Parametrization of five classical plant growth models applied to sugar beet and comparison of their predictive capacities on root yield and total biomass

Charlotte Baey, Anne Didier, Sébastien Lemaire, Fabienne Maupas, Paul-Henry Cournède

Parametrization of five classical plant growth models applied to sugar beet and comparison of their predictive capacities on root yield and total biomass.

Charlotte Baey[a,*], Anne Didier[b], Sébastien Lemaire[b], Fabienne Maupas[b], Paul-Henry Cournède[a]

[a] *École Centrale Paris, Laboratoire de Mathématiques appliquées aux systèmes, Grande Voie des Vignes, 92290 Châtenay-Malabry, France*
[b] *Institut Technique de la Betterave, 45 rue de Naples, 75008 Paris*

Abstract

A wide range of models have been proposed and developed for modelling sugar beet growth, each of them with different degrees of complexity and modelling assumptions. Many of them are used to predict crop production or yield, even when they were not originally designed for this purpose, and even though their predictive capacity has never been properly evaluated.

In this study, we propose the evaluation and comparison of five plant growth models that rely on a similar energetic concept for the production of biomass, but with different levels of description (individual-based or per square meter) and different ways to describe biomass repartition (empirical or via allocation): Greenlab, LNAS, CERES, PILOTE and STICS. The models were all programmed on the same modelling platform, calibrated on a first set of data, and then their predictive capacities were assessed on an independent data set. First, a sensitivity analysis was carried out on each model to identify a subset of parameters to be estimated, to reduce the variability of the models. We were able to reduce the number of parameters from 10 to 4 for Greenlab, and from 16 to 1 for STICS. Three criteria were then used to compare the predictive capacities of the models: the root mean squared error of prediction and the modelling efficiency for the total dry matter production and the dry matter of root, and the yield prediction error.

All the models provided good overall predictions, with high values of the modelling

efficiency. The use of sensitivity analysis allowed us to reduce the variability of the models and to enhance their predictive capacities. Models based on an empirical harvest index gave good yield predictions, and similar results compared to allocation models for the total dry matter, but the harvest index might not be very robust. The crucial role of initiation was also pointed out, as well as the need for an accurate estimation and modelling of this early phase of growth.

Keywords: sugar beet; prediction; model evaluation; model comparison; RMSEP; modelling efficiency

1. Introduction

A wide range of plant growth models are available in the literature, either generic ones, that can be applied to different species, or more specific ones built for given plants or trees. Some of them are designed to predict yield or biomass production at field scale, and help management decisions, while some others are built for descriptive purposes, to enhance our understanding of plant functioning and simulate plant architecture (Fourcaud et al., 2008). One can also be confronted with the need to compare and choose between different versions of the same model, corresponding to different biological assumptions, for example, or to decide whether a given biological process should be accounted for or not.

Depending on their initial objective, these models can have different levels of complexity. For example, descriptive models would tend to be more complex than purely predictive ones as they would integrate more underlying eco-physiological processes (e.g. allocation processes, reaction to environmental stresses, ...). Often, this increase in the model complexity results in a higher number of parameters, and consequently in a decrease of the predictive capacity of the model due to a higher variability. This is the

*Corresponding author, Tel: (+33) 141 131 559
Email addresses: charlotte.baey@ecp.fr (Charlotte Baey), didier@itbfr.org (Anne Didier), lemaire@itbfr.org (Sébastien Lemaire), maupas@itbfr.org (Fabienne Maupas), paul-henry.cournede@ecp.fr (Paul-Henry Cournède)

well known bias/variance compromise. However, a lot of models are used as prediction tools even though they were not originally designed for this purpose. It is thus necessary, when using a model, to define precisely the context in which it will be used, and even more importantly, to evaluate properly its performance according to the objective of the study.

In this context, we propose a methodology to build and evaluate different models in a predictive perspective. We apply this approach to five plant growth models for sugar beet crops, with different levels of description and modelling scales: Greenlab (de Reffye and Hu, 2003; Yan et al., 2004), CERES (Jones and Kiniry, 1986; Leviel, 2000), Pilote (Mailhol et al., 1997; Taky, 2008), STICS (Brisson et al., 1998, 2008) and a fifth model named LNAS (Cournède et al., 2013), based on a global allocation of biomass to the leaves compartment or root. First elements of comparison for Greenlab, CERES and Pilote are available in Lemaire (2010).

The five models rely on a similar formulation for the production of biomass, based on Monteith's equation (Monteith, 1977) and on an extension of the Beer-Lambert law. The accumulated dry matter production is linearly related to the fraction of intercepted radiation, which can generally be expressed according to the leaf area index (LAI) or to the leaves biomass. From this common basis, the models then differ in their formulations of the LAI curve, either based on allocation processes (Greenlab, LNAS) or on empirical relationships (PILOTE, CERES). STICS can be seen as an intermediate between these two approaches, since the LAI is computed from an empirical function, but modulated by a source-sink ratio. Two modelling scales were also compared, with either individual-based models (CERES, Greenlab) in which the LAI was computed from the development of each individual leaf, or more classical crop models (PILOTE, STICS, LNAS) where the LAI was computed per square meter at field scale. The differences between the five models are summarized in Table A.1.

[Table 1 about here.]

All these models have already been tested and calibrated in the case of sugar beet, but

the purpose here is to evaluate their predictive capacities. In this perspective, to reduce the variability of the models which included a large number of parameters, a sensitivity analysis was computed for each of them. The parameters were ranked according to their influence on the model outputs, and then, the best subset of parameters to be estimated was identified, according to AIC and BIC criteria. They were first calibrated on the same set of data, and their predictive capacity was then evaluated and compared on an independent data set using three classical criteria: the root mean squared error of prediction (RMSEP), the modelling efficiency (EF), and the yield prediction error.

In the second section, we present the five models, along with the data and the criteria used for the calibration and the evaluation of their predictive capacity. The calibration process, and in particular the sensitivity analysis performed on each model, is described in section 2.3. Results from this sensitivity analysis are given in section 3.1, those from the comparison between the different versions of STICS, in section 3.2.1, and between the two data sets (calibration and validation sets), in section 3.2.

2. Material and methods

2.1. Models

The five models rely on the same concept for the energetic production of biomass, based on an extension of the Beer-Lambert law (Monteith, 1977). The biomass production in grams per square meter on day t, $Q(t)$ is proportional to the incoming photosynthetically active radiation PAR(t) (in MJ/m^2), to the fraction of intercepted radiation $I(t)$ (which depends on the leaf area index or on the dry matter of leaves) and to the radiation use efficiency RUE (in g/MJ) (Damay and Le Gouis, 1993):

$$Q(t) = 0.95 \cdot \text{RUE} \cdot \text{PAR}(t) \cdot I(t). \tag{1}$$

The leaf area index is defined as the one-sided green leaf area per unit ground surface (Watson, 1947), thus some adjustments were necessary for the two individual-based mod-

els Greenlab and CERES. In Greenlab, as the biomass production is computed at the individual plant level, a 'local' LAI (Cournède et al., 2008) was defined, corresponding to the leaf surface of the plant multiplied by a coefficient related to the two-dimensional projection of the space occupied by the plant on the ground (see 2.1.1). In CERES, as the biomass production is computed at the square meter level, a 'global' LAI was constructed from the individual leaf surfaces of the plant, by multiplying by the crop density (see 2.1.5).

2.1.1. GreenLab

GreenLab is a generic functional-structural plant model (FSPM), combining the description of the plant architecture and its physiological functioning (Vos et al., 2007; Sievänen et al., 2000). The model in its discrete version was introduced by de Reffye and Hu (2003), and was studied in the case of sugar beet by Lemaire et al. (2008).

In its first version, the time step chosen to compute the organogenesis and the ecophysiological processes was the growth cycle (i.e. the thermal time elapsing between the appearance of two successive metamers). However, for a better accuracy in the handling of continuous variations of environmental conditions, and consistency with the usual daily collection of climatic data, a continuous version of the Greenlab model was used, discretized with a daily time step (Li et al., 2009). Such formulation is also more consistent with the other plant growth models studied in this paper and that provide daily outputs (Mailhol et al., 1997; Guérif and Duke, 1998; Spitters et al., 1989).

In Greenlab, the biomass production on day t is computed at the individual plant level, thus some adjustments were made from equation (1):

$$Q_{pl}(t) = 0.95 \cdot \mathrm{RUE} \cdot \frac{PAR(t)}{d} \cdot \left(1 - \exp\left(-k_B \, \frac{Q_b(t)}{e_b \cdot S_p}\right)\right),$$

with $Q_{pl}(t)$ the biomass production of an individual plant (in g/pl), d the plant density (in pl/m^2), k_B the Beer-Lambert law extinction coefficient, $Q_b(t)$ the accumulated blade mass (in g/pl) at day t, e_b the mass per unit area of blade (in g/m^2), and S_p an empirical

coefficient related to the two-dimensional projection of the space occupied by the plant (in m^2/pl). The biomass production per square meter $Q(t)$ can be obtained by multiplying $Q_{pl}(t)$ by the crop density d. The biomass is then allocated to the different organs of the plant according to source-sinks relationships (we refer the reader to Yan et al. (2004); Lemaire et al. (2008) for more details).

2.1.2. LNAS

A simplified model called LNAS (Cournède et al., 2013) was elaborated, where the biomass allocation is done globally for the whole leaves compartment, instead as leaf by leaf as in the Greenlab model. The leaf area index was obtained by dividing the biomass of leaves by the mass per unit area. It is a generic daily time-step model, presented here in the case of sugar-beet, but that can be easily extended to other plants.

The biomass production on day t is given by (1), with:

$$I(t) = 1 - \exp\left(-k_B \cdot \frac{Q_g(t)}{e_g}\right),$$

where $Q_g(t)$ is the dry matter of green leaves at day t (in g/m^2), k_B the extinction coefficient and e_g the mass per unit area of leaf (in g/m^2). Note that here, the quantity e_g is different from the quantity e_b used in the Greenlab model, as in LNAS we are dealing with the dry matter of leaves, considering blades and petioles together, whereas in Greenlab we are only dealing with the dry matter of blade.

Then, the produced biomass is allocated to the different organs compartments. Only two compartments are considered in the case of sugar beet: leaves and root. We denote by $Q_l(t)$ and $Q_r(t)$ respectively the total mass of leaves and the total mass of root on day t. At the beginning of day $t+1$, the masses of leaves and root are given by:

$$\begin{aligned} Q_l(t+1) &= Q_l(t) + \gamma(t) \cdot Q(t) \\ Q_r(t+1) &= Q_r(t) + (1-\gamma(t)) \cdot Q(t) \end{aligned}$$

where function γ is defined as:

$$\gamma(t) = \gamma_0 + (\gamma_f - \gamma_0) \cdot G_a(\tau(t)),$$

with G_a the cumulative distribution function of a log-normal law, parametrized by its median μ_a and its standard deviation σ_a, $\tau(t)$ the thermal time on day t, and γ_0 and γ_f respectively the initial and final proportion of biomass allocated to the leaves.

The proportion of non-senescent leaves is given according to the following equation:

$$Q_g(t) = (1 - G_s(\tau(t) - \tau_{sen}))\, Q_l(t)$$

where G_s is the cumulative distribution function of a log-normal law, parametrized by its median μ_s and its standard deviation σ_s, and τ_{sen} is the thermal time at which the senescence starts.

2.1.3. STICS

STICS (Simulateur mulTIdisciplinaire pour les Cultures Standard (Brisson et al., 1998, 2008)) is a generic daily time-step model, which has already been applied to a large variety of crops (maize, tomato, wheat, sugarbeet, ...). It is organized into seven modules corresponding to the different mechanisms involved in plant growth.

In the original formulation of STICS, the relation between biomass production and intercepted radiation is not linear as in (1), but quadratic, with the introduction of a saturation coefficient, and a radiation use efficiency that could vary according to the development stage $s(t)$. However, since the main differences between the models concern the modelling scale and the biomass repartition, we choose a simplified version of the production function, with a constant RUE and a linear relationship with radiation interception, so that the five models share the same formulation for biomass production. A complementary study is conducted in 3.2.1 in order to check if such simplifications alleviate the predictive capacity of STICS, and actually shows that the simplified version

161 performs better in our case.

The leaf area index is modelled with an empirical function as the net balance between growth and senescence, and is supposed to evolve in three phases: a first phase of logistic growth (from emergence to the maximal LAI point), a stability phase, and a senescent phase in which the LAI decreases linearly (Brisson et al., 2008). We have:

$$I(t) = \left(1 - \exp(-k_B \ \mathrm{LAI}(t))\right),$$

with k_B the Beer-Lambert law extinction coefficient, and:

$$\mathrm{LAI}(t) = \sum_{j=t_e}^{t} (\Delta\mathrm{LAI}(j) - \Delta\mathrm{LAI}_{\mathrm{sen}}(j))$$

162 where t_e is the day of emergence, $\Delta\mathrm{LAI}(j)$ is the net leaf area growth on day j and
163 $\Delta\mathrm{LAI}_{\mathrm{sen}}(j)$ is the leaf area senescence on day j.

The net leaf area growth on day j depends on the leaf development unit on day j, $u(j)$, which varies from 1 at emergence to 3 when the leaf area index is maximal. From emergence to the maximal LAI point, the LAI growth follows a logistic curve:

$$\Delta\mathrm{LAI}(j) = \frac{\alpha}{1 + \exp(\beta(u_{\mathrm{mat}} - u(j)))} \cdot d \cdot f_d(j) \cdot f_T(j) \cdot s(j), \quad \text{for } 1 \leq u(j) \leq 3$$

164 where u_{mat} is the leaf development unit at the end of the juvenile stage, d is the
165 plant density and f_d a density factor related to the competition between plants, f_T
166 is the effective crop temperature, and s is a trophic stress index. This trophic stress
167 is determined by a source-sink ratio, and thus induces a retroaction of the allocation
168 process on the LAI curve. With this formulation, the leaf area index stops growing all at
169 once after having reached its maximal point, but it is possible to introduce a progressive
170 decline of the LAI. We refer the reader to Brisson et al. (2008) for more detailed equations.
171 A lot of parameters are required for the model, but a list of recommended values for
172 different crops are available in Brisson et al. (2008).

As explained previously, in the original formulation of STICS, the relationship between biomass production and intercepted radiation is quadratic and not linear. Moreover, among the several modules available in STICS, some of them are dedicated to the management of environmental stresses, so that these processes can be easily implemented in the model. As a consequence, a complementary study was carried out on STICS to evaluate the effect of these different modelling assumptions, and four versions of STICS were compared (see section 3.2.1):

1. the initial version of STICS, with a varying radiation use efficiency and a quadratic relationship between biomass production and intercepted radiation, but no stresses
2. a linear version, with a varying radiation use efficiency but a linear relationship between production and radiation
3. a linear version with a constant radiation use efficiency (this version is the standard one used in the comparison with the other models)
4. the linear and constant RUE version, with the addition of thermal stress

The thermal stress was added in the modified version of the model because, as it will be shown latter in the paper, this modified version performed better than the initial one.

2.1.4. Pilote

Pilote is a crop-soil interaction model, which was first built for sorghum and sunflower (Mailhol et al., 1996, 1997), but that can be applied to a large variety of crops. It has been developped for sugar beet by Taky (2008). It is designed to predict the actual evapotranspiration and the yield of crops, through the modelling of the leaf area index. In this paper, we considered the version of Pilote without hydric stress. In such case, the biomass production per square meter at day t is given by (1), with:

$$I(t) = 1 - \exp(-k_B \cdot \mathrm{LAI}(t))$$

$$\mathrm{LAI}(t) = \mathrm{LAI}_{\max} \left(\frac{\tau(t) - \tau_e}{\tau_{\max}} \right)^{\beta} \exp \left[\frac{\beta}{\alpha} \left(1 - \left(\frac{\tau(t) - \tau_e}{\tau_{\max}} \right)^{\alpha} \right) \right]$$

with k_B the Beer-Lambert law extinction coefficient, $\mathrm{LAI}_{\max}$ the potential maximum value of LAI in non-limiting conditions, $\tau_{\max}$ the thermal time (in °Cday) necessary to reach this maximal LAI, τ_e the thermal time (in °Cday) of emergence, and α and β two parameters. It is also possible to model growth and senescence separately, using two different values α_1 and α_2 depending on whether we are before or after $\tau_{\max}$. Then, the biomass repartition to root and leaves is done with an empirical harvest index.

2.1.5. CERES

CERES (Crop Environment REsource Synthesis) was originally built on maize by Jones and Kiniry (1986), but a sugar beet version was developed by Leviel (2000). Effects of irrigation or nitrogen uptake can also be integrated in the model. In CERES, the biomass production is done at the square meter level from equation (1), from the individual foliar surfaces of the plant:

$$I(t) = 1 - \exp\left(-k_B \cdot d \cdot \sum_k S_k(t)\right),$$

with k_B the Beer-Lambert law extinction coefficient, d the plant density (in pl/m^2), and S_k the foliar surface of leaf k at time t (in m^2).

The foliar surface of leaf k, S_k is supposed to grow linearly from the thermal time of appearance of leaf k to the end of its expansion, then stay at its maximum surface $S_{k,\max}$ until the end of its lifetime:

$$S_k(t) = \begin{cases} 0 & \text{if } \tau(t) \in [0, \tau_k^i[\\ \frac{S_{k,\max}}{\tau_k^e - \tau_k^i}\ (\tau(t) - \tau_k^i) & \text{if } \tau(t) \in [\tau_k^i, \tau_k^e[\\ S_{k,\max} & \text{if } \tau(t) \in [\tau_k^e, \tau_k^i + \tau_k^s[\\ 0 & \text{otherwise} \end{cases}$$

with $\tau(t)$ the thermal time at time t, and τ_k^i, τ_k^e and τ_k^s respectively the thermal time

of initiation, the thermal time of end of expansion and the lifespan in thermal time of leaf k. The maximum foliar surfaces $S_{k,\max}$, as well as the thermal times of initiation, expansion and senescence, are thus needed to calibrate the model.

The biomass is then distributed to root and leaves thanks to an empirical harvest index, corresponding to the ratio between the dry matter of root and the total dry matter at harvest.

2.2. Data

We used a first dataset from 2010 experiments to calibrate the models. Field experiments took place at La Selve, France, N49°34'22", E3°59'24", on a sandy loam soil. A commercial variety, Python, was sown on April 15, with 45 cm between rows and 18 cm between seed-plots, and fertilized with 136 kg N ha^{-1}. The final plant density was estimated at 11.82 plants per square meter (pl/m^2). Dry matter of root and leaves (blades and petioles separately or altogether) were collected on 50 plants at fifteen different dates, and dry matter of individual blades and petioles, as well as blade surface areas, were measured on 10 plants at five different dates. Mean values were then used for the calibration process. The LAI was not measured directly on the field, but could be computed from the blade mass Q_b, the mass per unit area e_b and the plant density d: $\mathrm{LAI}_{\exp} = \dfrac{Q_b}{e_b}\, d$. The mass per unit area of blade e_b was obtained from a linear regression between leaf areas and blade masses on the five dates of individual measurements. The observed maximum surfaces of leaves were used for CERES model.

At first, the predictive capacity of the models was supposed to be tested on 2011 experiments, conducted on the same genotype near the 2010 site, at Bourgogne, France, N49°21'18", E4°4'12" on a calcareous loam soil. However, a hail episode occurred this year around day 100, and caused a lot of damages on leaves, with partial or total destruction of blades and petioles, and thus a much lower biomass production than in 2010, leading to over-estimated predictions for this variable. As a consequence, a second dataset was used to evaluate the predictive capacity of the models in the absence of any perturbing event. The field experiments in 2008 were conducted around 200km away from the 2010

site, in the Beauce plain near Bazainville, France, N48°11'15", E2°5'50" on a clay loam soil. A commercial variety, very similar to the Python genotype, Radar, was sown on March 11, with 50 cm between rows and 20 cm between seed-plots. The final measured density was 10.9 pl/m^2. Dry matter of root and leaves were measured at fifteen different dates on 30 plants.

Daily mean values of air temperature (°C) and solar radiation (MJ.m^{-2}) were obtained from French meteorological advisory services (Météo France) near the experimental site. Thermal time was computed using a base temperature of 0°C (Lemaire et al., 2008). All the experiments were conducted in non-limiting water conditions.

2.3. Calibration

2.3.1. Reduction of the variability

A very important part of this work consisted in the calibration of the five models. Indeed, parametric estimation is a crucial and yet tricky task, especially when the number of parameters is high compared to the amount of data available for the estimation. It can also be impossible to estimate simultaneously all the parameters of a model, due for example to numerical issues or high correlations between parameters (Wallach et al., 2006, Chapter 4). Moreover, estimating too many parameters can have an obnoxious effect on the predictive capacity of the model, due to overparametrization. If it is true that the estimation of a parameter can enhance the goodness of fit of the model, especially when the estimated value is far from the one given in the literature, or when no reliable information is available, the estimation process is also marred with errors, leading to a higher variance of the model. A good compromise has therefore to be made on the number of parameters to estimate.

A classical two-step approach to this issue is first to rank the parameters according to their importance, and then, to select the number of parameters to estimate. In the first step, parameters can be ordered for example according to their influence on the model outputs, or their ability to enhance the goodness of fit measured by a given criterion (Campolongo et al., 2007). In the second step, the final number of parameters to

estimate is determined using another appropriate criterion like the mean squared error of prediction (Wallach et al., 2001, 2006). In this paper, we adopted a very similar approach, with the use of sensitivity analysis (SA) to rank the parameters and model selection criteria such as AIC and BIC to choose the number of parameters that would be estimated, the others being set to recommended values given in the literature. A similar study has already been carried out for STICS, on wheat and maize, based on response surface method (Ruget et al., 2002).

A review on the role of sensitivity analysis can be found in Cariboni et al. (2007) or Saltelli et al. (2008). A first prerequisite to the use of SA is the definition of appropriate distributions for the parameters. Such information can be found in the literature, but may also not be easily available, in particular when a model is used for a species on which it has rarely or never been tested, or when a new model is considered. Uniform distributions were used in this study, and the corresponding variation intervals were defined according to the information available for each model as follow: (i) for STICS, we took a 10% variation around the recommended values given in the literature (Brisson et al., 2008), (ii) for Greenlab, we defined the variation intervals from a combination of the reference values given by Lemaire et al. (2008) for the discrete version of the model and results from previous fittings with the continuous version, (iii) for PILOTE, we defined variation intervals taking into account the different values of the parameters for different crops, including sugar beet (Mailhol et al., 1996, 1997; Taky, 2008), (iv) for LNAS, we built variation intervals using the similarities between LNAS and SUCROS (Guérif and Duke, 1998) or Greenlab, and the relationship between some of the parameters and well-known biological processes. Sensitivity analysis was not used on CERES, as there is only one parameter to calibrate in this model.

Sobol's method was used to compute the sensitivity indices for each parameter, each output, and each time of observation, with the estimator proposed by Wu et al. (2011). Then, sensitivity indices (SI) for each parameter and each output were obtained with a weighted sum over each time of observation using the variance of the outputs at each

time of observation, and finally, a global SI for each parameter was computed, taking into account all the model outputs. Parameters were ranked according to their global SI, and then, the models were calibrated with a growing number of parameters, introduced in the calibration process in the order determined by SA. AICc and BIC were then computed for each calibrated models. We used the corrected version of AIC, as it lessen the probability of overfitting when the sample size is too small, and as it converges towards AIC when the sample size gets large (Burnham and Anderson, 2002). These two criteria are defined as: AICc $= -2\log L + 2kn/(n-k-1)$ and BIC $= -2\log L + k\log n$, where L is the estimated log-likelihood of the model, k is the number of parameters and n the sample size. It is worth noting that, if these criteria are not relevant to assess the predictive capacity of a model, leading to the use of appropriate criteria as defined in section 2.4, there are nonetheless suitable for goodness-of-fit evaluations and can therefore be used in this calibration step.

A summary of the variation intervals adopted for each parameter can be found in Table A.2. To lighten the table, intervals are not shown for STICS, we thus refer the reader to Brisson et al. (2008) for a list of recommended values for the parameters on sugar beet.

[Table 2 about here.]

2.3.2. Fixed parameters

A first set of fixed parameters was identified, corresponding to those than can be measured directly or for which well documented values are given in the literature. These parameters included: the extinction coefficient of the Beer-Lambert law ($k_B = 0.7$ according to Andrieu et al. (1997), except for STICS when the relationship between biomass production and intercepted radiation was nonlinear, in which case $k_B = 0.58$), the mass per unit area in Greenlab (deduced from our field experiments, $e_b = 83\ g.m^{-2}$), and the maximum leaf surfaces for CERES (deduced from the individual measurements). The harvest index for Pilote and CERES was computed as the ratio between dry matter of

root (including crown) and total dry matter at harvest, and was estimated at 70% on 2010 data. This value is lower than the one found by Leviel (2000) for sugar beet (root + crown: 85%).

A very important parameter of the five models was the thermal time of initiation (or emergence, depending on the model). Indeed, as these models (except STICS) are not designed to predict the initiation time, this parameter must be given as an input. We used a nonlinear mixed model (Baey et al., 2013) to estimate it, along with the two phyllochrons and the rupture thermal time (Milford et al., 1985; Lemaire et al., 2008), to describe the rythm of leaf appearance. They were estimated using two datasets to ensure a more robust estimation (2010 and 2011 experiments, driven on the same genotype), as the data available for the estimation of these parameters in 2010 were not satisfying (small number of plants, and few measures before the rupture point). Indeed, as noticed by Lemaire et al. (2008) and Lemaire (2010), the phyllochrons and the duration of the first phase of development in sugar beet remain stable for a given genotype, and only the thermal time of initiation is subject to change, due for example to environmental conditions. A shift was thus progressively introduced between the two datasets, corresponding to the difference between the thermal times of initiation in the two different years, until the same values were obtained for the two phyllochrons and the rupture thermal time. The smallest AIC was obtained with a shift of 140°Cday.

The other parameters were included in the sensitivity analysis.

2.4. Criteria for evaluating the predictive capacity

The mean squared error of prediction (MSEP) is the standard criterion for evaluating the predictive capacity of a model (Wallach and Goffinet, 1987; Wallach et al., 2006). It measures the distance between observations and predictions, and its square root is used to obtain the same units as the observed and predicted values.

A first dataset is used for the model calibration, and a second one is used to compute

the criterion:

$$\hat{\text{RMSEP}}(\hat{\theta}) = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(Y_i - \hat{Y}_i\right)^2}$$

where Y_i are the observed values, $\hat{Y}_i$ are the predicted values and n is the number of observations in the second dataset.

The modelling efficiency (EF), as defined by Mayer and Butler (1993), is a dimensionless quantity which measures the overall goodness of fit between predictions and observations. It is similar to the coefficient of determination in linear regression.

$$\text{EF} = 1 - \frac{\sum_{i=1}^{n}(Y_i - \hat{Y}_i)^2}{\sum_{i=1}^{n}(Y_i - \overline{Y}_i)^2}$$

where $\overline{Y}_i$ is the mean of observed values. The modelling efficiency ranges from $-\infty$ to 1. In case of a perfect fit, i.e. when predicted and observed values are equal, the modelling efficiency is equal to 1. A value of 0 corresponds to the case where the model predictions are not better than the mean of the observed values, and a negative value is obtained when the predictions perform worse than the mean.

The two criteria defined above give a ranking of the models, the best model being the one with the smallest RMSEP and a modelling efficiency EF as close to 1 as possible. However, as the RMSEP share the same units as the observations, it can be difficult to compare the predictive capacity of a model on variables with different units, whereas the use of EF makes these comparisons easier.

The models were compared on two variables: the dry matter of root and the total dry matter. This allowed for a comparison of biomass production, and a comparison of the biomass allocation to the root. However, for the two models that rely on a constant harvest index for the biomass repartition (Pilote and CERES), the computation of the criteria for the whole time period for the dry matter of root did not make sense, as the harvest index is not supposed to be valid throughout the plant development, but only at harvest. Thus, a third criterion was introduced for the dry matter of root: the yield

prediction error. It is defined as follows:

$$YPE = \frac{|Y_{r,n} - \hat{Y}_{r,n}|}{Y_{r,n}}$$

where $Y_{r,n}$ is the observed root biomass and $\hat{Y}_{r,n}$ the predicted root biomass at harvest.

The criteria were calculated using the vectors of parameters found at the calibration stage for each model. Only the plant density and the thermal time of crop initiation were adapted to 2008 data.

2.5. Modelling platform

The five models presented above were programmed on the C++ modelling platform PyGMAlion (Cournède et al., 2013) developed in the team. It contains all the necessary tools for parametric estimation, sensitivity analysis, and model selection.

3. Results

3.1. Sensitivity analysis and parameter selection

Table A.3 gives the results of the parameter selection procedure. AICc and BIC are provided for the four models included in the analysis. They both selected the same number of parameters, except for Greenlab, where AICc included more parameters than BIC. We finally chose to include 7 parameters in the calibration process, as the fitted values were closer from the corresponding biological values found in the literature. In particular, the RUE was estimated at 5.93 g/MJ when only 4 parameters were included in the calibration process, which is higher than all reference values of RUE for sugar beet reported in the literature.

[Table 3 about here.]

The list of parameters finally estimated for each model is given in Table A.4. The remaining parameters were fixed to the mean value of their corresponding variation interval in the sensitivity analysis (see Table A.2). Expectedly, the radiation use efficiency

had the biggest influence on the model outputs, and was top-ranked by the sensitivity analyses for all models. For LNAS and STICS, calibrating only this parameter was enough to ensure a satisfying goodness of fit, and the calibration of more parameters did not allow for a sufficient decrease of the likelihood. For Greenlab and Pilote, more parameters were necessary to calibrate the models, but we were able to reduce the total number of parameters from 10 to 7 for Greenlab.

[Table 4 about here.]

3.2. Predictive capacity

As stated previously, different versions of a given model can be available, corresponding to different modelling assumptions and their corresponding biological interpretations. It is the case of the STICS model, in our study, where we are using a modified version of the original published STICS model. It is therefore necessary to compare these different models to check whether the simplifications adopted in the paper lessen the predictive capacity of the model.

3.2.1. Comparaison of different versions of STICS

As detailed in section 2.1.3, we adopted in this paper a version of STICS, where the relationship between intercepted radiation and biomass production is linear instead of quadratic, and with a constant efficiency coefficient. We provide here a comparison of the predictive capacity of the different versions of STICS listed in section 2.1.3 (see Table A.5).

[Table 5 about here.]

Interestingly, the modified versions of STICS performed better than the original one, suggesting that the linear formulation of biomass production and the use of a constant efficiency coefficient is more appropriate for prediction in the case of sugar beet. The introduction of thermal stress in the model slightly improved the performances for the dry matter of root prediction, but not for the total dry matter. These results supported

our use of the STICS version with a linear biomass production, a constant efficiency and no consideration of thermal stress for the comparison with the other models.

3.2.2. Comparison on 2008 data

In this section, we present the results for the comparison of the predictive capacity of the five models on 2008 data. We used for STICS the modified version defined in section 2.1.3. A comparison of the different versions of STICS is available in 3.2.1.

The best predictions for the total dry matter were given by LNAS (see Table A.6), followed by CERES. The five models provided good overall predictions, as indicated by the high values of modelling efficiency (above 0.95 for each of them). However, predictions were less accurate at the end of the growth period, as it can be seen on Figure A.1. In the two empirical models Pilote and CERES, the initiation began earlier than in the other models, or more specifically, the produced biomass grew faster, leading to better predictions during the first growth period (until 1000°Cday approximately), compared with other models. However, due to a higher RUE value than in LNAS for example (see Table A.4), they tended to produce overestimated predictions from approximately 1000°Cday. In LNAS, STICS and Greenlab, the initiation was slower, especially for Greenlab, which resulted in underestimated predictions, at least at the beginning of the plant growth. For STICS though, this effect was partly compensated by a very high RUE, and the model provided overestimated predictions from approximately 1500°Cday. LNAS, for its part, was able to better predict the total biomass all along the growth thanks to a lower RUE, even if we observed a decreasing trend at the end. For Greenlab, the underestimations provided by the model seemed to be partly due to a slower initiation that delayed the growth of the simulated plant.

The models performed well also for the root biomass prediction, with high modelling efficiency values. We recall that for Pilote and CERES, the comparison of observed and simulated data throughout the plant development did not make sense as there are no dynamic allocation processes in these models. The empirical harvest index which is used is only supposed to be valid at harvest. However, for the sake of illustration,

we plotted the simulations of these two models according to thermal time, to have an idea of their behaviours. STICS performed outstandingly well on this variable, which means, given its performance on the total dry matter, that the dry matter of leaves was highly underestimated by the model. For LNAS and Greenlab, the dry matter of leaves is certainly over-estimated by these models, given the fact that root biomass was largely under-estimated while total dry matter was only slightly under-estimated. These two models provided highly biased yield predictions, around 15%.

[Table 6 about here.]

[Figure 1 about here.]

[Figure 2 about here.]

4. Discussion

This study is a first attempt to develop a benchmarking approach in a research domain were a lot of models coexist. In this paper, we compared five plant growth models for sugar beet on their capacities to predict the total dry matter and the dry matter of root. The five models shared a comparable energetic production of biomass, but differed in complexity level, modelling scale and handling of biomass repartition.

More generally, the approach developed in this paper can be used not only for the comparison of different models, but also for the comparison of different versions of the same model as presented for STICS, and help to choose between different formulations.

The use of sensitivity analysis allowed us to reduce the variability and to enhance the predictive capacity of the models. They all provided good overall predictions for the total dry matter and the root biomass, with high modelling efficiency values. For Pilote and CERES, a relatively good handling of the plant initiation, coupled with a slightly too high RUE (for Pilote and CERES) resulted in good also overestimated predictions on this data set. For LNAS and Greenlab, a later initiation and a slower RUE also resulted in good but overestimated predictions for the total biomass. In STICS the initiation was

also delayed, even if in a less obvious way than LNAS and Greenlab, but it was partly compensated by a very high RUE value.

Models that relied on a harvest index for the repartition of biomass to root and leaves provided good yield predictions, and total dry matter predictions in the same range as the allocation models. However, the harvest index was deduced from 2010 measurements and estimated around 70%, which is lower than the recommended values found in the literature for these models, around 85%, and slightly lower than the value measured in 2011 of 75%, suggesting that this index might not be very robust. It should also be noted that for Pilote, the LAI curve used for model calibration in 2010 was constructed from mass measurements, while it is normally based on Licor LAI-2000 measurements. This could have probably improved its performance regarding total dry matter production. For CERES, the maximal surfaces $S_{k,\max}$ were computed from individual blade masses and mass per unit area, and may therefore be very variable from one year to another. Moreover, these data are not always available, which can prevent from using this model.

From a practical point of view, if one is only interested in yield and root biomass prediction (which is coherent in the case of sugar beet), STICS can be seen as a good candidate. Only one parameter needs to be estimated after a sensitivity analysis, and moreover, reliable recommended values are available in the literature (Brisson et al., 2008). Environmental stresses can also be easily introduced in the model. Current work on a modified version of STICS based on source-sink relationships suggests that this model is indeed very robust. However, if STICS provided good results for the root biomass in our study, its performance was less good on the total dry matter, due to a high underestimation of leaf biomass.

Results on the total dry matter also suggested that the initiation is a very crucial phase of the plant growth, as a too early or a too fast initiation could lead to overestimated predictions, whereas a too late or a too slow initiation could produce underestimated predictions. More accurate estimations of this parameter can be found using models designed to predict seed germination and seedling emergence (Forcella et al., 2000), or

by introducing the corresponding processes directly in the plant growth models. In any case, a careful attention must be paid to the simulation of this delicate growth phase, both by estimating precisely the thermal time of initiation or emergence, and by modelling adequately the dynamics of the plant during its early phase of growth.

This study is only a first attempt to evaluate the predictive capacities of plant growth models for sugar beet, and these results should be confirmed on other datasets, corresponding for example to more various situations (different genotypes, stressed environmental conditions, ...). Intuitively, we could imagine that the parameters of more mechanistic models would be more genotype-dependent than that of more empirical models (Tardieu, 2003; Letort, 2008; Yin et al., 2004).

Moreover, the biological validity of our results must be further explored. Indeed, the fitted RUE are quite different from one model to another, and slightly higher that the ones that can be found in the literature and that rarely exceed 4 g.MJ^{-1} : Damay and Le Gouis (1993), between 2.96 and 3.76 g.MJ^{-1}, Milford and Riley (1980), between 3.16 and 4.12 g.MJ^{-1}, Biscoe and Gallagher (1977), 3.5 g.MJ^{-1}, ... even if in their book on STICS, Brisson et al. (2008) provided a value of 4.8 g.MJ^{-1} for sugar beet, which is closer from the values that we obtained. This 'overestimation' of the RUE parameters can be due to the fact that not all of the models parameters were estimated in the end, thanks to the sensitivity analysis that allowed us to reduce the dimension of the parameter space, but that also lead to some compensations between the parameter values. This compensation can also arise when the initiation time is not properly estimated, as it can lead to a too high RUE value, as detailed above. In our study, the calibration of more parameters in the Greenlab model allowed us to obtain a more biologically sound value for this parameter. More generally, one should bear in mind that results from sensitivity analysis and model selection as proposed in the first part of our paper are only a guide for the modeller, and the biological interpretation of the fitted models should be checked for carefully.

Finally, the proposed approach is an illustration of "the good modelling practice" (Vos

et al., 2007) that should be implemented in plant growth modelling. When modelling complex systems like plants, it is important to rely on a rigorous methodology and particularly, to define quantitative or qualitative objectives to the models, to find the proper balance between model complexity and robustness, and to rely on objective criteria to evaluate (and compare) the models.

Bibliography

B. Andrieu, J.-M. Allirand, and K. W. Jaggard. Ground cover and leaf area index of maize and sugar beet crops. *Agronomie*, 17:315–321, 1997.

C. Baey, A. Didier, S. Lemaire, F. Maupas, and P.-H. Cournède. Modelling the interindividual variability of organogenesis in sugar beet populations using a hierarchical segmented model. *Ecological Modelling*, 263:56–63, 2013.

P. V. Biscoe and J. N. Gallagher. Weather, dry matter production and yield. In J. Landsberg and C. V. Cutting, editors, *Environmental effects on crop physiology*. Academic Press, 1977.

N. Brisson, B. Mary, D. Ripoche, M.-H. Jeuffroy, F. Ruget, B. Nicoullaud, P. Gate, F. Devienne-Barret, R. Antonioletti, C. Durr, G. Richard, N. Beaudoin, S. Recous, X. Tayot, D. Plenet, P. Cellier, J.-M. Machet, J.-M. Meynard, and R. Delécolle. STICS: a generic model for the simulation of crops and their water and nitrogen balances. I. Theory and parametrization applied to wheat and corn. *Agronomie*, 18(5-6):36, 1998.

N. Brisson, M. Launay, B. Mary, and N. Beaudoin. *Conceptual Basis, Formalisations and Parameterization of the Stics Crop Model*. Quae, 2008.

K. Burnham and D. Anderson. *Model selection and multimodel inference: a practical information-theoretic approach*. Springer Verlag, 2nd edition, 2002.

F. Campolongo, J. Cariboni, and A. Saltelli. An effective screening desing for sensitivity analysis of large models. *Environmental Modelling and Software*, 22:1509–1518, 2007.

J. Cariboni, D. Gatelli, R. Liska, and A. Saltelli. The role of sensitivity analysis in ecological modelling. *Ecological Modelling*, 203(1-2):167–182, 2007.

P.-H. Cournède, A. Mathieu, F. Houllier, D. Barthélémy, and P. de Reffye. Computing competition for light in the Greenlab model of plant growth: a contribution to the study of the effects of density on resource acquisition and architectural development. *Annals of Botany*, 101(8):1207–1219, 2008.

P.-H. Cournède, Y. Chen, Q. Wu, C. Baey, and B. Bayol. Development and Evaluation of Plant Growth Models: Methodology and Implementation in the PyGMAlion platform. *Mathematical Modelling of Natural Phenomena*, 8(4):112–130, 2013.

N. Damay and J. Le Gouis. Radiation use efficiency of sugar beet in northern France. *European Journal of Agronomy*, 2:179–184, 1993.

P. de Reffye and B.-G. Hu. Relevant qualitative and quantitative choices for building an efficient dynamic plant growth model: GreenLab case. In B. G. Hu and M. Jaeger, editors, *First International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications (PMA), Beijing, China*, pages 87–107. Tsinghua University Press and Springer, 2003.

F. Forcella, R. L. B. Arnold, R. Sanchez, and C. M. Ghersa. Modeling seedling emergence. *Field Crops Research*, 67:123–139, 2000.

T. Fourcaud, X. Zhang, A. Stokes, H. Lambers, and C. Korner. Plant growth modelling and applications: The increasing importance of plant architecture in growth models. *Annals of Botany*, 101(8):1053–1063, 2008.

M. Guérif and C. Duke. Calibration of the SUCROS emergence and early growth module for sugar beet using optical remote sensing data assimilation. *European journal of agronomy*, 9:127–136, 1998.

C. A. Jones and J. Kiniry, editors. *CERES-Maize: a simulation model of maize growth and development.* Texas A&M University Press, College Station, Temple, TX, 1986.

S. Lemaire. *Système dynamique de la croissance et du développement de la betterave sucrière (Beta vulgaris L.)*. PhD thesis, AgroParisTech, 2010.

S. Lemaire, F. Maupas, P.-H. Cournède, and P. de Reffye. A morphogenetic crop model for sugar-beet (*beta vulgaris l.*). In *International Symposium on Crop Modeling and Decision Support: ISCMDS 2008, Nanjing, China*, April 19-22 2008.

V. Letort. *Adaptation du modèle de croissance GreenLab aux plantes à architecture complexe et analyse multi-échelle des relation source-puits pour l'identification paramètrique.* PhD thesis, École Centrale Paris, 2008.

B. Leviel. *Evaluation des risques et maîtrise des flux d'azote au niveau d'une parcelle agricole dans la plaine roumaine et bulgare. Application aux cultures de mais, blé, colza et betterave.* PhD thesis, Institut National Polytechnique de Toulouse, 2000.

Z. Li, V. Le Chevalier, and P.-H. Cournède. Towards a continuous approach of functional-structural plant growth. In B.-G. Li, M. Jaeger, and Y. Guo, editors, *3rd international symposium on Plant Growth and Applications(PMA09), Beijing, China*. IEEE, November 9-12 2009.

J.-C. Mailhol, P. Revol, and P. Ruelle. Pilote : un modèle opérationnel pour déceler l'apparition de stress hydrique. In *ICID 16th international congress on irrigation and drainage : workhop on crop-water-environment models*, Cairo, Egypt, July 1996.

J.-C. Mailhol, A. Olufayo, and P. Ruelle. Sorghum and sunflower evapotranspiration and yield from simulated leaf area index. *Agricultural Water Management*, 35:167–182, 1997.

D. Mayer and D. Butler. Statistical validation. *Ecological Modelling*, 68:21–32, July 1993.

G. Milford, T. Pocock, and J. Riley. An analysis of leaf growth in sugar beet. II. Leaf appearance in field crops. *Annals of Applied Biology*, 106:173–185, 1985.

G. F. J. Milford and J. Riley. The effects of temperature on leaf growth of sugar beet varieties. *Annals of Applied Biology*, 94(3):431–443, 1980. ISSN 1744-7348.

J. Monteith. Climate and the efficiency of crop production in Britain. *Proceedings of the Royal Society of London*, 281:277–294, 1977.

F. Ruget, N. Brisson, R. Delécolle, and R. Faivre. Sensitivity analysis of a crop simulation model, STICS, in order to choose the main parameters to be estimated. *Agronomie*, 22:133–158, 2002.

A. Saltelli, M. Ratto, T. Andres, and F. Campolongo. *Global sensitivity analysis: the primer*. Wiley, 2008.

R. Sievänen, E. Nikinmaa, P. Nygren, H. Ozier-Lafontaine, J. Perttunen, and H. Hakula. Components of functional-structural tree models. *Annals of Forest Science*, 57(5):399–412, 2000.

C. Spitters, H. van Keulen, and D. van Kraalingen. A simple and universal crop growth simulator: SUCROS87. In R. Rabbinge, S. Ward, and H. H. van Laar, editors, *Simulation and systems management in crop protection*, pages 147–181. Pudoc, Wageningen, 1989.

A. Taky. *Maîtrise des excès d'eau hivernaux et de l'irrigation et de leurs conséquences sur la productivité de la betterave sucrière dans le périmètre irrigué du Gharb (Maroc). Analyse expérimentale et modélisation*. PhD thesis, AgroParisTech, 2008.

F. Tardieu. Virtual plants: modelling as a tool for the genomics of tolerance to water deficit. *Trends in Plant Science*, 8(1):9–14, 2003.

J. Vos, L. Marcelis, P. De Visser, P. Struik, and J. Evers, editors. *Functional-structural plant modelling in crop production*, chapter 1. Springer, 2007.

D. Wallach and B. Goffinet. Mean Squared Error of Prediction in Models for Studying Ecological and Agronomic Systems. *Biometrics*, 43(3):561, Sept. 1987.

D. Wallach, B. Goffinet, J.-e. Bergez, P. Debaeke, and D. Leenhardt. Parameter Estimation for Crop Models : A New Approach and Application to a Corn Model. *Agronomy Journal*, 93(4):757–766, 2001.

D. Wallach, D. Makowski, and J. Jones. *Working with dynamic crop models: evaluation, analysis, parameterization, and applications*, chapter Evaluating crop models, pages 11–53. Elsevier Science Ltd, 2006.

D. Watson. Comparative physiological studies in the growth of filed crops. I. Variation in net assimilation rate and leaf area between species and varieties, and within and between years. *Annals of Botany*, 11:41–76, 1947.

Q.-L. Wu, P.-H. Cournède, and A. Mathieu. An efficient computational method for global sensitivity analysis and its application to tree growth modelling. *Reliability Engineering and System Safety*, July

2011.

H.-P. Yan, M.-Z. Kang, P. de Reffye, and M. Dingkuhn. A Dynamic, Architectural Plant Model Simulating Resource-dependent Growth. *Annals of Botany*, 93(5):591, 2004.

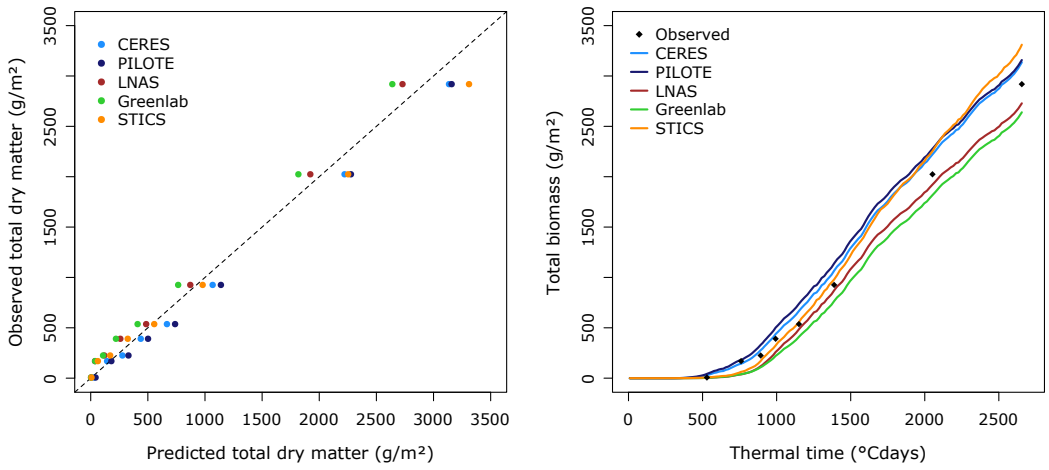
X. Yin, P. Struik, and M. Kropff. Role of crop physiology in predicting gene-to-phenotype relationships. *Trends in Plant Science*, 9(9):426–432, 2004.

AppendixA. Parameters

The models presented in this paper involve a lot of different parameters, that are presented in the table below.

[Table 7 about here.]

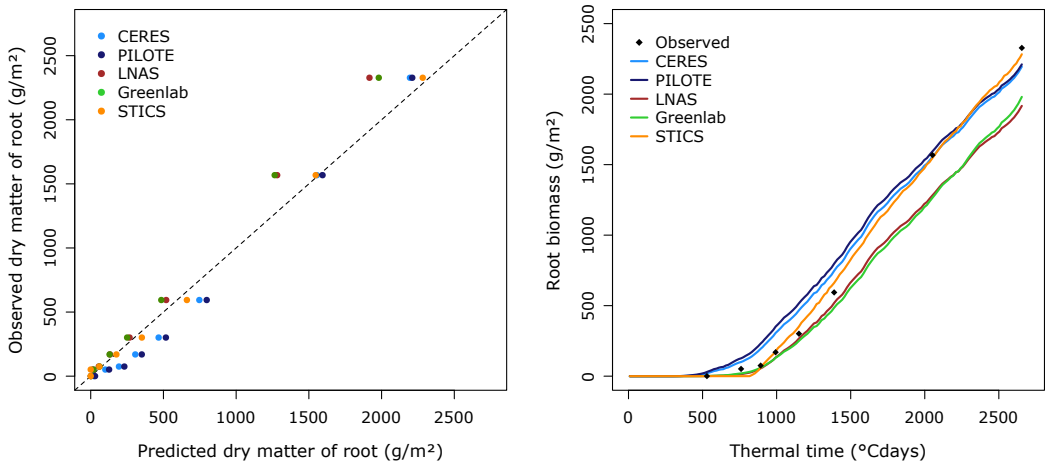
List of Figures



(a) Observed vs. predicted

(b) Over thermal time

Figure A.1: Models' predictions for the total dry matter in 2008



(a) Observed vs. predicted

(b) Over thermal time

Figure A.2: Models' predictions for the dry matter of root in 2008

List of Tables

Table A.1: Classification of the five models according to the modelling scale and the presence or absence of allocation processes.

	Modelling scale (per m^2)	**Modelling scale (per plant)**
No allocation	Pilote	CERES
Allocation	STICS LNAS	Greenlab

Greenlab	Pilote	LNAS
μ : [5;6]	RUE : [3;4]	RUE : [3;4]
S_p : [0.05;0.10]	α : [1;2]	e_g : [50;70]
p' : [0.3;0.45]	β : [0.5;3.5]	μ_a : [400;800]
p'' : [1.35;1.65]	τ_{max} : [800;1200]	σ_a : [200;2000]
α_b : [3.38;3.55]	LAI_{max} : [4;6]	μ_s : [2000;3000]
β_b : [5;5.45]		σ_s : [3000;6000]
α_p : [3.7;3.8]		γ_o : [0.7;1]
β_p : [5.25;5.40]		γ_f : [0;0.3]
α_r : [4;4.6]		
β_r : [2.1;2.8]		

Table A.2: Variation intervals for the model parameters. We refer to Brisson et al. (2008) for recommended values for STICS. The variation intervals were then defined with a 10% variation around these reference values.

Number of parameters		1	2	3	4	5	6	7	8	9	10
STICS	AICc	**4.63**	6.77	8.29	10.82	13.55	16.48	19.63	23.02	26.68	30.65
	BIC	**6.06**	9.50	12.18	15.70	19.26	22.81	26.37	29.92	33.48	37.03
Greenlab	AICc	166.91	74.40	74.10	59.46	60.77	62.54	**50.83**	51.20	53.25	55.31
	BIC	171.46	83.48	87.71	**77.59**	83.42	89.71	82.51	87.37	93.92	100.47
LNAS	AICc	**5.05**	7.12	9.44	11.77	14.34	17.05	19.92	22.95	-	-
	BIC	**6.71**	10.35	14.11	17.76	21.52	25.29	29.05	32.81		
PILOTE	AICc	47.08	45.50	16.15	17.63	**15.19**	-	-	-	-	-
	BIC	48.26	47.68	19.14	21.22	**19.13**					

Table A.3: AICc and BIC according to the number of parameters in the model. Results for STICS for more than 10 parameters are not shown as the log-likelihood is constant, and the two criteria are increasing.

Model	**Data used for the calibration**	**Estimated parameters**
Greenlab	Dry matter of root, blades and petioles Individual masses of blades and petioles	RUE = 4.03 g/MJ $\alpha_r = 3.16$ $\beta_r = 1.04$ $p_p = 0.0039$ $\alpha_b = 3.08$ $q_p = 1.70$
LNAS	Dry matter of root Dry matter of green leaves Dry matter of senescent leaves	RUE = 3.53 g/MJ
PILOTE	Total dry matter Leaf area index	RUE = 4.12 g/MJ $\alpha = 1.54$ $\beta = 1.92$ $\tau_{max} = 1830$ °Cday $\text{LAI}_{max} = 3.99$
CERES	Total dry matter	RUE = 4.37 g/MJ
STICS	Dry matter of root Dry matter of green blades Total dry matter	RUE = 4.76 g/MJ

Table A.4: Data used in the calibration step for the five models, and list of the parameters finally estimated.

STICS version	Total dry matter		Dry matter of root	
	RMSEP	**EF**	**RMSEP**	**EF**
Initial	271.24	0.9217	59.26	0.9945
Linear	192.3	0.9606	39.05	0.9976
Linear and cst RUE	168.87	0.9696	39.49	0.9976
Thermal stress	171.73	0.9686	38.61	0.9977

Table A.5: Evaluation criteria for the different versions of STICS on 2008 data

Table A.6: Evaluation criteria for the five models on 2008 data.

Model	**Total dry matter**		**Dry matter of root**		
	RMSEP	**EF**	**RMSEP**	**EF**	**YPE**
Greenlab	166.61	0.970	169.34	0.955	14.91%
LNAS	110.85	0.987	180.5	0.949	17.66%
CERES	127.02	0.983	-	-	5.70%
PILOTE	170.51	0.969	-	-	5.02%
STICS	168.87	0.970	39.5	0.998	1.92%

Model	Parameter	Unit	Meaning
All	RUE	$g.MJ^{-1}$	Radiation use efficiency
	PAR	$MJ.m^{-2}$	Photosynthetically active radiation
	k_B	-	Extinction coefficient of the Beer-Lambert law
Greenlab	μ	$g.MJ^{-1}.pl^{-1}$	Efficiency coefficient per plant related to the RUE in the following way : RUE $= \mu \cdot S_p \cdot d$
	S_p	m^2	Parameter related to the 2-D projection of the space occupied by the plant on the floor
	e_b	$g.m^{-2}$	Mass of blade per unit area of blade
	p'	-	Sink strength of petioles
	p''	-	Correction of p' according to a competition index between blades and petioles
	α_o, β_o	-	Parameters of beta law for the sink function of organ o ($o = b$ for blades, $o = p$ for petioles and $o = r$ for root)
LNAS	e_g	$g.m^{-2}$	Mass of leaf per unit area of leaf
	μ_a	-	Median of the lognormal law assumed for the allocation process
	σ_a	-	Standard error of the lognormal law assumed for the allocation process
	μ_s	-	Median of the lognormal law assumed for the senescence process
	σ_s	-	Standard error of the lognormal law assumed for the senescence process
	γ_0	-	Initial proportion of biomass allocated to leaves
	γ_f	-	Final proportion of biomass allocated to leaves
PILOTE	α, β	-	Empirical parameters for the LAI curve
	τ_{max}	°Cday	Thermal time at which the LAI is maximal
	LAI_{max}	-	Maximal value of the LAI
	τ_e	°Cday	Thermal time of emergence
STICS	α, β	-	Parameters of the logistic curve for the LAI growth
	u_{mat}	-	Leaf development unit at the inflexion point of the logistic curve