Evaluation of the Predictive Capacity of Five Plant Growth Models for Sugar Beet

Charlotte Baey, Li Song, Paul-Henry Cournède
Laboratory of Applied Mathematics and Systems
École Centrale Paris
Châtenay-Malabry, France
charlotte.baey@ecp.fr

Sébastien Lemaire, Fabienne Maupas
Institut Technique de la Betterave
Paris, France

Abstract—**A lot of plant growth models coexist, with different modelling approaches and levels of complexity. In the case of sugar beet, many of them are used as predictive tools, even when they were not originally designed for this purpose.**

We propose the evaluation and comparison of five plant growth models that rely on the same energetic production of biomass, but with different levels of description (per plant or per square meter) and different biomass repartition (empirical or via allocation): Greenlab, LNAS, CERES, PILOTE and STICS. The models were calibrated on a first set of data, and their predictive capacities were compared on an independent data set from the same variety and similar environmental conditions, using the root mean squared error of prediction (RMSEP) and modelling efficiency (EF) for the total dry matter production and the dry matter of root.

All the models tended to overestimate both the total dry matter and the dry matter of root. Greenlab gave the best predictions for the root biomass, and CERES the best total biomass predictions. The overestimation was partly explained by a hail episode that caused a lot of damages to the leaves in the validation year. The five models also provided similar yield prediction errors.

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

I. INTRODUCTION

Plant growth models can have many different applications. They can be used to enhance our understanding of the complex interactions between genotypes, cultural practices and environment, and to help management decisions. They can also be used to simulate plant architecture at organ level, or to predict yield at field level. Depending on the scale of interest, the level of description and the integration of environmental stresses, model complexity can increase dramatically, as well as the number of parameters.

However, depending on the objective of the study, some models can be more relevant than others. For example, complex models can be more suitable for descriptive purposes as they integrate in details the underlying eco-physiological processes. On the other hand, they can prove to be less robust than simpler models when used as prediction tools, due for example to a higher number of parameters leading to an increase in variability. It is therefore essential to define precisely the context in which a model is used, and to evaluate its performance according to the objective of the study.

In this paper, we address the question of root and biomass prediction in sugar beet crops. The predictive capacity and robustness of five models with different levels of description was evaluated: Greenlab [1], [2], CERES [3], [4], PILOTE

TABLE I
CLASSIFICATION OF THE FIVE MODELS ACCORDING TO THE MODELLING SCALE AND BIOMASS REPARTITION.

	Modelling scale (per m^2)	Modelling scale (per plant)
Harvest Index	PILOTE	CERES
Allocation	LNAS STICS	Greenlab

[5], STICS [6], [7] and a fifth model named LNAS, 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 [8].

The five models rely on an energetic production of biomass, based on Monteith's equation [9]. In such models, the total dry matter production is related to the fraction of radiation intercepted by the leaves, which can generally be expressed according to the leaf area index (LAI). The differences between the models lie in the way the evolution of this key variable is modelled, and in the handling of biomass repartition to the different organs of the plant. Indeed, two modelling scales were compared for the intercepted radiation: on one hand, we considered individual based models taking into account the development of each leaf individually (CERES, Greenlab), and on the other hand we considered classical crop models at the field scale, where foliage biomass is given per square meter (PILOTE, STICS, LNAS). Two approaches were also compared for the biomass repartition, either with empirical coefficients like the harvest index for the dry matter of root (PILOTE, CERES), or as a result of the allocation process (Greenlab, STICS, LNAS). In fact, STICS is an intermediate between the two approaches, since the LAI is computed from an empirical function, but modulated by a source-sink ratio. The differences between the five models are summarized in Table I.

The five models were programmed on a C++ modelling platform called PYGMALION, developed in the MAS laboratory of Ecole Centrale Paris, which provided all the necessary tools, including parametric estimation, sensitivity analysis and model selection. 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 the root mean squared error of prediction (RMSEP) and the modelling efficiency (EF).

IEEE PRESS

II. Material and Methods

A. Description of the models

The five models rely on an energetic production of biomass, based on an extension of the Beer-Lambert law. The biomass production on day t, $Q(t)$ is proportional to the incoming photosynthetically active radiation $PAR(t)$, 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 [10]:

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

The leaf area index is defined as the one-sided green leaf area per unit ground surface, some adjustments were thus necessary for the two individual-based models Greenlab and CERES. In Greenlab, as the biomass production is computed at the individual plant level, a 'local' LAI [11] is 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 II-A1). In CERES, as the biomass production is computed at the square meter level, a 'global' LAI is constructed from the individual leaf surfaces of the plant, by multiplying by the crop density (see II-A3).

1) GreenLab: GreenLab is a generic functional-structural plant model (FSPM), combining the description of the plant architecture and its physiological functioning [12], [13]. The model in its discrete version was introduced by [1], and has already been studied in the case of sugar beet by [14].

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, the choice of a discrete time step may involve several difficulties: for example, environmental data like temperature or global radiation vary continuously with time, whereas the growth cycle is supposed to be constant for a given plant, and could vary from several days to one year for trees. In this case, changes in the environment occurring during the growth cycle are not taken into account in the model. Moreover, environmental data are traditionally collected on a daily basis, and the other plant growth models studied in this paper provide daily outputs [5], [15], [16].

To overcome these difficulties and make the comparison easier, a continuous version of the model with a daily time-step was built, in which the functioning of the plant (biomass production and biomass allocation) was considered as a continuous process and was driven by a system of differential equations, solved with first-order Euler's method. This continuous version of Greenlab was used throughout the article. A continuous version of the GreenLab model has already been studied by [17] in the case of sugar beet, but with a continuous mechanism of senescence.

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

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

with d the plant density, k_B the Beer-Lambert law extinction coefficient, $Q_b(t)$ the accumulated blade mass at day t, e_b the specific blade mass, and S_p an empirical coefficient related to the two-dimensional projection of the space occupied by the plant. The biomass is then allocated to the different organs of the plant according to source-sinks relationships (we refer the reader to [1], [2] for more details).

Several assumptions are made in the Greenlab model. First, we assumed that the produced biomass is gathered in a common pool before being allocated to the different organs of the plant. Then, even if we adopted a continuous formulation of the model for the functioning part, the structural part of the model (i.e. organogenesis) still occurred at integer multiples of the phyllochron. The sink strengths of organs are also modelled as empirical functions (density of beta functions).

2) LNAS: A simplified model called LNAS (Log-Normal Allocation and Senescence) was elaborated, with a global biomass allocation for the foliage compartment, instead of leaf by leaf as in the Greenlab model. 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 per square meter at day t is given by (1), with:

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

where $Q_g(t)$ is the dry matter of green leaves at day t, k_B the extinction coefficient and e_l the specific green leaf mass.

The biomass produced at day t is then allocated to each organ compartment (foliage and root system in the case of sugar beet) at the beginning of day $t + 1$, according to the following equations:

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

with $Q_l(t)$ the total mass of leaves and $Q_r(t)$ the total mass of root at the beginning of day t.

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:

$$Q_g(t) = (1 - G_s(\tau(t))) Q_l(t),$$

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.

The main assumption of this model concerns the functional form of allocation and senescence, which are modelled by log-normal laws.

3) CERES: CERES (Crop Environment REsource Synthesis) was originally built on maize by [3], but a sugar beet version was developed by [4]. Effects of irrigation or nitrogen uptake can 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, and S_k the foliar surface of leaf k at time t.

The foliar surface of leaf k, S_k is supposed to grow linearly at rate GR 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:

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

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 inputs of the CERES model.

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

In CERES, the main assumptions concern the linear growth of individual leaf surfaces, before reaching a plateau, and the computation of the dry matter of root at harvest thanks to a harvest index.

4) PILOTE: PILOTE is a crop-soil interaction model, which was first built for sorghum and sunflower [5], [18], but that can be applied to a large variety of crops. It has been developped for sugar beet by [19]. 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 studied a simplified version of PILOTE, with no consideration of hydric stress. In such case, the biomass production per square meter at day t is given by (1), with:

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

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

with k_B the Beer-Lambert law extinction coefficient, $\text{LAI}_{\max}$ the maximum reachable value of LAI in non-limiting conditions, $\tau_{\max}$ the thermal time necessary to reach this maximal LAI, τ_e the thermal time of emergence, and α and β two

parameters. Then, the biomass repartition to root and leaves is done with an empirical harvest index.

The main assumptions of PILOTE come from the functional form of the LAI curve, and from the use of a harvest index to evaluate the dry matter of root at harvest.

5) STICS: STICS (Simulateur mulTIdisciplinaire pour les Cultures Standard [6], [7]) is a generic daily time-step model, which has already been applied to a large variety of crops (maize, tomato, wheat, ...). It is organized into seven modules corresponding to the different mechanisms involved in the 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 saturating coefficient, and the efficiency can vary according to the development stage. However, to ensure that the comparison will only concern the biomass repartition and the modelling scale, a linear relationship with a constant radiation use efficiency was modelled. The two versions of the model gave similar results in terms of RMSEP and EF. We have :

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

The leaf area index is modelled through 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 [7]. We have:

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

where t_e is the day of emergence, $\Delta\text{LAI}(j)$ is the net leaf area growth on day j and $\Delta\text{LAI}_{\text{sen}}(j)$ 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:

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

where u_{mat} is the leaf development unit at the end of the juvenile stage, d is the plant density and f_d a density factor related to the competition between plants, f_T is the effective crop temperature, and s is a trophic stress index. This trophic stress is determined by a source-sink ratio, and thus induce a retroaction of allocation on the LAI curve. With this formulation, the leaf area index stops growing all at once after having reached its maximal point, but it is possible to introduce a progressive decline of the LAI. We refer the reader to [7] for more detailed equations.

A lot of parameters are required for the model (such as the soil pH, the organic nitrogen content, the albedo, ...), but a

list of recommended values for different crops are available in [7].

The main assumptions of STICS include the quadratic relationship between biomass production and intercepted radiation, the modelling of the first phase of LAI growth by a logistic curve, and the fact that the ratio between blades and petioles is supposed to be constant. Petioles are also considered as stems in the STICS version for sugar beet, which means that they are not involved in the senescent process. Finally, the leaf mass variable is not directly related to the biomass production since it is determined by the LAI curve.

B. Calibration of the models

1) 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 squared 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 were collected on 10 plants at five different dates. Mean values were then used for the calibration process. Leaf area index was not measured directly on the field, but could be computed from the blade mass Q_b, the specific blade mass e_b and the plant density d: $\text{LAI}_{\text{exp}} = \dfrac{Q_b}{e_b} d$. This relation was used to compute the maximum LAI and the corresponding thermal time. The specific blade mass was obtained from a linear regression between leaf surfaces and blade masses on the five dates of individual measurements.

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 [14].

2) Calibration: The parameters of the five models were estimated using generalized least squares method on 2010 data, as described in [20], but not necessarily on the same subset of data, depending on the model formulation. For example, the individual masses of blades and petioles were used for the calibration of Greenlab, whereas only the total dry matter was needed for CERES. The list of variables used for the calibration of each model is given in table II.

For each model, the parameters were divided into two parts: those that can be deduced from the literature or measured directly, and those that have to be estimated by the models. The following parameters were not estimated: the extinction coefficient of the Beer-Lambert law ($k_B = 0.7$ according to [21]), the specific blade mass in Greenlab (deduced from our field experiments, see II-B1), the maximal LAI value and thermal time to reach this maximum value in PILOTE (deduced from our field experiments, see II-B1), and the maximum leaf surfaces for CERES (deduced from the individual

Model	Data used for the calibration	Nb of estimated parameters
Greenlab	Dry matter of root, blades and petioles Individual masses of blades and petioles Thermal times of leaf expansion and senescence	10
LNAS	Dry matter of root Dry matter of green leaves Dry matter of senescent leaves	6
PILOTE	Total dry matter Leaf area index	3
CERES	Total dry matter Dry matter of root Thermal times of leaf expansion Maximum leaf surfaces	1
STICS	Dry matter of root Dry matter of green blades Dry matter of senescent blades	4

measurements). The harvest index HI 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 [4] for sugar beet (root + crown: 85%). The thermal time of initiation, as well as the two phyllochrons and the rupture thermal time [14], [22] were estimated using the two datasets (calibration and test). Indeed, 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) and thus the two datasets were used to ensure a more robust estimation of these crucial parameters. Indeed, as noticed by [14] and [8], 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. We shifted the thermal time of initiation of 2010 data to obtain similar values of phyllochrons and rupture time for the two years. The smallest AIC was obtained with a shift of 140°Cd.

A preliminary study based on global sensitivity analysis (SA) was carried out on STICS, to identify the parameters that must be estimated, and those that could be fixed to reliable recommended values available in the literature [7]. Indeed, a large number of parameters drive the equations of STICS, and, even if recommended values are provided for the large majority of them [7], the use of estimated values can enhance the quality of the model (e.g. if our experimental conditions are too different from the ones on which the recommended values were measured). Yet, estimating too many parameters would increase the model's variability and could lead to a poor predictive capacity. As the objective of this study was to build predictive models (or, at least, to evaluate the predictive capacity of the models), the choice of a SA method seemed appropriate, to rank the parameters according to their influence on the model outputs. A review on the role of sensitivity analysis can be found in [23].

The standardized regression coefficient (SRC) method was used to rank the parameters according to their influence on the model outputs. Sensitivity indices were computed for each parameter, each output, and each time of observation. Then, sensitivity indices (SI) for each parameter and each output were obtained using a weighted sum over each time of observation, and finally, a global SI for each parameter was computed, taking into account all the model's outputs. Parameters were then ranked according to their SI. The use of SRC indices was validated by computing the linearity index of the model that was found to remain very high all along the crop growth. A similar study has already been carried out for STICS, on wheat and maize, based on response surface method [24].

In the simplified version of STICS used in this paper, environmental stresses were not taken into account, and not all of the STICS modules were included. For these reasons, some parameters were not included in the model, and we reached a total of 18 parameters. Three threshold parameters could not be estimated by the Gauss-Newton algorithm and were fixed to their recommended values; all the other parameters were included in the sensitivity analysis. The method was applied to three outputs: the dry matter of green blades, the dry matter of senescent blades, and the dry matter of root. Once the ranks of the parameters were obtained, the AIC criterion was used to determine the number of parameters to be estimated (the others being set to their recommended values).

The final numbers of parameters estimated for each model are given in table II. Greenlab is the model with the highest number of parameters to estimate, and needs 5 input variables for its calibration.

C. Evaluation of the models

1) Data: The predictive capacity of the models was evaluated on 2011 data. Fields experiments were conducted near the 2010 site, at Bourgogne, France, N49°21'18", E4°4'12" on a calcareous loam soil. The same commercial variety Python was sown on March 21, with 45 cm between rows and 18 cm between seed-plots, and fertilized with 113 kg N ha^{-1}. The final plant density was estimated at 10.89 pl/m^2.

Dry matter of root and leaves (blades and petioles separately or altogether) were measured at fifteen different dates.

Daily mean temperatures (°C) and solar radiation (MJ.m^{-2}) were obtained from French meteorological services. A hail episode occurred on June 29, which caused a lot of damages on leaves, with partial or total destruction of blades and petioles. The hail effect is discussed in the Results section.

2) Comparison criteria: Two criteria were used for the models comparison: the root mean squared error of prediction (RMSEP) and the modelling efficiency (EF). The mean squared error of prediction (MSEP) measures the distance between observed and predicted values, and its square root is used to obtain the same units as the observed and predicted

values [25]. It is given by [26], [27]:

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

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

The modelling efficiency (EF), as defined by [27], 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's 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 RMSEP gives an estimation of the mean error amplitude, whereas EF has no dimension and can therefore be more useful to compare variables with different units.

The models were compared on two variables: the dry matter of root and the total dry matter. This allowed for a comparison of the functions of biomass production, and a comparison of the biomass allocation to the root. However, for the two models that relied 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 HI is not supposed to be valid throughout the plant development, but only at harvest. Thus, RMSEP and EF for the dry matter of root were not computed for these models. Instead, we provided the yield prediction error (i.e. the relative error of prediction of the dry matter of root at the harvest date) for the five models.

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 initiation were adapted to 2011 data.

III. RESULTS

The root mean squared error and modelling efficiency for the five models are given in Table III, and predicted vs. observed plots can be seen on Figs. 1 and 2. By way of illustration, we added the predictions from CERES and PILOTE for the dry matter of root, considering a constant proportion HI of allocation to root.

The best predictions for the dry matter of root were given by Greenlab with a RMSEP of 154.9g/m^2, followed by STICS (RMSEP = 175.6 g/m^2) and LNAS (RMSEP = 229.6g/m^2). The modelling efficiency was high (above 90%) for these three models, indicating a good overall fit between predictions and observations. However, all the models provided over-estimated predictions for this variable.
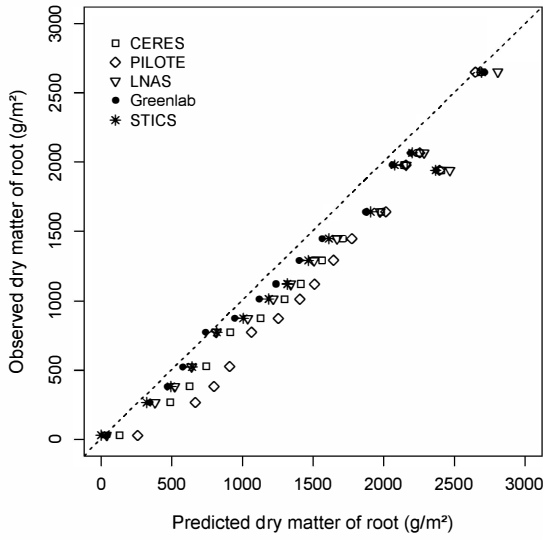
(a) Observed vs. predicted
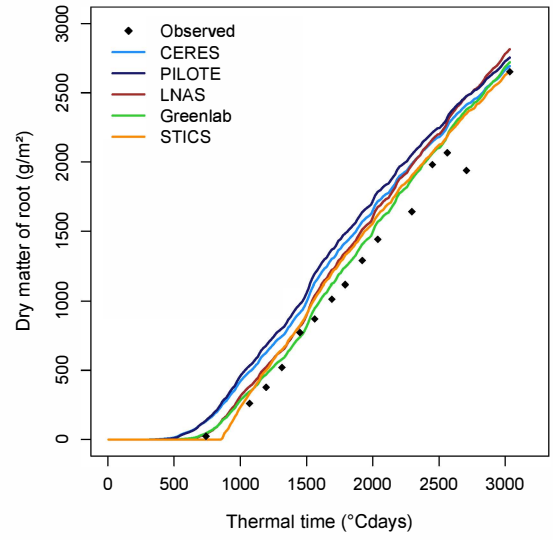
(b) Over thermal time

FIGURE 1. Models' predictions for the dry matter of root in 2011

TABLE III
EVALUATION CRITERIA FOR THE FIVE MODELS ON 2011 DATA

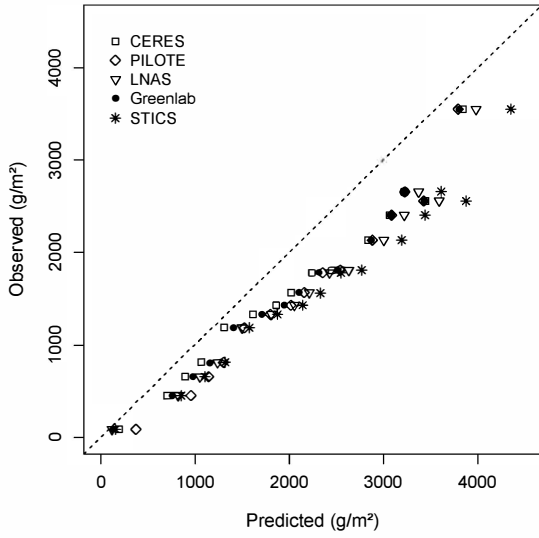| Model	Dry matter of root		Total dry matter	
	RMSEP	EF	RMSEP	EF
Greenlab	154.9	0.96	521.8	0.67
LNAS	229.6	0.90	632.6	0.51
CERES	-	-	483.5	0.71
PILOTE	-	-	564.2	0.61
STICS	175.6	0.94	670.0	0.45
STICS initial	168.1	0.95	664.8	0.46

Models based on allocation processes performed better at the time of initiation than models relying on a constant HI for biomass repartition (see Fig. 1). Indeed, the repartion of biomass to root and leaves is not constant over time, and could be influenced for example by environmental conditions, or by treatments applied to the crops [28]. However, these two models do not make the assumption of a constant repartition throughout the plant development, but only at harvest. Thus, we used a third criterion to compare the five models, by computing the yield prediction error, i.e. the dry matter of root at the harvest time: 2.6% for Greenlab, 6.1% for LNAS, 1.5% for CERES, 3.9% for PILOTE and 0.1 % for STICS (1.8% for the initial version of STICS). This results suggest that the use of a HI might be appropriate if one is only interested in the yield prediction, as all the models provided similar results.

The dry matter of root predicted by STICS increased slower than in the other models, when the predictions from LNAS increased at a higher rate. The three other models converged towards a similar value at harvest.
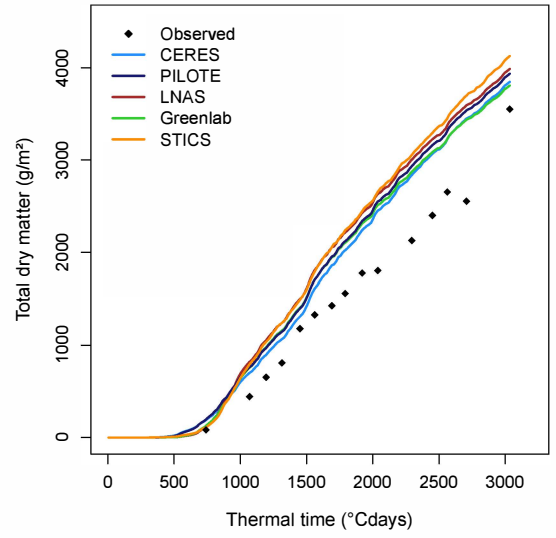
For the total dry matter, the best predictions were given by CERES with a RMSEP of $483.5 g/m^2$, followed by Greenlab (RMSEP = 521.8 g/m^2) and PILOTE (RMSEP = 564.2 g/m^2). The modelling efficiency was lower than for the dry matter of root, indicating that this variable was not well predicted by the models. Indeed, as it has been said before, a hail episode occurred at around $1500°$Cd, and caused a lot of damages to leaves. The dry matter of leaves was reduced drastically due to partial or total destruction of blades and petioles, which resulted in a decrease of the total dry matter. The impact of this climatic event can be clearly seen on Fig. 2, for a period of approximately $1000°$Cd, but the loss of leaf biomass and hence the induced deficit of photosynthesis was probably compensated by the remobilization of assimilates from root to leaves, as the observed total dry matter at harvest was more consistent with the models' predictions. However, all the models tended to overestimate the dry matter production even before this climatic episode, especially PILOTE and CERES, in which the initiation of biomass prediction took place earlier than in the other models (see Fig. 2).

Interestingly, the model with the fewest number of parameters (CERES) gave the best predictions and provided a good yield prediction, where a model like Greenlab gave less good results for biomass prediction, but with 9 parameters more to estimate.

Greenlab provided good results, specially for the total dry matter, but individual blades and petioles masses are needed for the calibration, and 10 parameters need to be estimated. LNAS, which is simpler than Greenlab in the formulation and thus in the number of parameters, did not give very good results for the dry matter of root, and a big yield

(a) Observed vs. predicted

(b) Over thermal time

FIGURE 2. Models' predictions for the total dry matter in 2011

prediction error compared to the other 4 models. It is easier to calibrate than Greenlab, and no individual measurements are needed for the calibration. CERES performed very well on the yield prediction, and also on the total biomass prediction. Only one parameter needs to be estimated, but the individual leaf surfaces are needed to build the model, and can be difficult to obtain. PILOTE gave good results on the yield prediction, and average results for the total biomass prediction. It is easy to calibrate, and no individual measurements are needed. Real LAI measurements based for example on Licor-2000 measurements should provide a better experimental LAI curve for calibration than the one constructed from mass measurements and used here. For STICS, even if it depends on a large number of parameters, recommended values are available for all of them, and a sensitivity analysis can help reduce the number of parameters to be estimated. The model gave good results for the dry matter of root, but less good results for the total biomass production.

IV. 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 on complexity level, modelling scale and handling of biomass repartition. Models that relied on a harvest index for the repartition of biomass to root and leaves provided good yield predictions, and similar total dry matter predictions as the allocation models. Indeed, the five models had similar performances, regardless of their complexity, and there is not one model

that performed outstandingly better than the others. However, the harvest index was deduced from 2010 measurements, and was found to be lower than the recommended values found in the literature for these models, and slightly lower than the value measured in 2011, suggesting that this index might not be very robust.

From an experimental point of view, i.e. with a need for a compromise between the ease of calibration and the robustness of the model, CERES can be seen as a good alternative to more complex models like Greenlab, but individual leaf surfaces are needed. In the same way, PILOTE is very easy to calibrate, but gave less good results than CERES.

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 probably improve its performance regarding total dry matter production. For CERES, the maximal surfaces $S_{k,\max}$ were computed from individual blade masses and specific blade mass, and may therefore be very variable from one year to another. Moreover, the computation of the thermal time of initiation was made on data from the two years of experiment, whereas in view of prediction, this variable should be estimated from calibration data only. This point can be enhanced. Also, data needed for the calibration of each model differ, which may affect the quality of the results.

The hydric stress was not taken into account in this study, even though a small water deficit occurred from 1700°Cd. The introduction of water stress in the models could also enhance their performances, especially for models like PILOTE which was specifically designed for the handling of crop-soil interactions and water budget. A study of the capacity of the

SUCROS model to predict sugar beet yields can be found in [29]. They computed the root mean squared error of final yield predictions of the initial SUCROS model and a modified version, in 20 field plots, taking into account hydric stress, and found that the integration of hydric stress in the model reduced the RMSEP by a half.

More generally, it would also be interesting to investigate the robustness of the models and their ability to adapt to more various situations, for example by evaluating their predictive capacity in different environmental or site conditions, or by considering other genotypes. Intuitively, we could imagine that the parameters of more mechanistic models would be more genotype-dependent than that of more empirical models [30]–[32].

REFERENCES

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

[2] H.-P. Yan, M.-Z. Kang, P. de Reffye, and M. Dingkuhn, "A Dynamic, Architectural Plant Model Simulating Resource-dependent Growth," *Annals of Botany*, vol. 93, no. 5, p. 591, 2004.

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

[4] 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 maïs, blé, colza et betterave." Ph.D. dissertation, Institut National Polytechnique de Toulouse, 2000.

[5] J.-C. Mailhol, A. Olufayo, and P. Ruelle, "Sorghum and sunflower evapotranspiration and yield from simulated leaf area index," *Agricultural Water Management*, vol. 35, no. 1-2, pp. 167–182, 1997.

[6] 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*, vol. 18, no. 5-6, p. 36, 1998.

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

[8] S. Lemaire, "Système dynamique de la croissance et du développement de la betterave sucrière (Beta vulgaris L.)," Ph.D. dissertation, AgroParisTech, 2010.

[9] J. Monteith, "Climate and the efficiency of crop production in britian," *Proceedings of the Royal Society of London B*, vol. 281, pp. 277–294, 1977.

[10] N. Damay and J. Le Gouis, "Radiation use efficiency of sugar beet in northern France," *European Journal of Agronomy*, vol. v. 2(3) p., 1993.

[11] 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*, vol. 101, no. 8, pp. 1207–1219, 2008.

[12] J. Vos, L. Marcelis, and J. Evers, *Functional-structural plant modelling in crop production*. Springer, 2007, ch. 1.

[13] R. Sievänen, E. Nikinmaa, P. Nygren, H. Ozier-Lafontaine, J. Perttunen, and H. Hakula, "Components of functional-structural tree models," *Annals of Forest Science*, vol. 57, no. 5, pp. 399–412, 2000.

[14] 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*, vol. 5. Springer, 2008, pp. 19–22.

[15] 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*, vol. 9, no. 2-3, pp. 127–136, 1998.

[16] C. Spitters, H. V. Keulen, and D. van Kraalingen, "A simple and universal crop growth simulator: SUCROS87," in *Simulation and systems management in crop protection*, R. Rabbinge, S. Ward, and H. H. van Laar, Eds. Wageningen: Pudoc, 1989, pp. 147–181.

[17] Z. Li, V. Le Chevalier, and P.-H. Cournède, "Towards a Continuous Approach of Functional-Structural Plant Growth," in *Third International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications (PMA), 2009*. Beijing, China: IEEE, Nov. 2010, pp. 334–340.

[18] 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, Jul. 1996.

[19] 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," Ph.D. dissertation, AgroParisTech, 2008.

[20] P.-H. Cournède, V. Letort, A. Mathieu, M.-Z. Kang, S. Lemaire, S. Trevezas, F. Houllier, and P. de Reffye, "Some Parameter Estimation Issues in Functional-Structural Plant Modelling," *Mathematical Modelling of Natural Phenomena*, vol. 6, pp. 133–159.

[21] B. Andrieu, J.-M. Allirand, and K. Jaggard, "Ground cover and leaf area index of maize and sugar beet crops," *Agronomie*, vol. 17, no. 6-7, pp. 315–321, 1997.

[22] 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*, pp. 173–185, 1985.

[23] J. Cariboni, D. Gatelli, R. Liska, and a. Saltelli, "The role of sensitivity analysis in ecological modelling," *Ecological Modelling*, vol. 203, no. 1-2, pp. 167–182, 2007.

[24] F. Ruget, N. Brisson, R. Delécolle, and R. F. C, "Sensitivity analysis of a crop simulation model, STICS, in order to choose the main parameters to be estimated," *Agronomie*, vol. 22, pp. 133–158, 2002.

[25] D. Wallach, D. Makowski, and J. Jones, *Working with dynamic crop models: evaluation, analysis, parameterization, and applications*. Elsevier Science Ltd, 2006, ch. Evaluating crop models, pp. 11–53.

[26] D. Wallach and B. Goffinet, "Mean Squared Error of Prediction in Models for Studying Ecological and Agronomic Systems," *Biometrics*, vol. 43, no. 3, p. 561, Sep. 1987.

[27] D. Mayer and D. Butler, "Statistical validation," *Ecological Modelling*, vol. 68, no. 1-2, pp. 21–32, Jul. 1993.

[28] C. Durr, K. Fares, N. Damay, A. Carrera, N. Beaudoin, J.-M. Machet, R. Duval, F. Maupas, and P. Postel, "A description of the development and growth of sugar beet for crop modelling," *Advances in Sugar Beet Research - Sugar Beet Growth and Growth Modelling*, vol. 5, pp. 71–85, 2003.

[29] M. Launay and M. Guérif, "Ability for a model to predict crop production variability at the regional scale: an evaluation for sugar beet," *Agronomie*, vol. 23, no. 2, pp. 135–146.

[30] F. Tardieu, "Virtual plants: modelling as a tool for the genomics of tolerance to water deficit," *Trends in Plant Science*, vol. 8, no. 1, pp. 9–14, 2003.

[31] "Quantitative genetics and functional-structural plant growth models: simulation of quantitative trait loci detection for model parameters and application to potential yield optimization." *Annals of botany*, vol. 101, no. 8, pp. 1243–54, May 2008.

[32] X. Yin, P. Struik, and M. Kropff, "Role of crop physiology in predicting gene-to-phenotype relationships," *Trends in Plant Science*, vol. 9, no. 9, pp. 426–432, 2004.