

Parameter identification of plant growth models with stochastic development

KANG Mengzhen^(1,2*), HUA Jing^(1,2)

⁽¹⁾ State Key Laboratory of Management and Control for Complex Systems, LIAMA, CASIA, Beijing, 100190 China

⁽²⁾ Qingdao Academy of Intelligent Sciences, Qingdao, China mengzhen.kang@ia.ac.cn

de REFFYE Philippe⁽³⁾, JAEGER Marc^(3,4)

⁽³⁾ CIRAD, AMAP Unit, Montpellier, F-34398, France

⁽⁴⁾ University of Montpellier, LIRMM, ICAR Montpellier, F-34095, France

Abstract—Plant architectures generally display structural variations among individuals. Stochastic FSPMs have been developed to capture such feature, but calibrating such models is a challenging issue. For GreenLab model, parameter identification has been achieved on several crops and trees, but the estimation of functional parameters is mostly limited to plants with deterministic development.

In this work, we propose a methodological framework allowing the efficient FSPM parameter estimation for stochastic ramified plants. We focus on the randomness in three kinds of meristem activities in plant development: growth, death and branching. Concepts of organic series and potential structure are introduced to build the fitting target as well as corresponding model output. We show that, with a limited set of sampled plants (here from simulation), using a few organic series, the inverse method retrieves well the parameter values (the original parameter set being known here).

Requiring the concept of physiological age and the assumption of common biomass pool, the proposed approach provides a solution of solving source-sink functions of complex plant architectures, with a novel simplified way of plant sampling. The proposed parameter estimation frame is promising, since this *in silico* process mimics the procedure of calibrating model for real plants in a stand. Estimating parameters on stochastic plant architectures opens a new range of coming applications.

Keywords— *GreenLab, inverse method, source-sink parameters, functional-structural plant model, stochastic development, parameter estimation, in silico*

I. INTRODUCTION

It is easily observable in nature that, even within the same stand under a homogeneous environment, the size and number of organs (leaves, flowers, internodes, *etc.*) vary from one plant to another, differing from the industrialised products. Functional-structural plant model (FSPM) deals with simulating and visualizing plants at the level of individual organ [1]. Inevitably, the variation in plant structures become the interest of researchers in this domain: what is their mechanism, how to simulate or evaluate, what is their effect, *etc.* Such examples include LIGNUM [2], L-peach [3], *etc.* However, calibrating such stochastic functional and structural plant models is a challenging issue, especially for the functional part.

Being a member of FSPM family, GreenLab model has been applied on crops and trees, either by choosing an average structure of the sample plants [4], or on a simple stochastic case of young pine trees [5], showing only branching variation evidence.

In this work, we consider the calibration of FSPM for plants with more generic stochastic development. We take the Roux architectural model as defined in Botanists F. Halle and R. Oldeman work [6], characterized by a continuous development and a monopodial branching pattern. We consider an immediate branching case with no branching delay. This pattern covers many herbaceous plants such as soybean and wheat, and some woody shrubs such as cotton and coffee tree. We introduce several important concepts to cope with the fitting of stochastic plant in a simplified way. The fitting process on real plants is shown by an *in silico* experiment, *i.e.*, by fitting on simulated stochastic plant samples, whose real parameters are known.

II. METHOD

A. Modeling stochastic development

1) Re-visiting the concept of chronological age (CA)

The concept of chronological age is used in numerous FSPM models, corresponding to one year (for trees with rhythmic development) or to the thermal time elapsed between two successive phytomers in a stem (for crop). On a leafy axis, although generally a linear relationship exists between the thermal time with the average number of phytomers, for multiple plants, the number of phytomers at a given time often can be described by a binomial distribution [7]. Therefore, we hypothesize that the appearance of successive phytomers are independent events, with a probability to occur in each cycle of development (CD). We define then the plant “chronological age” as the number of CDs that the plant has experienced, generally larger than the observable number of phytomers in the stem.

2) Re-visiting the concept of physiological age (PA)

Inspired from botanists’ observations [8], the concept of physiological age is used to classify branch typologies. It also expresses a hierarchy along a morphological growth gradient. For example, the main stem of a coffee tree has an orthotropic (erected) behaviour, while branches are plagiotropic: cuttings

from a branch cannot generate a coffee tree, but only a creeping plant. However, in this Roux model study, the physiological age, noted p in this paper, is almost equivalent to branching order, being one for main stem, two for branches of order two, etc.

3) Meristem probabilities

With the notion of the cycle of development as mentioned above, at each CD, we suppose that a meristem produces a phytomer or not with a growth probability, generating a typical “Bernoulli process”. The meristem growth probability is noted as b_p , where p is the physiological age of axis. At a given CD, a failed Bernoulli process test leads to a void entity, meaning that no phytomer is generated.

As an axis may end its development for many reasons (e.g., insufficient nutrition or light, insect attack), at each CD, we define a *reliability probability* of the terminal meristem, noted as c_p . Once the meristem dies at a given CD, no more phytomer will appear in future CDs, and void entities fill the whole axis from the current phytomer to the tip.

Considering now branching, we consider that there is a potential number of axillary branches on each phytomer, noted as m (being 1 or 2 typically). We consider that a ramification may occur with a *branching probability*, noted as a_p . Accordingly, the probability that the meristem never develops a branch is $1-a_p$. This assumption is based on the observation that the number of branches at the same site can be described by a branching ratio. On a single axis, the index p is often omitted for clarity.

According to the above definitions, at a given chronological age, without mortality, the number of phytomers in an axis shows a binomial distribution. With mortality, a compound law will be found, mixing dead axes with still living axes (Fig. 1). In such a compound law, the left part of the distribution are mainly composed by dead branches, while in the right part are living branches. In reality, the probabilities a , b , c may vary with the chronological age of the axis, but this pattern remains.

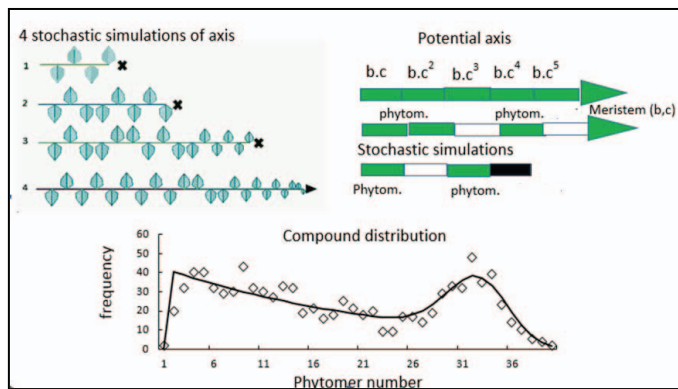


Fig. 1. Monte-Carlo simulation, probability distribution function and the potential structure of an axis with growth probability ($b=0.8$) and reliability probability ($c=0.95$). Topleft shows four topological axes from Monte-Carlo simulations; three of them did die (cross), one is still alive (arrow). Topright shows the potential axis at CD 5, and two

chronological axes with realized phytomers (in green), void entities (in white), and mortality event (in black). In the bottom graph, diamonds refer to a compound distribution from simulation of 500 dead and alive axes with 40 CDs, compared by the theoretical curve (solid line).

4) Chronological structure

A random chronological structure is composed by all the realized and void entities of the full structure with an explicit time representation. Two examples for single stem structure have been shown in Fig. 1, with alternations between green and blank entities. For branching structures, Fig. 2 gives an example with Roux model, in which each main stem phytomer can potentially bear one branch ($m=1$). Both void and realized entities contain temporal information, required when coupling the developmental model with the functional model. On real plants, the plant chronological structure is conceptual; it does not really exist, it aims to help understanding and quantifying plant development.

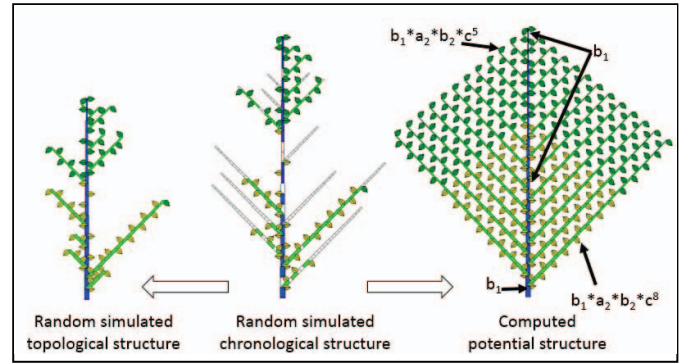


Fig. 2. The chronological, topological and potential structure of a Roux model ($c=1$). A chronological structure (middle) contains the result of a Monte-Carlo simulation, composed by realized phytomers and void entities. Suppressing the void entities gives the observable topological structure (left). The union set of all chronological structures gives the potential structure, each phytomer being associated with a probability of creation, depending on a, b, c respectively standing for branching, growth, and reliability probabilities.

5) Topological structure

Suppressing the void entities in the chronological structure results in the so-called topological structure (Fig. 2); it corresponds to the observable structure in nature. Note that the time sequence becomes fuzzy to observe: for example, a phytomer at rank five from tip can be of age 7 or 9, depending on its born date. From the topological structure it is thus hard to tell the organs' age, and this is the situation we meet in nature.

6) Potential structure

We define the potential structure as the union set of all chronological structures. It can be understood as the asymptotic upper structure limit of structure, defined by the development parameters without void entities. Besides, it is featured by a *probability of occurrence* at each potential phytomer site. This probability could be obtained from the chronological structures through multiple simulations. A more efficient way is to attribute each phytomer site with a theoretical probability, according to the probabilities of branching, growth, and reliability (a , b , and c) respectively. For example, for a

potential phytomer to appear at i -th CD, the probability of creation is $b \cdot c^i$ (Fig. 1). For the potential phytomer of physiological age 2 appearing at j -th CD, its occurrence probability is $b_1 \cdot c_1^i \cdot a_2 \cdot b_2 \cdot c_2^j$. The potential structure is very useful to compute the organic series (see below).

Note that the sum of all phytomer probabilities in the potential structure defines the theoretical mean number of phytomers produced by the structure. Since the numbers of each organ kind (leaf, internode, and fruit) by physiological age are known, this approach provides the potential organ production, from which the mean plant demand can be defined. Moreover, the potential structure is an efficient tool to compare randomly simulated structures with the theoretical one. We can also retrieve the variance, and the statistical generating function of the phytomer production [9].

7) Organic series

An organic series [10] is defined as the set of organs produced sequentially along an axis of a given physiological age. In a stochastic FSPM, an organic series corresponds to an average profile of individual organs along a stem, measured in weight or size, for a given physiological age.

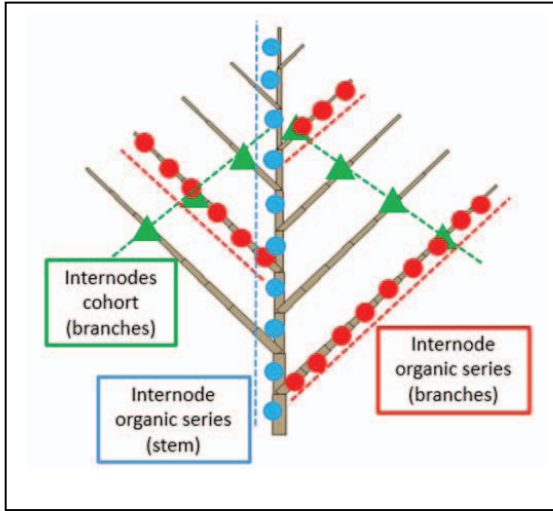


Fig. 3. The organic series (red and blue) and organ cohort (green) in a Roux model. They are orthogonal to each other. An organic series is the organ of same nature serial through the adjacent cohorts along time. According to the physiological age (same as branching order here), the same organic series can be found at numerous places within the plant architecture.

The organic series defines a frame for stochastic plant structure measurements. For example, for the structure in Fig. 3, two types of organic series are sufficient to feature the plant growth although its structure contains ten axes. The use of organic series eases the real plant structure sampling. Generally speaking, the number of organic series corresponds to the number of axis types multiplied by the number of different organ kinds.

8) Organ cohort

An organ cohort is defined as a batch of organs of same type appeared at the same time from the same kind of meristems. In the context of the GreenLab model, all phytomers within a cohort share the same meristem fate and

size (Fig. 3). It corresponds to the observation in nature that all the tips of shrubs look similar. By definition, organic series are thus orthogonal to organ cohorts.

B. Modelling stochastic plant growth

The functional mode of GreenLab inherits from classical crop models [11], with the common pool hypothesis for biomass allocation [12]. In the context of stochastic development, the plant growth can be computed by two different ways: Monte-Carlo simulations and the analytical computation. Here we focus on these two ways of obtaining plant growth, by simply reminding the model equations used in deterministic form without further justification.

1) Biomass production

An equation similar to the widely-used photosynthesis crop model is used in GreenLab. In case of plant with stochastic development, for a random sample plant s , in CD n , we have the following biomass production (1):

$$Q^s(n) = LUE \cdot PAR(n) \cdot S_p \cdot [1 - \exp(-k \cdot \frac{LA^s(n)}{S_p})] \quad (1)$$

Where s is the ID of stochastic plant sample, $Q^s(n)$ is the s plant's biomass increment at n -th CD ($\text{g plant}^{-1} \cdot \text{CD}^{-1}$). LUE is the light use efficiency, expressed in g/MJ . PAR is the photosynthetic active radiation above canopy. k is a cultural coefficient issued from leaf orientation. S_p is the projection surface. It is a hidden parameter, root of (1). $LA^s(n)$ is the s plant's total plant leaf area at n -th CD. Note that $LA^s(n)$ is a stochastic variable, differing from one plant to another.

The analytical computation case shares the same form. To distinguish, while a superscript s is used to indicate the variable of random plant sample s , such as $Q^s(n)$, the theoretical average one is noted using a superscript θ , such as $Q^\theta(n)$.

2) Sink functions

In the GreenLab model, the sink value of an individual organ at i -th cycle of its expansion is defined as the product of its sink strength and its sink variation function:

$$P_o^p(i) = p_o^p * F_o\left(\frac{i}{t_o}\right) \quad (2)$$

Where p_o^p is the dimensionless relative sink strength of the organ o (leaf, internode, and fruit) of physiological age p . The blade's sink strength of physiological age 1 (main stem) is set to one, as a reference. t_o is the duration of the expansion of organ o . F_o is an empirical function describing the sink strength evolution, modelled here by a Beta law (bell shaped) whose parameters are estimated by model inversion [13][14]. Sink function is regarded as the same both for the deterministic or stochastic model.

3) Simulated and theoretical plant demand

The plant demand at cycle n , i.e., its overall biomass request, is the sum of all the active sinks in the growing

structure. The number of organs in a random s plant sample of type o , physiological age p , appeared at plant age n , $N_o^{p,s}(n)$, can be computed from the stochastic simulation of plant development. The demand can be written as:

$$D^s(n) = \sum_{o,p} \left(\sum_{i=1}^n N_o^{p,s}(n-i+1) \cdot P_o^p(i) \right) \quad (3)$$

The plant demand is a convolution on all physiological ages and organ types. $N_o^{p,s}(n-i+1)$ is the number of organs o at cycle n , with i expansion cycles, born at cycle $n-i+1$.

The theoretical (or potential) demand is computed from the potential structure, by integrating the probabilities:

$$D^\theta(n) = \sum_{id=1}^M \sum_o \left(\pi(id) \cdot P_o^{PA(id)}(CA(id)) \right) \quad (4)$$

Where M is the total number of phytomers in the potential structure, $\pi(id)$ is the compound probability of organ existence, depending on the position of organ in plant structure, P_o is the sink strength of the organ. Notice that organs of the same chronological and topological age do not necessary share the same probability if their development history are different.

4) The chronological organic series

At each CD n , the incoming biomass $\Delta q_o^{p,i}(n)$ allocated to organ o , of age i , is proportional to its sink value $P_o^p(i)$ and the ratio between the plant production $Q^s(n-1)$, and plant demand $D^s(n)$:

$$\Delta q_o^{p,i}(n) = P_o^p(i) \cdot \frac{Q^s(n-1)}{D^s(n)} \quad (5)$$

According to this equation, if Q remains the same, less created organs will give smaller plant demand, and then bigger organs, although relative sink strength remains the same. The weight of an organ is the sum of its increments:

$$q_o^{p,i}(i, n)_{ch} = \sum_{j=1}^i P_o^p(j) \cdot \frac{Q^s(n-i+j-1)}{D^s(n-i+j)} \quad (6)$$

Considering organs of different ages, equation (6) can be written into a matrix form:

$$\begin{bmatrix} q_o^{p,s}(1, n) \\ q_o^{p,s}(2, n) \\ \vdots \\ q_o^{p,s}(n, n) \end{bmatrix}_{ch} = \begin{bmatrix} P_o^p(1) & 0 & \cdots & 0 \\ P_o^p(2) & P_o^p(1) & \cdots & 0 \\ \vdots & \vdots & \ddots & 0 \\ P_o^p(n) & P_o^p(n-1) & \cdots & P_o^p(1) \end{bmatrix} \cdot \begin{bmatrix} \frac{Q^s_{n-1}}{D_n} \\ \vdots \\ \frac{Q^s_1}{D_2} \\ \frac{Q_0}{D_1} \end{bmatrix} \quad (7)$$

$[q_o^{p,s}(*, n)]$ is a $(n \times 1)$ matrix standing for an organic series of organ o , in an axis of physiological age p .

$\begin{bmatrix} P_o^p(1) & \cdots & 0 \\ \vdots & \ddots & 0 \\ P_o^p(n) & \cdots & P_o^p(1) \end{bmatrix}$ is a triangular matrix containing the organs

sink evolution according to the phytomer ranks and chronological ages.

$\begin{bmatrix} \frac{Q^s_{n-1}}{D_n} \\ \vdots \\ \frac{Q_0}{D_1} \end{bmatrix}$ is a $(n \times 1)$ matrix representing the supply demand

ratio evolution along the n cycles. The organic series $[q_o^{p,s}(*, n)]$ contains the memory of the plant growth. Choosing multiple growth stages, it is possible to see the evolution of the organ weights along the ranks.

We can model the theoretical structure growth in a similar way, providing the theoretical biomass production $Q_\theta(t)$ at each growth cycle, and building the theoretical chronological organic series $[q_o^{p,\theta}(*, n)]$ of organ o . An example of the evolution of a theoretical leaf organic series is shown in Fig. 4.

$$\begin{bmatrix} q_o^{p,\theta}(1, n) \\ q_o^{p,\theta}(2, n) \\ \vdots \\ q_o^{p,\theta}(n, n) \end{bmatrix}_{ch} = \begin{bmatrix} P_o^p(1) & 0 & \cdots & 0 \\ P_o^p(2) & P_o^p(1) & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ P_o^p(n) & P_o^p(n-1) & \cdots & P_o^p(1) \end{bmatrix} \cdot \begin{bmatrix} \frac{Q_\theta^s_{n-1}}{D_n^\theta} \\ \vdots \\ \frac{Q_\theta^s_1}{D_2^\theta} \\ \frac{Q_\theta^s_0}{D_1^\theta} \end{bmatrix} \quad (8)$$

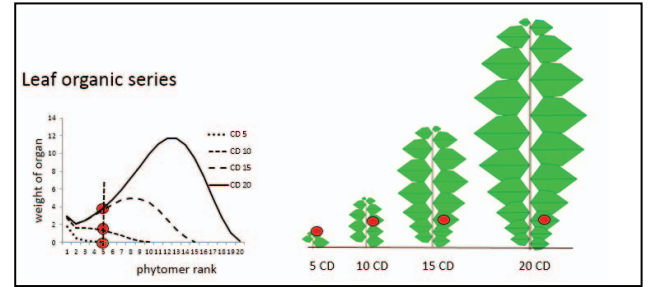


Fig. 4. Visualization of the leaf organic series evolution on a theoretical potential single-stem plant. Profiles are seen at 5, 10, 15, 20 CD. The rank 5 on the stem is selected. After 15 CD the leaf expansion ends.

From multiple stochastic chronological organic series (indexed here with ch), one can get the average of all organic series $\overline{q_o^{p,s}(*, n)_{ch}}$. It can be proved that, if the empirical $Q^s(t)$ and $D^s(t)$ distributions are bell or roughly symmetrical shaped (result of classical Bernoulli process), the following approximation is satisfied (and verified by stochastic simulations):

$$\overline{q_o^{p,s}(i, n)_{ch}} \approx q_o^{p,\theta}(i, n)_{ch} \quad (9)$$

5) Leaf area

Organ shape dimension can be retrieved from allometric rules. If the specific leaf weight is e and the weight of leaf is $q_d^{p,s}(i,n)$, its surface is then:

$$s_p(i,n) = \frac{q_d^{p,s}(i,n)}{e} \quad (10)$$

This enables to compute the individual plant leaf area $LA^s(n)$:

$$LA^s(n) = \frac{1}{e} \sum_p \sum_{i=1}^{\min(n,t_a)} N_a^{p,s}(n-i+1) \cdot q_d^{p,s}(i,n) \quad (11)$$

where t_a is the leaf functioning time, $N_a^{p,s}(n-i+1)$ is the number of functioning leaves of the corresponding cohort and $q_d^{p,s}(i,n)$ the individual leaf weight.

6) Recursive form of biomass production

Replacing $q_d^{p,s}(i,n)$ by its value in (6) gives finally the biomass production of plant at cycle n :

$$Q^s(n) = LUE \cdot PAR \cdot S_p \left(1 - \exp \left(- \frac{k}{e \cdot S_p} \sum_p \sum_{i=1}^{\min(n,t_a)} N_a^{p,s}(n-i+1) \cdot \sum_{j=1}^i \frac{p_a^p(j) Q^s(n-i+j-1)}{D^s(n-i+j)} \right) \right) \quad (12)$$

The system begins at $n=1$ from Q_0 provided by the seed. The theoretical biomass production shares the same form.

7) Topological organic series

The case of topological series (indexed with tp below) is more complex; time is not explicit (void entities are invisible). However, considering a phytomer of rank k from the tip, since the development is from a Bernoulli process with probability b , the chronological age of this phytomer follows a truncated negative binomial distribution (k, b), as follows:

$$\Pr_k(i) = \begin{cases} C_i^k \cdot b^k \cdot (1-b)^{i-k} & k \leq i < n \\ 1 - \sum_{i=k}^n C_i^k \cdot b^k \cdot (1-b)^{i-k} & i = n \end{cases} \quad (13)$$

The theoretical topological organic series is then extracted from the potential structure: the average organ weight located at rank k from axis tip is a compound law issued from the negative binomial distribution considering all $q_o^{p,\theta}(*,n)$ potentially seen at rank k , and clamped by the stem age.

$$q_o^{p,\theta}(k,n)_{tp} = \frac{\sum_{i=k}^n \Pr_k(i) \cdot q_o^{p,\theta}(i,n)_{ch}}{\sum_{i=k}^n \Pr_k(i)} \quad (14)$$

The simulated average organ weight at a given phytomer rank k from axis tip:

$$\overline{q_o^p(k,n)_{tp}} = \frac{1}{T} \sum_{s=1}^T q_o^{p,s}(k,n)_{tp} \quad (15)$$

where T is the number of simulations. We find the following tendency also by stochastic simulations:

$$q_o^p(k,n)_{tp} \approx q_o^{p,\theta}(k,n)_{tp} \quad (16)$$

As a summary, average series from stochastic simulations can be matched with the theoretical ones both for chronological

and topological series. This provides basis for the following fitting exercise.

C. Fitting exercise

As mentioned previously, the organ functioning parameter estimation from the measurements on plant architectures has been worked out on plant with deterministic development (single-stem plants like maize [14], sunflower [15], and beetroot [16]). The principle is to use the organic series as target data to fit, since they contain all the necessary and sufficient information about growth. The target for fitting includes $maxp$ types of axis and t (2-4) types of organs have $maxp \times t$ organic series, which are few.

Before dealing with real stochastic plants, we make *in silico* fitting process to make sure that software are working properly for both simulation and theoretical computation. This is not trivial since computational errors may not be observable as the Monte-Carlo simulation itself is noisy. The simulated topological organic series are sorted and averaged from the top to the bottom axis, which are regarded as the observed axes. While building the organic series, dead branches are skipped, since they do not contribute any more to the growth process (their demand is null). The workflow is as shown in Fig. 5.

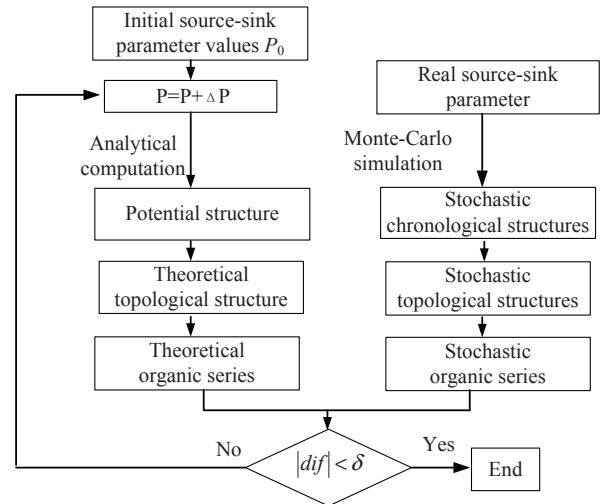


Fig. 5. Flowchart of the fitting process using simulated topological organic series as fitting target.

Starting from a known set of parameters values, the average topological series can be obtained through simulation. This can be regarded as a virtual sampling on the simulated plants. On the other hand, from a set of initial parameter values (P_0), theoretical topological series can be computed and compared to the simulated ones. The aim of least-squares fitting is to minimize their difference for all measurement dates t by searching for best set of parameter values.

$$|dif| = \sum_t \sum_{o,p} \left| \overline{q_o^{p,s}(*,n)}_{tp} - [q_o^{p,\theta}(*,n)]_{tp} \right| \quad (17)$$

The organic series contains a list of M observations y_i that match M corresponding outputs of the model with associated

functions f_i , whose parameters depend of the set of the m hidden source –sink parameters p_k to identify. All f_i depends on the probabilities for plant development that gives the demand, the allometry of organs, and the environmental parameters (light use efficiency, solar radiation at each step, etc) at each step of growth. The system can also be written as:

$$\begin{aligned} y_1 &= f_1(p_1, \dots, p_k, \dots, p_m) + e_1 \\ &\dots \\ y_M &= f_M(p_1, \dots, p_k, \dots, p_m) + e_n \end{aligned} \quad (18)$$

The values $y_i = f_i(p_1, \dots, p_k, \dots, p_m) + e_i$ give the weights of organs in the organic series. The e_i values stand for the deviation between the observed and computed data. We use the nonlinear generalized least square method to achieve the estimations.

III. CASE STUDY

A. Single stem plant

Let us first consider T topological stems issued from T single stem plant simulations. The reliability probability is set to 1. The development probability is set to $b=0.7$, and the chronological age of the stem is 30 CDs. Sinks are identical for leaves and internodes; both the functioning and expansion durations are set to 5 cycles.

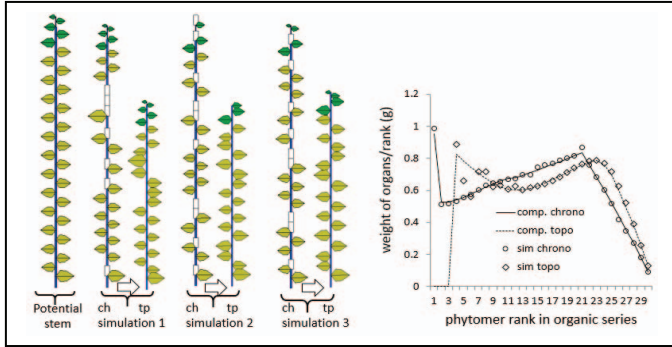


Fig. 6. Potential, simulated and theoretical (computed) chronological and topological organic series for a single stem developed from Bernoulli process ($b=0.7$, $n=30$, $T=1000$). The left is the potential structure, followed by three pairs of stochastic simulations in chronological mode (ch) and corresponding topological mode (tp); yellow leaves are no more functional. In the right is the corresponding organic series curves, sorted from the stem tip (right side). On chronological series, circles stand for the average values from simulation, to compare with the theoretical values (solid line). On topological organic series, diamonds stand for the simulation averages to compare with the theoretical values (dotted line).

Fig. 6 shows the potential stem, the visualization of three simulated chronological and topological stems, the simulated and theoretical chronological series, as well as the simulated and theoretical topological series. Note that the chronological and topological organic series show different patterns. Moreover, the topological organic series is shorter than the chronological one, because some CDs are void, *i.e.*, no phytomer is created.

B. Monopodial branching plant

This is a general plant growth and development pattern, to be found on herbaceous, like wheat, soybean, woody plants like cotton or coffee plants. In this example we set following development parameters: $b_1=0.8$; $b_2=0.8$; $a_2=0.9$; $c_2=0.95$. Notice here that the reliability probability is no more one. Remind that these probabilities will provide the potential structure for computing the theoretical plant demand. For growth parameters we set the leaf functioning time to $ta=9$ cycles, the specific leaf weight to $e=0.05$ g cm⁻², and the organ's expansion to $tx=9$ cycles. Sinks for all organs are set to be equal and constant, being $p_o=1$. Light use efficiency ($r=1/LUE$) for leaves is 30, and projection surface is $S_p=2000$ cm².

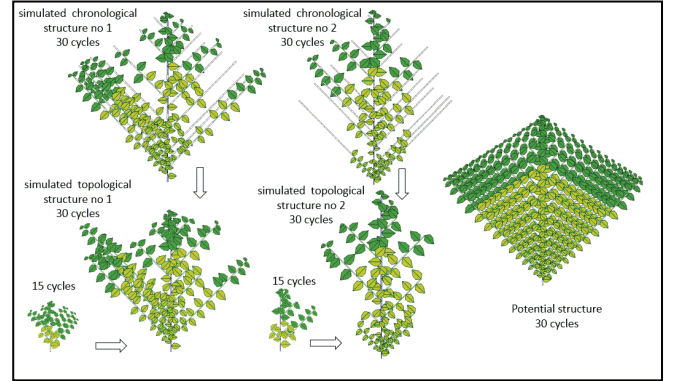


Fig. 7. Potential, simulated chronological and topological structures on a monopodial Roux architecture ($b_1=b_2=0.8$; $a_2=0.9$; $c_1=1$; $c_2=0.95$). Left: two simulations at CD 30, both in chronological and topological modes, display a large variability due to the stochastic meristem functioning. Right: the potential structure.

With the concepts of physiological age and organic series, only four organic series are to be considered: two for the stem (leaf and internode) and two for the branches (leaves and internodes). Since leaves and internodes share here the same sink functions, thus only two series are to consider, one for the stem and one for branches. The results are summarised in a target file (Fig. 7), where measurements on organ weights are included on specific cohorts and chronological series, corresponding to the observation stages. The theoretical topological organic series are computed according to the analytical computations.

Matching the average of simulated topological organic series with theoretical ones can be processed at the same time for several growth steps, defining a multi-fitting process. The fitting result is operative as shown in Fig. 8, where the theoretical organic series fit properly the corresponding average organic series, extracted from 50 stochastic structure simulations at 30 cycles of growth.

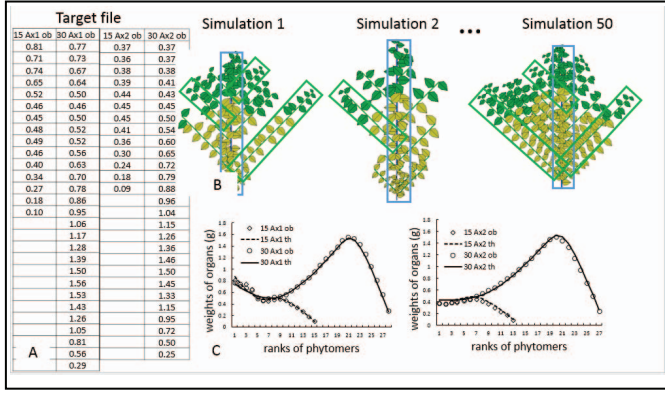


Fig. 8. The table, left (A) shows the target file built with the topological organic series at two development stages (15 and 30 cycles) on main stem (Ax1) and branches (Ax2). Topright (B): organic series for stems (blue) and branches (green). Bottom right (C): fitting curves between the average simulated organic series and the theoretical ones extracted from the potential structure.

Here the hidden parameters to identify from the target file are r , S_p , p_a for leaf, and p_i for internode. The nonlinear generalized least square method gives stable solutions: $r=29.7$; $S=1978 \text{ cm}^2$; $p_a=0.96$; $p_i=0.98$. These values are quite close from the original data use to simulate the computational plant. In Fig. 8, notice that branches with all leaves yellow are dead; they are ignored in the organic series analysis as well as branches with too few green leaves. Branches with small green leaves at the tip are still growing.

IV. DISCUSSION

A. New target, new sampling strategy

Sampling on real plants is closely related to the usage of the data. Observations (here from simulation) must be organized with the aim that all the necessary and sufficient information are included to estimate the hidden parameters. The measurements strategy must also be affordable and easy to enforce applications. Requiring the concept of physiological age and the assumption of the common biomass pool, the proposed approach is versatile; it doesn't rely on the plant architectural model complexity or in its specific stochastic expression.

In field conditions, sometimes lot of efforts are paid to measure the plants, even if the data cannot be fully used. Measuring on branching plant structures may lead to very tedious work that numerous researchers in FSPM domain have experienced, for example, measuring all the weight and size of individual organs in a wheat plant, at several stages. In the current work of solving the source-sink functions, the organic series are proposed as the fitting target. Similar to the

deterministic case, the organic series contains the history of biomass production and partitioning, through which the source-sink parameters can be estimated. To obtain the organic series, one need only to sample about dozens of (dependent on the variation of plant structure) stems and branches from different plants.

In this work, calibration of the model is made sequentially; first bud probabilities linked to the meristem functioning of each type of axis are estimated to retrieve plant development, then the source-sink parameters are estimated. Accordingly, record on topological structures and organ weight can be done separately. Retrieving topological parameters through crown analysis is not presented here, but can be inspired from [17]. Although it is hard to tell which is the cause between the stochastic development and the internal source-sink ratio of plant, solving the source-sink functions provides a way to understand its link with the development, such as the fruit set [18]. This could prepare for the further simulation based on the plant growth and development interaction [19].

B. Simplified, but not simple

Plant models generally provides a way to observe and understand the plants. So far GreenLab calibration is limited to single monoculm plants or simple deterministic structures plants. In this work, plant structure is seen as a set of stochastic ramified axes, composed of phytomers serials with their subsets of organs (leaves, internode, and fruit). By introducing a virtual cycle of development, each phytomer is given a probability of growth, a probability of reliability and a probability of branching. Here, although the presented type of plant development is simple (the Roux model with continuous development and monopodial branching); it still reflects the full complexity of the stochastic meristems activity, for growth, branching and death process. We also consider the axes death process, seldom covered in the literature on plant growth despite its importance when processing woody plants development.

We have introduced the concept of potential structure, as the theoretical structure which contains all possible plant realizations in which phytomers are weighted by their occurrence probability. We extended then the concept of organic series, the set of organs produced by a same apical meristem along an axis including pauses (chronological series) or not (topological series) as appeared on real plants. The method eases the plant measurement, but complex computation is performed in preparing the theoretical mapping in software.

Because of the high computational complexity in obtaining the theoretical or simulated, chronological or topological organic series, before coping with the architecture of real plants, *in silico* plant samples offer a versatile tool to work out a data assimilation method associated to an inverse method to identify the parameters of plant growth. On one hand, with known parameter values, the simulation platform gives the chronological and then topological organic series from very limited simulated plant samples, imitating the real data. On the other hand, the theoretical organic series are computed

independently, and starting from initial parameters values that differ from the real one, by fitting the theoretical and simulated organic series, the parameters are correctly identified. The parameter identification can be achieved because our GreenLab model gathers the main traits of plant architecture.

This work prepares the applications to real plants, replacing data in the target file. And this ease a lot the data organization in plant measurement, since the detailed architecture of the sample plants for destructive measurement is no more obliged to be fully recorded. The organic series provide a powerful tool to design a conceptual parameter estimation frame, validated on computational plants, then operative on coming studies related to real plant growth model assessments.

C. Future work

Applications on real plants (coffee, splilanthos, and soybean) are in progress, showing first interesting results on source and sink parameter estimation. The current work is also to be extended to trees of other architectural models, like Rauh models. Real plants are more complex, for example, the probabilities are not necessary stable, and there are delay in development and growth. To deal with such various cases, delicate software development is a key issue.

V. CONCLUSION

We proposed hereby a methodological framework allowing FSPM parameter estimation on stochastic ramified plants, extending structural parameter retrieval and functional parameter estimations. This work opens a window for FSPM to cope with more plants, whose structure shows strong variability. The organic series, computed from the potential structure, are introduced so that similar framework in deterministic plant analysis can be applied to stochastic plants. This method eases the sampling on stochastic branching plant structures, which could be a burden.

We show that, the theoretical mean of organ attributes values (such as organ weight), issued from the potential structure, mathematically converge to the experimental values issued from sets of simulations. We illustrated finally that, organic series are fitted with the corresponding series computed from the potential structure, using parameter identification provided by the generalized nonlinear least square method parameter estimation.

Acknowledgment

This work is supported in part by China 863 Program (#2012AA101906-2) and the National Science Foundation of China (#31170670). Authors thanks LIAMA for its support within cPlant project frame.

References

- [1] Vos J., Evers J.B., Buck-Sorlin G.H., Andrieu B., Chelle M., de Visser P.H.B. 2009. Functional-structural plant modelling: a new versatile tool in crop science. *Journal of Experimental Botany* 61: 2101–2115.
- [2] Perttunen, J., Siev nen, R., & Nikinmaa, E. (1998). LIGNUM: a model combining the structure and the functioning of trees. *Ecological modelling*, 108(1), 189–198.
- [3] Allen M., Prusinkiewicz P., DeJong T. 2005. Using L-systems for modeling source-sink interactions, architecture and physiology of growing trees: the L-peach model. *New Phytologist* 166: 869–880.
- [4] Kang, M. Z., Heuvelink, E., Carvalho, S. M. P., de Reffye, P. 2012. A virtual plant that responds to the environment like a real one: the case for chrysanthemum. *New Phytologist*. 195(2): 384–395
- [5] Wang, F., Kang, M. Z., Lu, Q., Han, H., Letort, V., Guo, Y., de Reffye, P., Li, B.G.. 2010. A stochastic model of tree architecture and biomass partitioning: application to Mongolian Scots pines. *Annals of Botany*. 107 (5): 781–792. doi: 10.1093/aob/mcq218 First published online: November 8, 2010
- [6] Hall  F., Oldeman R.A.A. & Tomlinson P.B. (1978). *Tropical Trees and Forests*. Springer Verlag, Berlin, Heidelberg, New-York, 441p.
- [7] de Reffye, P., Elguero, E., & Costes, E. (1991). Growth units construction in trees: a stochastic approach. *Acta Biotheoretica*, 39(3–4), 325–342.
- [8] Barth l my, D., Caraglio, Y. 2007. Plant Architecture: A Dynamic, Multilevel and Comprehensive Approach to Plant Form, Structure and Ontogeny. *Annals of Botany*, 99 (3) : pp. 375–407 19
- [9] Kang, M. Z., Courn de, P. H., De Reffye, P., Auclair, D., & Hu, B. G. (2008). Analytical study of a stochastic plant growth model: application to the GreenLab model. *Mathematics and Computers in Simulation*, 78(1), 57–75.
- [10] Buis R. et H. Barthou, 1984.- Relations dimensionnelles dans une s rie organique en croissance chez une plante sup rieure. *Rev. Biomath.*, 85,1-19
- [11] Brisson, N., Gary, C., Justes, E., Roche, R., Mary, B., Ripoche, D., ... & Bussiere, F. (2003). An overview of the crop model STICS. *European Journal of agronomy*, 18(3), 309–332.
- [12] Heuvelink, E. (1995). Dry matter partitioning in a tomato plant: one common assimilate pool? *Journal of Experimental Botany*, 46(8), 1025–1033.
- [13] Zhan ZG, de Reffye P, Houllier F, Hu BG. 2003. Fitting a functional-structural growth model with plant architectural data. In: Hu BG, Jaeger M, eds. *Plant growth modeling and applications*. Beijing: Tsinghua University Press/Springer-Verlag, 108–117.
- [14] Guo, Y. A. N., Ma, Y., Zhan, Z., Li, B., Dingkuhn, M., Luquet, D., & De Reffye, P. (2006). Parameter optimization and field validation of the functional-structural model GREENLAB for maize. *Annals of botany*, 97(2), 217–230.
- [15] Rey H., Courn de P. H., de Reffye P., Lecoeur J. 2006. Sunflower (*Helianthus annuus* sp.) growth analysis by source - sink relationships study with GreenLab model. Poster in: PMA06: The Second International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications, Beijing, P. R. China, 13–17/11/2006.
- [16] S. Lemaire, F. Maupas, P.-H. Courn de, 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*, April 19–22, 2008, Nanjing, China, 2008.
- [17] Diao J., De Reffye P., Lei X., Guo H., Letort V.. 2012. Simulation of the topological development of young eucalyptus using a stochastic model and sampling measurement strategy. *Computers and electronics in agriculture*, 80 (1) : 105–114
- [18] Ma Y. T., Wubs A. M., Mathieu A., Heuvelink E., Zhu J. Y., Hu B. G., Courn de P. H., de Reffye P. 2011. Simulation of fruit-set and trophic competition and optimization of yield advantages in six *Capsicum* cultivars using functional-structural plant modelling. *Annals of Botany*, 107, (5): 793–803.
- [19] Mathieu, A., Courn de, P.-H., Letort, V., Barth l my, D., De Reffye, P. A dynamic model of plant growth with interactions between development and functional mechanisms to study plant structural plasticity related to trophic competition.. 2009. *Annals of Botany*, 103 : 1173–11860