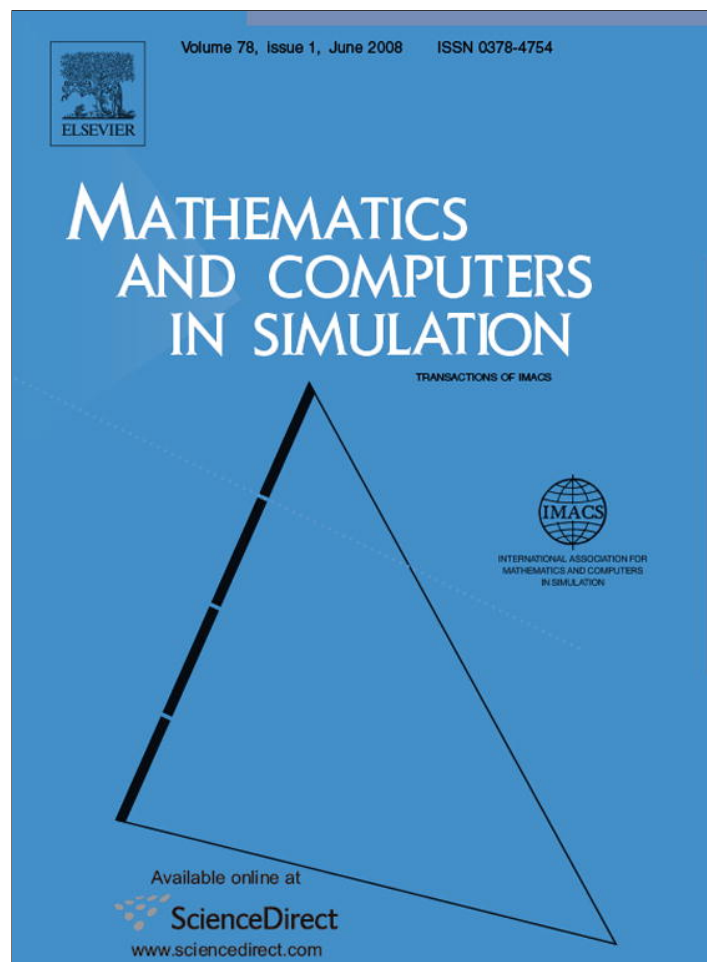


Provided for non-commercial research and education use.  
Not for reproduction, distribution or commercial use.



This article appeared in a journal published by Elsevier. The attached copy is furnished to the author for internal non-commercial research and education use, including for instruction at the authors institution and sharing with colleagues.

Other uses, including reproduction and distribution, or selling or licensing copies, or posting to personal, institutional or third party websites are prohibited.

In most cases authors are permitted to post their version of the article (e.g. in Word or Tex form) to their personal website or institutional repository. Authors requiring further information regarding Elsevier's archiving and manuscript policies are encouraged to visit:

<http://www.elsevier.com/copyright>



# Analytical study of a stochastic plant growth model: Application to the GreenLab model

M.Z. Kang<sup>a,c,\*</sup>, P.H. Cournède<sup>b</sup>, P. de Reffye<sup>c,d,e</sup>, D. Auclair<sup>f</sup>, B.G. Hu<sup>c</sup>

<sup>a</sup> Capital Normal University, Beijing 100037, China

<sup>b</sup> Ecole Centrale Paris, Laboratory of Applied Mathematics, Châtenay-Malabry F-92295, France

<sup>c</sup> Chinese Academy of Sciences, Institute of Automation, LIAMA, Beijing 100080, China

<sup>d</sup> INRIA, DigiPlante, Rocquencourt F-78150, France

<sup>e</sup> CIRAD, UMR AMAP, Montpellier F-34000, France

<sup>f</sup> INRA, UMR AMAP, Montpellier F-34000, France

Received 11 June 2007; accepted 13 June 2007

Available online 3 July 2007

## Abstract

A stochastic functional–structural model simulating plant development and growth is presented. The number of organs (internodes, leaves and fruits) produced by the model is not only a key intermediate variable for biomass production computation, but also an indicator of model complexity. To obtain their mean and variance through simulation is time-consuming and the results are approximate. In this paper, based on the idea of substructure decomposition, the theoretical mean and variance of the number of organs in a plant structure from the model are computed recurrently by applying a compound law of generating functions. This analytical method provides fast and precise results, which facilitates model analysis as well as model calibration and validation with real plants. Furthermore, the mean and variance of the biomass production from the stochastic plant model are of special interest linked to the prediction of yield. In this paper, through differential statistics, their approximate results are computed in an analytical way for any plant age. A case study on sample trees from this functional–structural model shows the theoretical moments of the number of organs and the biomass production, as well as the computation efficiency of the analytical method compared to a Monte-Carlo simulation method. The advantages and the drawbacks of this stochastic model for agricultural applications are discussed.

© 2007 IMACS. Published by Elsevier B.V. All rights reserved.

**Keywords:** GreenLab; Analytical moments; Recurrent computation; Bud probability; Stochastic functional–structural plant model

## 1. Introduction

As shown by Hallé et al. [18], the development of plants follows particular patterns that can be determined by the combination of a small number of characters (*e.g.* type of branching, flowering, and axis differentiation). The number of possible patterns observed, namely *architectural models*, is 23. Each of these patterns, named after a botanist, covers a large majority of species. For example the Corner architectural model is designed for un-branched plants such

\* Corresponding author at: Capital Normal University, Beijing 100037, China.

*E-mail addresses:* [mzkang@liama.ia.ac.cn](mailto:mzkang@liama.ia.ac.cn) (M.Z. Kang), [paul-henry.cournede@ecp.fr](mailto:paul-henry.cournede@ecp.fr) (P.H. Cournède), [philippe.de-reffye@inria.fr](mailto:philippe.de-reffye@inria.fr) (P. de Reffye), [auclair@cirad.fr](mailto:auclair@cirad.fr) (D. Auclair), [hubg@nlpr.ia.ac.cn](mailto:hubg@nlpr.ia.ac.cn) (B.G. Hu).

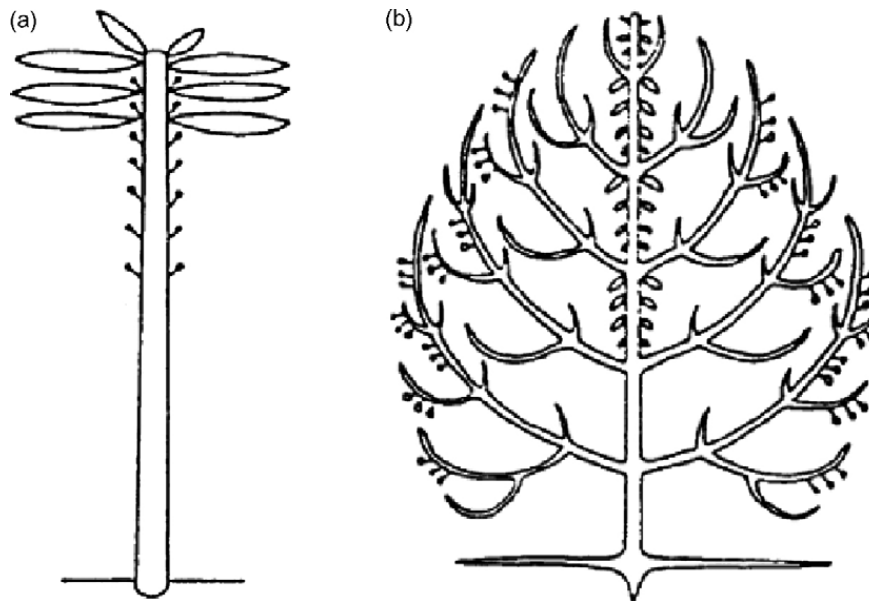


Fig. 1. Architectural models illustrated with drawings by botanists [18]. (a) Corner model and (b) Rauh model.

as palm trees, and the Rauh architectural model displays features of rhythmic growth, such as pine trees, shown in Fig. 1.

The term “model” has different meanings according to its context. In the area of computer sciences, plant architectural models are described by Fisher [14] as *algorithms of simulation that produce the forms and patterns of plants*. Sometimes they are also termed *geometrical models* (GM) or *structural models* (SM). These kinds of models appeared in the 1960s with the development of modern computers, for example, by Cohen [5], Lindenmayer [29] and Honda [23]. Since then, many computer simulation algorithms have been developed aiming at producing tree-like geometrical structures, more or less faithful to botany: L-systems [33], fractals [37], particle systems [34], ramified matrix [38], reference axis [1], and voxel space automata [15]. A critical review was made in 1992 by Fisher [14], discussing the ability of existing computer models to test hypotheses or to predict mature tree forms. Although some of them can test scientific hypotheses, e.g. Honda and Fisher [22], architectural models are mostly aimed at generating a realistic shape of trees (e.g. Bloomenthal [3]), and are often used for entertainment, computer games or landscape design. To simulate the dynamic branching pattern of real plants, however, integrating botanical knowledge into the models is more important than generating static and/or artistic plant geometry.

The topological structure of real plants combines the genetically determined form (described by the 23 architectural models) and the deviations due to environmental factors (noise). It can be described in terms of the dynamics of buds, with the fate of buds depending on their relative position in space and time [4]. Studies on the stochastic behaviour of buds in real plants began in the 1970's on coffee trees by de Reffye [13]. Using parameters derived from observation, realistic structures faithful to botany were simulated [10], where a bud can die, rest, or create a variable number of metamers. Besides, Semi-Markov chains have been used as a statistical model to quantify the flowering sequences [16], and the branching patterns for apple cultivars [6] and red oak trees [20]. It was shown that branches borne by one growth unit are organized as a succession of homogeneous zones, according to a physiological gradient of branches. This gradient has led Barthélémy et al. to define the *physiological age* of the meristem [2], by describing the differentiation patterns of shoots and branches.

Pure geometrical models themselves are not sufficient to describe the underlying mechanisms of plant growth as they do not take into account the plant functioning processes. In agricultural or forestry applications, there is another type of model, namely *process-based models* (PBM), which considers biomass production through the photosynthetic process and the global biomass partitioning among organs, e.g. Tomsim [21] or ROSGRO [9]. As commented in [30], such models are limited by the absence of dynamic development of plant architecture, thus it is difficult to integrate the effect of plant architecture on plant growth, for example leaf area. Between process-based models and architectural models, *functional–structural plant models* (FSPM) emerged since the end of the 1990's, such as LIGNUM [32], GROGRA [28], AMAPHydro [11], and have received increasing attention in recent years [39]. They seek to combine

both architectural development and the photosynthetic process. The high computation complexity and difficulties in model calibration are problems inherent to such kinds of models, as shown by Sievänen et al. [36].

Not only for pure architectural studies, but also in the context of functional–structural models, the quantity, size and weight of individual plant organs (internodes, leaves or fruits) are of special interest. In the case of stochastic FSPMs, Monte-Carlo simulation can be applied to obtain statistical results. However, it can be very time consuming to simulate a population of trees of complex structure. Moreover, the simulation software are usually so complex that a great amount of bug-proofing is necessary before the results can be trusted. On the other hand, analytical results are both fast and accurate in computation, but the main drawback is that they are usually difficult to obtain for stochastic models.

GreenLab is one of the few functional–structural models that can be calibrated with a mathematical approach [41,27]. Such calibration has been performed based on real data from agricultural or horticultural crops, for example cotton [42], maize [17], as well as chrysanthemum [26]. The architectural model of GreenLab inherits the main concepts of AMAPSim [1]. It can simulate the 23 botanical architectural models [44]. The model is organized according to physiological age, requiring only a limited number of parameters for simulating a complex structure. This model provides a simple way for users to construct even the complex structures of trees. The framework of the GreenLab model was built by coupling an architectural model with sink–source functions of individual organs [41]. The organogenesis, photosynthesis and morphogenesis processes were formalized in [12] for a deterministic model (GL1), and in [8] for stochastic (GL2) and feedback (GL3) cases, respectively.

The aim of this paper is to present the theoretical computation of mean and variance of the number of organs, and the biomass production resulting from the stochastic GreenLab model GL2. It is very important to have such fast calculation of moments when estimating the bud probabilities and functional parameters from the measured data with a non-linear least square method. Moments of the number of organs are obtained by applying compound rules from probability generating functions (PGFs), combined with substructure decomposition [8]. PGFs can serve as analytical tools to develop precise estimates for quantities of interest; they have been used intensively in algorithm analysis [35]. The method is presented here for the GreenLab model, but it is also applicable for other plant architectural models. The moments of biomass production of a plant are computed at each cycle through induction, using differential statistics [24]. The covariance matrix between the number of metamers and biomass production is computed.

The content of this paper is organized as follows: in Section 2, the botanically based architectural model and the probabilities of bud functioning of GL2 are presented; Section 3 gives the recurrent formulae for computing mean and variance of the number of organs from the architectural model. In Section 4 the theoretical work extends to approximate analytical formulae of mean and variance of the biomass production computed with a simplified GreenLab equation by using differential statistics. Section 5 shows an example of results from Sections 3 and 4, obtained on a Rauh architectural tree model. Section 6 discusses the methods presented.

## 2. The stochastic plant architectural model

### 2.1. Related botanical concepts

The main botanical concepts underlying the present study have been described by Barthélémy et al. [2]. During a growth cycle, a plant sets a *growth unit* (GU) rhythmically along an axis by its *apical* bud. The duration of a growth cycle can vary from several days (for crops) to 1 year (for temperate trees). A growth unit is a succession of *metamers*, each metamer is the basic botanical unit containing different types of individual organs: one internode, one or more leaves, one or more axillary buds and possibly flowers. The succession of growth units forms an *axis* (or *shoot*). The axillary buds are the starting states of new axes. The full set of organs issued from a bud is called a *structure*, or a *substructure* when it is a part of another structure.

The different categories of shoots in trees can be distinguished by *physiological age* (PA), a dimensionless number. Being 1 for the main stem, generally the PA of axillary branches (or axillary buds) is higher than that of the growth unit that bears these branches. In some cases they can be identical or smaller, which corresponds to a botanical phenomenon called *reiteration*. In the present paper we shall not consider the reiteration process. Generally, the highest physiological age of plants is less than five. Although they are close, physiological age is not necessarily the same as branching order. For example, a structure of PA 3 can appear as of branching order 3, but it can also appear as of branching order 2 grown from the main stem. The apical bud of an axis can also mute into another higher physiological age. The succession of GUs from apical buds of the same PA is called a *bearing axis*. The whole plant structure can thus be viewed as a stack of

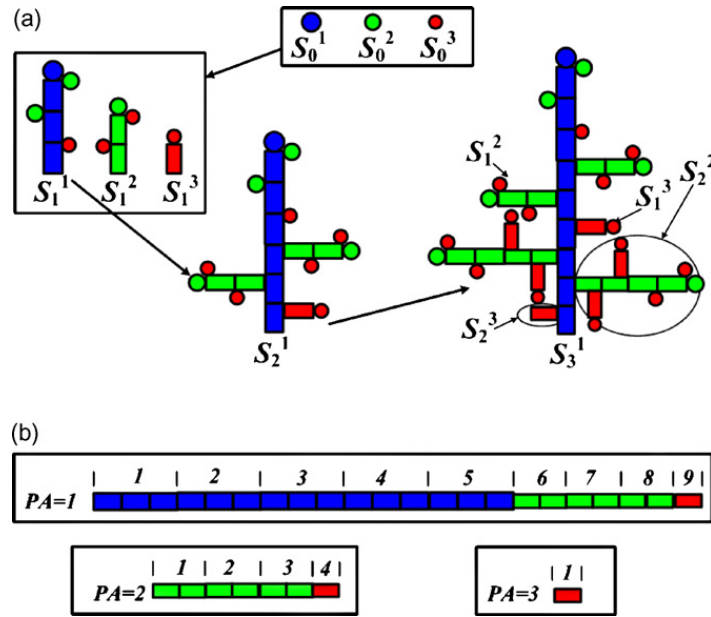


Fig. 2. Illustration of the GreenLab architectural model. Here  $P_m = 3$ .  $S_n^p$  represents a substructure of physiological age  $p$  and chronological age  $n$ ,  $S_0^p$  (circle) being the bud of PA  $p$ . In (a), buds of PA 1, 2 and 3 can develop into growth units of 3, 2 and 1 metamers ( $S_1^1$ ,  $S_1^2$ ,  $S_1^3$ ) respectively, which define the one-step transition of the branching process. The circles on the top of growth units are apical buds while the others are axillary buds. When the plant ( $S_n^1$ ) grows from cycle 1 to 3, all buds develop into the corresponding growth units. In  $S_3^1$  it is shown that the plant structure can be decomposed into axes and the substructures borne by the axes. In (b), the mutation of axes is shown: axes of PA 1, 2 and 3 grow 5, 3 and 1 cycles respectively before the terminal bud mutes into another PA. Axes of PA 3 have no more apical bud after the cycle.

substructures of different physiological ages. Organs, axes and substructures are characterised also by *Chronological Age* (CA), which corresponds to the number of cycles that the part of plant has undergone since it was born. CA can be converted to calendar time (days or years) as in functional plant models, when the *thermal time* (in degree-days) is known for each cycle. The relationship between the number of cycles and the thermal time since the emergence of a plant is often linear, as shown in [17] and [26].

### 2.2. The parameters of the architectural model

Let the maximum physiological age of the plant be  $P_m$ . The architectural model is designed according to the bud activity, where buds can develop into the same kind of growth unit, each containing an *apical* bud and one or several metamers that bear new *axillary* buds, see the first step in Fig. 2(a). It is in fact a multi-type *branching process* [19] where the buds produce a new generation during a cycle according to their own PAs. Fig. 2 shows an example of architectural model with three steps of transition from the seed to plant age 3. Let  $\mathcal{M}_I$  be a matrix defining the number of metamers in growth units,  $\mathcal{M}_I^{p,k}$  being the number of metamers that bear axillary buds of PA  $k$  in the growth unit of PA  $p$ .  $\mathcal{M}_I$  is a triangle matrix since there is no reiteration and the PA of axillary buds is not less than that of the metamer that it grows from. Let  $\mu^p$  be the total number of metamers in a growth unit of PA  $p$ ,  $\mu^p = \sum_{k=p}^{P_m} \mathcal{M}_I^{p,k}$ . Let  $\mathcal{N}_B$  be a matrix defining the number of axillary buds per metamer,  $\mathcal{N}_B^{p,k}$  being the number of axillary buds of PA  $k$  in a metamer of PA  $p$ . Similarly,  $\mathcal{N}_O$  is a matrix defining the number of organs per metamers,  $O = I, L, F$  for internodes, leaves and fruits respectively. For the example of Fig. 2, a bud of PA 1 produces two metamers containing an axillary bud of PA 2 and one metamer containing an axillary bud of PA 3. A bud of PA 2 produces two metamers containing an axillary bud of PA 3. A bud of PA 3 produces one metamer. This metamer can bear only axillary buds of PA 3 as  $P_m = 3$ , in the case of reiteration. But in this example no axillary bud is drawn for PA 3, so  $\mathcal{N}_B^{3,3} = 0$ . Finally the values of  $\mathcal{M}_I$  and  $\mathcal{N}_B$  are

$$\mathcal{M}_I = \begin{bmatrix} 0 & 2 & 1 \\ 0 & 0 & 2 \\ 0 & 0 & 1 \end{bmatrix}, \quad \mathcal{N}_B = \begin{bmatrix} 0 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}$$

The apical bud can change its physiological age at a certain cycle when the axis develops, as shown in Fig. 2(b). Let  $\mathcal{M}_A$  be a vector defining the number of cycles before a bud mutes,  $\mathcal{J}_A$  be a vector defining the new PA of the apical bud after mutation. For the example in Fig. 2, their values are:

$$\mathcal{M}_A = [5 \quad 3 \quad 1], \quad \mathcal{J}_A = [2 \quad 3 \quad 0]$$

The valid values of  $\mathcal{J}_A$  must be between 1 and  $P_m$ . Here  $\mathcal{J}_A^3 = 0$  means that there is no more apical bud when the development of the axis ends.

With the notion of physiological age, these four parameters,  $\mathcal{M}_A$ ,  $\mathcal{J}_A$ ,  $\mathcal{M}_I$  and  $\mathcal{N}_B$ , are sufficient to simulate plant architectures ranging from the Corner model (e.g. palm tree) to complex structures such as the Rauh model (e.g. pine tree). It is unnecessary to provide a large number of rewriting rules when simulating plants of high branching order, instead, one only needs to modify the dimensions and values of the parameters. However, some extra parameters are needed to simulate different flowering sequences, reiteration, and inflorescence structures. Here we present only the basic part of the model.

### 2.3. Bud functioning activities in GL2

In reality, even in a homogeneous environment, plants are topologically stochastic: the buds may die, rest or grow at each growth cycle [10], according to observations on real plants. The following probabilities are thus introduced into the GreenLab architectural model to simulate this stochastic development. For a specific plant in a given environment, some probabilities may be one if the corresponding event always happens. Each kind of probability is a vector of length  $P_m$  if it is constant over time. Otherwise it is a matrix of size  $P_m \times N$ ,  $N$  being the age of plant, as for coffee tree [13]. Without losing generality, the following presentation focuses on the case of constant probability and the theoretical distribution under this hypothesis, which is based on botanical observations [16,20,2].

- *Survival probabilities of buds.* The apical bud of an axis may die because of its genetic behaviour or environmental accident (for example, insect damage). Let  $\mathcal{P}_C$  be the probability that a bud survives at each cycle. In that case, the number of cycles before the apical bud is dead follows a truncated geometrical law ( $\mathcal{M}_A$ ,  $\mathcal{P}_C$ ).
- *Branching probabilities of buds.* Particularly for an axillary bud, which is the initial state of an axis, the probability that the bud can start or breakout is different (often lower) from the survival probability of subsequent apical buds of this axis since it is influenced by more factors (insufficient assimilate supply, or poor light condition). A failure of starting after bud creation leads to bud abortion. Let  $\mathcal{P}_B$  be the branching probabilities of buds.  $\mathcal{P}_B^1$  is then the germination probability of the seed. This probability can explain the changing number of branches of the same kind along an axis. Given the potential number of buds  $\mathcal{N}_B$ , the number of appeared branches then follows a binomial law ( $\mathcal{N}_B$ ,  $\mathcal{P}_B$ ) if  $\mathcal{P}_B$  is the same for each potential bud.
- *Growth probability of apical buds.* Even when no bud death occurs, because of an insufficient biomass supply, a living bud may stay dormant for one or several cycles but keep alive. Let  $\mathcal{P}_A$  be the probability that a living bud grows in a cycle. When both  $\mathcal{P}_C$  and  $\mathcal{P}_B$  are equal to 1, the number of growth units in an axis follows a binomial law ( $\mathcal{M}_A$ ,  $\mathcal{P}_A$ ).
- *Appearance probability of metamers.* The number of metamers inside a growth unit varies. Let  $\mathcal{P}_I$  be the probability that each possible metamer appears. Thus, in a growth unit, the number of metamers follows a binomial law of parameters ( $\mu$ ,  $\mathcal{P}_I$ ).

These probabilities can effect together and make the one-step transition as from  $S_0$  to  $S_1$  in Fig. 2 stochastic, as well as the resulting substructures. Take the bud of PA=1, we show in Fig. 3 how the probabilities act on the one-step transition of buds. One-step transition of buds of other physiological ages is similar. However, it is supposed that the bud behaviour is independent for different physiological age.

### 3. Moments of the numbers of metamers

In FSPMs, the number of organs are important state variables for computing biomass production and its allocation [41]. They are linked to the development stage of the plant, the vigor of the plant, the yield, etc. They are also the

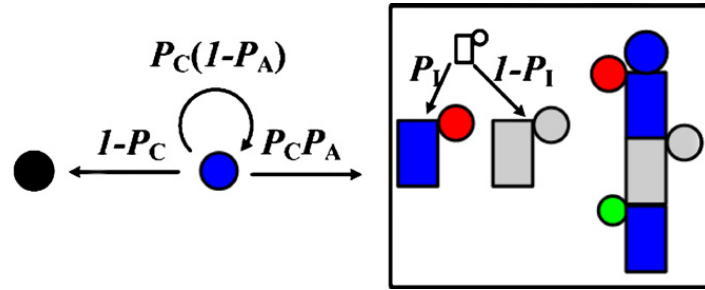


Fig. 3. One-step transition with bud probabilities, illustrated with the bud of PA 1 from Fig. 2. The black circle in the left means a dead bud. In each growth cycle, a bud survives with probability  $\mathcal{P}_C$  (if it is the first cycle of an axis, the bud first starts with probability  $\mathcal{P}_B$ ). A living bud either stays in bud state, or develops into a growth unit with probability  $\mathcal{P}_A$ . The number of metamers in the growth unit varies. Each possible metamer (the small white one) appears with probability  $\mathcal{P}_I$ . Here, given the maximum number of metamers in the growth unit  $\mu = 3$ , finally two metamers appear.

reference of model computation complexity. It is interesting to know them when rendering the 3D plant. Moments of variables from a discrete stochastic model can be obtained through Monte-Carlo simulation by counting them from a large number of random samples. This procedure may be quite expensive due to slow convergence, while analytical results are immediate and precise. As the metamer is the basic botanical unit, we are interested here in the mean and variance of the numbers of metamers. We suppose that there is no reiteration (that is to say that matrix  $\mathcal{N}_B$  is strictly upper triangular).

Since several kinds of probabilities occur together in GreenLab, the plant development can be regarded as a compound process (Appendix A). In this section, several simple and compound variables are presented, for counting the number of metamers in a plant. The numbers of metamers are computed based on substructure factorization, as well as their mean and variance.

### 3.1. Stochastic variables linked to the bud activities

#### 3.1.1. Simple variables

The following random variables are linked to one of the bud probabilities. To distinguish them from the following compound variables that are results of several probabilities, they are named *simple* variables. Some of them are indicator variables taking value of either 1 or 0.

- $x_i^p$ : if an apical bud of bearing axis of PA  $p$  is still alive at cycle  $i$ ,  $x_i^p=1$ ; otherwise  $x_i^p=0$ . The probability that the apical bud still survives at cycle  $i$  is  $(\mathcal{P}_C^p)^i$ . So the mean and variance of  $x_i^p$  are

$$M_{x_i^p} = (\mathcal{P}_C^p)^i, \quad V_{x_i^p} = (\mathcal{P}_C^p)^i [1 - (\mathcal{P}_C^p)^i] \quad (1)$$

Furthermore, the covariance between  $x_i^p$  and  $x_j^p$  is given as follows:

$$\text{Cov}_{x_i^p, x_j^p} = (\mathcal{P}_C^p)^j [1 - (\mathcal{P}_C^p)^i], \quad (j > i) \quad (2)$$

As the apical bud mutes into another PA after  $\mathcal{M}_A^p$  cycles, one has  $x_i^p = 0, i > \mathcal{M}_A^p$

- $y^p$ : if a living bud of PA  $p$  produces a growth unit during a cycle,  $y^p = 1$ ; otherwise  $y^p = 0$ . The mean and variance of  $y^p$  are

$$M_{y^p} = \mathcal{P}_A^p, \quad V_{y^p} = \mathcal{P}_A^p (1 - \mathcal{P}_A^p) \quad (3)$$

- $v^p$ : the total number of metamers in a growth unit of PA  $p$ . It is a binomial variable, whose mean and moments are

$$M_{v^p} = \mu^p \mathcal{P}_I^p, \quad V_{v^p} = \mu^p \mathcal{P}_I^p (1 - \mathcal{P}_I^p) \quad (4)$$

- $z^{p,k}$ : the number of metamers that bear axillary buds of PA  $k$  in a growth unit of PA  $p$ . Similar to  $v^p$ , the mean and variance of  $z^{p,k}$  are

$$M_{z^{p,k}} = \mathcal{M}_I^{p,k} \mathcal{P}_I^p, \quad V_{z^{p,k}} = \mathcal{M}_I^{p,k} \mathcal{P}_I^p (1 - \mathcal{P}_I^p) \quad (5)$$

- $w^{p,k}$ : the number of buds that break out to grow from a metamer of PA  $p$  that bears  $\mathcal{N}_B^{p,k}$  buds of PA  $k$ . The mean and variance of  $w^{p,k}$  are

$$M_{w^{p,k}} = \mathcal{N}_B^{p,k} \mathcal{P}_B^k, \quad V_{w^{p,k}} = \mathcal{N}_B^{p,k} \mathcal{P}_B^k (1 - \mathcal{P}_B^k) \quad (6)$$

It should be noticed that  $x_i^p$  and  $x_i^q$  ( $p \neq q$ ) are independent, i.e.  $\text{Cov}_{x_i^p, x_i^q} = 0$ . It is the same for the other variables of different PA.

### 3.1.2. Compound variables

The following variables are linked to more than one of the bud probabilities. Their mean and variance are computed from those of simple variables according to the compound law in Eq. (A.4).

- $\gamma^p$ : the number of metamers produced by a living bud of PA  $p$ . One has  $\gamma^p = \sum_{z=0}^{y^p} v_p$ . According to Appendix A,  $\gamma^p$  is a compound variable of  $y^p$  and  $v^p$ , noted as  $\gamma^p = y^p \circ v^p$ . By applying Eq. (A.4), the mean and variance of  $\gamma^p$  can be computed from those of  $y^p$  and  $v^p$

$$M_{\gamma^p} = M_{y^p} \cdot M_{v_p}, \quad V_{\gamma^p} = M_{y^p} \cdot V_{v_p} + V_{y^p} \cdot M_{v_p}^2 \quad (7)$$

- $q^{p,k}$ : the number of axillary substructures of PA  $k$  in a growth unit of PA  $p$ . It is a compound variable:  $q^{p,k} = \sum_{z=0}^{y^p} \sum_{m=0}^{z^{p,k}} w^{p,k} = y^p \circ z^{p,k} \circ w^{p,k}$ . Applying Eq. (A.4) twice, one gets

$$M_{q^{p,k}} = M_{y^p} \cdot M_{z^{p,k}} \cdot M_{w^{p,k}}, \quad V_{q^{p,k}} = M_{y^p} [M_{z^{p,k}} \cdot V_{w^{p,k}} + V_{z^{p,k}} \cdot M_{w^{p,k}}^2] + V_{y^p} \cdot M_{z^{p,k}}^2 M_{w^{p,k}}^2 \quad (8)$$

- $u_i^p$ : the number of metamers produced at  $i$ th cycle in a bearing axis of PA  $p$ . One has  $u_i^p = \sum_{y^p=0}^{x_i^p} \gamma^p = x_i^p \circ \gamma^p$ . Then the mean and variance of  $u_i^p$  are

$$M_{u_i^p} = M_{x_i^p} \cdot M_{\gamma^p}, \quad V_{u_i^p} = M_{x_i^p} \cdot V_{\gamma^p} + V_{x_i^p} \cdot M_{\gamma^p}^2 \quad (9)$$

The covariance between  $u_i^p$  and  $u_j^p$  is

$$\text{Cov}_{u_i^p, u_j^p} = \text{Cov}_{x_i^p \circ \gamma^p, x_j^p \circ \gamma^p} = M_{\gamma^p}^2 \text{Cov}_{x_i^p, x_j^p} \quad (10)$$

### 3.2. Counting numbers of metamers based on substructure factorization

Now let  $s_n^p$  be the number of metamers that appear at CA  $n$  inside a substructure of PA  $p$ . They are vectors:  $s_n^1 = [s_n^{1,1}, s_n^{1,2}, \dots, s_n^{1,P_m}]$ ,  $s_n^2 = [0, s_n^{2,2}, \dots, s_n^{2,P_m}]$ , ...,  $s_n^{P_m} = [0, 0, \dots, s_n^{P_m, P_m}]$ .  $s_n^{p,k}$  is the number of metamers of PA  $k$  in the substructure of PA  $p$ . When  $k < p$ , one has  $s_n^{p,k} = 0$ , from the definition of physiological age.  $s_n^p$  are random variables because of bud probabilities.

From Fig. 2 one can see that in the deterministic case, the parts of plant structure borne by a bud of the same PA and CA are topologically the same. The substructures are organized by hierarchical level, that is, substructures contain substructures of higher PA. A structure can be decomposed into a bearing axis, the axillary substructures along the axis, and the terminal substructure borne by the apical bud after mutation. This gives the recurrent equation for counting the number of metamers in the structures, as shown in the following equation:

$$s_i^{p,m} = \begin{cases} u_i^p = x_i^p \circ \gamma^p, & p = m \\ \sum_{a=1}^{i_0} x_a^p \circ \sum_{k=p+1}^{P_m} q^{p,k} \circ s_{i-a}^{k,m} + s_{i-i_0}^{\mathcal{J}_A^p, m}, & p < m \end{cases} \quad (11)$$

$i_0 = \min(i, \mathcal{M}_A^p)$ ,  $q^{p,k}$  is a compound variable as introduced in Section 3.1.2.



The accumulated number of metamers inside a substructure of PA  $p$  and CA  $n$ ,  $S_n^p$ , is the sum of those that appeared at each cycle, as follows:

$$S_n^{p,m} = \sum_{i=1}^n s_i^{p,m} \tag{12}$$

To clarify the formula expression, we suppose that  $n \leq \mathcal{M}_A$ , thus no mutation of axis may happen.

### 3.3. Moments of the numbers of metamers produced at each cycle $s_n^p$

- Case  $p = m$

These metamers are from the bearing axis if there is no reiteration. The mean and variance of  $s_i^{p,m}$  in this case are as in Eq. (13). The moments of  $x_i^p$  and  $\gamma^p$  are computed in Eqs. (1), (2) and (7)

$$M_{s_i^{p,m}} = M_X^p(i)M_{\gamma^p}, \quad \text{Cov}_{s_i^{p,m}, s_j^{p,m}} = V_X^p(i, j)M_{\gamma^p}^2 \tag{13}$$

where

$$M_X^p = \begin{bmatrix} M_{x_1^p} & M_{x_2^p} & \cdots & M_{x_n^p} \end{bmatrix}, \quad V_X^p = \begin{bmatrix} \text{Cov}_{x_1^p, x_1^p} & \text{Cov}_{x_1^p, x_2^p} & \cdots & \text{Cov}_{x_1^p, x_n^p} \\ \text{Cov}_{x_2^p, x_1^p} & \text{Cov}_{x_2^p, x_2^p} & \cdots & \text{Cov}_{x_2^p, x_n^p} \\ \cdots & \cdots & \cdots & \cdots \\ \text{Cov}_{x_n^p, x_1^p} & \text{Cov}_{x_n^p, x_2^p} & \cdots & \text{Cov}_{x_n^p, x_n^p} \end{bmatrix}$$

- Case  $p < m$

These are metamers from the substructures. Let  $t_{i-a}^{p,m} = \sum_{k=p+1}^{P_m} q^{p,k} \circ s_{i-a}^{k,m}$ , which is the number of metamers of PA  $m$  inside substructures of CA  $i - a$  attached to a growth unit of PA  $p$  that appeared at cycle  $a$ . Then one has  $s_i^{p,m} = \sum_{a=1}^i x_a^p \circ t_{i-a}^{p,m}$  according to Eq. (11). The mean of  $t_{i-a}^{p,m}$ , as well as the covariance between  $t_{i-a}^{p,m}$  and  $t_{j-b}^{p,m}$ , are computed as follows:

$$M_{t_{i-a}^{p,m}} = \sum_{k=p+1}^{P_m} M_{q^{p,k}} M_{s_{i-a}^{k,m}}, \quad \text{Cov}_{t_{i-a}^{p,m}, t_{j-b}^{p,m}} = \sum_{k=p+1}^{P_m} [M_{q^{p,k}} \text{Cov}_{s_{i-a}^{k,m}, s_{j-b}^{k,m}} + V_{q^{p,k}} M_{s_{i-a}^{k,m}} M_{s_{j-b}^{k,m}}] \tag{14}$$

where  $M_{q^{p,k}}$  and  $V_{q^{p,k}}$  are computed in Eq. (8). Notice that  $\text{Cov}_{s_{i-a}^{k,m}, s_{j-b}^{l,m}} = 0, l \neq k$ , as the bud behaviour of different PA are independent. Using the bilinearity of covariance, the mean of  $s_i^{p,m}$ , as well as the covariance between  $s_i^{p,m}$  and  $s_j^{p,m}$  ( $p < m$ ), are computed as follows:

$$M_{s_i^{p,m}} = \sum_{a=1}^i M_X^p(a) M_{t_{i-a}^{p,m}},$$

$$\text{Cov}_{s_i^{p,m}, s_j^{p,m}} = \sum_{a=1}^i \sum_{b=1}^j [M_X^p(\min(a, b)) \text{Cov}_{t_{i-a}^{p,m}, t_{j-b}^{p,m}} + V_X^p(a, b) M_{t_{i-a}^{p,m}} M_{t_{j-b}^{p,m}}] \tag{15}$$

### 3.4. Moments of the accumulated number of metamers $S_n^p$

The number of accumulated metamers of all cycles,  $S_n^{p,k}$ , shows the complexity of the resulting plant structure. Its mean and variance can be computed from those of  $s_n^{p,k}$ . The explicit recurrent form of the results is shown in Eqs. (16) and (17)

- Case  $p = m$

$$\begin{aligned}
 M_{S_n^{p,m}} &= \frac{\mathcal{P}_C^p [1 - (\mathcal{P}_C^p)^n]}{1 - \mathcal{P}_C^p} \mathcal{P}_A^p \mu^p \mathcal{P}_I^p, \\
 V_{S_n^{p,m}} &= \frac{\mathcal{P}_C^p [1 - (\mathcal{P}_C^p)^n]}{1 - \mathcal{P}_C^p} [\mathcal{P}_A^p \mu^p \mathcal{P}_I^p (1 - \mathcal{P}_I^p) + \mathcal{P}_A^p (1 - \mathcal{P}_A^p) (\mu^p \mathcal{P}_I^p)^2] \\
 &\quad + \frac{\mathcal{P}_C^p [1 - (2n + 1)(1 - \mathcal{P}_C^p)(\mathcal{P}_C^p)^n - (\mathcal{P}_C^p)^{2n+1}]}{(1 - \mathcal{P}_C^p)^2} (\mathcal{P}_A^p \mu^p \mathcal{P}_I^p)^2
 \end{aligned} \tag{16}$$

- Case  $p < m$

$$\begin{aligned}
 M_{S_n^{p,m}} &= \sum_{i=1}^n M_{x_i^p} \sum_{k=p+1}^{P_m} M_{q^{p,k}} M_{S_{n-i}^{k,m}}, \\
 V_{S_n^{p,m}} &= \sum_{i=1}^n \left\{ M_{x_i^p} \sum_{k=p+1}^{P_m} [M_{q^{p,k}} V_{S_{n-i}^{k,m}} + V_{q^{p,k}} M_{S_{n-i}^{k,m}}^2] + V_{x_i^p} \left[ \sum_{k=p+1}^{P_m} M_{q^{p,k}} M_{S_{n-i}^{k,m}} \right]^2 \right\} \\
 &\quad + 2 \sum_{i=1}^n \sum_{j=i+1}^n \text{Cov}_{x_i^p, x_j^p} \left[ \sum_{k=p+1}^{P_m} M_{q^{p,k}} M_{S_{n-i}^{k,m}} \right] \left[ \sum_{r=p+1}^{P_m} M_{q^{p,r}} M_{S_{n-j}^{r,m}} \right]
 \end{aligned} \tag{17}$$

$M_{x_i^p}$  and  $V_{x_i^p}$  are computed in Eq. (1). Both Eqs. (16) and (17) can be expressed in a matrix form of Eq. (A.4). Take the case  $p = m$  as an example. Let  $M_G^p = [M_{\gamma^p} \dots M_{\gamma^p}]$ ,  $V_G^p = [V_{\gamma^p} \cdot V_{\gamma^p}]$  (see Eq. (7) for  $M_{\gamma^p}$  and  $V_{\gamma^p}$ ). Then Eq. (16) can be written as follows:

$$M_{S_n^{p,p}} = M_G^p \cdot M_X^{p'}, \quad V_{S_n^{p,p}} = V_G^p \cdot M_X^{p'} + M_G^p \cdot V_X^p \cdot M_G^{p'} \tag{18}$$

When all probabilities are equal to 1, which is the case of the deterministic growth, the variances become zeros, and one obtains the same results as in the deterministic case [40], as shown in the following equation:

$$S_n^{p,m} = \begin{cases} n \cdot \mu^p, & p = m \\ \sum_{i=1}^n \sum_{k=p+1}^{P_m} \mathcal{M}_I^{p,k} \mathcal{N}_B^{p,k} S_{n-i}^{k,m}, & p < m \end{cases} \tag{19}$$

### 3.5. Moments of the number of organs $O_n^p$

Let  $O_n^p$  be the number of organs that appear at CA  $n$  inside a substructure of PA  $p$ .  $O$  is  $I$  for internode,  $L$  for leaf, and  $F$  for flower or fruit. Similar to  $s_n^p$ ,  $O_n^p$  are vectors. As there is one and only one internode per metamer, the number of internodes is the same as the number of metamers, i.e.,  $I_n^p = s_n^p$ . In GreenLab, the number of leaves per metamer,  $\mathcal{N}_L$ , are constant values. This is botanically meaningful as the number of leaves per metamer is very stable in real plants, for example, one for cotton plant and two for coffee tree. The procedure for obtaining the mean and variance of the number of leaves in plant structures,  $L_n^p$ , is the same as for  $s_n^p$  except that the variable  $v^p$  in Eq. (7) is to be replaced with  $v_L^p$ ,  $v_L^p = \sum_{k=p+1}^{P_m} z^{p,k} \cdot \mathcal{N}_L^{p,k}$ , the mean and variance of  $z^{p,k}$  being computed in Eq. (5). Then the mean and variance of  $v_L^p$  are

$$M_{v_L^p} = \sum_{k=p+1}^{P_m} M_{z^{p,k}} \mathcal{N}_L^{p,k}, \quad V_{v_L^p} = \sum_{k=p+1}^{P_m} V_{z^{p,k}} (\mathcal{N}_L^{p,k})^2 \tag{20}$$

The number of fruits per metamer, however, is not a constant due to flower bud abortion. Given a certain number of flower buds per metamer  $\mathcal{N}_F$ , suppose the probability that a flower bud gives eventually a fruit is  $\mathcal{P}_F$ . Then the

number of fruits in a metamer that bears axillary buds of PA  $k$  in a growth unit of PA  $p$ ,  $n_F^{p,k}$ , follows a binomial law  $(\mathcal{N}_F^{p,k}, \mathcal{P}_F^p)$ , whose mean and variance are

$$M_{n_F^{p,k}} = \mathcal{N}_F^{p,k} \mathcal{P}_F^p, \quad V_{n_F^{p,k}} = \mathcal{N}_F^{p,k} \mathcal{P}_F^p (1 - \mathcal{P}_F^p) \quad (21)$$

To obtain the mean and variance of the number of fruits  $F_n^p$ , the variable  $v^p$  in Eq. (7) is to be replaced with  $v_F^p$ , the total number of fruits inside a growth unit of PA  $p$ . Obviously  $v_F^p = \sum_{k=p+1}^{P_m} z^{p,k} \circ n_F^{p,k}$ . According to Eq. (A.4), the mean and variance of  $v_F^p$  are

$$M_{v_F^p} = \sum_{k=p+1}^{P_m} M_{z^{p,k}} M_{n_F^{p,k}}, \quad V_{v_F^p} = \sum_{k=p+1}^{P_m} (M_{z^{p,k}} V_{n_F^{p,k}} + V_{z^{p,k}} M_{n_F^{p,k}}^2) \quad (22)$$

The mean and variance of the number of fruits in plant structure can be computed following the same equations as for  $s_n^p$ .

### 3.6. Distribution of number of metamers in a bearing axis

In Eq. (16) the moments of the number of metamers are computed without knowing its probability distribution function. For a bearing axis of PA  $p$  and CA  $n$ , however, the distribution of number of metamers is given in Eq. (23). Here the index  $p$  is omitted as it is the same for each parameter

$$P(S_n^{p,p} = k) = \sum_{x=0}^{n-1} (1 - \mathcal{P}_C) (\mathcal{P}_C)^x \sum_{i=0}^x C_x^i \mathcal{P}_A^i (1 - \mathcal{P}_A)^{x-i} C_{i,\mu}^k \mathcal{P}_I^k (1 - \mathcal{P}_I)^{i,\mu-k} + \mathcal{P}_C^n \sum_{i=0}^n C_n^i \mathcal{P}_A^i (1 - \mathcal{P}_A)^{n-i} C_{i,\mu}^k \mathcal{P}_I^k (1 - \mathcal{P}_I)^{i,\mu-k}, \quad k \in [0, n \cdot \mu] \quad (23)$$

## 4. Plant biomass production

### 4.1. Biomass production in the GreenLab model GL2

In the GreenLab model, based on physiological assumptions [41], the biomass production of a single plant from photosynthesis at plant age  $n$ ,  $Q_n$ , is expressed as a recurrent function of the number of organs and production in the previous cycles, as shown in the following equation:

$$Q_n = E_n \sum_{i=1}^{t_a} \frac{N_{n-i+1}^L \sum_{j=1}^i p_L(j) Q_{n-(i-j)-1} / D_{n-(i-j)}}{\alpha + \beta \sum_{j=1}^i p_L(j) Q_{n-(i-j)-1} / D_{n-(i-j)}} \quad (24)$$

$Q_0$  is the seed biomass.  $E_n$  is the environmental factor providing a growth potential,  $N_{n-i+1}^L$  is the number of leaves of CA  $i$  at plant age  $n$ . It is assumed here that leaves of different physiological ages have the same functioning parameters, thus  $N_i^L = \sum_{k=1}^{P_m} L_i^{1,k}$ .  $p_L(j)$  is the sink strength of a leaf of CA  $j$ , and  $t_a$  is the leaf functioning duration.  $\alpha = r_1 e$ ,  $\beta = r_2$ , where  $e$  is the leaf thickness,  $r_1$  and  $r_2$  are two resistance terms linked to the blade and petiole of the leaves. The formulation of photosynthetic production for plants from a population is described in [17].  $D_n$  is the demand of the plant at age  $n$ , being the sum of sink strengths of all expanding organs, as in Eq. (25). In case of stochastic growth ( $N^O$  being random), both the plant biomass production and demand are random variables

$$D_n = \sum_O \sum_{i=1}^n p_i^O \cdot N_{n-i+1}^O \quad (25)$$

The accumulated plant biomass,  $W_n = \sum_{i=1}^n Q_i$ , is of special interest for agricultural or forestry applications.

For deciduous plants, leaf functioning time is one cycle ( $t_a=1$ ), and organs have immediate expansion. To illustrate biomass production in a simplified situation, let us suppose organs (only leaves and internodes here) have the same sink strength ( $p_L = p_I = 0.5$ ), and each metamer bears one leaf, then the demand of the plant is  $D_n = p_L N_n^L + p_I N_n^I = N_n^L$ . If we assume the climate condition is constant, i.e.,  $E_n \equiv 1$ , then Eq. (24) can be simplified to the following equation:

$$Q_n = \frac{N_n^L Q_{n-1}}{\alpha' N_n^L + \beta Q_{n-1}} = f(N_n^L, Q_{n-1}) \quad (26)$$

where the biomass production at cycle  $n$  depends only on the number of new leaves in the current cycle and the biomass production of the previous cycle.

#### 4.2. Computing moments of biomass production ( $Q_n$ )

The method of differential statistics described by Johnson and Kotz [24] is used to obtain the approximate and analytical mean and variance of biomass production. A brief introduction to differential statistics is given in Appendix B.

As  $Q_n = f(N_n^L, Q_{n-1})$ , according to Eq. (B.1), mean and variance of  $Q_n$  can be computed if those of  $N_n^L$  and  $Q_{n-1}$ , and  $\text{Cov}(N_n^L, Q_{n-1})$  are known. The first and second order derivatives of function  $f$  are computed using symbolic computation software.  $N_n^L$  is a random variable resulting from bud functioning, whose mean, variance and covariance between different cycles,  $M_i^L, V_i^L, \text{Cov}_{i,j}^L$  ( $1 \leq i, j \leq n$ ), are computed with the formulae given in Section 3. Being a function of  $N_n^L$ ,  $Q_n$  depends on  $N_n^L$ . However,  $N_n^L$  does not depend on  $Q_n$ , as in GL2, there is no feedback from biomass production on plant development. Notice that  $\text{Cov}(N_i^L, Q_j) \neq \text{Cov}(N_j^L, Q_i)$  as  $N$  and  $Q$  are two stochastic vectors.

The mean and variance of  $Q_i, \text{Cov}(Q_i, Q_j), \text{Cov}(N_i^L, Q_j)$  and  $\text{Cov}(N_j^L, Q_i)$  are computed by induction as follows:

- (1)  $\text{Cov}(N_i^L, Q_0) = 0$ , i.e., the number of leaves at cycle  $i$  and seed biomass are independent. Suppose  $\text{Cov}(N_i^L, Q_{j-1})$  ( $1 \leq j < i$ ) is known,  $\text{Cov}(N_i^L, Q_j)$  is computed according to Appendix C.
- (2)  $\text{Cov}(N_0^L, Q_i) = 0$  as  $N_0^L = 0$ . Suppose  $\text{Cov}(N_{j-1}^L, Q_i)$  ( $1 \leq j < i$ ) is known,  $\text{Cov}(N_j^L, Q_i)$  is computed according to Appendix C.
- (3)  $M_{Q_0} = 0$  and  $V_{Q_0} = 0$  as  $Q_0$  is a known variable. Suppose  $M_{Q_{i-1}}$  and  $V_{Q_{i-1}}$  are known.  $\text{Cov}(N_i^L, Q_{i-1}), M_{N_i}$  and  $V_{N_i}$  are already computed. Then  $M_{Q_i}$  and  $V_{Q_i}$  are computed according to Eq. (26) and Eq. (B.1).
- (4) We have  $\text{Cov}(Q_0, Q_i) = 0$ . Suppose  $\text{Cov}(Q_{j-1}, Q_i)$  ( $1 \leq j < i$ ) is known.  $\text{Cov}(Q_j, Q_i)$  is computed according to Appendix C.

As mean and covariance of  $Q_i$  ( $1 \leq i \leq n$ ) are all computed, mean and variance of  $W_n$  are easily computed. For agricultural applications, these results are important to evaluate the expected yield of a field. The production depends on the probabilities, which in turn can be strongly influenced by the environment. Thus one can analyze the effect of external conditions (pests, fertilizer, planting density, etc.) on the production variation once the model is well calibrated. Above we have shown the computation procedure for the simple case where the organ functioning time is only one cycle, which is the case of Eq. (26). For crops, where often  $t_a > 1$ , like in Eq. (24), the principle is the same. For example, if  $t_a = 2$ , we have  $Q_n = f(N_n^L, N_{n-1}^L, Q_{n-1}, Q_{n-2})$ , and a similar formula like Eq. (B.1) can be used to obtain its mean and variance.

## 5. Case study

All computations in this section were done by the GreenScilab software (<http://www.greenscilab.org>). It was developed by the authors in the open source software for scientific computation *Scilab*. The GreenScilab toolbox can be used not only for simulating and visualizing plants, but also for estimating model parameters from measurements made on real plants.



Fig. 4. Deterministic (complete) tree growth at plant age (a) 10, (b) 15 and (c) 20, where all probabilities are 1s.

### 5.1. Model parameters

Here we study an example of a stochastic plant simulated with the GreenLab model, to show the usage and advantage brought by the analytical results. We study the Rauh architectural tree model with the following parameters ( $P_m = 4$ ). This architectural model is chosen as its structure is complex enough to be representative (for example the *Ginkgo Biloba* tree)

$$\mathcal{M}_A = [20 \quad 10 \quad 5 \quad 3], \quad \mathcal{J}_A = [2 \quad 3 \quad 4 \quad 0], \quad \mathcal{M}_I = \begin{bmatrix} 0 & 2 & 3 & 4 \\ 0 & 0 & 2 & 3 \\ 0 & 0 & 1 & 2 \\ 0 & 0 & 0 & 2 \end{bmatrix},$$

$$\mathcal{N}_B = \begin{bmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$$

Let all organs (internodes and leaves here) be of the same functioning time ( $t_a=1$ ) and the same sink strength ( $p_L = p_I = 0.5$ ),  $r_1 = 17$ ,  $r_2 = 0.4$ ,  $e = 0.05$  (thus  $\alpha = 0.85$ ,  $\beta = 0.4$ ),  $E = 1$ ,  $Q_0 = 10$ . If all bud probabilities are 1s, together with some given geometrical parameters (not shown), the simulated plant structures at different ages are as shown in Fig. 4.

Now let the values for probabilities for each PA be different from 1, which is the main interest of this paper. The values are given here arbitrarily

$$\mathcal{P}_C = [.99 \quad .97 \quad .98 \quad 1], \quad \mathcal{P}_A = [.94 \quad .91 \quad .92 \quad 1], \quad \mathcal{P}_I = [.86 \quad .9 \quad .83 \quad 1],$$

$$\mathcal{P}_B = [1 \quad .9 \quad .85 \quad 1]$$



Fig. 5. Samples of stochastic plant structures at plant age 20.

Three plant structures at plant age 20 were simulated with the stochastic substructure algorithm [25], as shown in Fig. 5. The resulting plant structures are then stochastic and less dense, as some buds did not grow, and some buds produced less metamers. But they present statistical similarity as they result from the same distribution.

### 5.2. Distribution of the number of metamers in an axis

It is interesting to see how the compound distribution of the number of metamers results from  $\mathcal{P}_C$ ,  $\mathcal{P}_A$  and  $\mathcal{P}_I$ . Both the theoretical and simulated results are given, shown in Fig. 6. It can be seen that the distribution is irregular because  $\mathcal{P}_C$  are not 1s. Otherwise the distribution can be well approximated by a normal law.

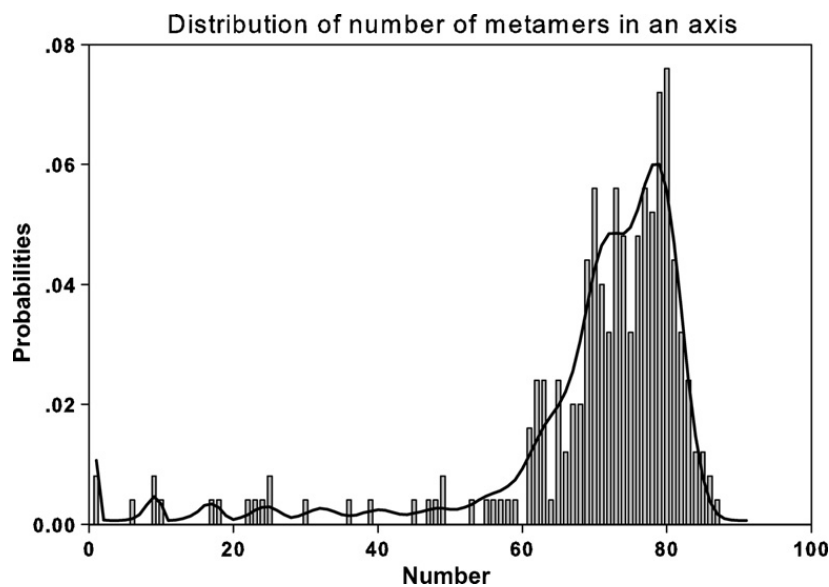


Fig. 6. Probability distribution function of the number of metamers in the bearing axis of PA 1 at age 10. Parameters are from Section 5.1. Solid line: the theoretical distribution, from Eq. (23). Histogram: results of a Monte-Carlo simulation of 250 samples.

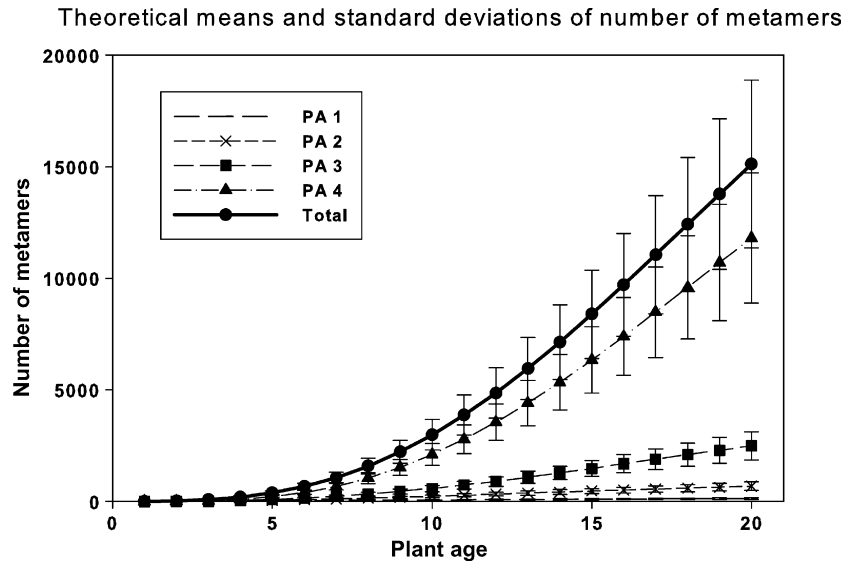


Fig. 7. Mean and standard deviation of number of metamers of each PA and their sum during plant growth, with parameters from Section 5.1.

### 5.3. Computation of the theoretical moments of number of metamers

The moments of the accumulated number of metamers at a given plant age show the complexity of the structure. They are computed with the formulae in Section 3.4. In the case when all probabilities are set to 1s, as in Fig. 4, at plant age 20, the number of metamers of PA 1–4 are 180, 1450, 7875, 33,102, respectively (without pruning of dead organs). In the case of stochastic development, the mean of the total number of metamers at plant age 20 is only 35.5% of that of the complete structure because of bud activity (Fig. 7).

While the mean of the total number of metamers is the sum of that of each PA, this is not the case for the variance, as the numbers of metamers of different PA are not necessarily independent variables. It can be seen that both mean and standard deviation increase when this virtual plant develops. It means the individual plants show more absolute variation when they become older.

### 5.4. Computation efficiency

It is obvious that to compute the moments with a formula is much more efficient compared to getting statistical results from simulated samples. A direct comparison is that the mean and variance of a variable from the binomial distribution  $(N, b)$  are immediate to compute using formula  $M = N \cdot b$ ,  $V = N \cdot b \cdot (1 - b)$ , while to get them through Monte-Carlo simulation is time consuming: each sample concerns  $N$  Bernoulli trials, and a big enough sample size  $T$  is needed to obtain desirable results.

Nevertheless, to provide a figure concerning the efficiency of the analytical method in this paper, the computation time was compared to simulation results. To simulate a branching plant structure, either the prefixed order algorithm [10] can be taken, where each metamer in each sample plant is generated one by one, or the stochastic substructure algorithm [25], where a library of stochastic substructure samples is built from  $P_m$  to PA 1, each sample being simulated by generating an axis and picking substructures from the library randomly, until PA 1 (the plant itself). While with prefixed order algorithm the time to simulate a plant is proportional to the number of metamers in plant, with the stochastic substructure algorithm, time to build a plant is proportional to  $P_m \times n^2 \times T$ ,  $T$  being the sample size of substructures [8]. For a plant of complex structure, the stochastic substructure algorithm can decrease dramatically the simulation time compared to the prefixed order algorithm [25] as it does not depend on the number of metamers to create. It is specially interesting to simulate numerous plant samples as the library can be reused once it has been built. In GreenScilab software, the stochastic substructure algorithm is used. For Fig. 5, the sample sizes are  $T = [3 \ 4 \ 4 \ 4]$ . Such small sample sizes are enough for visual output. Higher sample sizes are required for satisfactory statistical results, meanwhile the simulation gets more time-consuming. Table 1 shows the computation time of theoretical results and

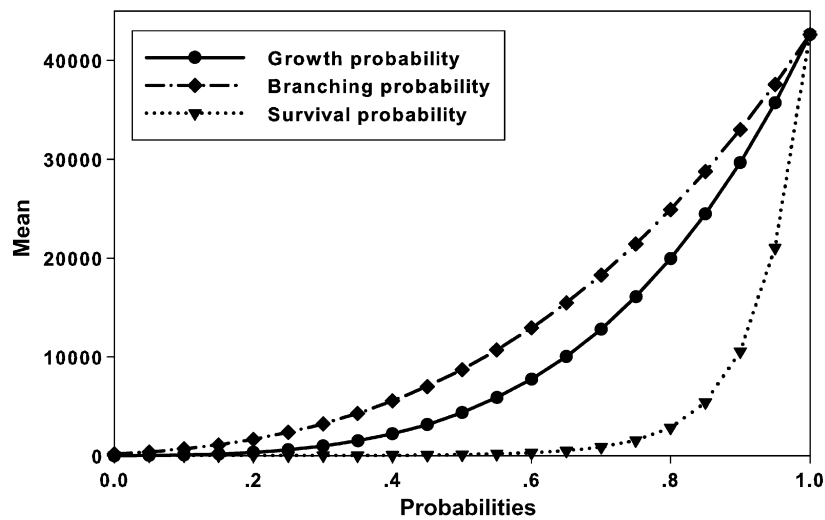
Table 1

Computation time in GreenScilab for theoretical and simulated results (sample size  $x$ ) respectively, as well as the corresponding mean and variance

	Time (s)	Mean	Variance
Theoretical	0.45	15120.1	14181092.6
Simulated ( $x=100$ )	44.25	16475.1	19391034.4
Simulated ( $x=500$ )	223.05	15409.2	17278379.5
Simulated ( $x=1000$ )	450.41	15158.5	12620310.0

the simulated results with the stochastic substructure algorithm at plant age 20. Computation time of the theoretical results is equivalent to the time of computing a deterministic plant, which is proportional to  $P_m \times n^2$ . For simulation, the sample size is set to  $T = [x \ x \ 1]$ . For PA 4 only one sample is needed as all the probabilities for PA 4 are 1. A higher sample size  $x$  is more time-consuming, but the results are closer to the theoretical values. It is obvious from Table 1 that the computation time for the theoretical results is much lower than those of the simulated results. The latter increase linearly with the sample size  $x$ .

(a) Effect of bud probabilities on mean of number of metamers



(b) Effect of bud probabilities on variance of number of metamers

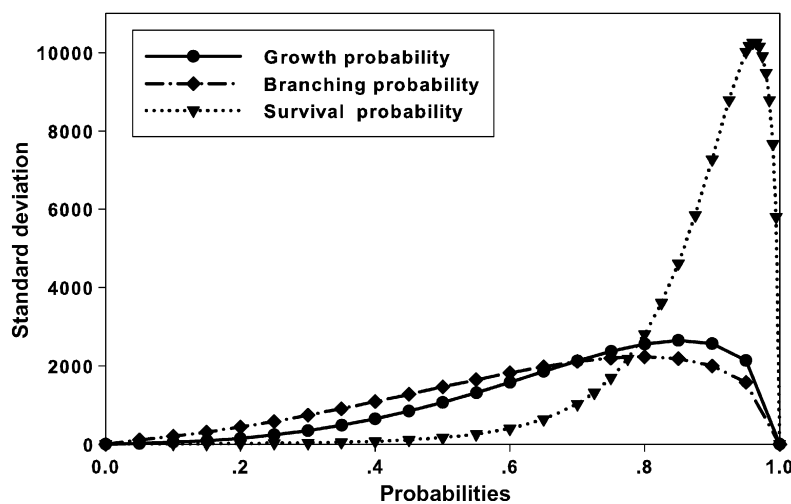


Fig. 8. Effects of different kinds of probabilities on (a) mean and (b) standard deviation of the numbers of metamers at plant age 20, with parameters from Section 5.1.



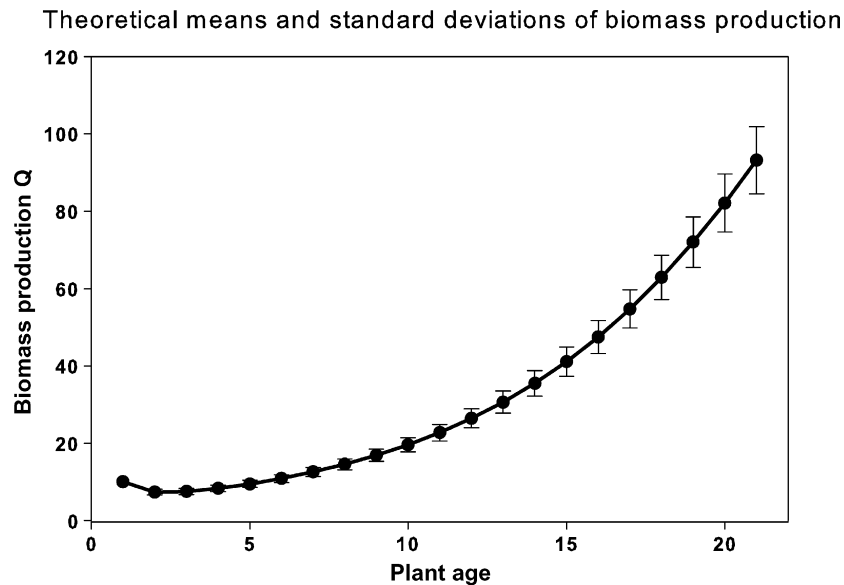


Fig. 9. Evolution of the moments of the biomass production at each cycle during plant growth, with parameters from Section 5.1.

### 5.5. Analysis of effects of the bud probabilities

Direct computation of means and variances facilitates the parametric analysis of the model. We can compare the effects of different probabilities on the number of metamers. We have set here the same values for each type of probability. When the value of one kind of probability varies from 0 to 1, the others are set to 1 to distinguish clearly the effect of that probability. Results are shown in Fig. 8. It is obvious that the survival probability of buds  $\mathcal{P}_C$  has a dramatic influence on mean and variance as it can interrupt the growth.

### 5.6. Efficient computation of biomass production

The analytical mean and variance of biomass production were computed using the method in Section 4.2, as shown in Fig. 9. As for the number of metamers, the inductive computation of the moment of biomass production is much quicker than the computation by Monte-Carlo simulation. The total biomass production (the plant weight) is the sum of the biomass productions at all cycles. Thus, its moments can be computed using the covariance between the biomass production at all cycles, as explained in Section 4.2.

## 6. Conclusion and discussion

Generally, stochastic models are more relevant than deterministic ones as they allow a representation of variability among plants, even though deterministic models are useful in showing clearly the effects of changes on plant development [14]. In this work, the theoretical results provide a fast and accurate way to predict the model outputs, and to analyze a stochastic growth model, as shown in Section 5. For example, in Fig. 8, it is shown how the survival probability of buds influences strongly the number of metamers without resorting to simulation. These results also provide a strong support for simulation software by allowing to check their validity, as the software that simulate the functional–structural plant models are becoming more and more complex [41]. When calibrating this probabilistic model for a real plant, these theoretical results are very useful in obtaining the values of probabilities with the maximum likelihood estimation method. For example, such an application has been used for modeling the root system of wheat seedlings [43].

The probabilistic model can be partially modified according to the object of interest. From observations on coffee trees, de Reffye [13] showed that the probabilities are not stable during development. In that case, the computing procedure presented in Section 3 keeps that the same except that some moments will be time-dependent. For example, if the branching probabilities of buds  $\mathcal{P}_B$  varies along an bearing axis, only a simple variable  $w^{p,k}$  and then a compound

variable  $q^{p,k}$  are relevant, and the  $M_{q^{p,k}}$  and  $M_{q_i^{p,k}}$  in the general formula Eq. (14) are to be replaced by  $M_{q_i^{p,k}}$  and  $M_{q_i^{p,k}}$ . Likewise, if the distribution law of the number of metamers in a growth unit is not a binomial one as presented in this article, but a negative binomial instead, as observed by Costes et al. [7], only part of the computation will be changed. More generally, it was shown in [4] that with increasing branching orders, and/or age of the shoot, the frequency function changes to Poisson and finally to geometrical laws. In such cases, the mean and variance of the compound variables in Eqs. (13) and (15) need to be updated, but the frame of the computation remains the same.

In this paper, we also present an analytical way to compute the moments of biomass production for the stochastic GreenLab model GL2. It can be seen that the biomass production depends on the plant architecture. However, this method is an approximate one in nature as shown in Appendix B because of the truncation of Taylor series. Moreover, the error is accumulated when computing the results from plant age 1 to  $n$ . Further work is under way to obtain the theoretical probability distribution function of biomass production. Differing from the number of metamers, the biomass production is a continuous variable. On the other hand, in computing biomass production of crops, new biomass formulae based on Beer's law can be used, as in [17] and [26]. It is a negative exponential form, different from the one in Eq. (24). This should be taken into account when predicting production for agricultural fields.

In the stochastic model presented here, the resulting plant architectures are not influenced by biomass production. In reality, it is generally expected that the plant architecture is the result of both genes and environmental effects. The physiological influence on branching is underlined in [4]. Thus the stochastic bud behaviour must be highly related to biomass production, which is in turn dependant on the environment and the plant architecture. In a recent thesis on the GreenLab model by Mathieu [31], the mathematical formulae that describe the interactions between plant organogenesis and photosynthesis were presented, and it was shown how the bud activity of trees is highly related to the light conditions in a deterministic way. A new perspective is to implement these latest results in the stochastic model and make the probabilities depend on dynamic sink–source relationships (GL4). It will then be very promising to help in prediction of the growth and development of plant populations when reacting to the environment.

### Acknowledgements

The authors would like to thank Prof. Philippe Flajolet (INRIA) for his talk on generating functions in 2002, which played an important role in completing this work. Help from Jonathan le Roux (ENS) is gratefully acknowledged. This work is supported in part by LIAMA (Sino-French Laboratory in Information, Automation and Applied Mathematics), Natural Science Foundation of China (Nos. 60073007, 60473110), and China 863 Hi-Tech Programs (No. 2006AA10Z229, 2006AA01Z301).

### Appendix A. The compound process

Given a random variable  $X$  that takes only non-negative integer values, the probability generating function (PGF) of  $X$  is defined as

$$g_X(u) = \sum_{k \geq 0} P(X = k)u^k \tag{A.1}$$

$P(X = k)$  is the probability that the variable  $X$  takes the value  $k$ .

The probability generating function provides a way to get the mean and variance without the tedious calculations involving discrete sums. According to the definition of moments, it is easy to prove that the mean and variance of  $X$  can be computed by differentiating the PGF in the following equation:

$$M_X = g'_X(1), V_X = g''_X(1) + g'_X(1) - g'_X(1)^2 \tag{A.2}$$

More interesting, the generating function satisfies a compound law. Suppose a stochastic variable  $Z = X_1 + X_2 + \dots + X_Y$ , where  $X_j (1 \leq j \leq Y)$  are independent and identically distributed variables,  $Y$  is an integer random variable. Then the generating function of variable  $Z$  is

$$g_Z(u) = g_Y(g_X(u)) = g_Y \circ g_X(u) \tag{A.3}$$

Note  $Z = Y \circ X$ . Through Eq. (A.2), one derives the mean and variance of  $Z$ ,  $M_Z$  and  $V_Z$ , from those of  $X$  and  $Y$  directly

$$M_Z = M_X \cdot M_Y, \quad V_Z = M_Y \cdot V_X + V_Y \cdot M_X^2 \quad (\text{A.4})$$

More details can be found in [35].

### Appendix B. The differential statistics [24]

Let  $X, Y, Z$  be stochastic variables such that  $Z = f(X, Y)$ . If the first two moments of variables  $X$  and  $Y$ ,  $M_X, M_Y, V_X, V_Y$ , and their covariance  $\text{Cov}_{X,Y}$  are known, the moments of  $Z$  can be approximated as follows:

$$\begin{aligned} M_Z &\approx f(M_X, M_Y) + \frac{1}{2}[f''_X(M_X, M_Y)V_X + f''_Y(M_X, M_Y)V_Y] + 2f''_{X,Y}(M_X, M_Y)\text{Cov}_{X,Y}, \\ V_Z &\approx f'_X(M_X, M_Y)^2 V_X + f'_Y(M_X, M_Y)^2 V_Y + 2f'_X(M_X, M_Y)f'_Y(M_X, M_Y)\text{Cov}_{X,Y} \end{aligned} \quad (\text{B.1})$$

Similarly, for a function  $Z = g(X_1, X_2, X_3)$ , the mean of the stochastic variable  $Z$  can be approximated as

$$M_Z \approx g(M_{X_1}, M_{X_2}, M_{X_3}) + \frac{1}{2} \sum_{i=1}^3 \sum_{j=1}^3 g''_{X_i}(M_{X_1}, M_{X_2}, M_{X_3}) \text{Cov}_{X_i, X_j} \quad (\text{B.2})$$

### Appendix C. Computing $\text{Cov}(X_j, Q_i)$ based on differential statistics

$X_j$  represents either  $N_j^L$  or  $Q_j$ . According to the definition of covariance

$$\text{Cov}(X_j, Q_i) = M_{X_j \cdot Q_i} - M_{X_j} M_{Q_i} \quad (\text{C.1})$$

To obtain  $M_{X_j \cdot Q_i}$ , define the function as in the following equation:

$$X_j \cdot Q_i = X_j \cdot E \frac{N_i Q_{i-1}}{\alpha N_i + \beta Q_{i-1}} = g(X_j, N_i, Q_{i-1}) \quad (\text{C.2})$$

Then according to Eq. (B.2),  $M_{X_j \cdot Q_i}$  can be computed if mean, variance and covariance between any two variables of  $X_j, N_i, Q_{i-1}$  are known.

### References

- [1] J. Barczi, P. de Reffye, Y. Caraglio, Essai sur l'identification et la mise en oeuvre des paramètres nécessaires à la simulation d'une architecture végétale: le logiciel amapsim, in: J. Bouchon, P. de Reffye, D. Barthélémy (Eds.), Modélisation et Simulation de l'Architecture des Végétaux, INRA, Science Update, Paris, France, 1997.
- [2] D. Barthélémy, Y. Caraglio, E. Costes, Architecture, gradients morphogénétiques et âge physiologique chez les végétaux, in: J. Bouchon, P. de Reffye, D. Barthélémy (Eds.), Modélisation et Simulation de l'Architecture des Végétaux, INRA, Science Update, Paris, France, 1997.
- [3] J. Bloomenthal, Modeling the mighty maple, *Comput. Graph.* 19 (3) (1985) 305–311.
- [4] G. Buck-Sorlin, A. Bell, Models of crown architecture in *Quercus petraea* and *Q. robur*: shoot lengths and bud numbers, *Forestry* 73 (2000) 1–19.
- [5] D. Cohen, Computer simulation of biological pattern generation processes, *Nature* 216 (21) (1967) 246–248.
- [6] E. Costes, Y. Guédon, Modelling branching patterns on 1-year-old trunks of six apple cultivars, *Ann. Bot.* 89 (2002) 513–524.
- [7] E. Costes, P. Lauri, Y. Guédon, P. de Reffye, Modelling growth of peach trees using the renewal theory, *Acta Horticult.* 349 (1993) 253–258.
- [8] P.-H. Cournède, M.-Z. Kang, A. Mathieu, H.-P. Yan, B.-G. Hu, P. de Reffye, Structural factorization of plants to compute their functional and architectural growth, *Simulation, Trans. Soc. Modell. Simul. Int.* 82 (7) (2006) 427–438.
- [9] E. Dayan, E. Presnov, M. Fuchs, Prediction and calculation of morphological characteristics and distribution of assimilates in the rosgro model, *Math. Comput. Simul.* 65 (2004) 101–116.
- [10] P. de Reffye, C. Edelin, J. Francon, M. Jaeger, C. Puech, Plant models faithful to botanical structure and development, *Comput. Graph.* 22 (4) (1988) 151–158.
- [11] P. de Reffye, T. Fourcaud, F. Blaise, D. Barthélémy, F. Houllier, A functional model of tree growth and tree architecture, *Silva Fennica* 31 (1997) 297–311.
- [12] P. de Reffye, M. Goursat, J. Quadrat, B.-G. Hu, The dynamic equations of the tree morphogenesis GreenLab model, in: B.-G. Hu, M. Jaeger (Eds.), *Plant Growth Modeling and Applications*, Springer/Tsinghua University Press, Beijing, China, 2003.

- [13] P. de Reffye, Modélisation de l'architecture des arbres par des processus stochastiques. simulation spatiale des modèles tropicaux sous l'effet de la pesanteur. application au coffea robusta, Ph.D. thesis, Université de Paris-Sud centre d'Orsay, 1979.
- [14] J. Fisher, How predictive are computer simulations of tree architecture, *Int. J. Plant Sci.* 152 (1992) 137–146.
- [15] N. Greene, Voxel space automaton: modeling with stochastic growth processes in voxel space, *Comput. Graph.* 23 (3) (1989) 175–184.
- [16] Y. Guédon, D. Barthélémy, Y. Caraglio, E. Costes, Pattern analysis in branching and axillary flowering sequences, *J. Theoret. Biol.* 212 (2001) 481–520.
- [17] Y. Guo, Y.-T. Ma, Z.-G. Zhan, B.-G. Li, M. Dingkuhn, D. Luquet, P. de Reffye, Parameter optimization and field validation of the functional–structural model GreenLab for maize, *Ann. Bot.* 97 (2006) 217–230.
- [18] F. Hallé, R. Oldeman, P.B. Tomlinson, *Tropical Trees and Forests: an Architectural Analysis*, Springer-Verlag, New York, 1978.
- [19] T. Harris, *The Theory of Branching Processes*, Springer-Verlag, 1963.
- [20] P. Heuret, Y. Guédon, N. Guérard, D. Barthélémy, Analyzing branching pattern in plantation of young red oak trees (*Quercus rubra* L., Fagaceae), *Ann. Bot.* 91 (2003) 479–492.
- [21] E. Heuvelink, Evaluation of a dynamic simulation model for tomato crop growth and development, *Ann. Bot.* 83 (1999) 413–422.
- [22] H. Honda, J.B. Fisher, Tree branch angle: maximizing effective leaf area, *Science* 199 (24) (1978) 888–890.
- [23] H. Honda, Description of the form of trees by the parameters of the tree-like body: Effects of the branching angle and the branch length on the shape of the tree-like body, *J. Theoret. Biol.* 31 (1971) 331–338.
- [24] N. Johnson, S. Kotz, *Discrete distributions*, John Wiley & Sons., 1969.
- [25] M.-Z. Kang, P. de Reffye, J. Barczy, B.-G. Hu, F. Houllier, Stochastic 3D tree simulation using substructure instancing, in: B.-G. Hu, M. Jaeger (Eds.), *Plant Growth Modeling and Applications*, Springer/Tsinghua University Press, Beijing, China, 2003.
- [26] M.-Z. Kang, E. Heuvelink, P. de Reffye, Building virtual chrysanthemum based on sink–source relationships: preliminary results, *Acta Horticult.* 718 (2006) 129–136.
- [27] M.-Z. Kang, P. de Reffye, A mathematical approach to the estimation of sink strength of competing organs, in: J. Vos, L.F.M. Marcelis, P.H.B. de Visser, P.C. Struik (Eds.), *Functional-structural plant modelling in crop production*, Kluwer Academic Publishers, 2007, pp. 64–74.
- [28] W. Kurth, B. Sloboda, Growth grammars simulating trees—an extension of L-systems incorporating local variables and sensitivity, *Silva Fennica* 31 (3) (1997) 285–295.
- [29] A. Lindenmayer, Mathematical models for cellular interactions in development. Part I and II, *J. Theoret. Biol.* 18 (1968), 280–299, 300–315.
- [30] L. Marcelis, E. Heuvelink, J. Goudriaan, Modeling biomass production and yield of horticultural crops: a review, *Sci. Horticult.* 74 (1998) 83–111.
- [31] A. Mathieu, Essai sur la modélisation des interactions entre la croissance et le développement d'une plante: cas du modèle GreenLab, Ph.D. thesis, Ecole Centrale de Paris, 2006.
- [32] J. Perttunen, R. Sievänen, E. Nikinmaa, H. Salminen, H. Saarenmaa, J. Väkevä, Lignum: a tree model based on simple structural units, *Ann. Bot.* 77 (1996) 87–98.
- [33] P. Prusinkiewicz, J. Hanan, Developmental models of herbaceous plants for computer imagery purpose, *Comput. Graph.* 22 (4) (1988) 141–150.
- [34] W. Reeves, R. Blau, Approximate and probabilistic algorithms for shading and rendering structured particle systems, *Comput. Graph.* 19 (3) (1985) 313–322.
- [35] R. Sedgewick, P. Flajolet, *An Introduction to the Analysis of Algorithms*, Addison-Wesley, 1996.
- [36] R. Sievänen, E. Nikinmaa, P. Nygren, H. Ozier-Lafontaine, J. Perttunen, H. Hakulai, Components of functional–structural tree models, *Ann. Forest Sci.* 57 (2000) 399–412.
- [37] A. Smith, *Plants, fractals and formal languages*, *Comput. Graph.* 18 (3) (1984) 1–10.
- [38] X. Viennot, G. Eyrolles, N. Janey, D. Arquès, Combinatorial analysis of ramified patterns and computer imagery of trees, *Comput. Graph.* 23 (3) (1989) 31–39.
- [39] J. Vos, L. Marcelis, P. de Visser, P. Struik (Eds.), *Functional–Structural Plant Modelling in Crop Production*, Kluwer Academic Publishers, 2007.
- [40] H.-P. Yan, J. Barczy, P. de Reffye, B.-G. Hu, Fast algorithms of plant computation based on substructure instances, in: *Proceedings of the International Conferences in Central Europe on Computer Graphics, Visualization, Computer Vision*, vol. 3, 2002.
- [41] H.-P. Yan, M.-Z. Kang, P. de Reffye, M. Dingkuhn, A dynamic, architectural plant model simulating resource-dependent growth, *Ann. Bot.* 93 (2004) 591–602.
- [42] Z.-G. Zhan, P. de Reffye, F. Houllier, B.-G. Hu, Fitting a functional–structural growth model with plant architectural data, in: B.-G. Hu, M. Jaeger (Eds.), *Plant Growth Modeling and Applications*, Springer/Tsinghua University Press, Beijing, China, 2003.
- [43] B.-G. Zhang, P. de Reffye, L. Liu, M.-Z. Kang, B. Li, Analysis and modeling of the root system architecture of winter wheat seedling, in: B.-G. Hu, M. Jaeger (Eds.), *Plant Growth Modeling and Applications*, Springer/Tsinghua University Press, Beijing, China, 2003.
- [44] X. Zhao, P. de Reffye, F. Xiong, B.-G. Hu, Z.-G. Zhan, Dual-scale automaton model of virtual plant growth, *Chin. J. Comput.* 24 (6) (2001) 608–615.