Structural Factorization of Plants to Compute their Functional and Architectural Growth

Paul-Henry Cournède, M.Z. Kang, A. Mathieu, Jean-François Barczi, H.P. Yan, B.G. Hu, Philippe de Reffye

Structural Factorization of Plants to Compute their Functional and Architectural Growth

Cournède P.H.[1,2], Kang M.Z.[2,3,4], Mathieu A.[1,2,5], Barczi J.F.[5,1], Yan H.P.[3], Hu B.G.[3], de Reffye P.[5,2]

[1] *Laboratory of Applied Mathematics, Ecole Centrale Paris, 92295 France*
[2] *Digiplante, INRIA - Rocquencourt, 78153 France*
[3] *LIAMA, Institute of Automation, Chinese Academy of Sciences Beijing 100080, P.R China*
[4] *Capital Normal University, Beijing, P.R China*
[5] *AMAP, CIRAD Montpellier, 34398 France*

Abstract

Numerical simulation of plant growth has been facing a bottleneck due to the cumbersome computation implied by the complex plant topological structure. In this paper, we present a new mathematical model for plant growth, GreenLab, overcoming these difficulties. GreenLab is based on a powerful factorization of the plant structure. Fast simulation algorithms are derived for deterministic and stochastic trees. The computation time no longer depends on the number of organs and grows at most quadratically with the age of the plant. This factorization finds applications in the context of Geometric Models, to build trees very efficiently, and in the context of Functional Structural Models, to compute biomass production and distribution.

1 Introduction

Virtual plants are more and more used for different applications, mainly in computer graphics (urbanism and landscaping) and environmental sciences (agronomy and forestry). Sievänen in [Sievänen et al., 2000] distinguishes three kinds of models for virtual plants: Process Based Models (PBM), Geometric Models (GM) and Functional Structural Models (FSM).

Process Based Models focus on ecophysiological processes and are used to answer agronomic questions (see for example GOSSYM for cotton [Baker et al., 1983], CERES for Maize [Jones and Kiniry, 1986] or [Marcelis et al., 1998] for a review). They typically describe the flux of external resources through a system (the plant or the crop) into a pool (the yield). A set of rules and parameters govern resource acquisition and conversion on a field area basis, according to a phenological timetable. This approach is particularly robust for field crops forming homogeneous canopies. It is no longer true for heterogeneous populations or culture systems in which plant architecture and geometry are of specific interest to the producer. The development and management of these systems require models providing the plant architectural description, in addition to resource flows and yield. Such models should also integrate the plant phenotypic plasticity resulting from feedbacks between functional growth and architectural development.

Geometric Models refer to plant architectural models built from simulation algorithms for computer graphics needs, (for pioneering works, see [Aono and Kunii, 1984], [Reffye (de) et al., 1988] or [Prusinkiewicz and Lindenmayer, 1990]). They are only based on a topological description of the plant and some geometric rules. Handled by good designers, these models produce authentic looking pictures of plants. Recent interesting examples can be found in

[Deussen and Lintermann, 2005]). If it works quite well for little plants (at least from a visual point of view), limitations occur so far when it comes to the representation of trees, as their botanical developments is too complex and strongly influenced by physiological functioning. Evolved software built on this type of model is commonly used in landscaping, advertising or video games but is of little interest for agro-ecological research, as no biomass is computed.
Developed recently, Functional Structural Models are mainly devoted to agronomy and forestry. They attend to combine Process Based Models and Geometric Models in order to control both photosynthesis and organogenesis. It is generally done by performing a dynamic simulation of plant morphogenesis. The plant architectural development provides the sources according to which the photosynthesis is determined and the compartments among which the produced biomass is distributed, see for example LIGNUM model [Perttunen et al., 1996] and AMAPhydro [Reffye (de) et al., 1997].
The need for fast simulation and visualization tools is obvious for computer graphics purposes. It is also crucial for agro-ecological research. As a matter of fact, for applications, plant growth models have to be calibrated on real plant measurements. Parameter identification can only be performed by running a very large number of times the growth simulation process. The same issue is raised when solving optimal control or optimization problems, see [Wu, 2005].
Most simulation algorithms rely on the complete construction of the plant topology by mimicking the simultaneous bud functioning, see for example [Prusinkiewicz and Lindenmayer, 1990] and other related works based on L-systems or automata. For Geometric Models, Sievänen observed empirically that the computational time is proportional to the square of the number of items, see [Sievänen et al., 2000]. When it comes to Functional Structural Models, the simulation of source and sink processes, biomass acquisition and distribution requires exploring several times the plant topological structure. It increases heavily the computational time (up to 8 times). In AMAPhydro model, the total biomass production is obtained by summing the photosynthetic production of every leaf and biomass partitioning is computed by determining the individual demand of all organs and then allocating biomass to each of them, see [Blaise et al., 1998]. But for branching plants, the number of organs grows exponentially, which leads to cumbersome computations.
As a consequence, the algorithms can be drastically improved by taking into account the fact that a plant is often composed of a juxtaposition of similar substructures. In this paper, we will present an improved method of plant simulation based on a powerful factorization of the plant architectural organization. This factorization method is applied to the functional structural model GreenLab detailed in [Reffye (de) and Hu, 2003] and [Yan et al., 2004], for deterministic and stochastic simulations. A real breakthrough is achieved as the bottleneck linked to complex tree architectures is suppressed.
In section 2, we will present the main botanical and ecophysiological concepts used in the GreenLab model and build the dynamical system of plant growth. The factorization method is introduced in section 3. A new mathematical formalism is presented, which allows rewriting the equations of GreenLab and deriving a fast algorithm for plant growth simulation. In section 4, the advantages and drawbacks of the method are discussed, as well as the perspectives it opens for applications in both computer graphics and environmental sciences.

2 Basic Biological Concepts Used in the GreenLab Model

In this section, we present the botanical and ecophysiological concepts from which the GreenLab model was built. A consistent time unit for architectural growth and photosynthetic production is defined, which allows us to derive the discrete dynamical system of growth. A more detailed biological presentation of the GreenLab model can be found in [Reffye (de) and Hu, 2003] or

[Yan et al., 2004].

2.1 Architectural Organization

As explained in [Barthélémy et al., 1997], the architecture of a plant can be seen as a hierarchical branching system in which the axes can be grouped into categories characterized by a particular combination of morphological parameters. Thus, the concept of physiological age was introduced to represent the different types of axes. For instance, on coffee trees, there are two types: orthotropic trunk and plagiotropic branches. Usually, we need less than 5 physiological ages to describe the axis typology in a given tree, see [Sabatier and Barthélémy, 1999] and the example of Cedrus Atlantica. The main trunk has a physiological age of 1 and the oldest physiological age is the ultimate state of differentiation for an axis, it is usually short, without branches.
Moreover, botanists generally agree on the fact that plants are modular organisms that develop by the repetition of elementary botanical entities or constructional units, see for example [Bell, 1991] or [Barthélémy, 1991]. The most elementary of these entities is the metamer. It is composed of an internode bearing organs: buds, leaves, flowers. Depending on plants, metamers are set in place rhythmically or continuously.
In the rhythmic case, see Figure 1, the plant grows by successive shoots of several metamers produced by buds. The apparition of these shoots defines the architectural Growth Cycle. A Growth Unit is the set of metamers built by a bud during a growth cycle. These metamers can be of different kinds and ordered according to botanical rules, like acrotony. To give a simple example, most temperate trees grow rhythmically, new shoots appearing at spring. If we do not consider polycyclism, their architectural growth cycle correspond to a year.
Plant growth is said continuous when meristems keep on functioning and generate metamers one by one, see Figure 1. The number of metamers on a given axis (that is to say generated by the same meristem) is proportional to the sum of daily temperatures received by the plant. It is called the Law of the Sum of Temperatures in Agronomy, see [Jones, 1992] and [Gao et al., 1992]. The Growth Cycle is defined as the thermal time unit necessary for a bud to build a new metamer. Tropical trees, bushes or agronomic plants have a continuous growth.
In both continuous and rhythmic cases, the Chronological Age of a plant (or of an organ) is defined as the number of growth cycles it has been existing for.

2.2 Ecophysiological Concepts

GreenLab aims at providing a general framework for plant growth models and their applications to agronomy and forestry. For this purpose, it is necessary to add to the architectural model some ecophysiological submodels to compute the photosynthetic biomass production and its distribution among organs. Here, we present the simple hypotheses used in GreenLab to describe plant functioning. However, it is important to note that more complex models could be implemented in the same framework, see [Yan et al., 2004].
We determine the photosynthetic biomass production by the law of Water Use Efficiency, see [Howell and Musick, 1984]: the plant dry matter acquisition is approximately proportional to the plant transpiration. To be able to build the plant architecture, we need a notion of volume and weight of the wet plant biomass. For this purpose, in GreenLab, we convert the dry matter production into an available fresh biomass by assuming that the moisture content is constant. Therefore, in GreenLab, the water use efficiency will be considered as the conversion rate of water transpiration into fresh biomass. The fresh biomass thus computed is used to build organs. Each leaf will have a transpiration proportional to a potential transpiration rate and to its conductance depending on its surface area, see [Jones, 1992].

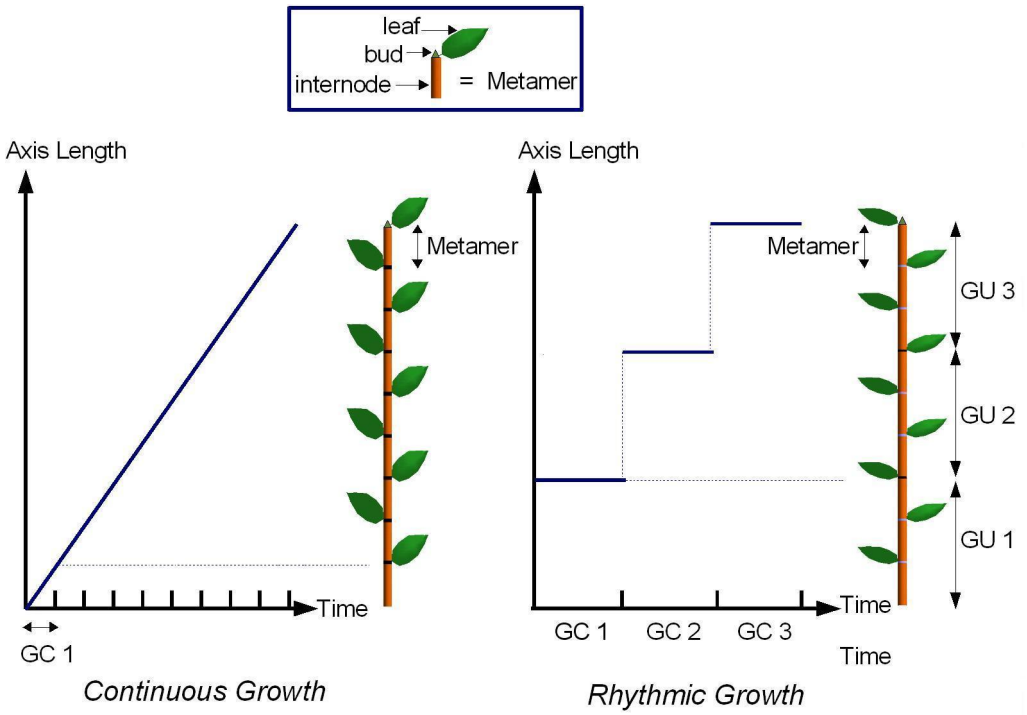

Figure 1: *The axis length is given as a function of time for continuous and rhythmic growths. The Growth Cycle is deduced in both cases: it is defined as the time unit necessary for a new metamer to be set in place when the growth is continuous, or the period at which the Growth Units (GU) appear when the growth is rhythmic. For the sake of simplicity, metamers are drawn here without lateral buds.*

Moreover, we do not consider any local diffusion process and make the hypothesis that the biomass produced by each leaf is stored in a common pool of reserves and redistributed among all organs according to their sink strengths. The initial seed and the leaves are sources. Leaves, internodes, fruits and rings (resulting from the plant secondary growth) are sinks. The root system is not described but is considered as a big sink at all stages of the plant growth. Secondary growth is the process controlling the increase of branch diameters. To describe it, we use the classical pipe model [Shinozaki et al., 1964]: the ring volume of a branch at a given spot in the plant architecture is proportional to the number of leaves seen above.

2.3 GreenLab Growth Cycle

The time unit for organogenesis and photosynthesis must be consistent, insuring a good synchronization between the two processes. In GreenLab, we consider that the new growth units appear at the beginning of the growth cycle and that, during this one, the plant architecture remains constant. As illustrated by Figure 1, it is a good approximation for rhythmic growth. In the case of continuous growth, it corresponds to a discretization of the system with respect to the growth cycle. Growth units are thus simply composed of one metamer. The discretization is reasonable in the continuous case as the cycle duration is usually quite short compared to that of the whole growth process. Consequently, we will no longer distinguish rhythmic and continuous growths.

We suppose as well that at the level of growth cycle, the fluctuations of water use efficiency and of potential transpiration rate are reasonable, and we use their average values at growth cycle n. Also knowing the architecture at growth cycle n, and thus leaf surface areas, we can determine their transpiration and the total amount of produced biomass, denoted $Q(n)$. At the end of the cycle, fresh matter is allocated to organs according to their relative demands (sinks): old organs in expansion, new organs preformed in buds (they will appear at the beginning of growth cycle $n+1$), root system, rings... Empirical allometric rules are used to determine the dimensions and shapes of organs from their volumes, and the plant architecture at growth cycle $n+1$ is thus obtained. Figure 2 illustrates the different steps to compute the plant functioning during its growth cycle. From this flowchart is derived the algorithms to simulate plant growth and development.

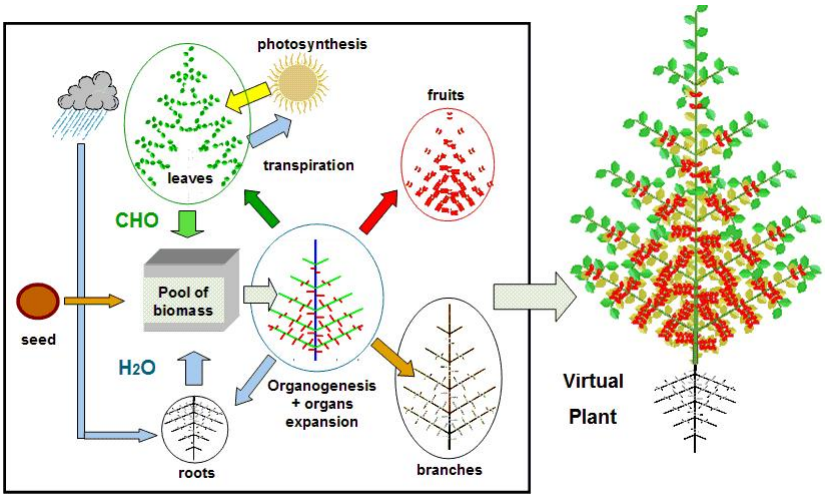

Figure 2: *Plant functioning flowchart: the seed gives the initial pool of biomass. It is used to build organs (internodes, roots, leaves, fruits) and thus the plant architecture. Roots allow water assimilation used for leaf transpiration and photosynthetic biomass production. This biomass is stored in the common pool of reserves and is then distributed among organs, which ends the growth cycle.*

3 A New Way to Simulate Plant Growth Based on its Structural Factorization

It is possible to simulate tree development and growth thanks to computer simulation techniques that imitate closely processes of plant functioning, i.e. handle the growth of each organ individually. Every bud organogenetic production is simulated, which means that all organs are created one by one. Likewise, the production and the circulation of fresh biomass use the paths of the topological and geometrical structures in the tree architecture, all organs have to be reached. Such simulation process was implemented in AMAPhydro Software, see [Reffye (de) et al., 1997] and [Blaise et al., 1998]. But this method leads to cumbersome computations.

The factorization method we present here strongly relies on the plant botanical organization. It takes into account the fact that identical organs and structures are repeated a large number of times in the plant architecture. We derive an efficient method of simulation based on plant instantiations, instead of simulating all organs.

3.1 Organogenesis

3.1.1 Botanical Instantiations and Notations

Metamers and buds: Let P be the maximum number of physiological ages in the plant. It is generally very small, inferior to 5, [Sabatier and Barthélémy, 1999]. From the botanical description of the plant, we get that metamers and buds are the elementary bricks of the plant structure.
At growth cycle t, a metamer is characterized by its physiological age p, the physiological age of its axillary branches q (with $q \geq p$; if there is no axillary branch, we take $q = P+1$) and its chronological age n. It will be denoted by $m^t_{pq}(n)$ or simply $m_{pq}(n)$ if there is no ambiguity on the growth cycle. These 3 indices p, q, n are sufficient to describe all the metamers composing a plant of chronological age t and we have at most $1/2\,P(P+3)t$ different metamers.
A bud is only characterized by its physiological age p and will be denoted by s^t_p or s_p. We have of course at most P different buds in a plant.

Substructures: The terminal bud of a plant axis produces different kinds of metamers bearing axillary buds of various physiological ages. These buds themselves give birth to axillary branches and so on... A substructure is the complete plant structure that is generated after one or several cycles by a bud. In the deterministic case, all the substructures with the same physiologicaland chornological ages are identical if they have been set in place at the same moment in the tree architecture. At cycle t, a substructure is thus characterized by its physiological age p and its chronological age n. It will be denoted by $S^t_p(n)$. We recall here that the physiological age of the main trunk is 1. Thus, at growth cycle t, the substructure of physiological age 1 and of chronological age t, $S^t_1(t)$, represents the whole plant.
Figure 3 illustrates the way substructures are organized. The total number of different substructures in a plant of chronological age t is thus Pt. This number is very small, usually less than 30, even if the total number of organs is high.
Substructures and metamers will be repeated a lot of times in the tree architecture, but they need to be computed only once for each kind.

3.1.2 Mathematical Formalism

This part is based on [Reffye (de) et al., 2003]. We develop and adapt their formalism. Without loss of generality and in order to simplify the notations, we will suppose that all metamers bear axillary buds and we will no longer consider the virtual physiological age $P+1$. We will also suppose that there is no mutation (the apical bud does not change its physiological age) while the axis grows.
Let $\mathcal{G}^t$ be the set of metamers and buds:

$$\mathcal{G}^t \;=\; \left\{m^t_{pq}(n),\; 1 \leq p \leq P,\; p \leq q \leq P,\; 1 \leq n \leq t\right\} \;\cup\; \left\{s^t_p,\; 1 \leq\; p \leq P\right\}\;. \qquad (1)$$

We consider the monoid $\mathcal{S}^t$ generated by $\mathcal{G}^t$ for the concatenation internal operator. This operator will be represented by the multiplicative sign. Substructures are thus elements of $\mathcal{S}^t$. They are built by induction, as follows:

- Substructures of chronological age 0 are buds:

$$S^t_p(0) = s^t_p\;, \qquad (2)$$

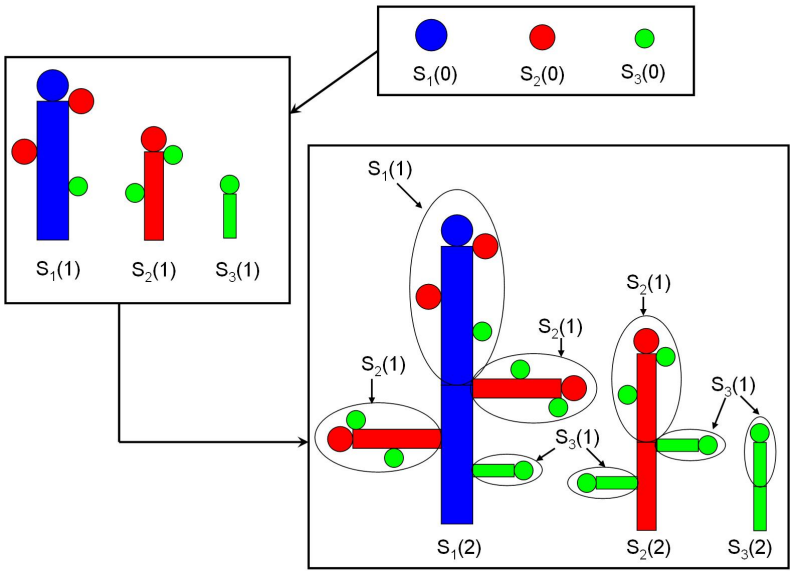

Figure 3: *Substructures of physiological ages 1, 2, 3 at chronological ages 0, 1, 2 and their organization:* $S_1(0)$, $S_2(0)$, $S_3(0)$ *are buds of physiological age* 1, 2, 3 *respectively. In this example, a growth unit of physiological age 1 is composed of 2 metamers of type* m_{12} *and 1 metamer of type* m_{13}*; a growth unit of physiological age 2 is composed of 2 metamers of type* m_{23}*; a growth unit of physiological age 3 is composed of 1 metamer of type* m_{34} *(without axillary bud).*

- and if we suppose built all the substructures of chronological age $n-1$, we deduce the substructures of chronological age n:

$$S_p^t(n) = \left[\prod_{p \le q \le P} \left(m_{pq}^t(n)\right)^{u_{pq}(t+1-n)} \left(S_q^t(n-1)\right)^{b_{pq}(t+1-n)} \right] S_p^t(n-1) \quad . \tag{3}$$

For all (p,q) (such that $1 \le p \le P,\ p \le q \le P$), $(u_{pq}(t))_t$ and $(b_{pq}(t))_t$ are sequences of integers that are characteristic of the plant organogenesis: $u_{pq}(t)$ corresponds to the number of metamers m_{pq} in growth units of physiological age p appearing at growth cycle t; $b_{pq}(t)$ is the number of axillary substructures of physiological age q in growth units of physiological age p that appeared at growth cycle t.
In equation (3), substructure $S_p^t(n)$ is decomposed into :
- its oldest growth unit, called base growth unit:

$$\prod_{p \le q \le P} \left(m_{pq}^t(n)\right)^{u_{pq}(t+1-n)}$$

- the lateral substructures borne by the base growth unit (they are one cycle younger):

$$\prod_{p \le q \le P} \left(S_q^t(n-1)\right)^{b_{pq}(t+1-n)}$$

- the substructure grown from the apical bud of the base growth unit (also one cycle younger):

$$S_p^t(n-1)$$

This decomposition is illustrated on $S_1(2)$ in Figure 3.

The GreenLab organogenesis model has been derived in 3 forms:

- GL1 corresponds to the deterministic organogenesis model, without influence of the plant functioning. Mathematically, it corresponds to u_{pq} and b_{pq} constant, see [Yan et al., 2004].
- GL2 corresponds to a stochastic model of organogenesis, u_{pq} and b_{pq} are stochastic variables. As a consequence, substructures of the same chronological and physiological ages can be very different: $S_p^t(n)$ is a stochastic variable with values in $\mathcal{S}_t$. More details are given in section 3.1.4 and in [Kang et al., 2004].
- GL3 corresponds to a deterministic model with total retroaction between organogenesis and photosynthesis. $u_{pq}(t)$ and $b_{pq}(t)$ are functions of the biomass produced by the plant at growth cycle $t-1$ and t respectively, see [Mathieu et al., 2004].

From equation (3), we can derive the algorithms to build the topological structure of the plant at any growth cycle t, in both deterministic and stochastic cases. The numbers of organs can also be computed as a result of organogenesis. These numbers will prove of crucial importance in section 3.2 to compute plant functioning.

3.1.3 Deterministic plants

Topology: The inductive construction of section 3.1.2 gives the algorithm to build the topological structure of the plant at any growth cycle t. We suppose that the sequences $(u_{pq}(t))_t$ and $(b_{pq}(t))_t$ are known. To get the full plant structure at cycle t, we need to build all the substructures $S_p^t(n)$, for all $p \in [1;P]$ and for all $n \in [0;t]$. It is simply done using equation (3) and by induction on n, as illustrated by Figure 3.
If $(u_{pq}(t))_t$ and $(b_{pq}(t))_t$ are constant (GL1 case), $S_p^t(n)$ are (topologically) independent of t, and the construction does not have to be done at each growth cycle but only once.

Numbers of Organs: Without factorization of the plant structure, counting the number of organs is a typical bottleneck; the computing time can be tremendous for big trees and forests. Here, we overcome this difficulty.
In order to determine the number of organs, we count the metamers. For this purpose, we define the morphism Ψ of $\mathcal{S}_t$ into $(\mathbb{R}^P, +)$. If S is in $\mathcal{S}_t$, $\Psi(S)$ is the vector of the number of metamers of each physiological age in S.
On the generating set $\mathcal{G}_t$:

$$\Psi(m_{pq}^t(n)) = I_p \quad \text{and} \quad \Psi(s_p^t) = (0) \tag{4}$$

where:

$$I_p = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow \text{rank p} \tag{5}$$

Then we deduce $\Psi(S_p^t(n)$ for all p, n using the morphism property and by induction on n. From equation (3):

$$\begin{aligned}\Psi\left(S_p^t(n)\right) &= \Psi\left(\left[\prod_{p\leq q\leq P}\left(m_{pq}^t(n)\right)^{u_{pq}(t+1-n)}\left(S_q^t(n-1)\right)^{b_{pq}(t+1-n)}\right]S_p^t(n-1)\right)\\ &= \sum_{p\leq q\leq P} u_{pq}(t+1-n)\ I_p + \sum_{p\leq q\leq P} b_{pq}(t+1-n)\Psi\left(S_q^t(n-1)\right) + \Psi\left(S_p^t(n-1)\right)\end{aligned} \tag{6}$$

An interesting result can be derived for the GL1 model, that is to say when u_{pq} and b_{pq} are constant, and if there is no mutation. In this case, $\Psi(S_p^t(n))$ is independent of t. Thus, if we introduce the square matrix of order P:

$$U^n = (\Psi(S_1(n)), \cdots, \Psi(S_P(n))) \quad , \tag{7}$$

we obtain the following system:

$$U^n = V + U^{n-1}(I + N) \quad , \tag{8}$$

where I is the identity matrix of order P, V is a diagonal matrix, $V_{ij} = \delta_{ij}\sum_{i\leq q\leq P} u_{iq}$ and N is an inferior triangular matrix, $N_{ij} = b_{ji}$. We obtain:

$$U^n = V\left(\sum_{i=0}^{n-1}(I+N)^i\right) \quad . \tag{9}$$

Without reiteration (if $b_{pp} = 0$ for all p), N is nilpotent and the equation simplifies.

This inductive plant construction is very fast. The computation time is proportional to Pt^2 (Pt for GL1) and not to the number of metamers. As the number of organs (internodes, leaves, fruits...) per metamer is botanically known, GreenLab provides a mathematical tool that enables to compute the organ production of a virtual plant very quickly and thus suppresses the drawback of counting the organs one by one by simulation.

Geometric Representation: If we add geometric rules (internode lengths, branching angles, phyllotaxy) to the construction algorithm, we will obtain the 3D architecture of a geometrical tree, [Reffye (de) et al., 2003]. We can extend the role of substructures to be sets of polygons storing their geometric shapes into a library. They become meta-organs, and positioning a substructure in the plant architecture needs the same operations as for a simple 3D organ. They can also be displayed separately, which can be useful. The tree of Figure 3.1.3 built with AMAPsim software can be computed either with the classical method simulating every organ or with substructure factorization, see [Yan et al., 2003]. Resulting computational times are compared: more than one hour for the total simulation of the 15 year old plant, less than one second with the substructure factorization! The mathematical algorithm proposed totally suppresses the simulation cost of tree architecture.

3.1.4 Stochastic plants - GL2 model

Topology: The inductive construction algorithm can be extended to stochastic plants, when $(u_{pq}(t))_t$ and $(b_{pq}(t))_t$ are stochastic variables controlling growth, branching and death processes, see [Kang et al., 2004]. $S_p^t(n)$ are also stochastic variables in $\mathcal{S}_t$. Equation (3) can be

Plant age	Substructure method	Total simulation
5 years	0.1	3.3
10 years	0.3	437.2
15 years	0.7	4743.0

Figure 4: *Comparison of computing times (in seconds) between the simulation based on substructure factorization and the classical method simulating every organ (total simulation) at different growth cycles - Visualization of the plant at growth cycles 5 and 15*

adapted as follows:

$$\left[S_p^t(n)\right] = \prod_{p \le q \le P} \left(m_{pq}^t(n)\right)^{u_{pq}(t+1-n)} \prod_{1 \le i \le b_{pq}(t+1-n)} \left[S_q^t(n-1)\right]_i \ \left[S_p^t(n-1)\right] \quad , \qquad (10)$$

where $\left[S_p^t(n)\right]$ denotes a realization of the stochastic variable $S_p^t(n)$. In the case of deterministic growth, for given chronological and physiological ages, there is only one substructure. For the stochastic growth, all the realizations of $S_p^t(n)$ could be different. As a consequence, we lose the advantage of the substructure decomposition as no factorization is possible.
For simulation, we overcome this difficulty by taking a limited number of samples of substructures that are realizations of the same stochastic distribution. These samples of substructures in a given set (at growth cycle t, for given physiological and chronological ages p and n) are obtained from Equation (10): for each sample, we get values for $\left(u_{pq}(t)\right)_t$ and $\left(b_{pq}(t)\right)_t$ and choose $\left[S_q^t(n-1)\right]$ in the appropriate sets. The same loop as in the deterministic case is used, which insures that these sets have already been built. On Figure 5, we give examples of these sampling sets and how samples are used in structures of higher chronological ages.

The size R of the substructure sets can vary. Of course, the accuracy will increase with R. This method to simulate stochastic trees is very efficient. The time necessary to build the first tree is proportional to $P \times t^2 \times R$. Once again, it does not depend on the number of organs to create. It is important to note that the algorithm is even more interesting to build stochastic forests, as the sets of substructures built for the first stochastic tree can be used to build other trees of the same kind that are realizations of the same stochastic distribution.

Numbers of Organs: Similarly to the deterministic case, the numbers of organs can be deduced from the number of metamers. We will thus compute its stochastic distribution.

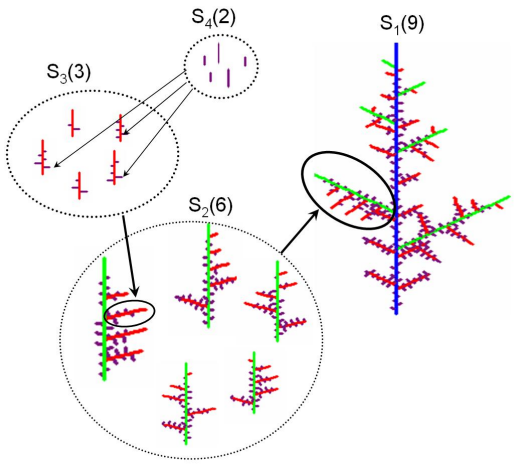

Figure 5: *Sampling sets and Stochastic Construction: for example, to build the set of $S_3(3)$ each time we need a substructure of type $S_2(2)$, we draw a sample in the corresponding set previoulsy built.*

Without bud death processes, the metamer distribution results from parallel stochastic processes and it is very close to a normal distribution. So the mean and variance are sufficient to obtain the distribution shape with accuracy. Using compound law formulas (see [Sedgewick and Flajolet, 1996]), it is possible to compute the theoretical mean, variance and covariances of the number of metamers in a plant whose topological structure is given by Equation (10). The theoretical formulas avoid heavy Monte-Carlo simulations to get the distribution shape as it used to be done with several days of computation! It is also possible to get a very good approximation of the mean and variance of the produced biomass using differential statistics applied to the production equation. The detailed calculations and formulas are given by [Kang et al., 2004]. These results are synthetic and do not need plant reconstruction. For many experiments they can be quite sufficient to assess the mean and variance of the yield.

3.2 Plant Functioning

The plant structural factorization given in Equation (3) also allows us a fast computation of the plant functioning. As explained in section 2.2, we consider two steps at each growth cycle: first, the biomass is produced by leaf photosynthesis and stored in a common pool, and then, it is redistributed among all organs according to their sink strengths. Both steps strongly rely on the numbers of organs.

3.2.1 The Operator of Growth

The growth of a plant from cycle t to cycle $t+1$ is given by the growth of all its substructures. We define the evolution operator ϕ^t of $\mathcal{S}^t$ into $\mathcal{S}^{t+1}$ as a morphism for the concatenation law. As $\mathcal{G}^t$ is a generating set of $\mathcal{S}^t$, the morphism ϕ^t is fully determined by its values on $\mathcal{G}^t$. We

have:

$$\begin{cases} \phi^t(m_{pq}^t(n)) = m_{pq}^{t+1}(n+1) \ , \text{ with } 1 \leq p \leq P,\ p \leq q \leq P,\ 1 \leq n \leq t \\ \\ \phi^t(s_p^t) = \left[\prod_{p \leq q \leq P} \left(m_{pq}^{t+1}(1)\right)^{u_{pq}(t+1)} \left(s_q^{t+1}\right)^{b_{pq}(t+1)} \right] s_p^{t+1} \ , \text{ with } 1 \leq p \leq P \end{cases} \tag{11}$$

The initial condition is given by s_1^0, which is the seed.
If we only consider organogenesis without functioning, we have $m_{pq}^t(n) = m_{pq}$ for all t, n, and this formalism is equivalent to the dual-scale automaton, see [Hu et al., 2003], and to grammars or L-Systems, see [Prusinkiewicz and Lindenmayer, 1990]. However, in their original forms these formalisms can not take properly into account functioning. A notable improvement on this regard was the development of sensitive growth grammars, see [Kurth, 1994] and [Kurth and Sloboda, 1997], attaching local variables to plant organs. In a similar way, substructures can integrate all necessary variables for plant functioning, but the main advantage is that the plant factorization allows a very quick treatment of these variables.

3.2.2 Biomass Acquisition - Photosynthesis

Every green leaf i produces fresh biomass that fills the pool of reserves. Its production is given by the water use efficiency times its water transpiration, and the leaf transpiration is obtained by the potential transpiration rate multiplied by the leaf conductance, see [Jones, 1992] or [Allen et al., 1998]. E^t denotes the water use efficiency times the potential transpiration rate at growth cycle t and sums up all the environmental effects. In its extremely simple version presented here, the model applies a constant E^t to all leaves. An empirical nonlinear function f depending on the leaf surface A_i and on physiological parameters r_1, r_2 gives the leaf conductance. The plant biomass production is thus:

$$Q^t = E^t \sum_i f(A_i, r_1, r_2) \ , \tag{12}$$

The empirical function f chosen in GreenLab is:

$$f(A_i, r_1, r_2) = \frac{1}{r_1/A_i + r_2} \ . \tag{13}$$

This function can be easily changed according to modellers' choices.
We define the conductance of a metamer as the sum of conductances of all its leaves and we introduce the morphism Γ of $\mathcal{S}_t$ into $(\mathbb{R}, +)$ on the generating set $\mathcal{G}_t$: $\Gamma(m_{pq}^t(n))$ gives the conductance of the metamer and $\Gamma(s_p^t) = 0$. Then, we can deduce $\Gamma(S_p^t(n))$ by induction on n.
First, we know that for all p, $1 \leq p \leq P$, $\Gamma(S_p^t(0)) = \Gamma(s_p^t) = 0$. If $\Gamma(S_p^t(n-1))$ are known for all p, we use equation 3:

$$\begin{aligned} \Gamma\left(S_p^t(n)\right) &= \Gamma\left(\left[\prod_{p \leq q \leq P} \left(m_{pq}^t(n)\right)^{u_{pq}(t+1-n)} \left(S_q^t(n-1)\right)^{b_{pq}(t+1-n)} \right] S_p^t(n-1) \right) \\ &= \sum_{p \leq q \leq P} \left[u_{pq}(t+1-n)\Gamma\left(m_{pq}^t(n)\right) + b_{pq}(t+1-n)\Gamma\left(S_q^t(n-1)\right) \right] + \Gamma\left(S_p^t(n-1)\right) \end{aligned} \tag{14}$$

Finally, as the structure $S_1^t(t)$ corresponds to the full plant, the production of biomass at growth cycle t is:

$$Q^t = E^t \ \Gamma\left(S_1^t(t)\right) \ . \tag{15}$$

3.2.3 Biomass Distribution

Each organ o has a potential biomass attraction value that we name sink or organ demand. This sink p_o depends on the type of organ, its physiological age and its chronological age n. $n = 0$ corresponds to organs that will appear at the next cycle (generated by the structures $S_p^t(0)$, *cf.* Section 3.2.1), and $n \geq 1$ to organs in expansion. The shape chosen for sink functions are left to the user's choice, but they should be able to fit properly any kind of sink variations, as a function of the organ chronological age. It must be flexible enough to give bell shapes, c or s shapes, *etc.*. More details on sink variation functions are given in [Reffye (de) and Hu, 2003]. We define the plant demand at growth cycle t as the total biomass attraction of all organs (leaves, internodes, fruits, layers, roots...):

$$D^t = \sum_o p_o \ , \tag{16}$$

As for the computation of the numbers of organs or the biomass production, we introduce the morphism Λ of $\mathcal{S}_t$ into $(\mathbb{R}, +)$ and we define it on the generating set $\mathcal{G}_t$: $\Lambda(m_{pq}^t(n))$ gives the metamer demand, that is to say the sum of the sinks of all its organs for their expansion and secondary growth, and $\Lambda(s_p^t)$ is the demand of the growth unit that will be generated by the bud. We deduce $\Lambda(S_p^t(n))$ at every growth cycle t, for all p, n, by induction on n and we have:

$$D^t = \Lambda\left(S_1^t(t)\right) \tag{17}$$

The biomass q_o^t allocated to organ o at growth cycle t is thus:

$$q_o^t = p_o \frac{Q^t}{D^t} \ . \tag{18}$$

Eventually, the organ volume depends on its apparent density and its dimensions on allometric rules. All this features can be measured directly from the organ shape. Particularly, we compute leaf surface areas that are used to determine photosynthesis at the next growth cycle.

3.2.4 Retroaction of Photosynthesis on Organogenesis

From Equation (18), we see the importance of the ratio $\frac{Q^t}{D^t}$. It is the main regulating variable of plant growth. According to the environmental conditions and to the available biomass, the plant will adapt its topology along its growth. This phenomenon is taken into account in GL3, see [Mathieu et al., 2004] or [Mathieu, 2006] for an exhaustive study. Regarding the topological equation (3), it corresponds to non constant sequences $(u_{pq}(t))_t$ and $(b_{pq}(t))_t$:

$$u_{pq}(t) = u_{pq}\left(\frac{Q^{t-1}}{D^{t-1}}\right) \quad \text{and} \quad b_{pq}(t) = b_{pq}\left(\frac{Q^t}{D^t}\right) \ . \tag{19}$$

On Figure 6 appears very clearly the difference between GL1 and GL3. We simulate with DigiPlant software the growth of a tree of chronological age 15 in two different environmental conditions, $E^t = 0.5$ and $E^t = 1.$ for all t. The two trees on the left are obtained with GL1: the plant keeps the same topology but is shorter in bad conditions. On the contrary, the two trees on the right are built with GL3: the plant changes topology with high plasticity and adapts to its environment. Concerning computational times, it is faster to compute GL1 topology than GL3 one ($P \times t$ structures to compute compared to $P \times t^2$), but there is no difference concerning the computation of functioning and of organ dimensions. To give an idea, the computation time of the full GL3 tree with secondary growth (rings) is less than 0.1 second on standard PC.

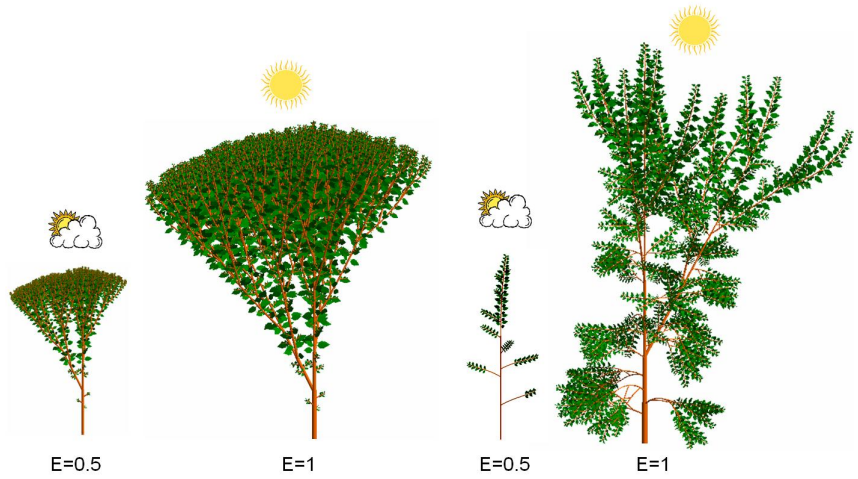

Figure 6: *Effects of climatic variations on the GL1 and GL3 organogenesis models: on the left side (GL1), topology is fixed, organs are simply smaller to adjust to the environment; on the right side (GL3), toplogy is variable and the plant adapts its development to the environmental conditions.*

4 Discussion

We proposed a topological factorization of plant architecture based on botanical instanciations, namely metamers and substructures. A proper mathematical formalism was adapted to this factorization and helped us derive a fast algorithm to simulate plant organogenesis (architectural development) in both deterministic and stochastic cases. Moreover, under the hypothesis of a homogeneous potential transpiration rate, we showed that the substructure factorization provides a very efficient framework to compute plant functioning (biomass production and distribution) when using ecophysiological submodels. The computational times varies at most quadratically with the plant chronological age. It would grow exponentially without structural factorization as the complexity would increase with the total number of organs.

The main drawback of the factorization method lies in the hypothesis that at a given growth cycle, all the substructures with the same physiological and chronological ages are identical and have the same functioning. It is too restrictive in numerous cases.

Firstly, the potential transpiration rate depends on the microclimatic conditions (temperature, light, humidity) perceived by each individual leaf. These conditions may vary significantly in the same plant. An average approach based on a homogeneous potential transpiration rate is generally sufficient to describe the biomass acquisition and distribution among organs, see [Allen et al., 1998] or [Guo et al., 2006]. However, it does not allow to take into account local effects which play a very important role for some applications, for example to study the development of pests (pathogens, insects), see [Chelle, 2005].

Another example of limitation concerns the biomechanical analysis of trees. An incremental finite element method was implemented in [Fourcaud et al., 2003], based on an exhaustive simulation of the growth of all organs. The complete computation is thus very heavy. Such analysis can not be directly adapted to a simulation based on substructure factorization as each substructure repetition has its own spatial position and their own gravitational constraints.

In both cases, an adaptation to our method would consist in discretizing the array of varying conditions (value of potential transpiration rate in the first example, orientation of the substructure main axis in the second one). As in the stochastic organogenesis case, see section

3.1.4, we overcome the difficulty by handling a limited number of samples of substructures for given physiological and chronological ages, corresponding to the discretization.
Fast plant growth simulation is very useful for graphical applications in urbanism or landscaping. Moreover, if plant geometric models are enriched with ecophysiological submodels, the visual impact is straightforward as plants look more authentic.
However, we think that the most promising applications of substructure factorization concerns agronomy and forestry. To derive prediction and optimization tools, it is first necessary to identify model parameters using data obtained by experiments made on real plants. Whatever the identification algorithm used (newtoninan method with numerical computation of the system partial derivatives or heuristic methods, see [Zhan et al., 2003]) the growth simulation process has to be run a very large number of times. The faster it is, the more efficient is the identification, particularly as speed allows a good interaction with the user.
Moreover, sophisticated ecophysiological models can be implemented in the framework we proposed to simulate complex biological phenomena. Currently, the validation of the GreenLab model is widely studied. Several agronomic plants have already been calibrated (see the example of Maize in [Guo et al., 2006] or of branching plants like coffee trees or young beech trees in [Mathieu, 2006]). With the same applicative purpose, optimal control techniques are now applied to the model in order to optimize water or nutriant supplies, see [Wu, 2005].
Finally, the substructure factorization open interesting prospects as GreenLab remains consistent when changing scales. The complete model can be simplified without major drawbacks. Instead of creating compartments like process based models, we transform branches into meta-organs (i.e. substructures) being at the same time sources and sinks, whose values depend on the cumulative actions of the underlying organs. The ultimate transformation is to consider the tree as a single structure that is source and sink. The changing of scales relies on the equations of the full model and can be properly written.

Acknowledgement:

This work is supported in part by LIAMA (Sino-French Laboratory in Information, Automation and Applied Mathematics), Natural Science Foundation of China (#60073007), and China 863 Program (#2002AA241221).

References

[Allen et al., 1998] Allen, R., Pereira, L., Raes, D., and Smith, M. (1998). Crop evapotranspiration. guidelines for computing crop water requirements. In *FAO Irrigation and Drainage*, 56. FAO (Rome, Italy).

[Aono and Kunii, 1984] Aono, M. and Kunii, T. (1984). Botanical tree image generation. In *Computer Graphics and Applications*, volume 4(5), pages 10–33. IEEE.

[Baker et al., 1983] Baker, D., Lambert, J., and McKinion, J. (1983). Gossym: a simulation of crop growth and yield. Technical Report 1089, Clemson University, South Carolina Experiment Station, USA.

[Barthélémy et al., 1997] Barthélémy, D., Caraglio, Y., and Costes, E. (1997). Architecture, gradients morphogénétiques et âge physiologique chez les végétaux. In Bouchon, J., editor, *Modélisation et simulation de l'architecture des végétaux*, Sciences Update, pages 89–136. INRA.

[Barthélémy, 1991] Barthélémy, D. (1991). Levels of organization and repetition phenomena in seed plants. *Acta Biotheoretica*, 39:309–323.

[Bell, 1991] Bell, A. (1991). *Plant Form. An illustrated guide to flowering plant morphology*. Oxford University Press.

[Blaise et al., 1998] Blaise, F., Barczi, J., Jaeger, M., Dinouard, P., and Reffye (de), P. (1998). Modeling of metamorphosis and spatial interactions in the architecture and development of plants. In T.L., K. and Luciani, A., editors, *Cyberworlds*, volume 6, pages 81–108. Springer Verlag Tokyo, ieee edition.

[Chelle, 2005] Chelle, M. (2005). Phylloclimate or the climate perceived by individual plant organes: What is it? how to model it? what for? *New Phytologist*, 3(166).

[Deussen and Lintermann, 2005] Deussen, O. and Lintermann, B. (2005). *Digital Design of Nature - computer generated plants and organics*. Springer Verlag New York.

[Fourcaud et al., 2003] Fourcaud, T., Blaise, F., Lac, P., Castéra, P., and Reffye (de), P. (2003). Numerical modelling of shape regulation and growth stresses in trees ii: Implementation in the amappara software and simulation of tree growth. *Trees*, 17:31–39.

[Gao et al., 1992] Gao, L., Jin, Z., Huang, Y., and Zhang, L. (1992). Rice clock model - a computer model to simulate rice development. *Agriculture and Forest Meteorology*, 60:1–16.

[Guo et al., 2006] Guo, Y., Ma, Y., Zhan, Z., Li, B., Dingkuhn, M., Luquet, D., and Reffye (de), P. (2006). Parameter optimization and field validation of the functional-structural model greenlab for maize. *Annals of Botany*, 97:217–230.

[Howell and Musick, 1984] Howell, T. and Musick, J. (1984). Relationship of dry matter production of field crops to water consumtion. In *Les besoins en eau des cultures, Paris*.

[Hu et al., 2003] Hu, B., Reffye (de), P., Zhao, X., Yan, H., and Kang, M. (2003). Greenlab: A new methodology towards plant functional-structural model – structural aspect. In *Plant Growth Models and Applications*. Tsinghua University Press and Springer.

[Jones and Kiniry, 1986] Jones, C. and Kiniry, J. (1986). *CERES - Maïze : A simulation model of maïze growth and development*. Texas A&M University Press.

[Jones, 1992] Jones, H. (1992). *Plants and Microclimate*. Cambridge University Press.

[Kang et al., 2004] Kang, M., Cournède, P., Le Roux, J., Reffye (de), P., and Hu, B. (2004). Theoretical study and numerical simulation of a stochastic model for plant growth. In *CARI04, Tunisia*.

[Kurth, 1994] Kurth, W. (1994). *Growth grammar interpreter GROGRA 2.4: A software tool for the 3-dimensional interpretation of stochastic, sensitive growth grammars in the context of plant modelling. Introduction and Reference Manual*. Berichte des Forschungszentrums Waldökosysteme der Universität Göttingen, Ser. B, Vol. 38.

[Kurth and Sloboda, 1997] Kurth, W. and Sloboda, B. (1997). Growth grammars simulating trees-an extension of l-systems incorporating local variables and sensitivity. *Silva Fennica*, 31(3):285–295.

[Marcelis et al., 1998] Marcelis, L., Heuvelink, E., and Goudriaan, J. (1998). Modelling of biomass production and yield of horticultural crops: a review. *Scientia Horticulturae*, 74:83–111.

[Mathieu, 2006] Mathieu, A. (2006). *Essai sur la modélisation des interactions entre la croissance d'une plante et son développement dans le modèle GreenLab*. PhD thesis, Ecole Centrale Paris.

[Mathieu et al., 2004] Mathieu, A., Cournède, P., and Reffye (de), P. (2004). A dynamical model of plant growth with full retroaction between organogenesis and photosynthesis. In *CARI04, Tunisia*.

[Perttunen et al., 1996] Perttunen, J., Sievänen, R., Nikinmaa, E., Salminen, H., Saarenmaa, H., and Väkevä, J. (1996). Lignum: a tree model based on simple structural units. *Annals of Botany*, 77:87–98.

[Prusinkiewicz and Lindenmayer, 1990] Prusinkiewicz, P. and Lindenmayer, A. (1990). *The Algorithmic Beauty of Plants*. Springer-Verlag, Berlin,.

[Reffye (de) et al., 1988] Reffye (de), P., Edelin, C., Françon, J., Jaeger, M., and Puech, C. (1988). Plant models faithful to botanical structure and development. In *Proc. SIGGRAPH 88, Computer Graphics*, volume 22(4), pages 151–158.

[Reffye (de) et al., 1997] Reffye (de), P., Fourcaud, T., Blaise, F., Barthélémy, D., and Houllier, F. (1997). A functional model of tree growth and tree architecture. *Silva Fennica*, 31(3):297–311.

[Reffye (de) et al., 2003] Reffye (de), P., Goursat, M., Quadrat, J., and Hu, B. (2003). The Dynamic Equations of the Tree Morphogenesis Greenlab Model. Technical Report 4877, INRIA.

[Reffye (de) and Hu, 2003] Reffye (de), P. and Hu, B. (2003). Relevant Choices in Botany and Mathematics for building efficient Dynamic Plant Growth Models: Greenlab Cases. In *Plant Growth Models and Applications*. Tsinghua University Press and Springer.

[Sabatier and Barthélémy, 1999] Sabatier, S. and Barthélémy, D. (1999). Growth dynamics and morphology of annual shoots according to their architectural position in young cedrs atlantica (endl.) manetti ex carrière (pinaceae). *Annals of Botany*, 84:387–392.

[Sedgewick and Flajolet, 1996] Sedgewick, R. and Flajolet, P. (1996). *Introduction to Algorithm Analysis*. Addison-Wesley.

[Shinozaki et al., 1964] Shinozaki, K., Yoda, K., Hozumi, K., and Kira, T. (1964). A quantitative analysis of plant form - the pipe model theory i. basic analysis. *Jpn J. Ecol.*, 14:97–105.

[Sievänen et al., 2000] Sievänen, R., Nikinmaa, E., Nygren, P., Ozier-Lafontaine, H., Perttunen, J., and Hakula, H. (2000). Components of a functional-structural tree model. *Annals of Forest Sciences*, 57:399–412.

[Wu, 2005] Wu, L. (2005). *Variational methods applied to plant functional-structural dynamics: parameter identification, control and data assimilation*. PhD thesis, Univeristé Joseph Fourier - Grenoble I.

[Yan et al., 2004] Yan, H., Kang, M., De Reffye, P., and Dingkuhn, M. (2004). A dynamic, architectural plant model simulating resource-dependent growth. *Annals of Botany*, 93:591–602.

[Yan et al., 2003] Yan, H., Reffye (de), P., Pan, C., and Hu, B. (2003). Fast construction of plant architectural models based on substructure decomposition. *Journal of Computer Science and Technology*, 18:780–787.

[Zhan et al., 2003] Zhan, Z., Reffye (de), P., Houllier, F., and Hu, B. (2003). Fitting a structural-functional model with plant architectural data. In *Plant Growth Models and Applications*. Tsinghua University Press and Springer.