

1

2 **Modeling sugar cane yield with a process-based**  
3 **model from site to continental scale:**  
4 **uncertainties arising from model structure and**  
5 **parameter values.**

6

7 Valade A.<sup>a</sup>, Ciais P.<sup>a</sup>, Vuichard N.<sup>a</sup>, Viovy N.<sup>a</sup>, Huth N.<sup>b</sup>, Marin F.<sup>c</sup>, Martiné J.-F.<sup>d</sup>

8

9 <sup>a</sup>LSCE, CEA-CNRS, Gif-sur-Yvette, 91191 France

10 <sup>b</sup>CSIRO. Ecosystem Sciences, PO Box 102, Toowoomba, Qld, 4350, Australia

11 <sup>c</sup>EMBRAPA Informatica Agropecuria, Barão Geraldo, 13083-886 Campinas SP, Brazil

12 <sup>d</sup>CIRAD, UR SCA, Saint-Denis, La Réunion, F-97408 France

13

1

## 2 **Abstract**

3 Agro-Land Surface Models (agro-LSM) have been developed from the integration  
4 of specific crop processes into large-scale generic land surface models that allow  
5 calculating the spatial distribution and variability of energy, water and carbon  
6 fluxes within the soil-vegetation-atmosphere continuum. When developing agro-  
7 LSM models, a particular attention must be given to the effects of crop phenology  
8 and management on the turbulent fluxes exchanged with the atmosphere, and  
9 the underlying water and carbon pools. A part of the uncertainty of Agro-LSM  
10 models is related to their usually large number of parameters. In this study, we  
11 quantify the parameter-values uncertainty in the simulation of sugar cane  
12 biomass production with the agro-LSM ORCHIDEE-STICS, using a multi-regional  
13 approach with data from sites in Australia, La Réunion and Brazil. In ORCHIDEE-  
14 STICS, two models are chained: STICS, an agronomy model that calculates  
15 phenology and management, and ORCHIDEE, a land surface model that calculates  
16 biomass and other ecosystem variables forced by STICS phenology. First, the  
17 parameters that dominate the uncertainty of simulated biomass at harvest date  
18 are determined through a screening of 67 different parameters of both STICS and  
19 ORCHIDEE on a multi-site basis. Secondly, the uncertainty of harvested biomass  
20 attributable to those most sensitive parameters is quantified and specifically  
21 attributed to either STICS (phenology, management) or to ORCHIDEE (other  
22 ecosystem variables including biomass) through distinct Monte-Carlo runs. The  
23 uncertainty on parameter values is constrained using observations by calibrating  
24 the model independently at seven sites. In a third step, a sensitivity analysis is  
25 carried out by varying the most sensitive parameters to investigate their effects  
26 at continental scale. A Monte-Carlo sampling method associated with the  
27 calculation of Partial Ranked Correlation Coefficients is used to quantify the  
28 sensitivity of harvested biomass to input parameters on a continental scale  
29 across the large regions of intensive sugar cane cultivation in Australia and  
30 Brazil. Ten parameters driving most of the uncertainty in the ORCHIDEE-STICS  
31 modeled biomass at the 7 sites are identified by the screening procedure. We  
32 found that the 10 most sensitive parameters control phenology (maximum rate

1 of increase of LAI) and root uptake of water and nitrogen (root profile and root  
2 growth rate, nitrogen stress threshold) in STICS, and photosynthesis (optimal  
3 temperature of photosynthesis, optimal carboxylation rate), radiation  
4 interception (extinction coefficient), and transpiration and respiration (stomatal  
5 conductance, growth and maintenance respiration coefficients) in ORCHIDEE.  
6 We find that the optimal carboxylation rate and photosynthesis temperature  
7 parameters contribute most to the uncertainty in harvested biomass simulations  
8 at site scale. The spatial variation of the ranked correlation between input  
9 parameters and modeled biomass at harvest is well explained by rain and  
10 temperature drivers, suggesting climate-mediated different sensitivities of  
11 modeled sugar cane yield to the model parameters, for Australia and Brazil. This  
12 study reveals the spatial and temporal patterns of uncertainty variability for a  
13 highly parameterized agro-LSM and calls for more systematic uncertainty  
14 analyses of such models.

15

## 16 **1 Introduction**

17

18 In the recent years, many governments have set targets in terms of biofuels  
19 consumption for transportation fuel (Sorda et al., 2010), resulting in a large increase  
20 in bioenergy cropping area around the world. Concerns about energy shortage, policy  
21 to reduce CO<sub>2</sub> emissions, and the search for new income for farmers can explain why  
22 energy policies have considered biofuels as a serious alternative to fossil fuel in many  
23 countries (Demirbas, 2008). Yet, the claimed benefits of biofuels for fossil fuel  
24 substitution have been questioned in terms of their net effect on atmospheric CO<sub>2</sub> and  
25 climate, and even of their economic return (Doornbosch and Steenblik; Naylor et al.,  
26 2007). In particular, the conditions of biofuel cultivation, such as the type of crop,  
27 practice, previous land use, and local climate, have emerged as key factors that  
28 determine the effectiveness of their carbon emissions reduction (Fargione et al., 2008;  
29 Hill et al., 2006; Searchinger et al., 2008). At the heart of biofuel cultivation is ethanol  
30 that represents today 74% of the energy content of the world production of liquid  
31 biofuels (Howarth et al., 2008) and whose production is expected to double between  
32 2011 and 2021 (OECD, 2012), hence the urgency to better quantify and understand

1 regional potentials of bioethanol crops. Based on recent life cycle analysis studies (de  
2 Vries et al., 2010; Schubert, 2006; von Blottnitz and Curran, 2007), ethanol from  
3 sugar cane is the most competitive in terms of energy use and net carbon balance and  
4 the energy use projections from the International Energy Agency foresee that by 2050,  
5 sugar cane is the only 1<sup>st</sup> generation biofuel that that will keep expanding (IEA, 2011).

6

7 The impact of sugar cane expansion on climate and carbon balance is under  
8 scrutiny with different approaches. Satellite observation data have been used to  
9 study biophysical effects of sugar cane expansion on local temperature in the  
10 Brazilian Cerrado (Loarie et al., 2011) Survey for agricultural and industrial  
11 performances from sugar cane mills have allowed Macedo et al. (2008) to establish  
12 the carbon balance of sugar cane ethanol production in the Center-South of Brazil.  
13 Georgescu et al. (2013) simulate the hydroclimatic impacts of sugar cane expansion  
14 by forcing sugar cane land cover characteristics into a regional climate model. All  
15 approaches provide useful information on impacts and potentials but are impractical  
16 to apply outside of the regions and conditions (climate, management) where they have  
17 been conducted.

18

19 In parallel with empirical approaches, significant progress has been made towards  
20 mechanistic modeling of sugar cane yields using models. Crop models are generally  
21 used to simulate sugar cane production at site scale, with specific parameters  
22 (Cheeroo-Nayamuth et al., 2000). Land surface models (LSM) are rather used to  
23 estimate the spatial distribution of crop productivity under different soil and climatic  
24 conditions, over a region or even over the globe but with a simpler and generic  
25 description of sugar cane plants (Black et al., 2012; Cuadra et al., 2012; Lapola et al.,  
26 2009). Agro-LSM models stand at the interface between plot-scale crop models and  
27 global LSMs. Yet, as highlighted by Surendran Nair et al. (2012) if the development  
28 of agro-LSM models for biofuels has been the subject of much interest recently,  
29 detailed parameterization, validation and uncertainty quantification is still very  
30 limited in regional and global applications, and efforts must be made in that direction.  
31 The importance of evaluating and communicating about global models uncertainty  
32 was as well emphasized within the framework of the model inter-comparison project

1 AgMIP - providing insights for IPCC AR5 report – in which crop models uncertainty  
2 is identified as a key theme of interest that was only little explored so far (Rosenzweig  
3 et al., 2013). ORCHIDEE-STICS (Gervois et al., 2004) is an agro-LSM model that  
4 has been developed from the coupling of the agronomical model STICS (Brisson et  
5 al., 1998) and the Land Surface Model ORCHIDEE (Krinner et al., 2005) and that has  
6 been applied for studies from site to continent mainly for temperate crops in Europe  
7 (Gervois et al., 2008) and has been recently adapted to sugar cane simulation (Valade  
8 et al., 2013).

9

10 Four uncertainty sources affect the simulation of sugar cane biomass with  
11 ORCHIDEE-STICS: 1) input uncertainty on boundary conditions used for climate  
12 drivers and soil properties, 2) structure uncertainty related to model equations and  
13 parameterizations, 3) parameters value uncertainty, and 4) uncertainty associated with  
14 the measurements used for model evaluation or calibration. Here we focus on  
15 structure and parameters uncertainty and try to estimate how these two sources of  
16 uncertainties affect the simulations of sugar cane harvest biomass. We want to  
17 determine which parameters are responsible for most of the uncertainty in harvest  
18 biomass (screening analysis) and to what extent this is related to the model's structure  
19 (uncertainty analysis). In addition, we want to quantify this uncertainty and examine  
20 its temporal and spatial variability (sensitivity analysis).

21

22 In the following, we first present the sites and regions considered in this study (section  
23 2.1) and the main features of the ORCHIDEE-STICS model (section 2.2). We then  
24 describe the screening algorithm used to sort the most important parameters (section  
25 2.3), and the uncertainty and the sensitivity analyses (sections 2.4 and 2.5). Then we  
26 discuss the results of the screening analysis, in terms of the parameters identified by  
27 the screening as the most important for controlling harvested sugar cane biomass  
28 (section 3.1). We describe the results for the measure of the uncertainty calculated for  
29 7 sites in section 3.2 to 3.4 and present maps of the sensitivity of the model to its main  
30 parameters in section 3.5.

31

## 1    **2    Materials and methods**

2    In this study, we aim to quantify the uncertainty related to the parameter values of a  
3    chain of two process-based models (STICS-ORCHIDEE) to simulate sugar cane yield  
4    (biomass at harvest date). This is a difficult task because this model is a detailed and  
5    complex model that contains over 100 plant specific parameters within the primitive  
6    equations of phenology, energy and water balance, photosynthesis and allocation. We  
7    perform the uncertainty analysis in three steps, illustrated in Figure 1 and consisting  
8    of screening, uncertainty and sensitivity analyses, all described in more details in  
9    section 2. These three steps are sequential and complementary. The first step is a  
10   screening to sort the most important parameters controlling yield, and to reduce the  
11   dimension of the parameter space from a large number of parameters to few key  
12   parameters, allowing a moderate number of sensitivity simulations. The screening  
13   allows the restriction of the two further steps to a smaller parameter subset. The  
14   second step is an uncertainty analysis that considers all retained parameters together  
15   with their probability distributions and determines the probability distribution for the  
16   output variable (biomass). The third step is a sensitivity analysis of the modeled  
17   spatial distribution of sugar cane yield to the model parameters for two large regions,  
18   in Brazil and Australia, at a spatial resolution of 0.7°. The sensitivity is established  
19   from the spatial distribution of ranked correlations between each parameter and yield  
20   in each grid point. Along the study steps, we address several problems inherent to  
21   uncertainty and sensitivity evaluation such as the determination of the uncertainty on  
22   the input parameters and the spatial (regional) differences of the sensitivity of the  
23   model to its key parameters.

### 24   **2.1   Sites and study areas**

25   This study is based on sugar cane field trials in three regions (figure 2) where sugar  
26   cane is of economical importance, Brazil (1 site), Australia (4 sites), and La Reunion  
27   Island (2 sites). These sites, already used by Valade et al. (2013) span different  
28   climatic conditions and agricultural practices, as shown in Table 1, which makes them  
29   useful for our purpose to provide continental-scale sugar cane yield uncertainty  
30   estimates. More details about the four sites from Australia and La Réunion can be  
31   found respectively in Keating et al. (1999); Muchow et al. (1994); Robertson et al.  
32   (1996) and in Martiné (unpublished). The site from Brazil is described in(Marin et al.,

1 2011). The sensitivity analysis of the yield spatial distribution to the model parameters  
2 is carried out for two continental-scale areas where sugar cane is cultivated at large  
3 scale. In Brazil, we consider the region encompassing partly the Sao Paulo and Mato  
4 Grosso states, and in Australia the sugar cane cultivation belt of the northeastern coast  
5 (Figure 2).

## 6 **2.2 Model & parameters considered**

7 We use the agro-Land Surface Model ORCHIDEE-STICS (Gervois et al., 2004) in a  
8 version that was already calibrated for sugar cane for Leaf Area Index at the same  
9 sites than used here (Valade et al., 2013). This model chains the crop model STICS  
10 with sugar cane specific phenology and management with the generic process-based  
11 land surface model ORCHIDEE that can be applied either at a site, or on a grid for  
12 regional runs.

13 STICS (Brisson et al., 1998) is an agronomical model designed for site-scale  
14 operational applications, which describes in details the soil and crop processes  
15 associated with specific crop varieties and with management practices, such as  
16 aboveground biomass, and biomass nitrogen content, water and nitrogen content in  
17 the soil, yield, root density. Yet, STICS is a generic crop model, because from a set of  
18 common equations it can describe a large number of crop species through specific  
19 parameterizations. Similarly, specific vectors of parameters define crop cultivars.  
20 STICS has been validated for a variety of cropping situations (Brisson et al., 2003)

21 ORCHIDEE (Krinner et al., 2005) is a land surface model developed for global  
22 applications, standing now as the land surface model of the IPSL Earth System  
23 Model. It has been developed from the association of a surface energy and water  
24 balance scheme (SECHIBA) with a biogeochemistry module (STOMATE) and as  
25 such simulates the short time scale exchanges of water and energy between the land  
26 surface and the atmosphere, as well as the processes of the carbon cycle including  
27 photosynthesis, respiration, carbon allocation, soil decomposition. The vegetation is  
28 represented in ORCHIDEE with the Plant Functional Type (PFT) concept, by  
29 grouping species into a few categories based on the similarities of their traits and  
30 resulting in an average plant. For example, sugar cane would fall in the generic 'C4  
31 crop' PFT in the standard version of ORCHIDEE, and this un-calibrated version of  
32 model fails to reproduce site-level phenology, as shown by Valade et al. (2013)

1 The chaining of STICS with ORCHIDEE was performed to improve the ability of  
2 ORCHIDEE to simulate specific crops, for which the PFT concept was not  
3 appropriate, as it lacks representation of crop phenology and crop management  
4 practices (Gervois et al., 2004). In the chain-like structure (Figure 3), STICS  
5 calculates phenology, water and nitrogen requirements, and passes the key variables  
6 of Leaf Area Index (LAI), root profile and nitrogen stress as well as the input data  
7 concerning irrigation requirements to ORCHIDEE that uses them to calculate carbon  
8 assimilation and allocation, water balance, and energy-related variables. The one-way  
9 coupling between the two models can generate some inconsistencies, such as the soil  
10 status that is different between ORCHIDEE and STICS. This type of inconsistencies,  
11 inherent to the structure of the model is considered as part of the structural uncertainty  
12 and is not covered in this study. However, this particular one-way structure will have  
13 a consequence in the uncertainty that we are analyzing in this study.

14 ORCHIDEE and STICS each have a large number of parameters involved at every  
15 step of a simulation over the course of a growing season. The values of these  
16 parameters - often empirically prescribed - are not easy to measure or are not  
17 measurable at all, calling in many cases for expert judgment to set their values, when  
18 it is impractical to find reference values. The uncertainty of these parameters is  
19 propagated onto the output variables of ORCHIDEE STICS and has impacts which  
20 strength depends on the structure of both STICS and ORCHIDEE. Because of the  
21 chain-type structure of ORCHIDEE-STICS (fig.3), the parameters from STICS that  
22 control LAI and nitrogen stress are expected to have a weaker and more indirect effect  
23 on downstream variables such as biomass compared with parameters from  
24 ORCHIDEE that directly control carbon assimilation processes and the development  
25 of biomass to produce yield at the date of harvest.

### 26 **2.3 Parameter screening**

27 In this section, we describe the screening step that allows us to select the most  
28 influential parameters upon which the model uncertainty is investigated. An initial set  
29 of 17 parameters from ORCHIDEE and 50 parameters from STICS is considered for  
30 the screening, according to their influence on the simulation of biomass production,  
31 based on expert knowledge and literature as listed in Table 2. The screening analysis  
32 procedure is the same as described in (Valade et al., 2013). It is based upon the

1 method of Morris (Campolongo et al., 2007; Morris, 1991; Pujol, 2009) often used to  
2 explore the parameters space for complex models with a large number of parameters.  
3 Like all screening methods, the Morris method gives qualitative information on the  
4 sensitivity of the output variables to the parameters, since it only discriminates  
5 parameters based on their importance, but does not provide information on the relative  
6 difference of importance (Cariboni et al., 2007). Its aim is to reduce the  
7 dimensionality of the problem for further use of quantitative, computationally heavier  
8 methods (Saltelli et al., 2004).

9 The advantage of the Morris method is that it is computationally efficient and easy to  
10 implement and interpret. It is based on a one-at-a-time approach, in which only one  
11 parameter is changed between two runs, allowing for the calculation of a local partial  
12 derivative of the output variable with respect to the input parameter, called an  
13 elementary effect. The Morris method is considered to be a “global” screening  
14 method, because the algorithm is repeated several times to calculate the elementary  
15 effects of each parameter in several locations of the parameters space so that the  
16 average and standard deviation of all elementary effects associated with each  
17 parameter are representative of the behavior of this parameter in its whole range of  
18 variation. The results of the Morris screening algorithm can be represented by a 2-D  
19 plot of standard deviation versus mean value of the elementary effects on the output  
20 variable (here harvested biomass) of each parameter. A parameter with a high mean  
21 elementary effect (called  $\mu$ , or  $\mu^*$  for mean of absolute values) is interpreted as a  
22 parameter with high influence on the output harvested biomass variable. A parameter  
23 with a high standard deviation of its elementary effects ( $\sigma$ ) is interpreted as inducing  
24 non-linearities in the model output, and/or as having interactions with other  
25 parameters.

26 Here, we apply the Morris method as implemented in the R 'sensitivity' package  
27 (Pujol et al., 2013) using site-scale simulations of ORCHIDEE STICS across the 7  
28 field trial sites listed in Table 1. For each site, we identify the most influential  
29 parameters for the output variable harvested biomass. The parameters identified as  
30 important at least at two sites are selected for the rest of the study.

31

## 1 **2.4 Uncertainty analysis (UA)**

2 The goal of the UA is to quantify the overall uncertainty in the harvested biomass  
3 output variable that results from uncertain input parameter values. Firstly, based on  
4 the a priori probability of each parameter's value, a Probability Density Function is  
5 assigned to each parameter in order to generate sample parameter sets according to the  
6 Latin Hypercube Sampling (LHS) method. Secondly, an ensemble of model runs is  
7 performed using those samples. Thirdly, the uncertainty on the output variables is  
8 obtained from the statistical properties of the distribution of simulated harvested  
9 biomass from the ensemble runs by defining the uncertainty as one standard deviation  
10 of the distribution.

11 The first step is thus to generate parameters samples constrained with prior parameters  
12 ranges and statistical distributions that are then used as inputs for ensemble  
13 simulations.

14 The parameters considered for the uncertainty (UA) for both STICS and ORCHIDEE  
15 are those selected by the screening analysis, allowing a reduction in the parameters  
16 space hypercube dimensionality and therefore in the required computing resources.  
17 Starting from the initial set of 17 and 50 parameters respectively for the screening of  
18 ORCHIDEE and STICS parameters, the Morris algorithm result (see Section 3.1)  
19 allows us to reduce the parameter numbers to 8 and 3 parameters for ORCHIDEE and  
20 STICS, respectively.

21 For the UA, we use Monte-Carlo methods, which are less computationally expensive  
22 than variance-based approaches (Marino et al., 2008), making them a frequent choice  
23 in environmental sciences (Poulter et al., 2010; Verbeeck et al., 2006; Zaehle et al.,  
24 2005). The Monte-Carlo sampling scheme used here is the stratified Latin Hypercube  
25 Sampling (LHS), which is an efficient scheme for generation of multivariate samples  
26 of statistical distributions (McKay et al., 1979) In LHS, the range of each of the  $k$   
27 parameters  $X_1, X_2, \dots, X_k$  included in the study is divided into  $N$  intervals of equal  
28 probability. One value is randomly selected from each interval. The  $N$  values obtained  
29 for the  $X_1$  parameter are then paired at random, without replacement, with the  $N$   
30 values obtained for the  $X_2$  parameter, then to the  $N$  values obtained for the  $X_3$   
31 parameter and so on until the  $k^{\text{th}}$  parameter. The procedure results in  $N$  sets of  $k$   
32 parameters, or samples, that can be used for input to the model. In this study, from the

1 11 parameters identified by the screening, the N value is set to 250 resulting in 250  
2 simulations for exploring the uncertainty around modeled biomass for each site.

3

4 In order to get insights on the part of the uncertainty attributable to each of the two  
5 models chained together, STICS and ORCHIDEE (fig.1), first, only the uncertainty  
6 coming from ORCHIDEE parameters is evaluated (fig.1), secondly, only the  
7 uncertainty propagated from STICS parameters (fig.1), and last, uncertainties  
8 propagated from both ORCHIDEE and STICS parameters are considered together  
9 through the chained model ORCHIDEE-STICS.

10 An important difficulty in the utilization of sampling-based UA methods is the lack of  
11 literature about a priori probability distribution of most parameters, given the  
12 dependency of output upon a priori assigned values (Marino et al., 2008) If most  
13 studies rely on a thorough literature search and expert judgment (Medlyn et al., 2005;  
14 Verbeeck et al., 2006; Wang et al., 2005), this approach might result in an  
15 overestimation of the model output uncertainty due to combinations of extreme  
16 parameters values that are not realistic and therefore excessively decrease the  
17 estimated reliability of the models. Some studies have addressed this issue by trying  
18 to rationalize the parameters ranges through benchmarking outputs (removing  
19 parameter sets resulting in values for output variables outside of a given benchmark  
20 range) or by prescribing hypothesized correlations between parameters (Poulter et al.,  
21 2010; Zaehle et al., 2005). Here, after a first estimation of uncertainty based on expert  
22 opinion for the a priori parameters range (overestimation of uncertainty), we propose  
23 a second approach to overcome the scarcity of information about parameters reference  
24 distributions by reducing the parameters a priori range based on site-optimized values,  
25 thus providing narrower and more realistic a priori ranges that are constrained by  
26 observations (likely underestimation of uncertainty).

27 For the first a priori estimation of parameters range, ranges and distributions are  
28 assigned to parameters based on expert knowledge and previous parameterization  
29 studies (Kuppel et al., 2012) and centered on their a priori values. The *a priori* ranges  
30 prescribed using this approach are considered as overestimations of the likely ranges  
31 for parameters' values for sugar cane because they are adapted from studies in which  
32 parameters' ranges were assigned for plant functional types instead of a single crop as

1 is the case here and sometimes used for optimization studies therefore requiring wide  
2 enough ranges within the model's domain of applicability (Groenendijk et al., 2011;  
3 Kuppel et al., 2012). By using overestimated ranges for input parameters, we estimate  
4 an upper bound for the value of the uncertainty on output variables.

5 The second (site-constrained) a priori estimation is a refinement of the uncertainty  
6 estimation based on the idea that the 'real' probability distribution of the parameters  
7 can be approached by the distribution of optimal parameters over all the possible case  
8 studies (sites, weather, management). It is of course not possible to determine the  
9 model's optimal parameters for an infinite number of eco-climatic and land-  
10 management conditions, but a sample of representative case studies can provide a  
11 rough estimate of the parameters plausible range. Building on this hypothesis, the  
12 model is calibrated independently at 7 sites using an iterative method, seeking to  
13 constrain the uncertainty analysis with observation-based parameters ranges. For this,  
14 we performed a Bayesian calibration of the model parameters, using a standard  
15 variational method based on the iterative minimization of a cost function that  
16 measures both the model data misfit as well as the parameters' deviations from a prior  
17 knowledge. The iterative scheme is described in (Tarantola, 1987) with the hypothesis  
18 of Gaussian error on the observations and the parameters. At each site, parameter  
19 values are varied iteratively until the best match between simulation and observation  
20 is found. More details on the calibration results can be found in the Supporting  
21 Information. We are aware that the optimization of the parameters at 7 sites only to  
22 obtain a representative a priori range of the parameters distributions likely results into  
23 an optimistic estimate of this range even though the sites chosen cover different  
24 climatic, edaphic and management conditions making them well suited for applying  
25 our method.

26 For both a priori parameters range estimations (expert judgment vs. site constrained),  
27 when no parameter value appears to be more likely than another, a uniform *a priori*  
28 uncertainty distribution is prescribed. When there is some level of confidence that the  
29 a priori value is more likely, we use a beta distribution. This type of distribution is  
30 often used for uncertainty analyses, because of its adjustable shape (parameterized  
31 equation) yet having the advantage of bounded tails (Monod et al., 2006; Wyss and  
32 Jorgensen, 1998). The successive analysis of both techniques provides an  
33 improvement in the estimation of the uncertainty from the first (expert-judgment

1 based, likely too pessimistic) to the second (observation-based, perhaps too  
2 optimistic) approach.

3

#### 4 **2.5 Spatial sensitivity analysis (SA)**

5 The first step in the sensitivity analysis also consists in generating parameters  
6 samples. The same parameters are considered for the SA as for the UA (section 2.4),  
7 i.e. the 11 parameters (8 parameters from ORCHIDEE and 3 parameters from STICS)  
8 selected by the screening analysis.

9 As opposed to the UA where all parameters are considered together for their effect on  
10 the distribution of the harvested biomass output variable, the goal of the sensitivity  
11 analysis is to rank the influence of parameters based on their impact on the biomass  
12 and its spatial distribution obtained in the continental-scale 0.7° runs. The partial  
13 correlation coefficient (PCC) measures the correlation between an output variable and  
14 a parameter after the correlation with other parameters has been eliminated (Marino et  
15 al., 2008). However, for monotonic but non-linear relationships, these measures  
16 perform poorly and a rank transformation needs to be applied to the data first to  
17 linearize the relationship. The correlation calculated between the rank-transformed  
18 data is then called partial rank correlation coefficients (PRCC). PRCC has been found  
19 to be an efficient indicator for the influence of parameters, because it is a measure of  
20 the sensitivity of the output to parameters (Saltelli and Marivoet, 1990). The larger the  
21 PRCC, the more important the parameter is with respect to the output variable. Here,  
22 the relationship between modeled biomass on a grid, and parameters is diagnosed  
23 through the calculation of the Partial Ranked Correlation Coefficients (PRCC) on  
24 each grid point between the output and parameter assuming a monotonic behavior of  
25 the model.

26 The SA is implemented from the results of the 0.7° simulations over Brazil and  
27 Australia (see fig.1 and section 3.5). In this regional sensitivity analysis, ORCHIDEE-  
28 STICS is run for each region on a grid of 20 by 15 grid points and 13 by 20 grid  
29 points respectively, driven by gridded climate forcing fields from the reanalysis  
30 products ERA-Interim (Dee et al., 2011), with varying parameter values from a  
31 sampling where only bounds and no distributions were assigned to the parameters.  
32 The management information (date of planting, date of harvest, fertilization,

1 irrigation) and the soil properties (as described in Valade et al. (2013)) are assumed to  
2 be uniform across each region and were defined as typical of each area. The a priori  
3 bounds used for the parameters in the SA correspond to the first version of the  
4 parameters ranges considered in the uncertainty analysis (i.e. derived from expert  
5 knowledge). As cited by Wang et al. (2005), for sensitivity analyses, Bouman (1994)  
6 advises to use parameters ranges as broad as possible within the limits of the model  
7 validity domain. Once the parameters' a priori bounds have been set, ensemble runs  
8 are performed with all the parameter sets. From the distributions of input parameters  
9 and output variables obtained at each pixel, a spatial distribution of PRCC is obtained,  
10 which is interpreted in section 3.5 in terms of regional differences of each parameter  
11 on modeled sugar cane yield.

12 The interest of carrying out such a regional sensitivity analysis is that it provides maps  
13 of the geographic patterns of the importance of each parameter, leading to a better  
14 comprehension of the mechanisms behind the parameter-related model sensitivity.  
15 These results can be very useful for planning purposes, for instance to quantify what  
16 are the different factors that control sugar cane yield and ethanol production over a  
17 large region under future climatic conditions as compared to present-day conditions.

## 18 **3 Results and discussion**

19

### 20 **3.1 Screening**

21 From the Morris screening method, we obtain for each parameter two indices  $\mu^*$  and  
22  $\sigma$ , that measure the influence of each parameter and its degree of involvement in non-  
23 linearities and interactions with other parameters, respectively. We first made sure  
24 that no parameter with a significant value for  $\mu^*$  was above the line  $\sigma=2\mu^*$  which  
25 would imply that non-linearities and/or interactions would be so strong that the  
26 uncertainty propagation from the parameter to the model output could not be clearly  
27 established. None of our parameters selected for their significant values of  $\mu^*$  was  
28 above this line (Supporting information figure 2). From  $\mu^*$  and  $\sigma$  values, we establish  
29 a ranking of the parameters by only considering parameters involved in limited  
30 interactions and/or non-linearities ( $\sigma<2\mu^*$ ) and then we rank the remaining parameters  
31 based on their  $\mu^*$  index, a larger  $\mu^*$  being interpreted as a more influential parameter.

1 The Morris parameters ranks for ORCHIDEE and STICS are respectively shown in  
2 Figure 5a and 5b where each radar plot corresponds to one model. The axes refer to  
3 the parameters and the line colors to the sites. For STICS, for the sake of readability,  
4 not all of the initially selected 50 parameters are represented on the radar plot but only  
5 those parameters that pertain to the 10 top-ranked parameters at least at one site. The  
6 maximum number of 10 parameters was fixed based on examination of Morris indices  
7  $\mu^*$  and  $\sigma$  at individual sites that only revealed 3 to 5 sensitive parameters each time.  
8 The positions and roles in the model of the parameters identified as most important  
9 are shown in Figure 3. Figure 4 gives more details, with the main equations through  
10 which these parameters affect the output variables of STICS and of ORCHIDEE.

11

12 The 3 most influential parameters of STICS (fig.3a) reflect the way STICS and  
13 ORCHIDEE are chained (fig.3). Indeed, from the chained model structure, the  
14 indirect impact of STICS parameters on harvested biomass occurs through their effect  
15 on processes related to LAI, root growth and nitrogen stress, the only STICS variables  
16 passed to ORCHIDEE for calculating biomass. This chaining of the models through  
17 three variables is reflected in the identification of the 3 most important STICS  
18 parameters, which control the daily maximum rate of foliage production  $\delta_{LAI}^{max}$ , the  
19 growth rate of the root front,  $\kappa_{root}$  and the threshold of nitrogen nutrition index  
20  $INN_{min}$ .  $\delta_{LAI}^{max}$  and  $INN_{min}$  parameters are both involved in LAI calculation. Indeed,  
21 the LAI equation has four members describing four processes of the sugar cane  
22 foliage development. First, the LAI-development ( $\Delta_{LAI}^{dev}$  in fig.4) describes the  
23 potential LAI increase through the scaling of the daily maximum rate of foliage  
24 production by a function of the development stage ( $k_{LAI}$ ), and is logically directly  
25 controlled by the value of parameter  $\delta_{LAI}^{max}$ . The second member in equation (\*)  
26 represents the temperature effect on LAI growth through the accumulation of degrees  
27 above a temperature threshold ( $T_{min}$  in fig.3). The last two members of the equation  
28 represent processes that can limit LAI development, competition for light between  
29 plants due to planting density ( $\Delta_{LAI}^{dens}$  in fig.4) and a limitation from trophic stress  
30 emerging from competition between plant components for nitrogen based in  
31 calculation of a nitrogen nutrition index limited by parameter  $INN_{min}$ . The root  
32 growth rate  $\kappa_{root}$  has a less direct impact on LAI since it intervenes in the calculation

1 of the root front depth, which then impacts the availability of nitrogen and water and  
2 therefore the stress status of the crop (impact on  $C_N^{plant}$  and  $W_s$  in fig.4).

3

4 The 8 most influential parameters that control harvested biomass in ORCHIDEE, are  
5 identical for all sites except at the Colimaçons site (where only 7 parameters are  
6 identified as influential by the Morris method). The Morris top ranked parameters of  
7 ORCHIDEE control photosynthesis and water budget equations as well as respiration  
8 processes (fig.4). Three of those (the minimum and optimal temperatures for  
9 photosynthesis,  $T_{min}$ ,  $T_{opt}$ , the maximum rate of carboxylation  $V_{Cmax}^{opt}$ ) affect directly  
10 the rate of carboxylation  $V_c$  that is calculated from the maximum rate of carboxylation  
11 weighted by a mean leaf efficiency and scaled by a limiting factor depending on the  
12 optimum and minimum temperatures for photosynthesis. The stomatal conductance  
13  $g_s$  that links assimilation and transpiration is defined by the Ball-Berry equation (Ball  
14 et al., 1987) as a function of assimilation and depends on the air relative humidity and  
15  $CO_2$  concentration, scaled by a slope factor, called the Ball-Berry slope ( $\beta$ ). The root  
16 profile constant ( $\kappa_{hum}$ ) describes the exponential distribution of root density in the  
17 soil and is involved in the definition of available water and root temperature. Finally,  
18 the extinction coefficient ( $k_{ext}$ ) intervenes in an equation derived by Monsi and Saeki  
19 (1953), similar to Beer's law, which describes the attenuation of light with depth in  
20 the canopy.

21

22 Two ORCHIDEE parameters controlling autotrophic respiration also stand out, with  
23 the maintenance respiration coefficient ( $\alpha_{Mresp}$ ) and the fraction of biomass allocated  
24 to growth respiration ( $f_{Gresp}$ ). The  $leaf_{age}^{crit}$  parameter that is involved in the biomass  
25 allocation also ranked high (5<sup>th</sup> most important) but only for one site and is therefore  
26 not retained for the rest of the study.

27

28 For the chained model STICS-ORCHIDEE, the 11 most influential parameters show a  
29 good agreement between sites for the most important parameters as seen on fig.5  
30 where ranking lines overlap for most of the parameters. Building on the results of the  
31 Morris screening analysis, we select the 8 top ranked parameters for ORCHIDEE and

1 3 for STICS that were revealed as influential for biomass for further uncertainty and  
2 sensitivity analysis.

3

4

5

6

### 7 **3.2 Uncertainty analysis: Parameters controlling biomass**

#### 8 ***uncertainty at a typical site***

9 In this section, we attribute the harvested biomass uncertainty to the uncertainty of the  
10 ORCHIDEE vs. STICS parameters. The simulated biomass uncertainty is a function  
11 of time during the growing season, and it differs between sites. In Figure 6, we show  
12 the contributions of ORCHIDEE and STICS parameters respectively to the total  
13 uncertainty for one typical site, Grafton, Australia, during the 1994-95 growing  
14 season, which has climate conditions within the range of other sites. Fig.6 a-c displays  
15 the normalized frequency distributions of simulated biomass obtained from ensemble  
16 runs for three times in the growing season: 1) very early in the cycle in fig.6a, at 100  
17 days after planting (DAP), 2) during the peak growing season in fig.6b, at 200 DAP  
18 and 3) short before harvest in fig.6c, at 350 DAP. We distinguish between the  
19 normalized frequency distributions of simulated biomass when considering the  
20 uncertainty propagated from STICS parameters alone (green), ORCHIDEE  
21 parameters alone (yellow), and from ORCHIDEE and STICS parameters together  
22 (brown), along with their best-fit normal distributions overlaid. These distributions  
23 were obtained by Monte Carlo LHS ensemble runs (section 2.4) with a sampling of  
24 parameters of STICS alone, ORCHIDEE alone and of both models together. We  
25 consider uncertainties starting from the time when biomass reaches  $50 \text{ gC.m}^{-2}$  in order  
26 to discard the emergence phase during which biomass is very low and uncertainties  
27 are therefore not significant.

28

29 At 100 DAP (Fig 6a), the uncertainty distribution of biomass related to ORCHIDEE  
30 parameters U(O) spans a slightly larger range than the distribution related to STICS,  
31 U(S), and it has more extreme values. The U(O) distribution is symmetrical around

1 the mean value, with a standard deviation of 86.9 gC.m<sup>-2</sup>. The U(S) distribution is  
2 non-symmetric, skewed towards larger values of biomass, and it has a slightly smaller  
3 standard deviation (76.5 gC.m<sup>-2</sup>) than that of U(O). Combining U(O) and U(S) in  
4 Monte Carlo runs by varying the parameters of both models at the same time gives the  
5 total uncertainty distribution, U(O+S), shown in brown in fig.6. This distribution has  
6 more extreme values and a higher standard deviation (112.0 gC.m<sup>-2</sup>), i.e. U(O+S) >  
7 U(O) + U(S).

8

9 At 200 DAP (Fig 6b), and later at 350 DAP (Fig 6c), the picture has changed. First, all  
10 uncertainties distributions are wider than at 100 DAP. Secondly, the means of U(O)  
11 and U(S) are no longer in agreement, with the asymmetric U(S) distribution being  
12 even more shifted towards high values of the harvested biomass. The reason for this  
13 shift is that among the variables transmitted from STICS to ORCHIDEE in the chain  
14 of models, the only one that can act to increase the biomass calculated by ORCHIDEE  
15 in the later phase of the growing season, near 350 DAP, is LAI. This is because a  
16 higher LAI will result into increased photosynthesis and therefore biomass in  
17 ORCHIDEE. However, passed a certain threshold, the LAI impact saturates when the  
18 foliage is sufficient for all incoming light to be captured, and therefore, uncertainty on  
19 the STICS parameters that impact LAI will not increase the uncertainty of biomass  
20 any longer. Unlike LAI, the nitrogen stress and root profile variables controlled by the  
21 parameters of STICS continue to act as limiting factors on biomass throughout the  
22 peak and late growing season. The saturation of the biomass uncertainty associated  
23 with STICS parameters is stronger at 200 DAP than at 300 DAP, when biomass  
24 increase has slowed down and the role of LAI for driving biomass is less important.

25

26 On fig.6d, the total uncertainty U(O+S) is given for the reference simulation (with  
27 parameters at their maximum likelihood values, red line) and the uncertainty on  
28 harvested biomass can be defined as a percentage of the harvested biomass in the  
29 reference simulation. For the Grafton site, at harvest, the overall uncertainty is 26%.  
30 The relative contributions of ORCHIDEE and STICS to the total uncertainty,  $\alpha_O$  and  
31  $\alpha_S$  respectively, are defined by  $\alpha_O = \frac{U(O)}{U(O+S)}$ ,  $\alpha_S = \frac{U(S)}{U(O+S)}$ . The evolution of these  
32 contributions to the total uncertainty is shown in fig.6e. We can see in this example

1 that  $U(O) > U(S)$  during the entire growing season, but with a decrease of  $U(S)$ , and  
2 an increase of  $U(O)$  such that the increase in biomass uncertainty seen on fig.6d  
3 becomes increasingly dominated by uncertain ORCHIDEE parameters. The  
4 progressive increase in the weight of ORCHIDEE parameters uncertainties is due to  
5 the reduction in the role played by LAI for biomass increase along the growing  
6 season. Indeed, if early in the season the foliage is crucial to allow photosynthesis and  
7 carbon allocation, later in the cycle, other processes become important as well and  
8 passed a certain LAI for which all incoming light is captured, it might not even play a  
9 role anymore and then the STICS parameters only impact biomass accumulation  
10 through nitrogen stress index and root depth.

11

### 12 **3.3 Uncertainty analysis: role of ORCHIDEE vs. STICS parameters in** 13 **controlling biomass uncertainty at 7 sites**

14 Table 3 summarizes the results of the overall parametric uncertainty analysis at the 7  
15 sites, including Grafton. The total uncertainty  $U(O+S)$  ranges between 25.5% of  
16 biomass at Piracicaba, Brazil during 2004-05 and 44.26% of harvested biomass at  
17 Tirano, La Réunion in 1998-99 yielding an average uncertainty on biomass at harvest  
18 due to uncertain parameter values of the chained model ORCHIDEE-STICS of 34.0%  
19 of harvested biomass across the 7 sites, in the order of previous results on different  
20 variables in similar studies using process-based models such as (Dufrêne et al., 2005)  
21 who found an uncertainty of 30% on modeled NEE for a forest sites in France with  
22 the CASTANEA model.

23

24 As for the ORCHIDEE vs. STICS relative contributions to the uncertainty of  
25 simulated biomass at all sites, the results at each site are not identical but display a  
26 similar general pattern shown by figure 7. For all sites, the ORCHIDEE parameters  
27 contribution to total uncertainty increases during the cycle, or remains approximately  
28 constant for Ingham in 1992-93, and increases during the growing cycle to dominate  
29 entirely the total uncertainty at the end of the cycle compared to STICS parameters.  
30 The STICS contribution to overall uncertainty decreases during the growing season to  
31 reach a minimum by the end of the growing season. For sites Piracicaba during 2004-  
32 05, Tirano in 1998-99 and Colimaçons during 1994-95, during the beginning of the

1 cycle the U(S) is even larger than U(O). The results for Ayr in 1991-92 display a less  
2 clear pattern. Indeed, at the end of the cycle, the contributions of ORCHIDEE and  
3 STICS to the total uncertainty are almost equal, due to an increase in STICS  
4 contribution during the second half of the cycle. This result confirms a hypothesis  
5 made in Valade et al. (2013) where the difficult calibration of LAI at this site was  
6 attributed to the simulation by STICS of an important stress. Indeed if a large stress is  
7 simulated by the phenological module, this can impede ORCHIDEE processes of  
8 biomass growth and therefore increases the weight of STICS parameters with respect  
9 to ORCHIDEE ones.

10

### 11 **3.4 Uncertainty analysis: constraining uncertainty from sites** 12 **optimization**

13 Optimizing the 11 ORCHIDEE-STICS parameters selected from the screening  
14 analysis at 7 sites leads to a reduction of the width of the a priori uncertainty  
15 distribution of the parameters (Table 2). Carrying out the same uncertainty analysis  
16 with a narrower uncertainty range of parameters (thanks to their site calibration) leads  
17 to an important reduction of uncertainties of biomass both for the STICS and  
18 ORCHIDEE components of uncertainty. This can be seen by comparing Figure 6  
19 (initial range of parameters) with figure 8 (narrower range after parameters calibration  
20 at the sites). For site Grafton during 1994-95 for example, U(O+S) gets reduced from  
21 26% to 17% of the reference harvested biomass, U(O) from 24% to 15% and U(S)  
22 from 14% to 10%. Figure 9 and Table 3 (bottom section) show the uncertainty  
23 contributions and overall uncertainty estimates for the 7 sites after observation-based  
24 reduction of the a priori uncertainty on parameters. The overall parametric uncertainty  
25 of biomass defined as the 1-sigma standard deviation of the (O+S) distribution has  
26 thus been reduced to 21% in average, to 11.48% when attributed to STICS alone, and  
27 to 17.15% when attributed to ORCHIDEE alone, (Table 3).

28

29 The ORCHIDEE and STICS contributions to the total uncertainty keep the same  
30 general pattern as with the initial parameters uncertainty distribution, with a  
31 domination of ORCHIDEE parameters in the uncertainty towards the end of the  
32 growing season (fig.9). Compared with the first uncertainty budget with expert-based

1 parameters uncertainties (fig.8), there is generally a slight decrease in the STICS  
2 contribution at the end of the season.

3

4 We have thus established full uncertainty budgets for the two components of the  
5 ORCHIDEE-STICS chain of models, which has revealed variations in the uncertainty  
6 in the biomass simulation from site to site. The next step is to discriminate between  
7 the different parameters the ones that contribute most to the overall uncertainty  
8 through a sensitivity analysis at regional scale.

### 9 **3.5 Spatial sensitivity analysis: sensitivity of sugar cane yields to the** 10 **model parameters for Brazil and Australia**

11 The overall parametric uncertainties have been quantified at 7 sites and attributed to  
12 either STICS or ORCHIDEE. The sensitivity analysis (SA) in this section will go a  
13 step further and leads to discriminate the different parameters that contribute to the  
14 spatial distribution of uncertainty over the two regions considered. This sensitivity  
15 analysis is performed at regional scale because from the previous section, we have  
16 seen that the uncertainty in the biomass simulation varies from site to site.

17

18 Ensemble runs at regional scale were realized over Brazil and Australia each with  
19 different value combinations for the 11 parameters previously selected through the  
20 Morris screening analysis (Table 1). The Partial Rank Correlation Coefficients  
21 (PRCC) were then calculated for each pixel in each of the two regions (see section  
22 2.5), and the SA results are discussed for two dates during the growing season, 200  
23 and 350 days after planting (DAP). The SA results express the strength of the  
24 relationship between an uncertain parameter and the simulated biomass at harvest at  
25 each pixel. The statistical significance of the PRCC calculated for each grid cell is  
26 tested with the associated p-values, and non-significant PRCC are removed (p-  
27 value<0.05). The first date 100 DAP examined for site scale UA studies (section 2.3)  
28 is not shown here, because no statistical significance was found in the correlations  
29 between the parameters and the harvested biomass at 100 DAP. Then, the pixels  
30 statistically significant PRCC calculated for each parameter can be analyzed both in a  
31 geographical projection (latitude, longitude) (fig. 11 & 12, columns 1-2 and 4-5) and

1 in a (Temperature, Precipitation) climatic space projection (fig 11 & 12, columns 3  
2 and 6). The regional sensitivity analysis thus carried out for sugar cane growing areas  
3 in Brazil and Australia shows the magnitude, spatial distribution and climatic  
4 dependency of the sensitivity of harvested biomass to the 11 parameters previously  
5 selected through the Morris screening analysis (Table 2).

6

7 Across both regions in Brazil and Australia, we find that the sensitivity of biomass to  
8 the model parameters is not uniformly distributed. This means that the simulated yield  
9 depends on different parameters within different parts of the same region. This result  
10 shows that applying a model at one site to determine the most important parameters,  
11 and generalizing its conclusion across a region generates biased conclusions.  
12 Considering only the first most important parameter in each pixel (fig. 10), we can see  
13 that early in the cycle (200 DAP, Figure 10a) four parameters dominate the spatial  
14 distribution of the U(O+S) uncertainty of biomass at 200 DAP, both over Brazil and  
15 Australia. These parameters are three ORCHIDEE parameters involved in the  
16 photosynthesis process, the minimum and optimum temperature for photosynthesis  
17  $T_{min}$ ,  $T_{opt}$ , and the maximum rate of carboxylation  $V_{Cmax}^{opt}$ , and one parameter from  
18 STICS  $\delta_{LAI}^{max}$ , defining the maximum rate of increase of LAI and only appearing in  
19 the Australian region. In Brazil, the parameter  $V_{Cmax}^{opt}$  is the first most important  
20 parameter for 93% of the area, whereas the optimum and minimum photosynthesis  
21 temperatures parameters only dominate in respectively 3 and 4% of the area. In  
22 Australia, the parameters' domination is more balanced with 37.5% for each of  
23  $V_{Cmax}^{opt}$  and  $\delta_{LAI}^{max}$  and 25% for  $T_{min}$ .

24 Later in the growing season (350DAP, fig.10b), consistently with the results of the  
25 site-scale uncertainty analysis, the influence of the STICS parameters decreases until  
26 no STICS parameters appear any longer as a dominant parameter in any of the  
27 regions. At this later stage in the season, two parameters stand out as explaining most  
28 of the uncertainty in most pixels of both regions,  $V_{Cmax}^{opt}$  and  $T_{min}$ . In Brazil,  $V_{Cmax}^{opt}$  is  
29 still the most sensitive parameter for most of the region, but  $T_{opt}$  disappeared and the  
30 area dominated by  $T_{min}$  expanded and now covers the cooler area of the southeast  
31 coastal zone, which is likely to result from the growing calendar of sugarcane in  
32 Brazil since the later part of the growing season takes place during winter in this

1 region. In Australia, the area dominated by  $V_{Cmax}^{opt}$  expanded into most of the region  
2 and now covers 83% of the area. In the coolest pixels, the soil-related parameters  
3 appear with the two root profile parameters from STICS and from ORCHIDEE,  $\kappa_{root}$   
4 and  $\kappa_{hum}$ .

5

6 Figures 11 and 12 focus on the values of the PRCC for each parameter as well as their  
7 spatial distribution. Their projection in a Temperature-Precipitation space for a given  
8 time (fig.11 for 200 DAP, fig.12 for 350 DAP) give more insights on the dependency  
9 of the sensitivity to the climatic conditions along the growing cycle. As an example,  
10 the sensitivity of the simulated biomass to  $T_{min}$  is highly sensitive to the average  
11 temperature of the location. At low-temperature sites, where temperature is a limiting  
12 factor for crop growth (below 17°C), the PRCC is higher than 0.8, whereas at high-  
13 temperature sites (above 22°C) the PRCC is below 0.3. Sites with temperatures above  
14 25°C do not even show significant correlations (grey symbols on the scatter plot).

15

16 For the parameter  $\kappa_{hum}$ , which describes the root profile of the cane (inverse of root  
17 depth), the dependency is most obvious on precipitation amount. For annual  
18 precipitations above 2500mm, no significant correlation is found.

19

20 Comparing the regional sensitivities at two times in the growing season shows again  
21 the decrease in the importance of STICS parameters whereas all of most important  
22 ORCHIDEE parameters have larger RPCC than earlier in the season.

23

## 24 **4 Concluding remarks**

25 In the perspective of applying spatially explicit mechanistic vegetation models such as  
26 ORCHIDEE-STICS to biofuel yield simulations we have sought the quantification  
27 and understanding of parametric uncertainty propagation in the model, both at site  
28 level and at sub-continental scale over two large regions, Australia and Brazil. For  
29 this, a rigorous analysis of the uncertainty budget of simulated sugar cane biomass has  
30 been established, using a step by step tracking of uncertainty in the model.

1 The main parameters from the two chain components of the model responsible for  
2 most of the uncertainty propagation have been identified through a Morris screening  
3 analysis. For the ORCHIDEE carbon, water and energy model, the most influential  
4 parameters are those involved in photosynthesis equations,  $T_{min}$ ,  $T_{opt}$ ,  $V_{Cmax}^{opt}$ , the  
5 radiation interception parameter  $k_{ext}$ , the root profile constant  $\kappa_{hum}$ , the parameters  
6 for respiration, slope of the Ball-Berry relation  $\beta$ , maintenance and growth  
7 respiration parameters  $f_{Gresp}$  and  $\alpha_{Mresp}$ . For the STICS model, the most influential  
8 parameters are those responsible for simulation of phenology, nitrogen and water  
9 stress. The parameters describing the maximum rate of carboxylation, the maximum  
10 growth rate of the root front and the threshold for nitrogen stress have been found to  
11 have the greatest role. The parameters identified are closely related to the structure of  
12 the coupling since the key variables transmitted from STICS to ORCHIDEE each  
13 convey one key parameter.

14 We used two approaches for estimating the total uncertainty propagated from the  
15 parameters into the model by assigning uncertainties on parameters with two methods,  
16 one 'pessimistic', in which a-priori parameter uncertainty bounds are set based on  
17 expert judgment, and one optimistic where smaller uncertainty is derived by an  
18 optimization of the model parameters at several sites thus providing a smaller,  
19 arguably more realistic, a-priori uncertainty range.

20 We found that all these parameters together contribute to an overall uncertainty of  
21 21% on sugar cane biomass simulations with an agro-LSM model and that this  
22 amount is variable among sites with different climatic, edaphic and management  
23 situations. We also analyzed this uncertainty separately for each component of the  
24 model and found that whatever estimate chosen for the parameters uncertainty, by the  
25 end of the growing season, the uncertainty propagated from the phenology module  
26 STICS decreases and the overall uncertainty is almost totally explained by the  
27 ORCHIDEE uncertainty. The lower uncertainty from STICS parameters compared to  
28 ORCHIDEE ones is likely related with the lower number of processes solved by  
29 STICS in its configuration with ORCHIDEE, and to some extent to the lower number  
30 of parameters propagating their uncertainties. The decrease in the weight of the  
31 STICS' parameters to the overall uncertainty is linked to the canopy closure (LAI  
32 sufficient to capture all incoming light) and would therefore probably happen at a  
33 different timing in the growing season for different crops. For example, soybean

1 experiences a later canopy closure and would probably show a later diminution of the  
2 STICS contribution to overall uncertainty, therefore remaining relatively high by the  
3 end of the cycle.

4 The overall origin of uncertainty has then been diagnosed in even more detail through  
5 a regional sensitivity analysis allowing the identification of the parameter for which  
6 harvested biomass is most sensitive for each pixel within regions of Australia and  
7 Brazil. We revealed a strong heterogeneity of the results based on climatic conditions  
8 and also variability in time that confirms the results of the uncertainty analysis, by  
9 showing a decrease in the importance of the STICS parameters along the growing  
10 season.

11 We believe that our results for the sugar cane crop simulated with the model  
12 ORCHIDEE-STICS are relevant to other agro-LSM with different crops. All these  
13 results prove the importance of establishing clear uncertainty budgets for highly  
14 parameterized models such as agro-LSM, especially when applying these models to  
15 answer questions related to political decisions such as biofuels burning topics.

16 As an example, combining our optimistic uncertainty estimation with the estimations  
17 from (Lapola et al., 2009) for irrigated sugar cane (obtained with the model LPJml,  
18 very similar to ORCHIDEE-STICS), we can evaluate the range assorted with their  
19 estimation of land requirements to fulfill the demand in ethanol in Brazil. Similarly to  
20 our study they use a multi-continental approach, focusing on Brazil and India. They  
21 simulate with a single parameterization the sugarcane productivity over both  
22 considered countries, spanning a wide range of climatic conditions. They found a  
23 mean yield of 68.8 t/ha over Brazil and 73.3 t/ha over India, and conclude that to  
24 fulfill government targets, the sugar cane areas would need to expand by 2.8 million  
25 hectares in Brazil and 1 million hectare in India. Because the yield estimates derived  
26 in (Lapola et al., 2009) are retrieved with an global agro-LSM parameterized for  
27 global applications and used in a range of climatic conditions (whole Brazil and  
28 India), we make the hypothesis that our uncertainty calculation is applicable to the  
29 LPJml results. We can then take into account the parametric uncertainty of the model  
30 and translate the potential mean production into a range of [54-83t/ha] for Brazil and  
31 [58-89t/ha] for India. The land requirements when including parameters uncertainty  
32 would then becomes [2.6– 3.9 million hectares],for Brazil and [0.9 – 1.4 million  
33 hectares] for India. To go further in the application of this result, and assuming that

1 sugar cane expansion results in deforestation through direct or indirect land use  
2 change, we can translate the land expansion of sugar cane for biofuels into carbon  
3 emissions from deforestation. Several estimates of carbon emissions associated with  
4 conversion of tropical forest to croplands have been published and their results span a  
5 large range revealing the large uncertainties in this area (BSI, 2008; Cederberg et al.,  
6 2011; Searchinger et al., 2008) Discussing the uncertainty on this estimate is beyond  
7 the scope of this paper so we will only consider the value from (Searchinger et al.,  
8 2008), of 604tCO<sub>2</sub>eq/ha. Using this conversion factor, the expansion of sugar cane  
9 calculated by (Lapola et al., 2009) would result in CO<sub>2</sub>eq emissions of 1,68GtCO<sub>2</sub>eq  
10 whereas including the parametric uncertainty of the model we obtain a range of 1,6 to  
11 2,4 GtCO<sub>2</sub>eq provoked by Brazilian government's ethanol targets with our calculation  
12 of uncertainty.

13 With the choice of the study from Lapola et al. (2009) to apply our uncertainty  
14 estimates on, we favored the closeness of the models over the full consistency of the  
15 methodologies. If the primary goal had been to calculate estimates of uncertainty of  
16 land requirements in the specific region of Brazil, we would have constrained our  
17 parameters ranges for conditions of this region, which would have resulted in lower  
18 uncertainty ranges for area requirements. However, we want to stress that agro-LSMs  
19 like ORCHIDEE-STICS or LPJml are designed for global studies and their  
20 parameters are therefore supposed to cover the full range of climatic conditions even  
21 when they are used for regional applications. This quick application of our uncertainty  
22 calculation proves how important it is to consider the uncertainty when addressing  
23 issues aimed at decision-makers.

## 24 **5 Acknowledgements**

25 This study was performed using HPC resources from GENCI- CCRT (Grant 2012-016328).  
26 We acknowledge Arnaud Caubel for his help with the experimental setup and the  
27 development of scripts for multiple runs and Patrick Brockmann for his help with data  
28 visualization.

29

30

31

1  
2  
3  
4  
5  
6  
7  
8  
9  
10  
11  
12  
13  
14  
15  
16  
17  
18  
19  
20  
21  
22  
23  
24  
25  
26  
27  
28  
29  
30  
31  
32  
33  
34  
35  
36  
37  
38  
39  
40  
41

## 6     **References**

Ball, J.T., Woodrow, I.E., Berry, J.A., (1987) A model predicting stomatal conductance and its contribution to the control of photosynthesis under different environmental conditions, *Progress in photosynthesis research*. Springer, pp. 221-224.

Black, E., Vidale, P.L., Verhoef, A., Cuadra, S.V., Osborne, T., Van den Hoof, C. (2012) Cultivating C4 crops in a changing climate: sugarcane in Ghana. *Environmental Research Letters* 7, 044027.

Bouman, B. (1994) A framework to deal with uncertainty in soil and management parameters in crop yield simulation: a case study for rice. *Agricultural Systems* 46, 1-17.

Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., Zimmer, D., Sierra, J., Bertuzzi, P., Burger, P. (2003) An overview of the crop model STICS. *European Journal of Agronomy* 18, 309-332.

Brisson, N., Mary, B., Ripoche, D., Jeuffroy, M.-H., Ruget, F., Nicoullaud, B., Gate, P., Devienne-Barret, F., Antonioletti, R., Durr, C., Richard, G., 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.

BSI (2008) PAS 2050: 2008 Specification for the assessment of the life cycle greenhouse gas emissions of goods and services. British Standards Institution.

Campolongo, F., Cariboni, J., Saltelli, A. (2007) An effective screening design for sensitivity analysis of large models. *Environmental Modelling & Software* 22, 1509-1518.

Cariboni, J., Gatelli, D., Liska, R., Saltelli, A. (2007) The role of sensitivity analysis in ecological modelling. *Ecological Modelling* 203, 167-182.

Cederberg, C., Persson, U.M., Neovius, K., Molander, S., Clift, R. (2011) Including carbon emissions from deforestation in the carbon footprint of Brazilian beef. *Environmental Science & Technology* 45, 1773-1779.

Cheeroo-Nayamuth, F.C., Robertson, M.J., Wegener, M.K., Nayamuth, A.R.H. (2000) Using a simulation model to assess potential and attainable sugar cane yield in Mauritius. *Field Crops Research* 66, 225-243.

1 Cuadra, S.V., Costa, M.H., Kucharik, C.J., Da Rocha, H.R., Tatsch, J.D., Inman-  
2 Bamber, G., Da Rocha, R.P., Leite, C.C., Cabral, O.M.R. (2012) A biophysical model  
3 of Sugarcane growth. *GCB Bioenergy* 4, 36-48.

4 de Vries, S.C., van de Ven, G.W.J., van Ittersum, M.K., Giller, K.E. (2010) Resource  
5 use efficiency and environmental performance of nine major biofuel crops,  
6 processed by first-generation conversion techniques. *Biomass and Bioenergy* 34,  
7 588-601.

8 Dee, D.P., Uppala, S.M., Simmons, A.J., Berrisford, P., Poli, P., Kobayashi, S., Andrae,  
9 U., Balmaseda, M.A., Balsamo, G., Bauer, P., Bechtold, P., Beljaars, A.C.M., van de  
10 Berg, L., Bidlot, J., Bormann, N., Delsol, C., Dragani, R., Fuentes, M., Geer, A.J.,  
11 Haimberger, L., Healy, S.B., Hersbach, H., Hólm, E.V., Isaksen, I., Kållberg, P.,  
12 Köhler, M., Matricardi, M., McNally, A.P., Monge-Sanz, B.M., Morcrette, J.J., Park,  
13 B.K., Peubey, C., de Rosnay, P., Tavolato, C., Thépaut, J.N., Vitart, F. (2011) The  
14 ERA-Interim reanalysis: configuration and performance of the data assimilation  
15 system. *Quarterly Journal of the Royal Meteorological Society* 137, 553-597.

16 Demirbas, A. (2008) Biofuels sources, biofuel policy, biofuel economy and global  
17 biofuel projections. *Energy Conversion and Management* 49, 2106-2116.

18 Doornbosch, R., Steenblik, R. Biofuels: Is the cure worse than the disease? *Revista*  
19 *Virtual REDESMA* 2, 63-100.

20 Dufrêne, E., Davi, H., François, C., Maire, G.I., Dantec, V.L., Granier, A. (2005)  
21 Modelling carbon and water cycles in a beech forest: Part I: Model description  
22 and uncertainty analysis on modelled NEE. *Ecological Modelling* 185, 407-436.

23 Fargione, J., Hill, J., Tilman, D., Polasky, S., Hawthorne, P. (2008) Land Clearing  
24 and the Biofuel Carbon Debt. *Science* 319, 1235-1238.

25 Georgescu, M., Lobell, D.B., Field, C.B., Mahalov, A. (2013) Simulated  
26 hydroclimatic impacts of projected Brazilian sugarcane expansion. *Geophysical*  
27 *Research Letters* 40, 972-977.

28 Gervois, S., Ciais, P., de Noblet-DucoudrÈ, N., Brisson, N., Vuichard, N., Viovy, N.  
29 (2008) Carbon and water balance of European croplands throughout the 20th  
30 century. *Global Biogeochem. Cycles* 22, GB2022.

31 Gervois, S., de Noblet-Ducoudre, N., Viovy, N., Ciais, P., Brisson, N., Seguin, B.,  
32 Perrier, A. (2004) Including Croplands in a Global Biosphere Model: Methodology  
33 and Evaluation at Specific Sites. *Earth Interactions* 8, 1-25.

34 Groenendijk, M., Dolman, A.J., van der Molen, M.K., Leuning, R., Arneeth, A.,  
35 Delpierre, N., Gash, J.H.C., Lindroth, A., Richardson, A.D., Verbeeck, H., Wohlfahrt,  
36 G. (2011) Assessing parameter variability in a photosynthesis model within and  
37 between plant functional types using global Fluxnet eddy covariance data.  
38 *Agricultural and Forest Meteorology* 151, 22-38.

39 Hill, J., Nelson, E., Tilman, D., Polasky, S., Tiffany, D. (2006) Environmental,  
40 economic, and energetic costs and benefits of biodiesel and ethanol biofuels.  
41 *Proceedings of the National Academy of Sciences* 103, 11206-11210.

42 Howarth, R.W., Bringezu, S., Martinelli, L.A., Santoro, R., Messer, D., Sala, O.E.  
43 (2008) Introduction: Biofuels and the Environment in the 21st Century. *Biofuels:*  
44 *environmental consequences and interactions with changing land use.*  
45 *Proceedings of the Scientific Committee on Problems of the Environment*  
46 *(SCOPE) International Biofuels Project Rapid Assessment*, 22-25.

47 IEA, (2004) *Biofuels for transport An international perspective.* International  
48 Energy Agency, p. 210.

49 IEA, (2011) *Technology Roadmap, Biofuels for Transport.*

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

4 Krinner, G., Viovy, N., de Noblet-Ducoudré, N., Ogée, J., Polcher, J., Friedlingstein,  
5 P., Ciais, P., Sitch, S., Prentice, I.C. (2005) A dynamic global vegetation model for  
6 studies of the coupled atmosphere-biosphere system. *Global Biogeochem. Cycles*  
7 19, GB1015.

8 Kuppel, S., Chevallier, F., Peylin, P. (2012) Quantifying the model structural error  
9 in carbon cycle data assimilation systems. *Geoscientific Model Development*  
10 *Discussions* 5, 2259-2288.

11 Lapola, D.M., Priess, J.A., Bondeau, A. (2009) Modeling the land requirements and  
12 potential productivity of sugarcane and jatropha in Brazil and India using the  
13 LPJmL dynamic global vegetation model. *Biomass and Bioenergy* 33, 1087-1095.

14 Loarie, S.R., Lobell, D.B., Asner, G.P., Mu, Q., Field, C.B. (2011) Direct impacts on  
15 local climate of sugar-cane expansion in Brazil. *Nature Clim. Change* 1, 105-109.

16 Macedo, I.C., Seabra, J.E.A., Silva, J.E.A.R. (2008) Greenhouse gases emissions in  
17 the production and use of ethanol from sugarcane in Brazil: The 2005/2006  
18 averages and a prediction for 2020. *Biomass and Bioenergy* 32, 582-595.

19 Marin, F.R., Jones, J.W., Royce, F., Suguitani, C., Donzeli, J.L., Wander Filho, J.P.,  
20 Nassif, D.S. (2011) Parameterization and evaluation of predictions of  
21 DSSAT/CANEGRO for Brazilian sugarcane. *Agronomy journal* 103, 304-315.

22 Marino, S., Hogue, I.B., Ray, C.J., Kirschner, D.E., Marino, S., Hogue, I., Ray, C.,  
23 Kirschner, D. (2008) A methodology for performing global uncertainty and  
24 sensitivity analysis in systems biology. *Journal of theoretical biology* 254, 178.

25 McKay, M.D., Beckman, R.J., Conover, W.J. (1979) Comparison of Three Methods  
26 for Selecting Values of Input Variables in the Analysis of Output from a Computer  
27 Code. *Technometrics* 21, 239-245.

28 Medlyn, B.E., Robinson, A.P., Clement, R., McMurtrie, R.E. (2005) On the  
29 validation of models of forest CO<sub>2</sub> exchange using eddy covariance data: some  
30 perils and pitfalls. *Tree Physiology* 25, 839-857.

31 Monod, H., Naud, C., Makowski, D., (2006) Uncertainty and sensitivity analysis for  
32 crop models, in: D. Wallach, D.M., J.W. Jones (Ed.), *Working with Dynamic Crop*  
33 *Models. Evaluation, Analysis, Parameterization and Applications*. Elsevier,  
34 Amsterdam, pp. 55-100.

35 Monsi, M., Saeki, T. (1953) Über den Lichtfaktor in den pflanzengesellschaften  
36 und seine bedeutung für die stoffproduktion. *Japanese Journal of Botany* 14, 22-  
37 52.

38 Morris, M.D. (1991) Factorial sampling plans for preliminary computational  
39 experiments. *Technometrics* 33, 161-174.

40 Muchow, R.C., Spillman, M.F., Wood, A.W., Thomas, M.R. (1994) Radiation  
41 interception and biomass accumulation in a sugarcane crop grown under  
42 irrigated tropical conditions. *Australian Journal of Agricultural Research* 45, 37-  
43 49.

44 Naylor, R.L., Liska, A.J., Burke, M.B., Falcon, W.P., Gaskell, J.C., Rozelle, S.D.,  
45 Cassman, K.G. (2007) The Ripple Effect: Biofuels, Food Security, and the  
46 Environment. *Environment: Science and Policy for Sustainable Development* 49,  
47 30-43.

48 OECD, (2012) *OECD-FAO Agricultural Outlook 2012-2021*, June 2012 ed. OECD  
49 Publishing and FAO.

1 Poulter, B., Hattermann, F., Hawkins, E.D., Zaehle, S., Sitch, S., Restrepo-Coupe, N.,  
2 Heyder, U., Cramer, W. (2010) Robust dynamics of Amazon dieback to climate  
3 change with perturbed ecosystem model parameters. *Global Change Biology* 16,  
4 2476-2495.

5 Pujol, G. (2009) Simplex-based screening designs for estimating metamodels.  
6 *Reliability Engineering & System Safety* 94, 1156-1160.

7 Pujol, G., Iooss, B., Janon, A., (2013) sensitivity: sensitivity analysis, 1.7 ed, p. R  
8 package.

9 Robertson, M.J., Wood, A.W., Muchow, R.C. (1996) Growth of sugarcane under  
10 high input conditions in tropical Australia. I. Radiation use, biomass  
11 accumulation and partitioning. *Field Crops Research* 48, 11-25.

12 Rosenzweig, C., Jones, J.W., Hatfield, J.L., Ruane, A.C., Boote, K.J., Thorburn, P.,  
13 Antle, J.M., Nelson, G.C., Porter, C., Janssen, S., Asseng, S., Basso, B., Ewert, F.,  
14 Wallach, D., Baigorria, G., Winter, J.M. (2013) The Agricultural Model  
15 Intercomparison and Improvement Project (AgMIP): Protocols and pilot studies.  
16 *Agricultural and Forest Meteorology* 170, 166-182.

17 Saltelli, A., Marivoet, J. (1990) Non-parametric statistics in sensitivity analysis for  
18 model output: a comparison of selected techniques. *Reliability Engineering &  
19 System Safety* 28, 229-253.

20 Saltelli, A., Tarantola, S., Campolongo, F., Ratto, M. (2004) Sensitivity Analysis in  
21 practice. A Guide to Assessing Scientific Models. Halsted Press New York, NY,  
22 USA.

23 Schubert, C. (2006) Can biofuels finally take center stage? *Nat Biotech* 24, 777-  
24 784.

25 Searchinger, T., Heimlich, R., Houghton, R.A., Dong, F., Elobeid, A., Fabiosa, J.,  
26 Tokgoz, S., Hayes, D., Yu, T.-H. (2008) Use of U.S. Croplands for Biofuels Increases  
27 Greenhouse Gases Through Emissions from Land-Use Change. *Science* 319, 1238-  
28 1240.

29 Sorda, G., Banse, M., Kemfert, C. (2010) An overview of biofuel policies across the  
30 world. *Energy Policy* 38, 6977-6988.

31 Surendran Nair, S., Kang, S., Zhang, X., Miguez, F.E., Izaurralde, R.C., Post, W.M.,  
32 Dietze, M.C., Lynd, L.R., Wullschlegel, S.D. (2012) Bioenergy crop models:  
33 descriptions, data requirements, and future challenges. *GCB Bioenergy* 4, 620-  
34 633.

35 Tarantola, A. (1987) Inverse problem theory : methods for data fitting and model  
36 parameter estimation. Elsevier ; Distributors for the United States and Canada,  
37 Elsevier Science Pub. Co., Amsterdam; New York; New York, NY, U.S.A.

38 Valade, A., Vuichard, N., Ciais, P., Ruget, F., Viovy, N., Gabrielle, B., Huth, N.,  
39 Martiné, J.F. (2013) ORCHIDEE-STICS, a process-based model of sugarcane  
40 biomass production: calibration of model parameters governing phenology. *GCB  
41 Bioenergy*.

42 Verbeeck, H., Samson, R., Verdonck, F., Lemeur, R. (2006) Parameter sensitivity  
43 and uncertainty of the forest carbon flux model FORUG: a Monte Carlo analysis.  
44 *Tree Physiology* 26, 807-817.

45 von Blottnitz, H., Curran, M.A. (2007) A review of assessments conducted on bio-  
46 ethanol as a transportation fuel from a net energy, greenhouse gas, and  
47 environmental life cycle perspective. *Journal of Cleaner Production* 15, 607-619.

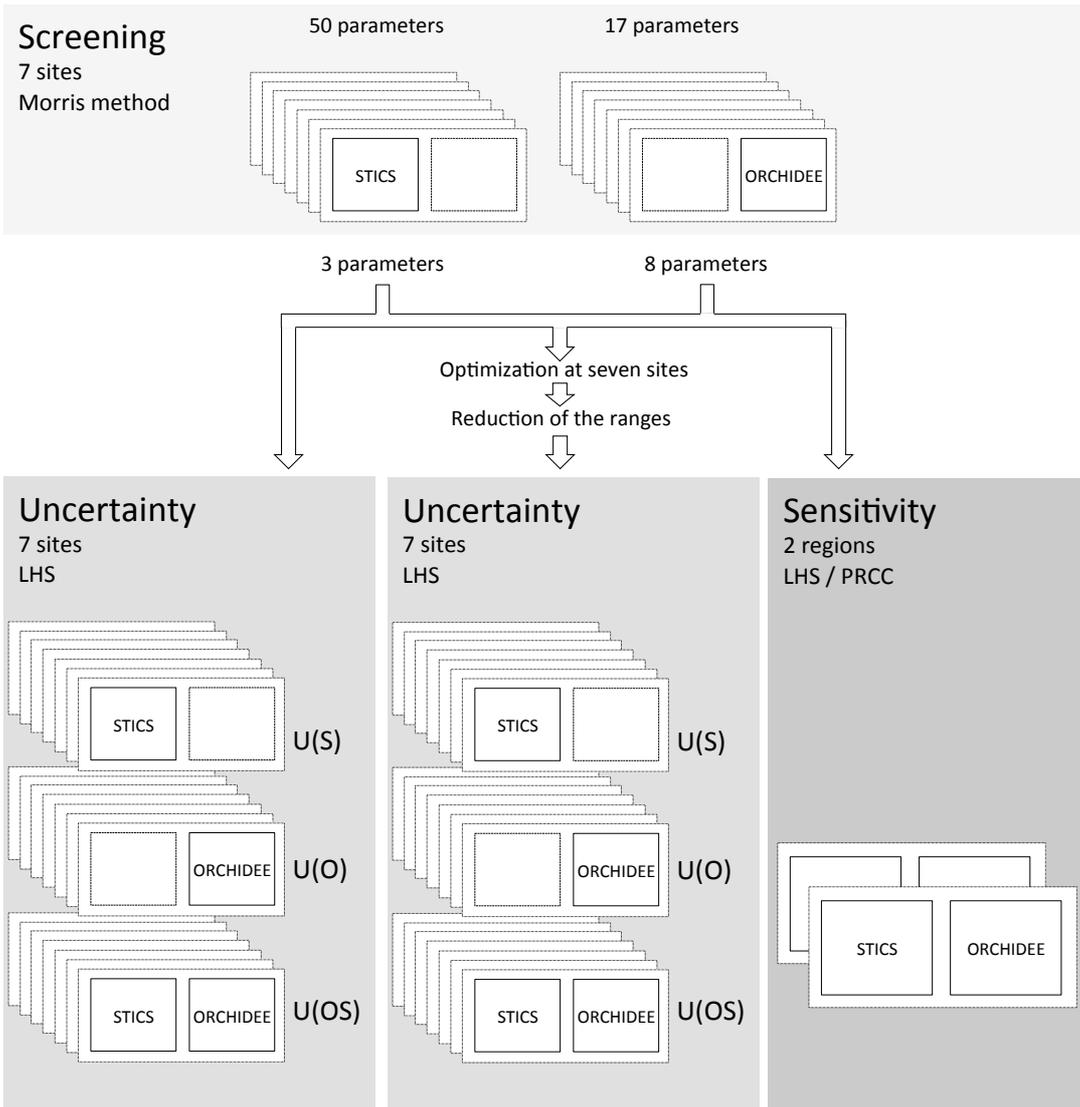
1 Wang, X., He, X., Williams, J., Izaurrealde, R., Atwood, J. (2005) Sensitivity and  
2 uncertainty analyses of crop yields and soil organic carbon simulated with EPIC.  
3 TRANSACTIONS-AMERICAN SOCIETY OF AGRICULTURAL ENGINEERS 48, 1041.  
4 Wyss, G.D., Jorgensen, K.H. (1998) A user's guide to LHS: Sandia's Latin  
5 hypercube sampling software. SAND98-0210, Sandia National Laboratories,  
6 Albuquerque, NM.  
7 Zaehle, S., Sitch, S., Smith, B., Hatterman, F. (2005) Effects of parameter  
8 uncertainties on the modeling of terrestrial biosphere dynamics. Global  
9 Biogeochemical Cycles 19, n/a-n/a.

10

11

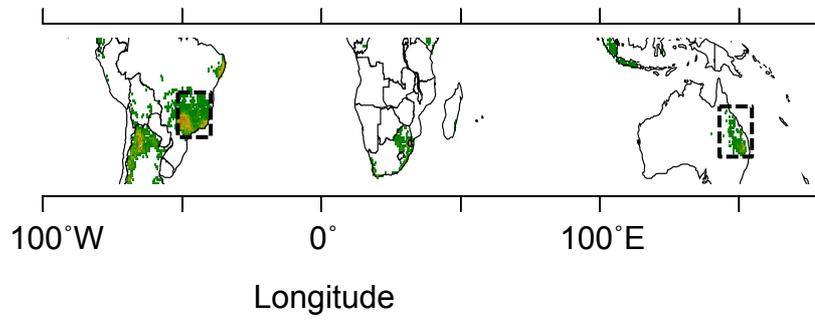
1 Figure 1: flowchart of the analysis carried out in this study. The first step is the  
2 separate screening for 7 sites of the *STICS* and *ORCHIDEE* parameters. The selection  
3 of parameters obtained from the screening are then used for two uncertainty analysis,  
4 one with the same parameters ranges of variation as for the screening, the other with  
5 parameters ranges of variation constrained by the optimization of the model at 7 sites.  
6 Each uncertainty analysis is decomposed in three parts, one including only  
7 *ORCHIDEE* parameters, one including only *STICS* parameters and one including  
8 parameters from both *ORCHIDEE* and *STICS*. Finally a sensitivity analysis is carried  
9 out for two small regions in Australia in Brazil for all parameters together.

10



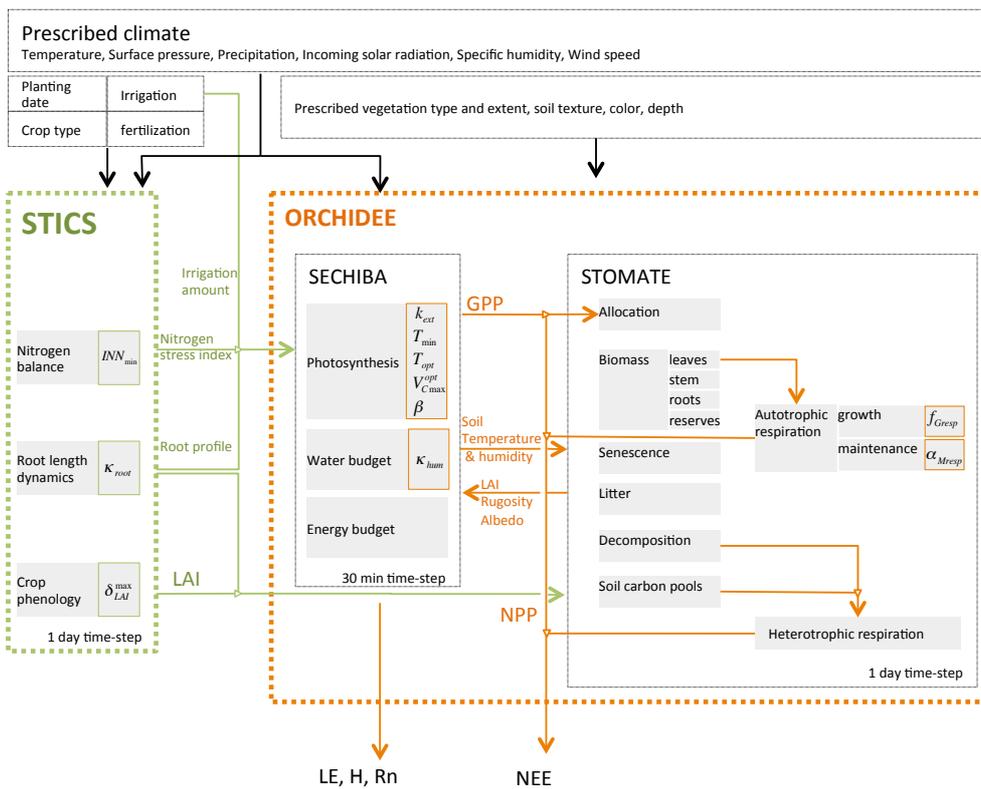
1

2 Figure 2: Spatial distribution of the sites (dots) and regions (dashed rectangles) used  
3 in this study overlaid on a map of the distribution of sugar cane growing areas  
4 indicated in green.



5

1 Figure 3: Structure of the *ORCHIDEE-STICS* chain model. *STICS* calculates the crop  
 2 phenology, water and nitrogen requirements and passes LAI, root profile, irrigation  
 3 and Nitrogen nutrition index to *ORCHIDEE*. *ORCHIDEE* consists in the coupling of  
 4 two module. *SECHIBA* simulates the photosynthesis process, water and energy  
 5 budgets, *STOMATE* is a carbon module and calculates carbon fluxes and to the  
 6 atmosphere (respiration) and carbon accumulation in the carbon pools (biomass  
 7 compartments, litter, soil).



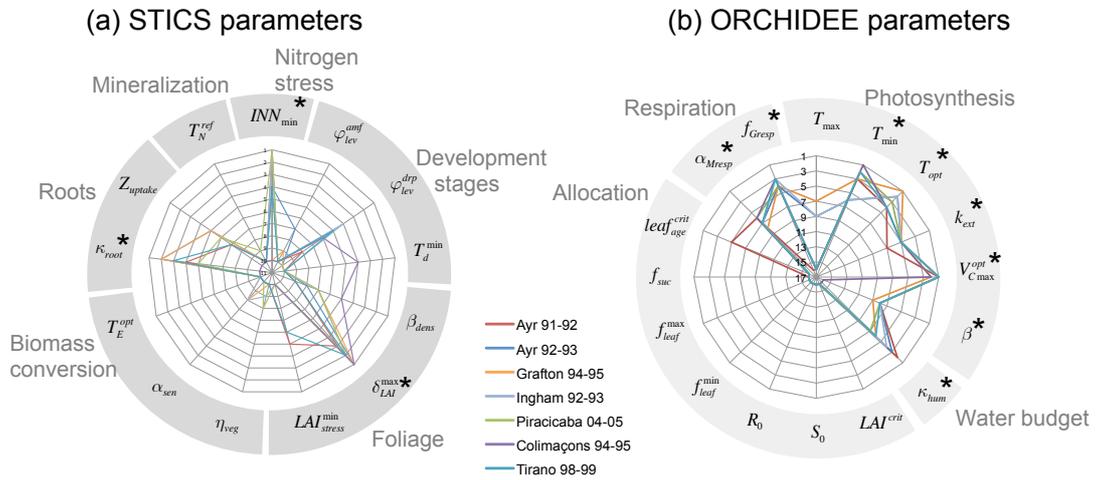
8

9

10

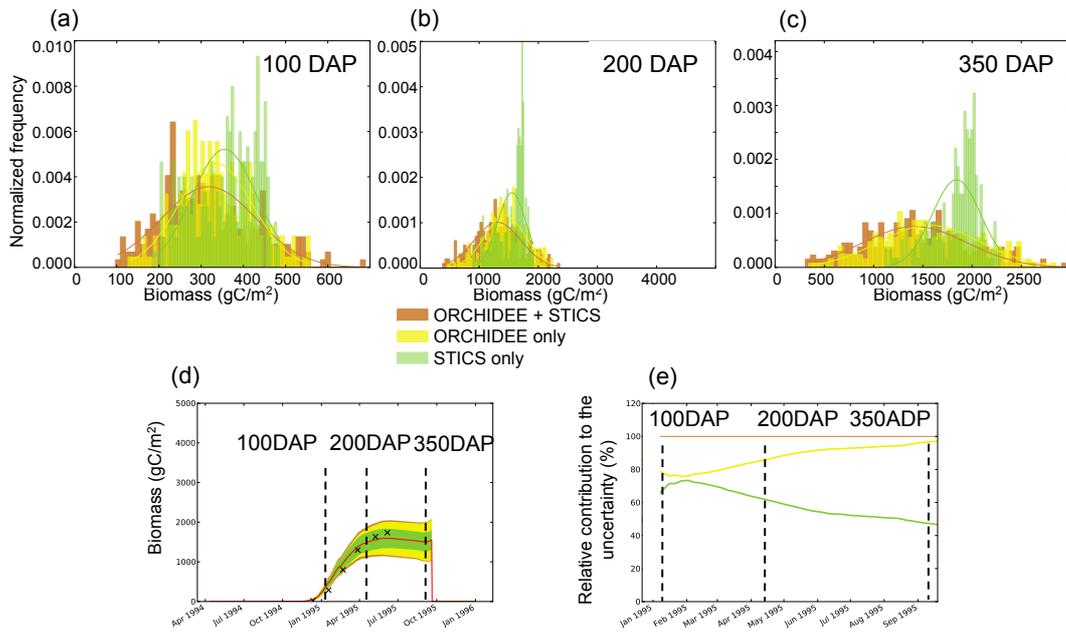


1  
 2 Figure 5: Parameters rankings derived from the Morris screening analysis for *STICS*  
 3 parameters (a) and *ORCHIDEE* parameters (b) for 7 sites (color lines). Each axis of  
 4 the radar plot corresponds to the rank of a parameter, the lower the rank, the more  
 5 important the parameter.



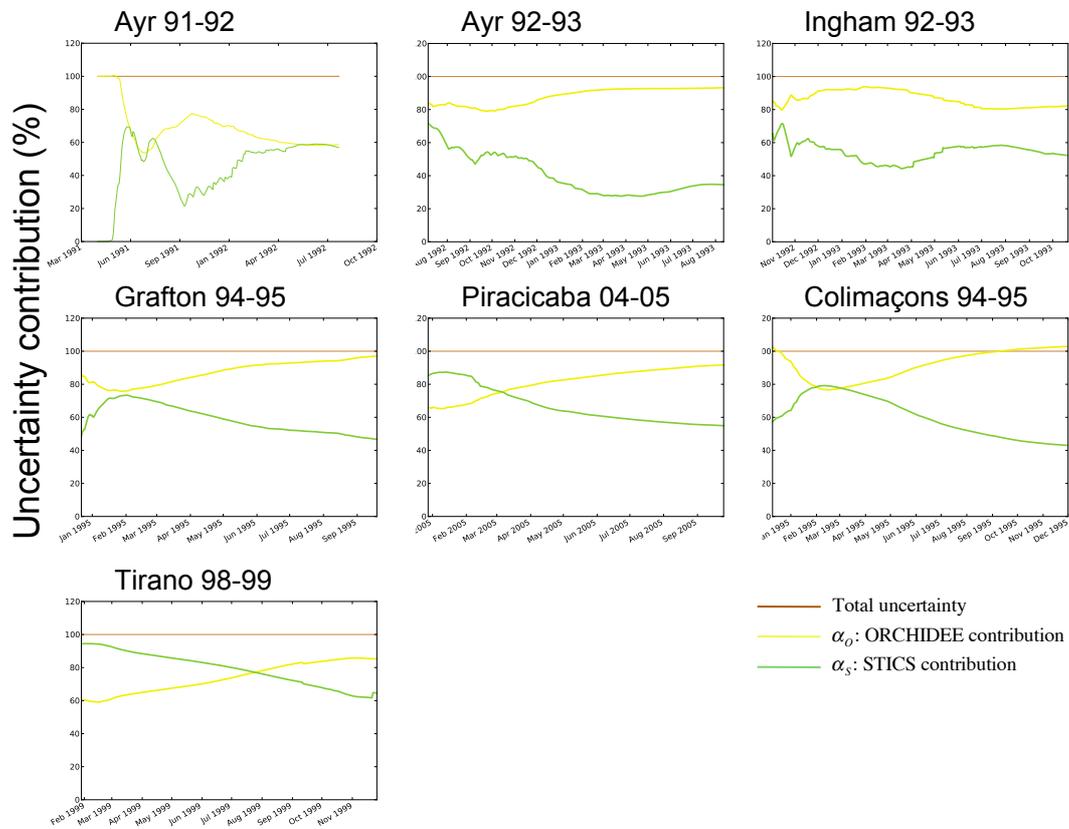
6

1  
 2 Figure 6 : Uncertainty analysis for the site Grafton 94-95. (a-c) probability  
 3 distributions of harvested biomass simulated after parameters uncertainty (from  
 4 *STICS*:green, from *ORCHIDEE*: yellow, from *ORCHIDEE+STICS*: brown) has been  
 5 propagated into the model. (d) reference simulation of harvested biomass (red) and  
 6 uncertainty from *ORCHIDEE*, *STICS*, *ORCHIDEE+STICS*. (e) Contribution (%)of  
 7 *ORCHIDEE* (yellow) and *STICS* (green) to the total uncertainty (brown) over the  
 8 length of the growing season.



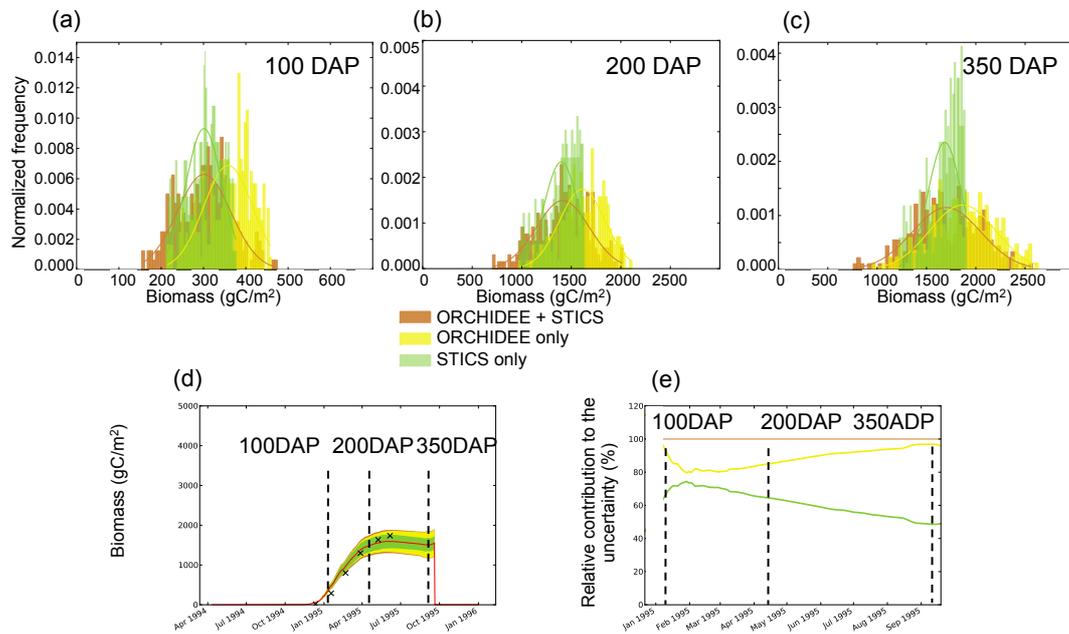
9  
 10

- 1 Figure 7 : Contribution (%) of *ORCHIDEE* (yellow) and *STICS* (green) to the total
- 2 uncertainty (brown) over the length of the growing season for 7 sites.



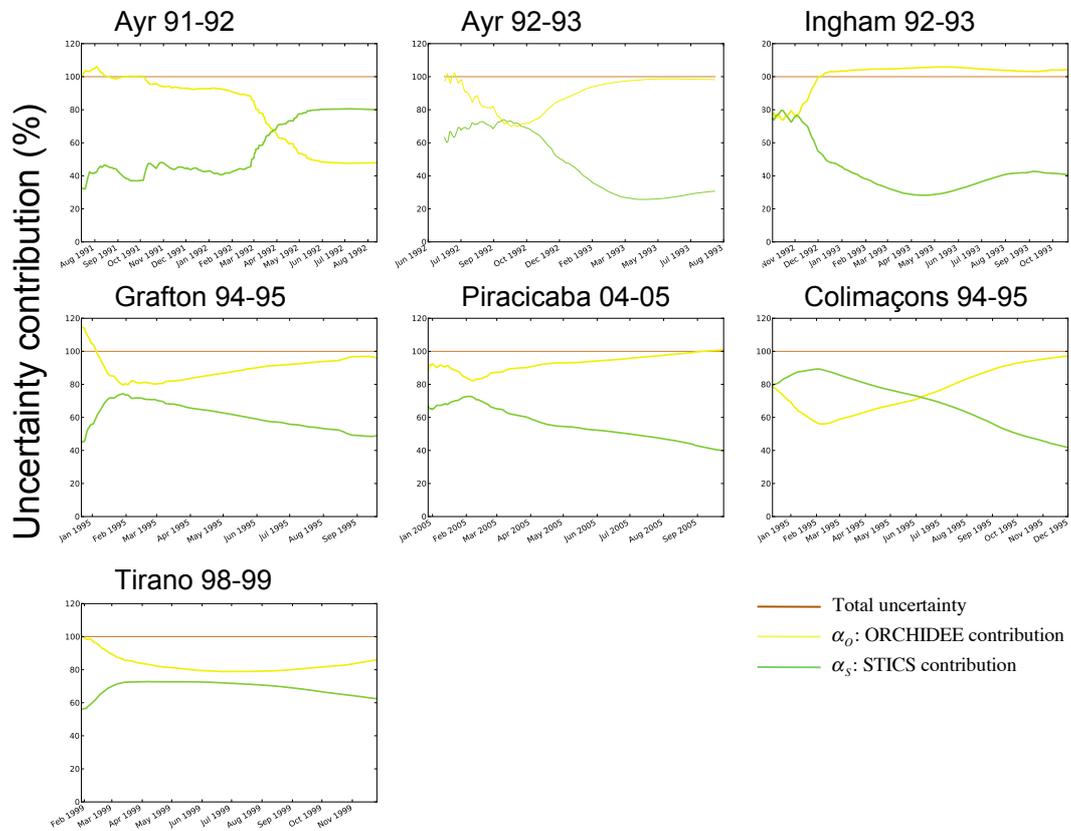
- 3
- 4
- 5

1 Figure 8: Uncertainty analysis for the site Grafton 94-95 after parameters uncertainty  
 2 ranges have been constrained through optimization at 7 sites. (a-c) probability  
 3 distributions of harvested biomass simulated after parameters un- certainty (from  
 4 *STICS*: green, from *ORCHIDEE*: yellow, from *ORCHIDEE+STICS*: brown) has been  
 5 propagated into the model. (d) reference simulation of harvested biomass (red) and  
 6 uncertainty from *ORCHIDEE*, *STICS*, *OR- CHIDEE+STICS*. (e) Contribution (%) of  
 7 *ORCHIDEE* (yellow) and *STICS* (green) to the total uncertainty (brown) over the  
 8 length of the growing season.



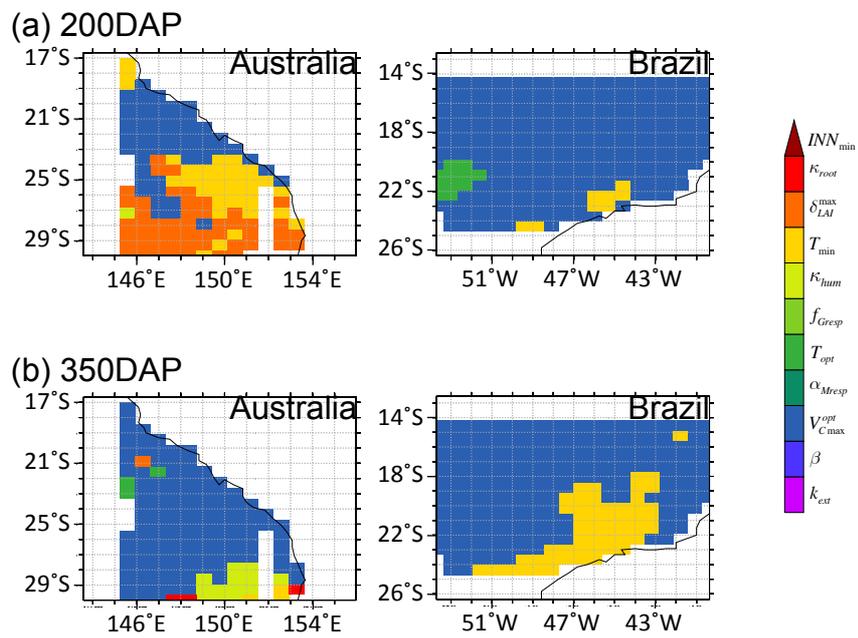
1

2 Figure 9: Contribution (%) of *ORCHIDEE* (yellow) and *STICS* (green) to the total  
3 uncertainty (brown) over the length of the growing season for 7 sites after parameters  
4 uncertainty ranges have been constrained through optimization at 7 sites.



5  
6

- 1
- 2 Figure 10: Spatial distribution of the most influential parameters for the simulation of
- 3 harvestable biomass for two milestones during the growing season, 200 days after
- 4 planting (DAP) and 350DAP



1

2 Figure 11: Sensitivity of *ORCHIDEE-STICS* to its main parameters at 200 days after

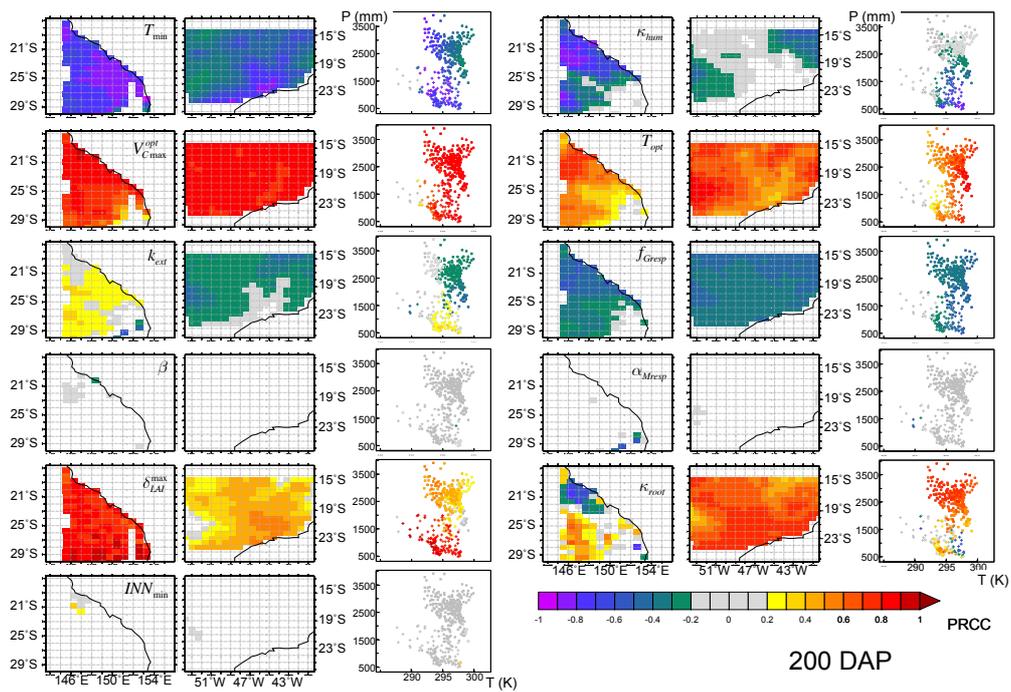
3 planting, as measured with Partial Ranked Correlation Coefficients (PRCC). The

4 color indicates the strength of the relation between the parameter and the harvestable

5 biomass, which is represented spatially (columns 1,2,4,5) and in a (Temperature,

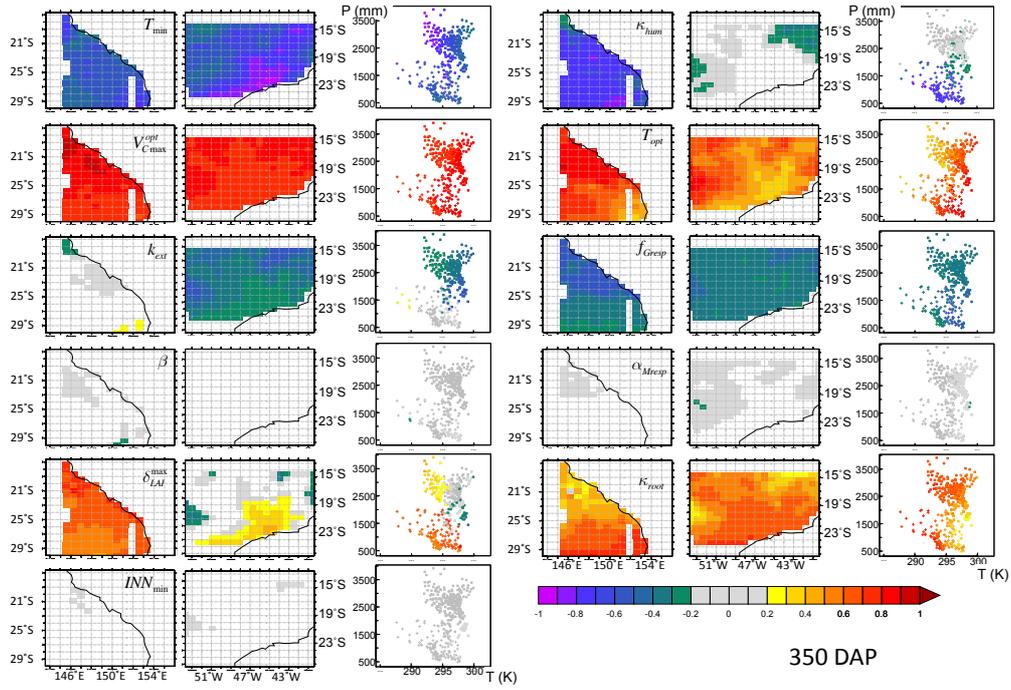
6 Precipitation) referential (columns 3,6).

7



1  
2

1 Figure 12: Sensitivity of *ORCHIDEE-STICS* to its main parameters at 350 days after  
 2 planting, as measured with Partial Ranked Correlation Coefficients (PRCC). The  
 3 color indicates the strength of the relation between the parameter and the harvestable  
 4 biomass, which is represented spatially (columns 1,2,4,5) and in a (Temperature,  
 5 Precipitation) referential (columns 3,6).



6

1

2 Table 1: Description of climate and management for the sites used in this study in  
3 Australia (Ayr, Ingham, Grafton), Brazil (Piracicaba) and La Runion (Colimaons,  
4 Tirano).

	Planting and harvest dates		Mean annual precipitation	Average temperature	irrigation	Fertilization
Ayr	4/19/1991	8/13/1992	964	23.4	irrigated	no
Ayr	4/22/1992	8/13/1993	560	23.6	irrigated	yes
Grafton	9/28/1994	9/19/1995	768	19.6	irrigated	yes
Ingham	7/23/1992	10/21/1993	1294	24.2	irrigated	yes
Piracicaba	10/29/2004	9/26/2005	1230	21.6	irrigated	
Colimacons	8/3/1994	12/1/1995	989.5	19	rainfed	yes
Tirano	11/26/1998	11/26/1999	813	22.34	irrigated	yes

5

1

2 Table 2: List of parameters from STICS and ORCHIDEE included in each step of the  
 3 analysis with their ranges of variation.

		expert judgment based ranges		Uncertainty analysis distribution	Observations constrained ranges	
STICS						
Water budget	absolute value for stomatic closure potential	psisto	5	15		
	Absolute value for start of reduction in cell expansion	psiturg	1	5		
Initial conditions	Table of initial humidity levels in 5 soil horizons for fine soil, % weighted	Hinitf1	11	22		
		Hinitf2	11	22		
		Hinitf3	10	21		
	Table of initial quantities of nitrogen in the 5 soil horizons for fine soil	Ninitf1	0	30		
		Ninitf2	0	30		
		Ninitf3	0	30		
Biomass conversion	Relative age of fruit when rate of growth is maximum	afpf	0.15	0.5		
	Maximum number of set fruits per inflorescence and by degree.day	afruitpot	0.0015	0.2		
	Maximum daily allocation of assimilates towards fruits	allocamx	0.63	0.86		
	Rate of maximum growth as a proportion of maximum fruit weight	bfpf	1	10		
	Radiative effect on conversion efficiency	coefb	0.0015	0.0815		
	Duration of growth of a fruit from setting to physiological maturity	dureefruit	2850	3000		
	Maximum growth efficiency during juvenile phase	efcroijuv	1.7	2.3		
	Maximum growth efficiency during grain filling phase	efcroirepro	2	6		
	Maximum growth efficiency during vegetative phase	efcroiveg	3.2	6		
	Number of age groups of fruits for fruit growth	nboite	12	25		
	Maximum weight of a grain (% water)	pgrainmaxi	1200	2000		
	Fraction of senescent biomass	ratiosen	0	1		
	Quantity of biomass exploited during the cycle	remobil	0.728	0.92		
	Development range between DRP and NOU stages	sdrpnou	552.5	747.5		
	Threshold to calculate trophic stress on LAI	splainin	0	0.3		
Time between emergence and senescence	stlevsenms	400	800			



ORCHIDEE							
Allocation		f fruit	0.05	0.5			
	Maximum LAI per PFT	lai_max	3	9			
	Average critical age for leaves	leaf_age_crit	30	200			
	Upper bounds for leaf allocation	max_lto_lsr	0.25	0.5			
	Lower bounds for leaf allocation	min_lto_lsr	0.05	0.24			
	Root allocation	R0	0.05	0.5			
	Sapwood allocation	S0	0.05	0.5			
Photosynthesis	<b>Extinction coefficient</b>	<b>ext_coef</b>	<b>0.5</b>	<b>0.9</b>	<b>uniform</b>	<b>0.5</b>	<b>0.72</b>
	<b>Slope of relationship between assimilation and stomatal conductance</b>	<b>gsslope</b>	<b>7</b>	<b>11</b>	<b>beta(2,2)</b>	<b>7.7</b>	<b>9.5</b>
	<b>Temperature at which photosynthesis is maximal</b>	tphoto_max	30	45			
	<b>Temperature at which photosynthesis is minimal</b>	tphoto_min_c	12	19	<b>uniform</b>	12	<b>16.7</b>
	<b>Temperature at which photosynthesis is optimal</b>	tphoto_opt	24	36	<b>uniform</b>	24	<b>36</b>
	<b>Maximum carboxylation rate</b>	vcmax_opt	40	100	<b>beta(2,2)</b>	64	<b>81.3</b>
Respiration	<b>Fraction of biomass available for growth respiration</b>	<b>frac_growth_resp</b>	<b>0.2</b>	<b>0.5</b>	<b>beta(2,2)</b>	<b>0.23</b>	<b>0.3</b>
	<b>Slope of the relationship between temperature and maintenance respiration</b>	<b>maint_resp_slope1</b>	<b>0.08</b>	<b>0.16</b>	<b>beta(2,2)</b>	<b>0.11</b>	<b>0.12</b>
Water budget	<b>Root profile to determine soil moisture content available to plants</b>	<b>humcste</b>	<b>0.8</b>	<b>7.2</b>	<b>uniform</b>	<b>3.2</b>	<b>4.1</b>

1

2 Table 3: Uncertainty associated with *STICS*, *ORCHIDEE*, or *ORCHIDEE+STICS*  
3 parameters uncertainties expressed as percentage of the reference harvested biomass  
4 for each site and for each of the two uncertainty analysis.

		Total Uncertainty (% of observed value)	ORCHIDEE Uncertainty (% of observed value)	STICS Uncertainty (% of observed value)
Expert-based parameters' uncertainties	Ayr 91-92	35.11	20.43	20.73
	Ayr 92-93	27.21	25.26	9.31
	Ingham 92-93	38.60	31.42	21.04
	Grafton 94-95	26.05	23.92	14.07
	Piracicaba 04- 05	25.49	23.36	14.00
	Colimacons 94-95	41.21	41.87	18.61
	Tirano 98-99	44.26	36.80	30.61
Optimization-based parameters' uncertainties	Ayr 91-92	31.20	14.01	25.64
	Ayr 92-93	15.84	15.60	4.58
	Ingham 92-93	21.66	22.35	9.19
	Grafton 94-95	16.84	15.25	9.81
	Piracicaba 04- 05	14.67	14.80	5.84
	Colimacons 94-95	21.31	20.01	10.28
	Tirano 98-99	22.26	18.06	15.03

5

6