsyst/index.html (CropSyst);
- <http://www.wur.nl/en/Research-Results/Research-Institutes/Environmental-Research/Facilities-Tools/Software-models-and-databases/WOFOST.htm> (WOFOST);
- https://www6.paca.inrae.fr/stics_eng (STICS).

2.3. Simulation experiment

The E-FAST and Morris methods were used to run SA experiments with the three crop models under temperate (northern Italy) and tropical (Philippines) conditions. Only parameters involved with crop growth and development under unlimited conditions for water and nitrogen were considered. The simulations carried out in Italy refer to experiments conducted in Vignate (45°29' N, 9°22' E, 121 m a.s.l.) and Opera (45°22' N, 9°12' E, 99 m a.s.l.) in 2002, where the tropical japonica-profile cultivars Thaibonnet and Sillaro were sown on 29 April. Crops were grown (continuous flooding) under different nitrogen treatments and harvested at physiological maturity on 29 (Vignate) and 20 (Opera) September. Simulations in the Philippines refer to experiments carried out in Los Baños (14°10' N, 121°15' E, 118 m a.s.l.) at the International Rice Research Institute (IRRI), with the Indica rice cultivar IR64 transplanted on 30 November 2011 (dry season) and 6 June 2012 (wet season). In this case, rice was grown under continuous flooding and different salinity levels (including a control treatment irrigated with fresh water) and harvested on 19 March 2012 (dry season) and 19 September 2012 (wet season). Further details on the experiments are provided by Confalonieri and Bocchi (2005) and Radanielson et al. (2018) for the experiments carried out in Italy and in the Philippines, respectively.

Given parameters involved with phenological development markedly affect also processes involved with crop growth, their sensitivity metrics (calculated on variables such as AGB or yield) can assume values much higher than those of parameters involved with other processes, like radiation interception, photosynthesis or leaf area expansion (Richter et al., 2010). This could decrease the capability of SA to discriminate among the relevance of parameters not involved with phenology. For this reason, although the variable analyzed was always AGB simulated at physiological maturity (synthetically representing all processes involved with crop growth and development), SA experiments were carried out separately for parameters involved with phenological development and growth (Confalonieri et al., 2010). This allowed increasing the discriminating power of SA while quantifying the impact of changes in parameter values on the variability of the model output. Parameter distributions and related sources of information are presented in Appendices A to C. For the parameters for which sources were not enough to check distribution, available values were averaged and distribution was assumed to be normal, with standard deviation equal to 5% of the mean (Richter et al., 2010).

According to Confalonieri et al. (2010), the Morris method was parameterized with six trajectories and six levels. For STICS, this corresponded to 78 and 156 model executions for the SA experiments carried out on parameters involved with phenology and growth, respectively. Corresponding values for CropSyst were 48 and 60, whereas they were 54 and 120 for WOFOST. Previous studies suggested to parameterize the E-FAST method to get a number of model executions of about 250 multiplied by the number of model parameters (Confalonieri et al., 2010). This led to 2988 and 6225 executions for phenology and growth in case of STICS, 1743 (phenology) and 2241 (growth) for CropSyst, and 1992 (phenology) and 4731 (growth) for WOFOST. These sample sizes are consistent with what suggested in the literature (e.g., Campolongo et al., 2007). However, their suitability was evaluated – for all the models and for sample environments and processes – using smaller (five trajectories for Morris and ~200 times the number of parameters for E-FAST) and larger (seven trajectories and ~300 times the parameters) sample sizes to verify the stability of SA results.

In order to compare the rankings obtained (i) using the Morris and E-FAST SA methods within each environment (concordance among SA methods) and (ii) using each of the two methods in different environments (concordance among environments), we used the top-down concordance coefficient (TDCC; 0 to 1, 1 indicating complete concordance; Iman and Conover, 1987). Compared to other concordance

metrics, TDCC emphasizes the agreement among high-ranked parameters and vice versa. The first step to calculate TDCC is to rank the model parameters according to their sensitivity metrics (rank 1 is assigned to the parameter that achieved the highest sensitivity metric, rank N to the parameter with the lowest sensitivity metric). This step was performed for each SA experiment j , i.e., for each combination environment \times SA method \times process (growth or phenological development). The second step was the calculation of the savage scores (Savage, 1956) $ss(S_{ij})$ from the $\{1 \dots i \dots N\}$ ranks $r(S_{ij})$ assigned to each parameter i :

$$ss(S_{ij}) = \sum_{i=r(S_{ij})}^N \frac{1}{i} \quad (6)$$

where S_{ij} is the sensitivity metric of the i th parameter (after parameters have been ranked according to their relevance) as achieved in the j th SA experiment.

The third step was the calculation of TDCC from the $ss(S_{ij})$ calculated for the different SA experiments (Eq. (7)):

$$TDCC = \frac{\sum_{i=1}^N \left[\sum_{j=1}^{n_{SA}} ss(S_{ij}) \right]^2 - n_{SA}^2 \cdot N}{n_{SA}^2 \cdot (N - \sum_{i=1}^N 1/i)} \quad (7)$$

where n_{SA} is the number of SA experiments for which rankings are compared (e.g., rankings obtained for parameters involved with phenological development in a certain environment using Morris and E-FAST SA methods).

Table 1

Results of sensitivity analysis on the STICS parameters performed with the E-FAST and Morris methods. Vi02 and Op02 refer to the rice crops sown in Vignate and Opera (Italy) in 2002; LB11 and LB12 to the rice crops sown in Los Baños in 2011 and 2012. Parameter acronyms are explained in [Appendix A](#).

Parameter	E-FAST St				Morris μ			
	LB11	LB12	Vi02	Op02	LB11	LB12	Vi02	Op02
Phenological development								
tdmin	0.0364	0.0379	0.0499	0.0483	1.310	1.203	1.983	1.950
tdmax	0.0058	0.0066	0.0081	0.0065	0.003	0.350	0.000	0.007
phosat	0.0060	0.0060	0.0087	0.0068	0.130	0.017	0.000	0.000
phobase	0.0062	0.0072	0.0063	0.0069	0.417	0.297	0.000	0.000
sensiphot	0.0094	0.0152	0.0117	0.0125	0.467	0.670	0.443	0.420
stpltger	0.0053	0.0055	0.0083	0.0069	0.000	0.000	0.003	0.207
tgmin	0.0056	0.0055	0.0114	0.0074	0.000	0.000	0.163	0.337
stlevamf	0.0298	0.0329	0.1501	0.0647	1.060	0.947	2.920	2.370
stamflax	0.0119	0.0145	0.01086	0.0111	0.067	0.077	0.077	0.053
stlevdrp	0.9463	0.9361	0.8129	0.9117	9.207	8.887	8.127	11.857
stdrpdcs	0.0053	0.0052	0.0048	0.0047	0.000	0.000	0.000	0.000
stdrpmat	0.0089	0.0089	0.0131	0.0108	0.683	0.580	0.993	0.947
Growth (light interception, photosynthesis, assimilate partitioning, leaf area, yield formation)								
bdens	0.0088	0.0104	0.0062	0.0064	0.103	0.107	0.157	0.150
tcxstop	0.0123	0.0176	0.0072	0.0071	0.000	0.000	0.000	0.000
tcmin	0.0296	0.0304	0.1481	0.1079	0.193	0.190	0.683	0.597
tcmax	0.0071	0.0095	0.0054	0.0059	0.000	0.000	0.000	0.000
dlaimax	0.0137	0.0137	0.0103	0.0094	0.267	0.270	0.360	0.410
pentlai	0.0107	0.0122	0.0079	0.0083	0.087	0.083	0.130	0.117
hautmax	0.0079	0.0126	0.0085	0.0092	0.000	0.003	0.067	0.000
extin	0.6991	0.7643	0.3523	0.4107	1.517	1.553	1.980	2.137
temin	0.0083	0.0106	0.0806	0.0475	0.000	0.000	0.877	0.663
teoptbis	0.0083	0.0118	0.0067	0.0077	0.007	0.020	0.000	0.000
teopt	0.0096	0.0125	0.2390	0.1618	0.000	0.000	1.100	0.873
temax	0.0064	0.0084	0.0062	0.0066	0.003	0.007	0.000	0.000
efcroijuv	0.0350	0.0382	0.0736	0.0610	0.340	0.333	0.700	0.643
efcroiveg	0.2082	0.1803	0.1574	0.1742	0.780	0.720	1.290	1.340
efcroirepro	0.0935	0.0739	0.0297	0.0598	0.633	0.530	0.710	0.970
slamin	0.0065	0.0083	0.0087	0.0076	0.000	0.000	0.000	0.000
slamax	0.0154	0.0174	0.0075	0.0090	0.150	0.153	0.113	0.183
tgefeuille	0.0098	0.0117	0.0066	0.0086	0.113	0.123	0.087	0.133
vitircarb	0.0060	0.0065	0.0058	0.0060	0.000	0.000	0.000	0.000
irmax	0.0077	0.0111	0.0070	0.0065	0.000	0.000	0.000	0.000
tmaxremp	0.0084	0.0103	0.0062	0.0060	0.000	0.000	0.000	0.000
tminremp	0.0069	0.0070	0.0053	0.0059	0.000	0.000	0.000	0.000
pgrain	0.0057	0.0064	0.0048	0.0052	0.000	0.000	0.000	0.000
adens	0.0131	0.0149	0.0131	0.0111	0.193	0.197	0.277	0.257
durvieF	0.0494	0.0494	0.0210	0.0630	0.437	0.213	0.230	0.543

The fourth (final) step was the evaluation of the significance of the concordance. Under the null hypothesis of no concordance among parameter rankings, the p -value for TDCC can be estimated using the statistics T , which approximates a χ^2 distribution with $N - 1$ degrees of freedom:

$$T = n_{SA} \cdot (N - 1) \cdot TDCC \quad (8)$$

with concordance among rankings considered non-statistically significant for p -values higher than 0.05.

3. Results

Table 1 shows the values of E-FAST St and Morris μ calculated for the STICS model. Both methods identified thermal time from emergence to beginning of grain filling (stlevdrp) as decidedly the most relevant among the parameters involved with phenological development (it explained on average 90% of the variability in model output according to E-FAST), followed by thermal time from emergence to end of juvenile phase (stlevamf) and minimum temperature below which development stops (tdmin). On average, the relevance of tdmin was larger for the two Italian environments (56% and 32% according to μ and St , respectively), likely because of the lower temperatures experienced by rice especially in the first part of the cycle.

Concerning the parameters involved with crop growth, both methods led canopy extinction coefficient (extin) and maximum RUE during vegetative phase (efcroiveg) to achieve the highest SA indices. For

efcroveg, higher values for the sensitivity metrics were obtained in the Italian environments, because changes in parameters involved with photosynthetic efficiency (especially after the close canopy stage) have a larger effect on productivity in temperate environments, where radiation may be suboptimal for rice. Compared to results of SA experiments performed on STICS parameters involved with phenology, both Morris and E-FAST provided metric values more homogeneous in case of parameters affecting crop growth, although a discontinuity was still observed between top-ranked parameter and the others (Fig. 1). Although with different ranks, indeed, the two methods identified as relevant also other parameters involved with photosynthetic efficiency (RUE in different phenological phases and parameters modulating the thermal limitation to photosynthesis).

As already observed for STICS, also the SA performed on CropSyst

(Table 2) indicated thermal time to complete the vegetative phase (GDDf) as the most relevant in affecting AGB at maturity among the parameters involved with phenological development. This was particularly evident for E-FAST, for which the parameter explained more than 90% of the total variance in the model output. Moreover, like for STICS, base temperature for thermal time accumulation (Tbasedev) resulted as one of the most important parameters, being ranked 3rd and 2nd according to Morris and E-FAST, respectively. On the contrary, Morris identified one of the parameters involved with photoperiod as highly impacting on AGB at maturity in the tropical environments, whereas photoperiod parameters were never high-ranked for STICS.

Both SA methods indicated canopy extinction coefficient (k) as the parameter with the largest impact on AGB accumulation also for CropSyst, followed by optimum temperature for biomass accumulation

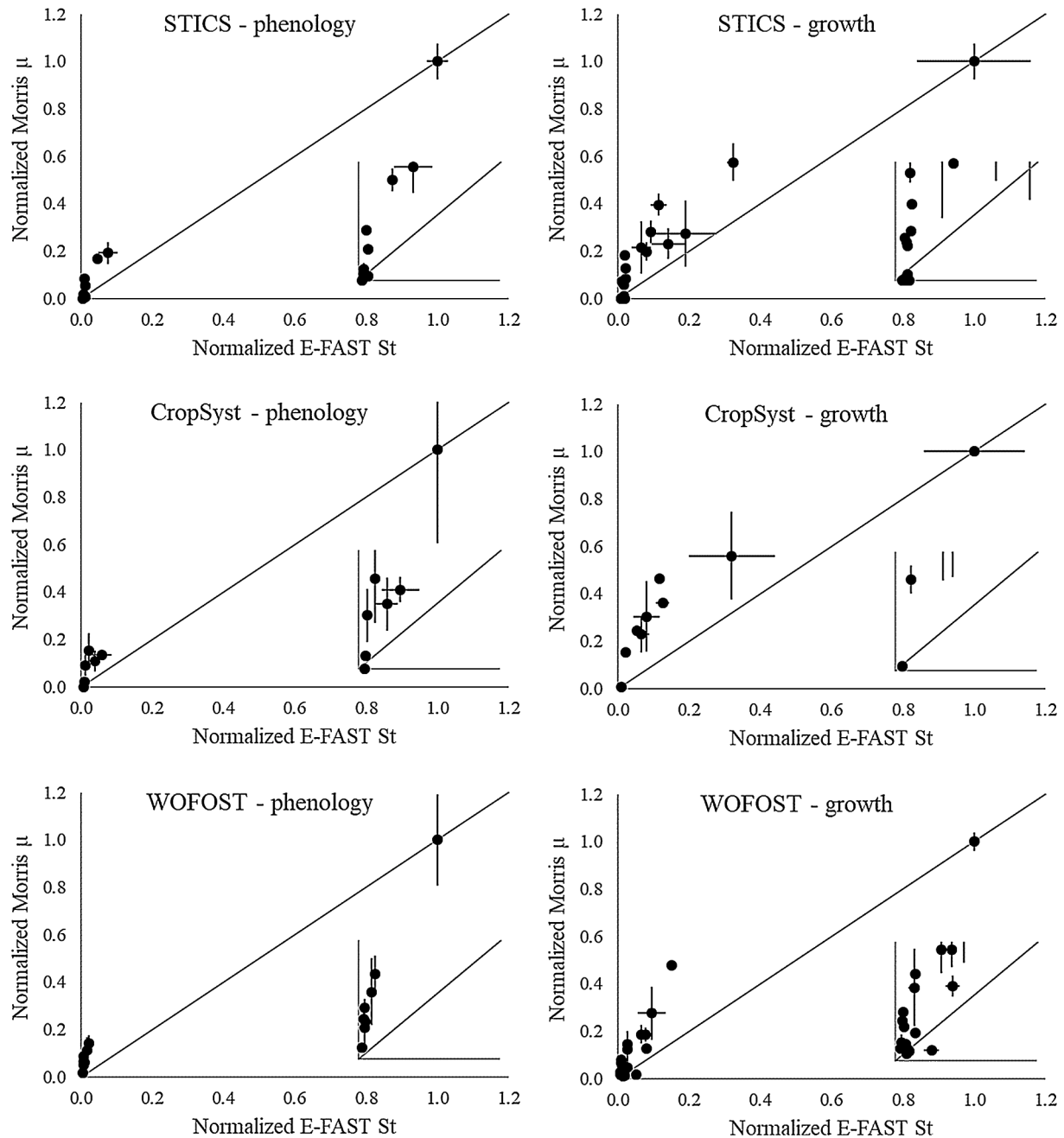


Fig. 1. Comparison between E-FAST \bar{S}_t and Morris μ . Data shown in Table 1 are here linearly normalized to 1 to increase readability. Bars (one standard deviation) represent the variability in the total order metrics obtained for simulations carried out in different environments. The small charts zoom in the range 0 to 0.2 to increase readability for normalized metrics close to zero. 1:1 lines represent perfect agreement between normalized \bar{S}_t and $\bar{\mu}$.

Table 2

Results of sensitivity analysis on the CropSyst parameters performed with the E-FAST and Morris methods. Vi02 and Op02 refer to the rice crops sown in Vignate and Opera (Italy) in 2002; LB11 and LB12 to the rice crops sown in Los Baños in 2011 and 2012. Parameter acronyms are explained in [Appendix B](#).

Parameter	E-FAST σ_t				Morris μ			
	LB11	LB12	Vi02	Op02	LB11	LB12	Vi02	Op02
Phenological development								
GDDmat	0.0136	0.0143	0.0850	0.040201	1040.000	1110.000	0.000	212.809
GDDfl	0.9321	0.9390	0.9625	0.9690	10,300.000	9790.000	377.381	1260.000
GDDem	0.0050	0.0048	0.0216	0.0088	129.816	91.658	101.278	128.310
Tbasedev	0.0166	0.0168	0.1371	0.0560	949.934	921.085	373.673	651.830
Toptdev	0.0048	0.0046	0.0160	0.0079	0.000	0.000	0.000	0.000
Phototh	0.0358	0.0286	0.0170	0.0079	1660.000	1620.000	4.046	15.729
Photosens	0.0139	0.0114	0.0169	0.0076	1010.000	930.836	2.423	9.819
Growth (light interception, photosynthesis, assimilate partitioning, leaf area, yield formation)								
SLA	0.0400	0.0400	0.0217	0.0233	505.945	525.050	532.549	531.1053
SLP	0.0115	0.0134	0.0122	0.0148	204.938	253.195	396.981	452.121
RUE	0.0127	0.0137	0.0681	0.0612	149.604	172.571	852.537	781.311
k	0.7494	0.7309	0.3975	0.4319	2050.000	2150.000	2190.000	2160.000
Tmin	0.0048	0.0047	0.0925	0.0839	24.085	20.823	1320.000	1240.000
LeafDuration	0.0052	0.0054	0.0060	0.0066	8.463	12.354	6.825	32.611
Topt	0.0407	0.0508	0.3390	0.3099	346.949	468.286	2060.000	1910.000
TUE	0.0867	0.0976	0.0431	0.0658	723.189	807.101	708.714	859.227
LAlini	0.0867	0.0839	0.0513	0.0510	924.677	955.405	1060.000	1030.000

(Topt), with the latter achieving in the temperate environments larger values for μ compared to what estimated in tropical ones, and vice versa for σ_t . This was one of the largest differences between the two SA methods in terms of values assumed by the SA metrics for high-ranked parameters in the different environments. Parameters involved with leaf area expansion were low-ranked, regardless of the SA method and environment, and RUE resulted decidedly less importance compared to what observed for STICS, given biomass accumulation in CropSyst is driven by RUE only for very low VPD values, otherwise net photosynthesis is driven by the water potentially transpired.

Coherently with the results achieved for the other two models, thermal time during the vegetative phase (TSUM1) resulted largely the most relevant parameter also for WOFOST regardless of the SA method

and the environment for which simulations were run ([Table 3](#)). Base temperature for development (Tbdev) was ranked 2nd according to both σ_t and μ , in line with the importance achieved by the corresponding parameters for STICS and CropSyst. Concerning this parameter, only the E-FAST method assigned higher sensitivity in the temperate environments, whereas the opposite was obtained using Morris. This is in line with the results of the SA experiments performed on CropSyst, whereas STICS was the only model for which both SA metrics were coherent in identifying the parameter as more important in case of lower thermal regimes. Although with values of the sensitivity metrics decidedly lower than those estimated for TSUM1, the 3rd-ranked parameter according to both methods was PHinib, involved with photoperiod response.

Contrarily to what obtained for STICS and CropSyst, the parameters

Table 3

Results of sensitivity analysis on the WOFOST parameters performed with the E-FAST and Morris methods. Vi02 and Op02 refer to the rice crops sown in Vignate and Opera (Italy) in 2002; LB11 and LB12 to the rice crops sown in Los Baños in 2011 and 2012. Parameter acronyms are explained in [Appendix C](#).

Parameter	E-FAST σ_t		Vi02	Op02	Morris μ		Vi02	Op02
	LB11	LB12			LB11	LB12		
Phenological development								
Tbdev	0.0149	0.0137	0.0282	0.0347	2710.000	2100.000	728.795	3340.000
Toptdev	0.0049	0.0062	0.0141	0.0130	662.486	830.998	959.006	1450.000
Tmaxdev	0.0049	0.0046	0.0088	0.0101	562.325	186.533	2100.000	1270.000
TSUM2	0.0087	0.0081	0.0096	0.0099	1770.000	1380.000	601.347	1590.000
TSUM1	0.9738	0.9806	0.9806	0.9753	21,600.000	20,600.000	7000.000	13,000.000
TSUMEM	0.0044	0.0045	0.0065	0.0056	38.981	243.222	544.052	358.753
PHinib	0.0335	0.0264	0.0066	0.0055	4030.000	2890.000	41.562	31.998
PHins	0.0120	0.0110	0.0057	0.0071	1990.000	1140.000	68.060	17.938
Growth (light interception, photosynthesis, assimilate partitioning, leaf area, yield formation)								
AMAX	0.0591	0.0426	0.0503	0.0622	530.969	393.760	1050.000	1130.000
Tb	0.0122	0.0119	0.0429	0.0203	13.598	7.885	1320.000	709.766
Topt	0.0175	0.0142	0.1581	0.1253	291.778	170.314	2200.000	1950.000
Tmax	0.0068	0.0067	0.0075	0.0075	15.832	9.153	264.367	212.387
RMR	0.0101	0.0104	0.0154	0.0156	32.956	36.935	56.584	57.305
RML	0.0191	0.0209	0.0122	0.0130	44.082	47.619	82.118	85.207
RMS	0.0278	0.0298	0.0173	0.0185	140.101	131.383	250.794	254.001
RMO	0.0125	0.0122	0.0146	0.0147	55.996	52.801	115.603	138.710
CVR	0.0264	0.0227	0.0558	0.0623	67.364	59.064	75.373	77.662
CVL	0.0106	0.0099	0.0099	0.0104	221.568	191.180	256.124	275.142
CVS	0.0523	0.0489	0.0851	0.0774	400.894	352.969	662.403	678.709
CVO	0.0664	0.0669	0.0520	0.0727	562.356	537.444	904.300	1100.000
SLAini	0.0124	0.0123	0.0115	0.0128	74.725	64.674	151.640	160.686
SLAtill	0.0107	0.0100	0.0086	0.0095	361.043	275.165	337.331	378.774
RipLO	0.0078	0.0062	0.0096	0.0090	105.733	40.402	504.609	451.557
LAIEM	0.8430	0.8242	0.7796	0.8110	3820.000	3850.000	4460.000	4510.000
SPAN	0.0059	0.0060	0.0062	0.0073	56.508	95.313	49.588	122.642
k	0.0088	0.0067	0.0295	0.0481	225.606	130.119	898.804	1180.000
RGRlAI	0.1314	0.1479	0.1063	0.1103	1770.000	1930.000	2110.000	2140.000

with the highest relevance according to both SA methods were involved with the simulation of leaf area expansion before the close canopy stage (LAIEM and RGRLAI) and not with photosynthetic efficiency (e.g., AMAX was ranked only fourth and sixth by Morris and E-FAST, respectively). However, the 3rd-ranked parameter was involved with the temperature response functions limiting CO₂ assimilation, with Morris results underlying the higher importance of thermal limitation in temperate environments already observed when the same SA method was applied to CropSyst. Another marked difference between the SA results obtained for WOFOST compared to what achieved for the other two models was related with extinction coefficient, which was ranked 1st in the two models based on net photosynthesis whereas it was only

6th (Morris) and 9th (E-FAST) in WOFOST.

Although both Morris μ and E-FAST St represent the overall influence of a parameter on the model output (total-order effect), they are not directly comparable, given they share neither the same conceptual meaning nor the same unit. For this reason, to investigate their relationships, we applied a linear normalization to the means and the standard deviations of the values of the metrics that were estimated, for each model and parameter, for the different environments. This allowed re-scaling all the mean values of μ ($\bar{\mu}$) and St (\bar{St}) in the range 0 to 1 and to get coherent information on their variability across environments (Fig. 1). Regardless of the model and of the process considered (phenological development or growth), it is possible to notice the

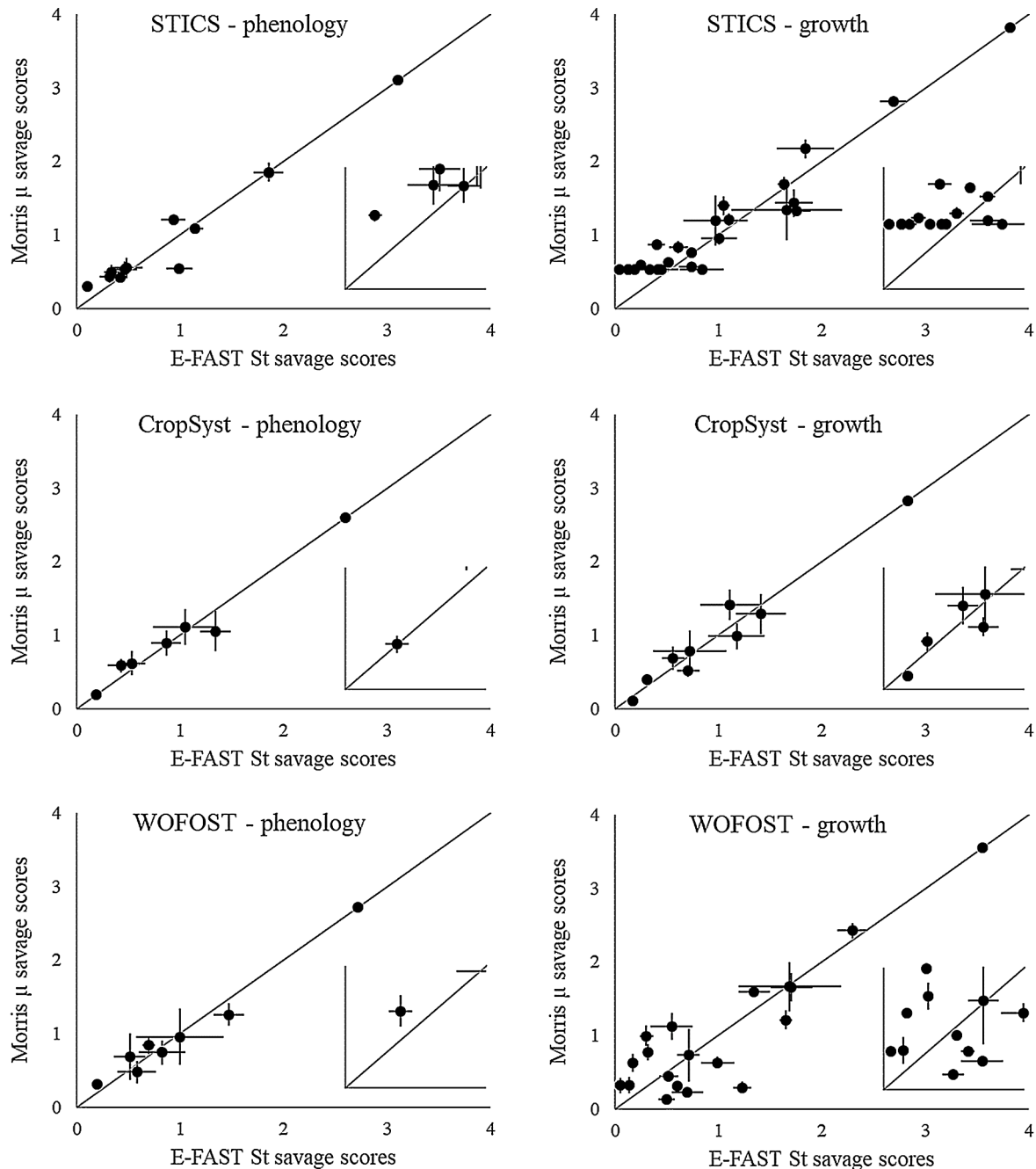


Fig. 2. Comparison between savage scores calculated on rankings resulting from sensitivity analyses performed using the E-FAST and Morris methods. Bars (one standard deviation) represent the variability in the savage scores obtained for simulations carried out in different environments. The small charts zoom in the range 0 to 0.5 for phenology and 0 to 1 for growth to increase readability for savage scores close to zero. 1:1 lines represent perfect agreement between savage scores calculated on the ranking obtained with the two sensitivity analysis methods.

absence of linearity in the relationships between the normalized values of the two sensitivity metrics. Given variance-based methods are considered as benchmark approaches for quantifying total-order effects (e.g., Saltelli and Sobol', 1995), Fig. 1 shows the low reliability of the Morris method for quantifying the absolute importance of parameters.

The savage scores of the sensitivity metrics estimated for the different environments were averaged to compare the overall capability of the two methods to rank parameters according to their total-order effect (Fig. 2). Mean savage scores highlighted the overall reliability of the Morris method in ranking model parameters, the main exception being represented by low-ranked parameters for processes involved with crop growth for STICS and WOFOST. In particular, mean savage scores estimated on the sensitivity metrics provided by the two methods were consistent for top-ranked parameters in most cases, despite a certain variability (bars in Fig. 2) among the values achieved for some parameters across different environments.

The comparison of the rankings obtained with the two methods for each environment always led to TDCC values higher than 0.90, with the concordance being significant ($p < 0.05$) regardless of the model and the environment for the parameters involved with crop growth and for the STICS parameters involved with phenological development. The concordance was instead significant only at $p < 0.10$ for the CropSyst and WOFOST parameters involved with phenology. The differences in the rankings obtained for different environments within the same SA method were always considered as minor according to TDCC (p -values always lower than 0.05), with E-FAST achieving the highest TDCC values for all the combinations process (development vs. growth) \times crop model, the only exception being represented by the CropSyst parameters involved with crop growth.

4. Discussion

For all crop models, parameters representing thermal time to complete the vegetative phase were decidedly the most relevant among those involved with phenological development. This is consistent with the great importance of the underlying plant trait within rice breeding programs, given this trait is among those for which breeders have the largest possibility to generate variability among genotypes to increase their suitability to different environments (e.g., Shrestha et al., 2014; Gao et al., 2014). Moreover, thermal time to complete the vegetative phase has a large effect on the amount of green leaf area produced by the crop before the reproductive phase, and thus it directly affects the amount of radiation that is intercepted and, in turn, the amount of biomass accumulated at the end of the crop cycle (i.e., the model output investigated in this study). However, besides the importance of related parameters for genetic and physiological reasons, the high sensitivity metrics achieved by the parameters representing thermal time to complete the vegetative phase was partly due to the fact that values found in literature led to uniform distributions, whereas normal ones were achieved for all the other parameters (Appendices A-C). Contrarily to what achieved for STICS and to what should be – to a certain extent – expected, the relevance of base temperature for development for CropSyst and WOFOST was not larger in the colder environments (Italy) according to the Morris method, thus suggesting that the different temperature response functions implemented in the three models led to compensations among parameters that were differently captured by the two SA methods.

Concerning crop growth, the most important parameters for STICS and CropSyst were involved with canopy structure (extinction coefficient ranked 1st regardless of the model and method) and photosynthetic efficiency (RUE for STICS, TUE for CropSyst), whereas for WOFOST the highest values for the sensitivity metrics were achieved by parameters involved with leaf area expansion before the close canopy stage. Besides the great conceptual difference in the way knowledge on processes involved with photosynthesis is formalized in WOFOST compared to the other two models – which could translate, *per se*, in

different models' behavior –, the lower importance of WOFOST parameters involved with photosynthetic efficiency can likely be due also to the number of parameters involved. For WOFOST, indeed, there are many parameters dealing with gross CO₂ assimilation as well as with growth and maintenance respiration that, at the end of each time step, contribute to the quantification of biomass accumulation. On the contrary, biomass accumulation is typically driven by few parameters in models based on the concept of net photosynthesis (Hammer et al., 2019). In practice, the portion of AGB variability due to photosynthesis is attributed to few parameters for STICS and CropSyst and to many parameters for WOFOST.

The lack of linearity in the relationships between the values of μ and St (Fig. 1) is in line with what achieved by Herman et al. (2013) during SA experiments carried out on distributed watershed models. These authors observed that the relationships between μ and St (the latter estimated using the Sobol' method) were approximatively linear for low-sensitive parameters, whereas they were nonlinear for high-sensitive ones. This pattern is partly reproduced in the relationships we found regardless of the process, model and environment (Fig. 1).

Differences between Morris and E-FAST mean savage scores (Fig. 2) increased with the number of parameters for which the SA was performed, with the lowest differences observed for phenology in CropSyst and WOFOST (7 and 8 parameters, respectively) and the largest for growth in STICS and WOFOST (25 and 19 parameters). Indeed, the linear relationships between the number of parameters and the relative root mean square error (RRMSE, %) calculated on mean St - and μ -savage scores was almost significant ($R^2 = 0.65$, p -value = 0.054). The same linear relationship was instead significant (p -value = 0.026) when St - and μ -savage scores were considered separately for the different environments. This finding could be considered as in line with what achieved by Silvestro et al. (2017), who reported that the correlation between Morris μ and E-FAST St was higher for the SAFYE crop model (17 parameters) than for Aquacrop (34 parameters) for water-limited wheat yield simulations in China and Italy. This consideration (higher Morris instability for large number of parameters) seems to be supported also by Hsieh et al. (2018), who discussed divergence between μ - and St -based rankings with a physiologically-based pharmacokinetic model with 58 parameters. However, this aspect needs further investigation, given the number of parameters, *per se*, is likely not enough to explain the performances of the Morris method when they are compared with variance-based approaches. Indeed, Herman et al. (2013) achieved good performance when Morris was applied to a spatially distributed model with 1092 parameters, whereas Menberg et al. (2016) considered the Morris method as unstable with a building energy model with only 11 parameters, as well as Vanrolleghem et al. (2015) with an urban water quality model with 17 parameters.

Despite minor differences in the parameter rankings obtained with the two SA methods, their concordance (TDCC) resulted always significant at the 0.10 level, and this is – to a certain extent – surprising in light of the number of model executions needed by the two methods to provide the sensitivity metrics, which for Morris was less than 3% of the number of executions needed by E-FAST.

The sample sizes we used for the two SA methods allowed achieving stable results. Indeed, the comparison between the rankings presented in Fig. 2 (obtained using six trajectories (Morris) and ~250 times the number of parameters (E-FAST)) and the rankings obtained with five and seven trajectories, and ~200 and ~300 times the parameters led to TDCC values ranging from 0.979 to 1.000, with concordance being always significant ($p < 0.05$).

Our results seem to contradict the findings from other authors (e.g., Sobol' and Kucherenko, 2009; Feng et al., 2018), who observed distortions in the relationships between sensitivity metrics from Morris and variance-based methods in case non-uniform distributions for input factors were used. Indeed, regardless of the crop model and of the environmental conditions explored, Morris and E-FAST results were

consistent despite the values for the parameters we collected from the literature led to normal distributions for all the parameters but one (thermal time to complete the vegetative phase).

5. Conclusions

In this study, a number of SA experiments was performed for the first time using different SA methods on different crop models at sub-model level (phenological development vs crop growth) while exploring different environmental conditions.

Results achieved by Morris seem – to a certain extent – to cast a shadow on the need of using SA methods more demanding in terms of model executions, especially in case of computationally expensive models or for high-resolution spatially distributed SA experiments (e.g., Herman et al., 2013; Koo et al., 2020). However, the solution is more complex, since it depends on the specific objectives of the SA study. Indeed, although SA is often performed to identify relevant parameters before calibration (e.g., Mathieu et al., 2018) or non-influential ones for model reduction or simplification (e.g., Hsieh et al., 2018), there are cases where the contribution of parameters to output variance needs to be quantified beyond what is needed to rank parameters. This means that the sensitivity metrics of the different parameters are important also in term of their absolute value, not only in relative terms as when ranking is the final goal. This is the case when sensitivity metrics are used for further analyses, like for deriving quantitative criteria to identify crop ideotypes (e.g., Paleari et al., 2017) or, in general, when an accurate estimate of the contribution of parameters to output variability is needed. Indeed, the lack of linearity between E-FAST St and Morris μ confirmed the unsuitability of the latter for quantifying parameter

relevance in absolute terms.

CRedit authorship contribution statement

Livia Paleari: Conceptualization, Methodology, Investigation, Formal analysis, Writing - original draft, Writing - review & editing. **Ernes Movedi:** Methodology, Investigation, Formal analysis, Writing - review & editing. **Michele Zoli:** Investigation, Formal analysis, Writing - review & editing. **Andrea Burato:** Investigation, Formal analysis, Writing - review & editing. **Irene Cecconi:** Investigation, Formal analysis, Writing - review & editing. **Jabir Errahouly:** Investigation, Formal analysis, Writing - review & editing. **Eleonora Pecollo:** Investigation, Formal analysis, Writing - review & editing. **Carla Sorvillo:** Investigation, Formal analysis, Writing - review & editing. **Roberto Confalonieri:** Conceptualization, Methodology, Investigation, Formal analysis, Writing - original draft, Writing - review & editing.

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.

Acknowledgements

We would like to gratefully acknowledge Dr. Tao Li from DNDI Applications Research and Training and Dr. Ando Radanielson from the University of Southern Queensland for providing information needed to run the simulations in the two environments in the Philippines.

Appendices

Appendix A. STICS parameters, distributions and sources of information

Parameter name (acronym, units)	Distribution	Source ^a
Phenological development		
Minimum temperature below which development stops (tdmin, °C)	Normal (11.5, 0.71)	1
Maximum temperature above which development stops (tdmax, °C)	Normal (30, 1.5)	1
Saturating photoperiod for development (phosat, h)	Normal (10, 0.5)	2
Base photoperiod for development (phobase, h)	Normal (14, 0.7)	2
Index of photoperiod sensitivity (1 = insensitive) (sensiphot, -)	Normal (0.9, 0.045)	2
Duration between sowing and emergence (stplgtger, °C-days)	Normal (50, 2.5)	2
Minimum temperature for emergence (tgmin, °C)	Normal (8.5, 0.43)	2, 3
Thermal time emerg. – end of juvenile phase (stlevamf, °C-days)	Normal (447.4, 146.1)	2, 4
Therm. time max. acceleration of leaf growth – max. LAI (stamflax, °C-days)	Normal (359, 44.4)	2, 4
Thermal time emergence – beginning of grain filling (stlevdrp, °C-days)	Uniform (330–1100)	2, 4
Duration grain filling – onset water dynamics in grains (stdrpdcs, °C-days)	Normal (250, 12.5)	2
Thermal time beginning of grain filling – maturity (stdrpmat, °C-days)	Normal (477.7, 35.2)	2, 4
Growth (light interception, photosynthesis, assimilate partitioning, leaf area, yield formation)		
Minim. density above which interplant competition starts (bdens, plants m^{-2})	Normal (34.5, 1.72)	2
Temperature beyond which foliar growth stops (tcxstop, °C)	Normal (36, 1.8)	2
Minimum temperature at which growth ceases (tcmin, °C)	Normal (12, 0.6)	2
Maximum temperature at which growth ceases (tcmax, °C)	Normal (30, 1.5)	2
Max. rate of the setting up of LAI (dlaimaxbrut, m^2 leaf plant $^{-1}$ °C-day $^{-1}$)	Normal (6×10^{-4} , 3×10^{-5})	2, 4
Parameter of the logistic curve of LAI growth (pentlai, -)	Normal (5.5, 0.275)	2
Maximum height of the crop (hautmax, m)	Normal (0.96, 0.12)	2
Extinction coefficient of photosynthetic active radiation (extin, -)	Normal (0.45, 0.08)	1, 5, 6
Minimum temperature for plant growth (temin, °C)	Normal (12, 0.6)	1, 2
Optimal temperature (2/2) for plant growth (teoptbis, °C)	Normal (28, 2)	1, 4
Optimal temperature (1/2) for plant growth (teopt, °C)	Normal (20.27, 1.01)	2, 4
Maximum threshold temp. for net photosynthesis (temax, °C)	Normal (38.15, 1.91)	2, 4
Max. RUE during the juvenile phase (efcroijuv, g MJ $^{-1}$)	Normal (2.95, 0.15)	1, 2, 4, 6
Max. RUE during the vegetative stage (efcroiveg, g MJ $^{-1}$)	Normal (3.2, 0.16)	1, 2, 4, 6
Max. RUE during the grain filling phase (efcroirepro, g MJ $^{-1}$)	Normal (2.8, 0.14)	1, 2, 4, 6
Minimum specific leaf area of green leaves (SLAmin, m^2 kg $^{-1}$)	Normal (18, 9)	1, 2
Maximum specific leaf area of green leaves (SLAmax, m^2 kg $^{-1}$)	Normal (500, 25)	1, 2
Stem/leaf proportion (tgefeuille, -)	Normal (3.5, 0.175)	2
Rate of increase of the C harvest index (vitircarb, g grain g plant $^{-1}$ day $^{-1}$)	Normal (0.01, 5×10^{-5})	2, 4
Maximum C harvest index (irmax, g grain g plant $^{-1}$)	Normal (0.65, 0.034)	2

(continued on next page)

(continued)

Parameter name (acronym, units)	Distribution	Source ^a
Max. temperature above which grain filling stops (tmaxemp, °C)	Normal (36, 1.8)	2
Minimal temperature below which grain filling stops (tminemp, °C)	Normal (12, 0.6)	2
Max. grain weight at 0% water content (pgrainmaxi, g grain ⁻¹)	Norm. (25×10^{-3} , 2.7×10^{-3})	2, 4, 7
Interplant competition parameter (adens, -)	Normal (-0.6, 0.03)	2, 4
Max. lifespan of adult leaves in sum. Q10=2 (2(T-Tbase)) (durvieF, -)	Normal (120, 6)	2

^a1: Confalonieri and Bocchi (2005); 2: model default; 3: Sanchez et al. (2014); 4: Bregaglio et al. (2017); 5: Kiniri et al. (2001); 6: Boschetti et al. (2006); 7: Timsina and Humphreys (2006).

Appendix B. CropSyst parameters, distributions and sources of information

Parameter name (acronym, units)	Distribution	Source ^a
Phenological development		
Thermal time from flowering to maturity (GDDmat, °C-days)	Normal (388, 61)	1, 2
Thermal time from emergence to flowering (GDDfl, °C-days)	Uniform (400–1560)	1, 2
Thermal time to reach emergence (GDDem, °C-days)	Normal (70, 3.5)	1, 2
Base temperature for thermal time accumulation (Tbasedev, °C)	Normal (11.5, 0.7)	1
Optimum temperature for thermal time accumulation (Toptdev, °C)	Normal (30, 1.5)	1
Critical day length duration (Phototh, h)	Normal (12.16, 1.13)	3
Photoperiod sensitivity (Photosens, -)	Normal (0.9, 0.05)	4
Growth (light interception, photosynthesis, assimilate partitioning, leaf area, yield formation)		
Specific leaf area at early stages (SLA, m ² kg ⁻¹)	Normal (28.05, 1.4)	1, 5
Stem-leaf partition coefficient (SLP, -)	Normal (1.45, 0.53)	1
Maximum radiation use efficiency (RUE, g MJ ⁻¹)	Normal (2.7, 0.1)	5, 6
Canopy light extinction coefficient (k, -)	Normal (0.45, 0.08)	1, 5, 6
Minimum temperature for growth (Tmin, °C)	Normal (10.9, 0.75)	1, 4
Optimum temperature for growth (Topt, °C)	Normal (28, 2)	1
Leaf duration (LeafDuration, °C day)	Normal (700, 80)	1, 4
Biomass-transpiration coefficient (BTR, kPa kg m ⁻³)	Normal (6.1, 0.3)	1
Leaf area index at emergence (LALini, m ² m ⁻²)	Normal (0.01, 0.005)	1, 4

^a 1: Confalonieri and Bocchi (2005); 2: Li et al. (2015); 3: Timsina and Humphreys (2006); 4: unpublished data; 5: Boschetti et al. (2006); 6: Kiniri et al. (2001).

Appendix C. WOFOST parameters, distributions and sources of information

Parameter name (acronym, units)	Distribution	Source ^a
Phenological development		
Base temperature for development (Tbdev, °C)	Normal (11.5, 0.7)	1
Optimum temperature for development (Toptdev, °C)	Normal (30, 1.5)	1
Maximum threshold temperature for development (Tmaxdev, °C)	Normal (42, 2.1)	1
Thermal time from flowering to maturity (TSUM2, °C-days)	Normal (388, 61)	1, 2
Thermal time from emergence to flowering (TSUM1, °C-days)	Uniform (400–1560)	1, 2
Thermal time from sowing to emergence (TSUMEM, °C-days)	Normal (70, 3.5)	1, 2
Critical day length (PHinib, h)	Normal (12.16, 1.13)	3
Photoperiod sensitivity (PHins, -)	Normal (0.9, 0.05)	4
Growth (light interception, photosynthesis, assimilate partitioning, leaf area, yield formation)		
Maximum leaf CO ₂ assimilation rate (AMAX, kg ha ⁻¹ h ⁻¹)	Normal (49.6, 2.48)	5, 6, 7
Lower threshold of temp. effects on CO ₂ assimilation (Tb, °C)	Normal (12, 0.6)	1, 7
Optimum temp. effects on CO ₂ assimilation (Topt, °C)	Normal (28, 2)	1, 7
Max. threshold of temp. effects on CO ₂ assimilation (Tmax, °C)	Normal (42, 2)	1, 7, 8
Rel. maintenance respiration rate for roots (RMR, kg CH ₂ O kg ⁻¹ day ⁻¹)	Normal (0.012, 0.0011)	5, 7, 9
Rel. maintenance respiration rate for leaves (RML, kg CH ₂ O kg ⁻¹ day ⁻¹)	Normal (0.028, 0.0005)	5, 7, 9
Rel. maintenance respiration rate for stems (RMS, kg CH ₂ O kg ⁻¹ day ⁻¹)	Normal (0.018, 0.001)	5, 7, 9
Rel. maint. resp. rate for storage organs (RMO, kg CH ₂ O kg ⁻¹ day ⁻¹)	Normal (0.01, 0.0005)	5, 7, 9
Efficiency of conversion into roots (CVR, -)	Normal (0.7125, 0.0356)	5, 7, 9
Efficiency of conversion into leaves (CVL, -)	Normal (0.7125, 0.0356)	5, 7, 9
Efficiency of conversion into stems (CVS, -)	Normal (0.7125, 0.0356)	5, 7, 9
Efficiency of conversion into storage organs (CVO, -)	Normal (0.675, 0.0338)	5, 7, 9
Specific leaf area at emergence (SLAini, kg m ⁻²)	Normal (0.0045, 0.0003)	5
Specific leaf area at tillering (SLAtill, kg m ⁻²)	Normal (0.003, 0.0002)	5
Partitioning of assimilates to leaves at emergence (RipLO, -)	Normal (0.7, 0.1)	5, 10
Leaf area index at emergence (LAIEM, -)	Normal (0.01, 0.005)	5, 10
Life span of leaves growing at 35 °C (SPAN, days)	Normal (35, 3.5)	7
Canopy light extinction coefficient (k, -)	Normal (0.45, 0.08)	7, 8, 11
Maximum relative increase in leaf area index (RGRLAI, -)	Normal (0.0085, 0.00048)	5, 7, 9

^a1: Confalonieri and Bocchi (2005); Li et al. (2015); 3: Timsina and Humphreys (2006); 4: unpublished data; 5: Stella et al. (2014); 6: Ziska and Teramura (1992); 7: Kropff et al. (1994); 8: Dingkuhn et al. (1999); 9: van Diepen et al., 1988; 10: Confalonieri et al. (2010); 11: Boschetti et al. (2006);

References

- Anderson, B., Borgonovo, E., Galeotti, M., Roson, R., 2014. Uncertainty in climate change modeling: can global sensitivity analysis be of help? *Risk Anal.* 34, 271–293.
- Borgonovo, E., Plischke, E., 2016. Sensitivity analysis: a review of recent advances. *Eur. J. Oper. Res.* 248, 869–887.
- Boschetti, M., Bocchi, S., Stroppiana, D., Brivio, P.A., 2006. Estimation of parameters describing morpho-physiological features of Mediterranean rice varieties for modelling purposes. *Ital. J. Agrometeorol.* 3, 40–49.
- Bregaglio, S., Hossard, L., Cappelli, G., Resmond, R., Bocchi, S., Barbier, J.-M., Ruget, F., Delmotte, S., 2017. Identifying trends and associated uncertainties in potential rice production under climate change in Mediterranean areas. *Agr. Forest Meteorol.* 237–238, 219–232.
- Brisson, N., Mary, B., Ripoche, D., Jeuffroy, M.H., Ruget, F., Nicoulaud, B., Gate, P., Devienne-Barret, F., Antonioletti, R., Durr, C., Richard, C., Beaudoin, N., Recous, S., Tayot, X., Plenet, D., Cellier, P., Machet, J.-M., Meynard, J.M., Delécolle, R., 1998. STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parameterization applied to wheat and corn. *Agronomie* 18, 311–346.
- Campolongo, F., Cariboni, J., Saltelli, A., 2007. An effective screening design for sensitivity analysis of large models. *Environ. Model. Softw.* 22, 1509–1518.
- Confalonieri, R., 2012. Combining a weather generator and a standard sensitivity analysis method to quantify the relevance of weather variables on agrometeorological models outputs. *Theor. Appl. Climatol.* 108, 19–30.
- Confalonieri, R., Bellocchi, G., Bregaglio, S., Donatelli, M., Acutis, M., 2010. Comparison of sensitivity analysis techniques: a case study with the rice model WARM. *Ecol. Model.* 221, 1897–1906.
- Confalonieri, R., Bocchi, S., 2005. Evaluation of CropSyst for simulating the yield of flooded rice in northern Italy. *Eur. J. Agron.* 23, 315–326.
- Degenring, D., Froemel, C., Dikta, G., Takors, R., 2004. Sensitivity analysis for the reduction of complex metabolic models. *J. Process Contr.* 14, 729–745.
- Dingkuhn, M., Johnson, D.E., Sow, A., Audebert, A.Y., 1999. Relationships between upland rice canopy characteristics and weed competitiveness. *Field Crop. Res.* 61, 79–95.
- Feng, K., Lu, Z., Yang, C., 2018. Enhanced Morris method for global sensitivity analysis: good proxy of Sobol' index. *Struct. Multidiscip. Opt.* 59, 373–387.
- Gao, H., Jin, M., Zheng, X.-M., Chen, J., Yuan, D., Xin, Y., Wang, M., Huang, D., Zhang, Z., Zhou, K., Sheng, P., Ma, J., Ma, W., Deng, H., Jiang, L., Liu, S., Wang, H., Wu, C., Yuan, L., Wan, J., 2014. Days to heading 7, a major quantitative locus determining photoperiod sensitivity and regional adaptation in rice. *P. Natl. Acad. Sci. USA* 111, 16337–16342.
- Hamby, D.M., 1995. A comparison of sensitivity analysis techniques. *Health Phys* 68, 195–204.
- Hammer, G., Messina, C., Wu, A., Cooper, M., 2019. Biological reality and parsimony in crop models – why we need both in crop improvement! *Silico Plants* 1, diz010.
- Her, Y., Chaubey, I., 2015. Impact of the numbers of observations and calibration parameters on equifinality, model performance, and output and parameter uncertainty: parameters, observations, and uncertainty. *Hydrol. Process.* 29, 4220–4237.
- Herman, J.D., Kollat, J.B., Reed, P.M., Wagener, T., 2013. Technical note: method of Morris effectively reduces the computational demands of global sensitivity analysis for distributed watershed models. *Hydrol. Earth Syst. Sci.* 17, 2893–2903.
- Homma, T., Saltelli, A., 1996. Importance measures in global sensitivity analysis of nonlinear models. *Reliab. Eng. Syst. Safe.* 52, 1–17.
- Hsieh, N.-H., Reifeld, B., Bois, F.Y., Chiu, W.A., 2018. Applying a global sensitivity analysis workflow to improve the computational efficiencies in physiologically-based pharmacokinetic modeling. *Front. Pharmacol.* 9, 588.
- Iman, R.L., Conover, W.J., 1987. A measure of top-down correlation. *Technometrics* 29, 351–357.
- King, D.M., Perera, B.J.C., 2013. Morris method of sensitivity analysis applied to assess the importance of input variables on urban water supply yield – A case study. *J. Hydrol.* 477, 17–32.
- Kiniry, J.R., Mc Cauley, G., Xie, Y., Arnorl, J.G., 2001. Rice parameters describing crop performance of four U.S. cultivars. *Agron. J.* 93, 1354–1361.
- Koo, H., Iwanaga, T., Croke, B.F.W., Jakeman, A.J., Yang, J., Wang, H.-H., Sun, X., Lü, G., Li, X., Yue, T., Yuan, W., Liu, X., Chen, M., 2020. Position paper: sensitivity analysis of spatially distributed environmental models- a pragmatic framework for the exploration of uncertainty sources. *Environ. Model. Softw.* 134, 104857.
- Kropff, M.J., van Laar, H.H., Matthews, R.B., 1994. ORYZA1: an ecophysiological model for irrigated rice production. *SARP Research Proceedings. IRRI, Los Baños, Philippines.*
- Li, T., Hasegawa, T., Yin, X., Zhu, Y., Boote, K., Adam, M., Bregaglio, S., Buis, S., Confalonieri, R., Fumoto, T., Gaydon, D., Marcaida III, M., Nakagawa, H., Oriol, P., Ruane, A.C., Ruget, F., Singh, B., Singh, U., Tang, L., Tao, F., Wilkens, P., Yoshida, H., Zhang, Z., Bouman, B., 2015. Uncertainties in predicting rice yield by current crop models under a wide range of climatic conditions. *Glob. Change Biol.* 21, 1328–1341.
- Ligmann-Zielinska, A., Siebers, P.-O., Magliocchia, N., Parker, D., Grimm, V., Jing Du, E., Cenek, M., Radchuk, V., Arbab, N.N., Li, S., Berger, U., Paudel, R., Robinson, D.T., Jankowski, P., An, L., Ye, X., 2020. JASSS-J. Artif. Soc. S. 23 (6), 2020.
- Makowski, D., Naud, C., Jeuffroy, M.-H., Barbotin, A., Monod, H., 2006. Global sensitivity analysis for calculating the contribution of genetic parameters to the variance of crop model prediction. *Reliab. Eng. Syst. Safe.* 91, 1142–1147.
- Martre, P., He, J., Le Gouis, J., Semenov, M.A., 2015. In silico system analysis of physiological traits determining grain yield and protein concentration for wheat as influenced by climate and crop management. *J. Exp. Bot.* 66, 3581–3598.
- Mathieu, A., Vidal, T., Jullien, A., Wu, Q., Chambon, C., Bayol, B., Cournède, P.-H., 2018. A new methodology based on sensitivity analysis to simplify the recalibration of functional-structural plant models in new conditions. *Ann. Bot.* 122, 397–408.
- Menberg, K., Heo, Y., Augenbroe, G., Choudhary, R., 2016. New extension of Morris method for sensitivity analysis of building energy models. *Proceedings of the Building Simulation & Optimization Conference. Newcastle, UK*, p. 1101 paper n.
- Morris, M.D., 1991. Factorial sampling plans for preliminary computational experiments. *Technometrics* 33, 161–174.
- Nguyen, A.-T., Reiter, S., 2015. A performance comparison of sensitivity analysis methods for building energy models. *Build. Simul.* 8, 651–664.
- Paleari, L., Confalonieri, R., 2016. Sensitivity analysis of a sensitivity analysis: we are likely overlooking the impact of distributional assumptions. *Ecol. Model.* 340, 57–63.
- Paleari, L., Movedi, E., Cappelli, G., Wilson, L.T., Confalonieri, R., 2017. Surfing parameter hyperspaces under climate change scenarios to design future rice ideotypes. *Glob. Change Biol.* 23, 4651–4662.
- Pastres, R., Chan, K., Solidoro, C., Dejak, C., 1999. Global sensitivity analysis of a shallow-water 3D eutrophication model. *Comput. Phys. Commun.* 117, 62–74.
- Pianosi, F., Sarrazin, F., Wagener, T., 2015. A Matlab toolbox for Global Sensitivity Analysis. *Environ. Model. Softw.* 70, 80–85.
- Pianosi, F., Beven, K., Freer, J., Hall, J.W., Rougier, J., Stephenson, D.B., Wagener, T., 2016. Sensitivity analysis of environmental models: a systematic review with practical workflow. *Environ. Model. Softw.* 79, 214–232.
- Radanielson, A.M., Gaydon, D.S., Li, T., Angeles, O., Roth, C.H., 2018. Modeling salinity effect on rice growth and grain yield with ORYZA v3 and APSIM-Oryza. *Eur. J. Agron.* 100, 44–55.
- Richter, G.M., Acutis, M., Trevisiol, P., Latiri, K., Confalonieri, R., 2010. Sensitivity analysis for a complex crop model applied to Durum wheat in the Mediterranean. *Eur. J. Agron.* 32, 127–136.
- Ryan, E., Wild, O., Voulgarakis, A., Lee, L., 2018. Fast sensitivity analysis methods for computationally expensive models with multi-dimensional output. *Geosci. Model Dev.* 11, 3131–3146.
- Saltelli, A., 2002. Sensitivity analysis for importance assessment. *Risk Anal* 22, 579–590.
- Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., Wu, Q., 2019. Why so many published sensitivity analyses are false: a systematic review of sensitivity analysis practices. *Environ. Model. Softw.* 114, 29–39.
- Saltelli, A., Sobol', I.M., 1995. About the use of rank transformation in sensitivity analysis model. *Reliab. Eng. Syst. Saf.* 50, 225–239.
- Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M., 2004. *Sensitivity Analysis in Practice*. Wiley, New York, NY, USA, p. 219.
- Saltelli, A., Tarantola, S., Chan, K.P.S., 1999. A quantitative model-independent method for global sensitivity analysis of model output. *Technometrics* 41, 39–56.
- Sánchez, B., Rasmussen, A., Porter, J.R., 2014. Temperatures and the growth and development of maize and rice: a review. *Glob. Change Biol.* 20, 408–417.
- Shrestha, R., Gómez-Ariza, J., Brambilla, V., Fornara, F., 2014. Molecular control of seasonal flowering in rice, arabidopsis and temperate cereals. *Ann. Bot.* 114, 1445–1458.
- Silvestro, P.C., Pignatti, S., Yang, H., Yang, G., Pescucci, S., Castaldi, F., Casa, R., 2017. Sensitivity analysis of the Aquacrop and SAFYE crop models for the assessment of water limited winter wheat yield in regional scale applications. *PLoS ONE* 12, e0187485.
- Sobol', I.M., 1993. Sensitivity estimates for nonlinear mathematical models. *Math. Model. Comput. Exp.* 14, 407–414.
- Sobol', I.M., Kucherenko, S., 2009. Derivative based global sensitivity measures and the link with global sensitivity indices. *Math. Comput. Simulat.* 79, 3009–3017.
- Stearns, S.C., 1992. *The Evolution of Life Histories*. Oxford University Press, Oxford.
- Stella, T., Frasso, N., Negrini, G., Bregaglio, S., Cappelli, G., Acutis, M., Confalonieri, R., 2014. Model simplification and development via reuse, sensitivity analysis and composition: a case study in crop modelling. *Environ. Model. Softw.* 59, 44–58.
- Stöckle, C.O., Martin, S.A., Campbell, G.S., 1994. CropSyst, a cropping systems simulation model: water/nitrogen budgets and crop yield. *Agric. Syst.* 46, 335–359.
- Tanner, C.B., Sinclair, T.R., 1983. Efficient water use in crop production: research or research? In: Taylor, H.M., Jordan, W.R., Sinclair, T.R. (Eds.), *Limitations to Efficient Water Use in Crop Production*. Amer. Soc. of Agron. Madison, pp. 1–27.
- Tarantola, S., Saltelli, A., 2003. SAMO 2001: methodological advances and innovative applications of sensitivity analysis. *Reliab. Eng. Syst. Saf.* 79, 121–122.
- Timsina, J., Humphreys, E., 2006. Performance of CERES-Rice and CERES-Wheat models in rice-wheat systems: a review. *Agr. Syst.* 90, 5–31.
- van Diepen, C.A., Rappoldt, C., Wolf, J., van Keulen, H., 1988. *Crop Growth Simulation Model WOFOST (Documentation Version 4.1)*. Centre for World Food Studies, Wageningen, p. 299.
- van Keulen, H., Wolf, J., 1986. *Modelling of agricultural production: weather soils and crops. Simulation Monographs*. Pudoc, Wageningen.
- Vanrolleghem, P.A., Mannina, G., Cosenza, A., Neumann, M.B., 2015. Global sensitivity analysis for urban water quality modelling: terminology, convergence and comparison of different methods. *J. Hydrol.* 522, 339–352.
- Whittaker, G., Confesor Jr., R., Di Luzio, M., Arnold, J.G., 2010. Detection of overparameterization and overfitting in an automatic calibration of SWAT. *T. ASABE* 53, 1487–1499.
- Xu, C., Gertner, G., 2011. Understanding and comparisons of different sampling approaches for the Fourier Amplitudes Sensitivity Test (FAST). *Comput. Stat. Data Anal.* 55, 184–198.
- Ziska, L.H., Teramura, A.H., 1992. CO₂ enhancement of growth and photosynthesis in rice (*Oryza sativa*). *Plant Physiol* 99, 473–481.