

Dynamic System of Plant Growth

Paul-Henry Cournède

▶ **To cite this version:**

Paul-Henry Cournède. Dynamic System of Plant Growth. Mathematics [math]. Université Montpellier II - Sciences et Techniques du Languedoc, 2009. tel-00377462v2

HAL Id: tel-00377462
https://theses.hal.science/tel-00377462v2

Submitted on 2 May 2009

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Habilitation à Diriger des Recherches
Université de Montpellier II

Dynamic System of Plant Growth

Paul-Henry COURNÈDE

Ecole Centrale Paris, MAS
INRIA Saclay Île-de-France, EPI DIGIPLANTE
paul-henry.cournede@ecp.fr

10 March 2009

Habilitation à Diriger des Recherches
Université de Montpellier II

Système Dynamique de la Croissance des Plantes

présentée par

Paul-Henry COURNÈDE

le 10 Mars 2009, devant le jury composé de :

M. Christian SAGUEZ	Professeur à l'Ecole Centrale de Paris	Président
M. Alain CHARCOSSET	Directeur de Recherche à l'INRA	Rapporteur
M. Jean FRANÇON	Professeur Honoraire à l'Université de Strasbourg	Rapporteur
M. Winfried KURTH	Professeur à l'Université de Göttingen	Rapporteur
M. Philippe DE REFFYE	Directeur de Recherche au CIRAD	Examinateur
M. Jérémie LECOEUR	Professeur à Montpellier SupAgro	Examinateur

Résumé :

Les travaux de recherche portent sur la modélisation mathématique de la croissance des plantes et le développement de méthodes mathématiques adaptées à l'étude des modèles. *In fine*, les objectifs sont d'une part la prévision quantitative et qualitative de la production végétale, et d'autre part l'optimisation et le contrôle optimal des cultures. L'étape préalable est la compréhension et l'analyse des interactions génotype x environnement.
Nous proposons une voie mathématique d'exploration de ces interactions basée sur l'écriture du système dynamique de croissance des plantes. Le modèle de base sur lequel repose notre étude est le modèle individu-centré GreenLab, combinant la description de la structure et du fonctionnement de la plante, à l'échelle de l'individu ou du peuplement.
L'architecture de la plante est le résultat complexe de cette interaction génotype x environnement. Nous montrons comment il est possible de mettre en œuvre des méthodes d'analyse de cette architecture pour expliquer le passage du génotype au phénotype. Mathématiquement, il s'agit :
– de décrire la mise en place dynamique des structures de la plante,
– de dériver de cette structure les équations fonctionnelles de la croissance et le système de rétroaction entre croissance et processus de développement,
– de mettre au point les méthodes d'estimation paramétrique à partir des données expérimentales sur l'architecture.
Une fois le modèle d'interaction génotype x environnement déterminé, nous illustrons la mise en œuvre de méthodes d'optimisation et de contrôle optimal à la résolution de problèmes applicatifs.

Table des matières

1 Introduction **7**

2 A Formal Language to Model Plant Development **11**
- 2.1 **Plant Development and Growth Cycle** . . . 11
- 2.2 **Alphabet** . . . 12
- 2.3 **L-Systems** . . . 13
- 2.4 **Structural Factorization** . . . 14
  - 2.4.1 **Substructures** . . . 14
  - 2.4.2 **Dynamic Equation of Development** . . . 15
  - 2.4.3 **Recursive Construction** . . . 17
  - 2.4.4 **Development Sequences** . . . 17
- 2.5 **Structural Morphism** . . . 18
- 2.6 **Numbers of Metamers** . . . 18
  - 2.6.1 **A structural morphism** . . . 18
  - 2.6.2 **GL1 case** . . . 19
  - 2.6.3 **Numbers of metamers and organs in the whole plant** . . . 19
- 2.7 **Stochastic Organogenesis** . . . 20
  - 2.7.1 **Stochastic L-system** . . . 20
  - 2.7.2 **Markov kernel and stochastic L-system** . . . 21
  - 2.7.3 **Multitype branching processes and stochastic L-systems** . . . 22
  - 2.7.4 **Application to stochastic plant development** . . . 24

3 The Equations of Plant Growth **27**
- 3.1 **Allocation** . . . 27
  - 3.1.1 **GreenLab fundamental equation** . . . 27
  - 3.1.2 **Primary and secondary growthes** . . . 29
  - 3.1.3 **Root System** . . . 30
  - 3.1.4 **Organ dimensions** . . . 30
- 3.2 **Sink variation function :** p^t . . . 30
- 3.3 **Number of organs** N^t **and topology variables** U^t, V^t . . . 32
- 3.4 **Control variable for biomass allocation :** u^t . . . 33
- 3.5 **Discussion and examples** . . . 36
  - 3.5.1 **First example : given numbers of organs, no influence of plant topology, biomass production as a feedback function** . . . 37
  - 3.5.2 **Second example : retroaction of biomass production on the numbers of organs (GL3)** . . . 39
- 3.6 **Stochasticity** . . . 41
  - 3.6.1 **Propagation of uncertainty** . . . 42
  - 3.6.2 **Application to the GreenLab equation of growth** . . . 44

4 System Identification and Control **47**
4.1 **Parameter Estimation** 47
4.1.1 **Estimation of Discrete Dynamic Systems** 48
4.1.2 **Application to the GreenLab model** 49
4.1.3 **An example on experimental data of sugar-beet** 52
4.2 **Optimization and control** 54

5 Perspectives **61**

Chapitre 1

Introduction

Comment un apprenti mathématicien spécialisé dans les équations aux dérivées partielles des écoulements bifluides a-t-il bien pu se découvrir une passion dévorante pour l'étude de la betterave ou du hêtre ? Ce cheminement du doctorat à la candidature au diplôme d'habilitation à diriger des recherches a été marqué par une succession d'événements tous aussi peu probables les uns que les autres. Malgré tout, par l'expérience des compositions de fonctions génératrices, nous avons conscience que même si le chemin suivi avait une probabilité faible, celle du passage de l'état initial à l'état final ne l'était pas nécessairement, que différents chemins étaient possibles ... Je ne m'aventurerai cependant pas à l'évocation du destin ou de toute autre considération métaphysique, car tel n'est pas l'objet de ce manuscrit. Il nous suffira de préciser que ce chemin vers des sujets de recherche passionnants a été marqué par des rencontres de personnes formidables, et que sans ces personnes, je n'en serais pas là. D'une part, c'est avant tout le plaisir de travailler au laboratoire MAS de Christian Saguez avec les collègues et les étudiants qui a motivé mon investissement dans la recherche et d'autre part, le travail réalisé n'aurait pu l'être sans les efforts communs de tous, dans chacune des institutions partenaires, tant sont diverses les compétences nécessaires pour faire progresser la modélisation de la croissance des plantes, de l'expérimentation à la formalisation.

Cette expérience collective et internationale a laquelle je participe depuis 2002 concerne l'élaboration du modèle de croissance de plantes GreenLab. Elle a été initiée par Philippe de Reffye et peut être vue comme une synthèse originale de l'expérience qu'il a accumulée en modélisation mathématique des plantes depuis le début des années 1970, et des connaissances en botanique, ecophysiologie, agronomie et foresterie de différentes écoles de pensée.

L'histoire de la modélisation de la croissance des plantes est assez passionnante pour que le philosophe et historien des sciences Franck Varenne y ait consacré sa thèse de doctorat (de 900 pages ! Varenne (2003), Varenne (2007)). Il y retrace en particulier l'itinéraire de Philippe de Reffye, depuis sa thèse d'état en Côte d'Ivoire sur la modélisation de la croissance des arbres par des processus stochastiques, l'étude du caféier et pour la première fois la visualisation 3D des modèles botaniques de Hallé-Oldeman-Tomlison (Hallé et al. (1978)) sur table traçante. Nous rappelons ici quelques étapes importantes de cet itinéraire, marquée par la création de l'Atelier de Modélisation de l'Architecture des Plantes (AMAP) à Montpellier. Tout d'abord la thèse de Marc Jaeger (Jaeger, 1987) aboutissant à l'article du SIGGRAPH en 1988 (de Reffye et al., 1988), (Jaeger and de Reffye, 1992) a permis la traduction en Fortran des premiers programmes de Reffye et la visualisation de plantes botaniques sur écran graphique. La simulation de l'architecture s'est ensuite orientée vers une description botanique plus fine grâce aux concepts d'âge physiologique et d'axe de référence (de Reffye et al., 1990) (de Reffye et al., 1991b), (de Reffye et al., 1991a), (Barthélémy, 1991). Ceci a permis le développement du logiciel AMAPsim par Pierre Dinouard et Jean-François Barczi, dont l'histoire continue jusqu'à aujourd'hui (Barczi et al., 2008).

En parallèle, dés le début des années 1990, était poursuivi l'objectif de lier la structure botanique au fonctionnement écophysiologique, autour de Philippe de Reffye, Frédéric Blaise, François Houllier et Thierry Fourcaud. Le logiciel de simulation AMAPpara développé par Frédéric Blaise va servir de base pour l'allocation de matière dans la structure des arbres, *cf.* (de Reffye et al., 1997b), en particulier pour la croissance secondaire dans le tronc et les branches. AMAPpara est ensuite adapté de manière à y intégrer la loi de l'efficience de l'eau pour déterminer la production de matière à partir de la résistance hydraulique de la structure calculée à chaque cycle de croissance. L'allocation de matière entre organes se fait déjà en fonction de puits. Ce modèle prend le nom d'AMAPhydro, puis AMAPagro *cf.* (de Reffye et al., 1997a). Il sera calibré sur plantes réelles - cotonnier - dés 1999, (de Reffye et al., 1999), et servira de base pour l'étude des contraintes mécaniques dans les arbres par Thierry Fourcaud, (Fourcaud, 2004). L'élaboration de ce modèle combinant la description architecturale et son fonctionnement écophysiologique est alors en phase avec les avancées de la communauté scientifique. Plusieurs modèles sont développés à peu près simultanément et la communauté des "FSPM" (Functional Structural Plant Models) émerge, avec un premier colloque notable en Finlande et un numéro spécial de la revue Silva Fennica 31(3), *cf.* en particulier LIGNUM (Perttunen et al., 1996), GROGRA (Kurth and Sloboda, 1997), ou (Sievänen et al., 2000) pour une revue.

Mais la complexité de la simulation de l'architecture "en parallèle" comme elle est réalisée dans AMAPpara ou AMAPsim, c'est à dire la nécessité de simuler un par un tous les méristèmes et tous les organes, apparaît comme une très forte limitation lorsqu'il faut combiner structure et fonctionnement. Sievänen et al. (2000) évoquent un coût de simulation multiplié par 8 par rapport à la seule simulation de la structure. C'est pour faire face à ce problème que GreenLab va être développé, résultat des travaux de Philippe de Reffye avec ses doctorants et Hu Bao Gang au LIAMA, à partir du début des années 2000. Un nouveau modèle architectural simplifié est mis en oeuvre, revisitant les concepts d'unité de croissance et d'âge physiologique précédemment développés à AMAP. Il prend le nom d'automate double échelle, *cf.* Zhao et al. (2001). Même si le modèle est encore mal formalisé mathématiquement, les concepts sont d'une grande puissance et permettent le développement de la méthode des sous-structures, exploitant ainsi les répétitions de structures identiques dans la plante, à la fois pour le développement et pour le fonctionnement, *cf.* Yan et al. (2002a). Le temps de simulation de la croissance d'une plante devient négligeable. En parallèle, les concepts écophysiologiques du modèle structure-fonction s'affinent au contact des écophysiologistes et agronomes, Jérémie Lecoeur (UMR LEPSE, Montpellier), Michael Dingkuhn (CIRAD, Montpellier), Guo Yan et Zhang Bao Gui (CAU, Pékin) et Ep Heuvelink (Université de Wageningen). Le modèle est présenté en détails au premier colloque "Plant Models and Applications", (Hu and Jaeger, 2003), en particulier dans l'article fondateur (de Reffye and Hu, 2003). L'estimation des paramètres est également améliorée grâce à François Houllier, (Houllier, 1999). La méthode des moindres carrés généralisés remplace le recuit simulé des premiers ajustements. Cela ajouté aux progrès réalisés sur les temps de simulation rend la confrontation du modèle aux données expérimentales sur plantes réelles beaucoup plus interactive, ce qui lance véritablement la calibration du modèle GreenLab (Guo et al. (2003), Zhan et al. (2003)). C'est dans ce contexte, que Varenne appelle un début de "rematématisation" du modèle de simulation, que va commencer mon travail de recherche avec Philippe de Reffye en 2002. Rapidement, une nouvelle équipe de recherche a ainsi vu le jour, DigiPlante, commune à l'INRIA, à Centrale et au CIRAD, en association avec l'équipe GreenLab du LIAMA.

Dans cette oeuvre collective, j'ai tout d'abord participé au développement des dernières versions du modèle de croissance. Ces développements concernent principalement la modélisation des interactions architecture - fonctionnement dans la thèse d'Amélie Mathieu (2006), le modèle d'allocation de la croissance secondaire pour la thèse de Véronique Letort (2008) et le passage au peuplement, *cf.* (Le Chevalier et al., 2007a), (Cournède et al.,

2008). Ces aspects sont bien détaillés par ailleurs, et ne seront donc présentés ici que succintement dans le corps du manuscrit, sous un éclairage différent, pour illustrer le formalisme mathématique. 3 références importantes sur ces travaux de modélisation sont joints en annexes : (Mathieu et al., 2009) présente le modèle d'interaction fonctionnement - développement (GL3) avec comme propriété émergente la possible apparition de rythmes au cours de la croissance ; (Letort et al., 2008a) fait un point sur le modèle structure-fonction GreenLab appliqué aux arbres et introduit la méthodologie d'identification paramétrique pour les arbres ; (Cournède et al., 2008) modélise la compétition pour la lumière pour le passage au GreenLab peuplement.

D'autre part, étant donnée ma formation, ma principale contribution réside dans cette étape de "remathématisation", débutée au LIAMA et poursuivie à l'INRIA et à Centrale. Les premiers concepts de GreenLab au LIAMA ont ainsi fait apparaître un fort besoin en mathématiques. Il s'agissait de développer le formalisme pour la modélisation et les méthodes mathématiques pour les applications de ce formalisme, principalement l'identification paramétrique et le contrôle. GreenLab a ainsi pu être formalisé comme un système dynamique discret, tant pour la partie développement (organogenèse) que pour la partie fonctionnement (photosynthèse), avec interaction entre les deux, dans les cas déterministes et stochastiques. Là aussi, quelques références ont joué un rôle fondamental, tout d'abord la thèse de Philippe de Reffye (de Reffye (1979)), un excellent papier visionnaire de Jean Françon (Françon (1990)) pour le lien avec les grammaires formelles, une note didactique de François Houllier sur l'estimation des paramètres (Houllier (1999)), les travaux de thèse de Kang Meng Zhen sur le stochastique (Kang et al. (2003), Kang et al. (2004)) et surtout le papier résultant des travaux de Philippe de Reffye avec les automaticiens du projet METALAU, Maurice Goursat et Jean-Pierre Quadrat (de Reffye et al. (2003)).

Le formalisme développé s'est révélé très général, capable de s'adapter aux différentes versions du modèle. Il a aussi été à la base du développement du code de simulation et d'estimation paramétrique de la croissance des plantes en C++, `DigiPlante`, dont je suis le créateur et l'auteur principal, même si Pierre Labatut, Amélie Mathieu et Véronique Letort m'ont largement secondé pour son développement et son adaptation aux différentes plantes. Ce logiciel a été ou est utilisé par de nombreux partenaires pour leurs recherches (Institut Technique de la Betterave - Paris, UMR EGC - Paris Grignon, UMR AMAP - Montpellier, UMR LEPSE - Montpellier, Univ. de Wageningen - Pays-Bas, China Agriculture University - Pékin, China Academy of Forestry - Pékin).

Si les avancées en terme de modélisation ont pu être décrites de façon détaillée et didactique dans les deux thèses de doctorat d'Amélie Mathieu (2006) et Véronique Letort (2008), il manquait dans la bibliographie GreenLab une synthèse sur l'état d'avancement du développement du formalisme et des méthodes mathématiques. L'ambition de ce manuscrit intitulé "Système dynamique de la croissance des plantes" est d'y remédier, en présentant les "Maths de GreenLab" (ou plutôt leur état d'avancement). Par ailleurs, quelques originalités sur la façon de présenter le modèle et quelques nouveautés sur le calcul des incertitudes ou le contrôle optimal susciteront, je l'espère, l'intérêt auprès des plus fins connaisseurs de GreenLab.

Dans le premier chapitre, nous montrons comment le modèle d'organogenèse de GreenLab, basé sur les concepts de cycle de croissance et d'âge physiologique, se formalise naturellement dans le cadre de la théorie des L-systèmes, c'est à dire des grammaires à réécriture parallèle. Mais le principal résultat mis en évidence est la factorisation de la grammaire d'organogenèse conduisant à la construction des structures de la plante par récurrence. L'application de morphismes structuraux permet de déterminer de façon efficace certaines variables caractéristiques pour la croissance, comme le nombre d'organes ou la surface foliaire "vue" par une sous-structure. Le formalisme se généralise au cas stochastique, et le lien est fait avec les processus de branchement multi-type.

Dans le deuxième chapitre, nous présentons le modèle de fonctionnement autour de l'équa-

tion d'allocation. Considerer celle-ci comme centrale nous permet de présenter plusieurs déclinaisons du modèle selon une classification mathématique précise, en fonction du choix des trois types de variables intervenant dans l'équation d'allocation : le vecteur puits des organes, les variables topologiques (avec les nombres d'organes) et enfin la commande, c'est à dire la biomasse disponible pour la répartition. Le système peut être en boucle ouverte ou en boucle fermée, avec prise en compte de la topologie ou non, avec différents niveaux de rétroaction. Deux systèmes spécifiques sont étudiés en détails. D'une part, le modèle classique de GreenLab pour une plante de type Corner (avec topologie fixée et la production de biomasse fonction de l'état du système) et un modèle de type GL3, c'est à dire avec rétroaction du fonctionnement sur l'organogenèse. Dans le cas stochastique, nous montrons comment la théorie des incertitudes peut être utilisée pour propager les variabilités du système.

Enfin, dans un troisième chapitre, nous présentons la méthodologie de mise en oeuvre du modèle pour son application. Il s'agit d'une part de l'identification paramétrique du système plante à partir de données expérimentales, c'est à dire pour une variété spécifique, d'estimer les paramètres du modèle de croissance, paramètres endogènes et d'interaction avec l'environnement. La méthode est illustrée sur une application à la betterave sucrière. D'autre part, on montre comment la modélisation peut permettre la résolution de problèmes d'optimisation (par exemple pour la sélection variétale) ou de contrôle (détermination de modes de culture optimaux).

Chapitre 2

A Formal Language to Model Plant Development

In computational models, plants are generally represented as words in a formal language, more precisely a language based on a generative parallel rewriting grammar also called L-system, *cf.* Lindenmayer (1968), Smith (1984), Prusinkiewicz et al. (1988), Prusinkiewicz and Lindenmayer (1990). Even though the mathematical concepts underlying the GreenLab model of plant development were historically based on automaton theory (*cf.* Zhao et al. (2001), de Reffye and Hu (2003)), it is shown here that the model can be properly written in the L-System frame. Such a link was already made by Françon (1990) for AMAP architectural models (de Reffye et al. (1988)). Moreover, we show that the grammar formalism can be taken advantage of to derive the factorization of plant structure and to write in a compact form the recurrent equations of plant development. We will first recall the basic botanical hypotheses underlying the GreenLab model of plant development and derive the associated grammar, in both deterministic and stochastic cases.

2.1 Plant Development and Growth Cycle

As explained in Barthélémy and Caraglio (2007), organogenesis results from the functioning of undifferentiated cells constituting the apical meristem and located at the tip of axes. When in active phase, this meristem forms buds that will develop into new growth units composed of one or several metamers (also called phytomers). A metamer is a botanical entity chosen as the elementary scale to model plant architectural development in this study. It is composed of an internode bearing organs : axillary buds, leaves, flowers. Depending on species, metamers are set in place rhythmically or continuously.
In the rhythmic case, see Figure 2.1, the plant grows by successive shoots of several metamers produced by buds. The appearance 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. For example, most temperate trees grow rhythmically, new shoots appearing at spring. For such plants, and if we do not consider polycyclism and neoformation, the architectural growth cycle corresponds to one year.
Plant growth is said continuous when meristems keep on functioning and generate metamers one by one, see Figure 2.1. The number of metamers on a given axis (that is to say generated by the same meristem) is generally proportional to the sum of daily temperatures received by the plant, see Jones (1992). The growth cycle is defined as the thermal time unit necessary for a meristem to build a new metamer, it can be quite short, corresponding to a couple of days. The growth unit is thus simply composed of one metamer. The growth of tropical trees, bushes or agronomic plants is often continuous.
In this work, we do not consider time scales that are smaller than the architectural growth

cycle and we study the development of new growth units as a discrete process. The *Chronological Age* (CA) of a plant (or of an organ) is defined as the number of growth cycles it has existed for.

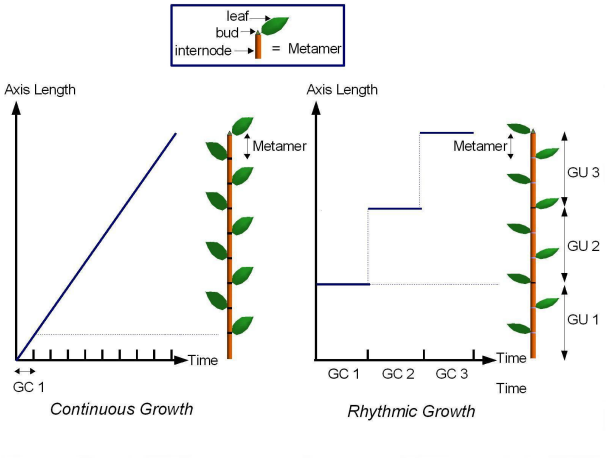


FIG. 2.1 – Growth Cycle (GC), Growth Unit (GU) and Metamer in GreenLab. The axis length is given as a function of time for continuous and rhythmic growths, see Cournède et al. (2006)

Since metamers may bear axillary buds, plant architecture develops into a hierarchical branching system. Barthélémy and Caraglio (2007) underlined that architectural units can be grouped into categories characterized by a particular combination of morphological parameters. Thus, the concept of *Physiological Age* (PA) was introduced to represent the different types of growth units and axes. For instance, on coffee trees, there are two types : orthotropic trunk and plagiotropic branches. The main trunk's physiological age is equal to 1 and the oldest physiological age denoted by P corresponds to the ultimate state of differentiation for an axis, it is usually short, without branches. We need less than 5 physiological ages to describe the axis typology of most trees. The apical meristem or bud of an axis is thus characterized by the physiological age of the growth unit that it may produce and a metamer is characterized by its physiological age i (which is the physiological age of the growth unit that it belongs to) and that of the buds that it bears j. Except in some very rare cases not considered in this study, we always have : $i \leq j$. Moreover, along an axis, the morphological features of the growth unit may evolve with the age of the apical meristem. This process is described as the meristem sequence of differentiation by Barthélémy and Caraglio (2007), and corresponds to a transition to a superior physiological age of the meristem.

2.2 Alphabet

In GreenLab, the alphabet $\mathcal{G}$ is given by the set of metamers $\mathcal{M}$ and buds $\mathcal{B}$, see de Reffye et al. (2003). A metamer is defined with four indices and is denoted by $m_{pq}^t(n)$:
– its chronological age : n,
– its physiological age : p,
– the physiological age of its axillary buds : q, $(q \geq p)$,
– the chronological age of the plant : t.
Note however that usually, we will consider that $m_{pq}^t(n)$ are the same for all q, except that they bear different types of axillary buds. $m_p^t(n)$ will thus be used when we consider metamers independently of their lateral productions.
A bud is defined by three indices and is denoted by $b_p^t(n)$:

– its physiological age : p,
– the number of growth cycles n for which bud's physiological age has been p - in the sequel, we will call it $ontogenic\ age$ of a bud,
– the chronological age of the plant : t.
The organogenesis alphabet is given by :

$$\mathcal{G} = \mathcal{M} \cup \mathcal{B} \tag{2.1}$$

with

$$\mathcal{M} = \{m_{pq}^t(n) \ / \ 1 \leq p \leq P,\ p \leq q \leq P,\ 1 \leq t \leq T, 1 \leq n \leq t\}\ , \tag{2.2}$$

and

$$\mathcal{B} = \{b_p^t(n) \ / \ 1 \leq p \leq P,\ 0 \leq t \leq T, 0 \leq n \leq t\}\ . \tag{2.3}$$

T denotes the maximum time of the simulation (from a mathematical point of view, it could be $+\infty$).
We do not consider symbols for organs since the constitution of a metamer is supposed fixed by botanical rules (an internode and a given number of leaves and fruits). If, for example, flowering is particularly studied, symbols denoting flowers would be introduced in the alphabet without difficulty.

Definition 2.1 (Set of words over an alphabet) *The set of words over an alphabet $\mathcal{A}$ is defined as the monoid generated for the concatenation operator ".", seen as an internal, non-commutative operation, by $\mathcal{A} \cup \{\mathbf{1}\}$, where $\mathbf{1}$ is the neutral element for the concatenation operator (which corresponds to the empty word). It is denoted $\mathcal{A}^*$.*

$\mathcal{G}^*$ will thus represent all the possible topological structures composed with buds and metamers. For example, $m_{12}^t(1)b_2^t(0)b_1^t(1) \in \mathcal{G}^*$ represents at growth cycle t a structure composed of an internode of physiological age 1, bearing a lateral bud of physiological age 2 and an apical bud of physiological age 1. We will see that of course the structures of interest, that is to say botanically relevant, form a small subset of $\mathcal{G}^*$ (corresponding to a language over $\mathcal{G}$). It is important to consider the concatenation operator as non-commutative when studying plant topology.

2.3 L-Systems

Since plant development can be seen as the result of bud functioning, generative grammars describing the production of buds have shown to be particularly adapted to model mathematically plant development (Prusinkiewicz et al. (1988), Françon (1990)). We recall below the definition of the simplest type of parallel generative grammars called 0L-System and first introduced by Lindenmayer (1968).
Let $\mathcal{V} = \{v_1, v_2, \cdots, v_m\}$ denote an alphabet and $\mathcal{V}^*$ the set of words over $\mathcal{V}$.

Definition 2.2 (0L-system) *A 0L-system is a construct $G = \langle \mathcal{V}, w_a, P_r \rangle$ where :*
– *$w_a \in \mathcal{V}^*$ is a word (called the* axiom*) which represents the structure initiating the growth.*
– *$P_r \subset (\mathcal{V} \cup \{\mathbf{1}\} \times \mathcal{V}^*)$ is a finite* set of productions *and corresponds to the graph of a mapping from $V \cup \{\mathbf{1}\}$ into $\mathcal{V}^*$. We necessarily have $(\mathbf{1}, \mathbf{1}) \in P_r$. A production $(s, \chi) \in P_r$ is written as $s \rightarrow \chi$ and represents the evolution of the symbol s into χ. s is called a* predecessor*.*

Directly inspired by this definition, a more general class of grammars was later introduced by Rozenberg and Salomaa (1980a), the F0L-systems.

Definition 2.3 (F0L-System) *A F0L-system is a construct $G = \langle \mathcal{V}, W, P_r \rangle$ where :*
– *$W \subset \mathcal{V}^*$ such that $Card\,(W)$ is finite, $W \neq \emptyset$ and $W \neq \{1\}$.*

– *for all* $w_a \in W$, $G[w_a] = \langle \mathcal{V}, w_a, P_r \rangle$ *is a 0L-system.*
W *is called the set of axioms and* $G[w_a]$ *is called a component system of* G.

As recalled in section 2.1, GreenLab organogenesis can be seen as the combination of two phenomena, branching and meristem differentiation. It can be easily modelled with the grammar formalism and we propose to define it as a F0L-system as follows :

Definition 2.4 (GreenLab Organogenesis) *GreenLab organogenesis is defined as a F0L-system* $\langle \mathcal{G}, \mathcal{B}, P_r \rangle$ *with the following production rules* P_r *:*
For all $(t, n, p) \in [0; T] \times [0; \min(\tau(p), t)] \times [1; P]$ *:*

$$b_p^t(n) \xrightarrow{\text{if } n<\tau(p)} \left(\prod_{q=p}^{P} \left(m_{pq}^{t+1}(1)\right)^{u_{pq}(t+1)} \left(b_q^{t+1}(0)\right)^{v_{pq}(t+1)} \right) b_p^{t+1}(n+1)$$
$$\xrightarrow{\text{if } n=\tau(p)} \left(\prod_{q=p}^{P} \left(m_{pq}^{t+1}(1)\right)^{u_{pq}(t+1)} \left(b_q^{t+1}(0)\right)^{v_{pq}(t+1)} \right) b_{\mu(p)}^{t+1}(0)$$

and for all $(t, n, p, q) \in [1; T] \times [1; t] \times [1; P] \times [p; P]$ *:*

$$m_{pq}^t(n) \longrightarrow m_{pq}^{t+1}(n+1)$$

with :
– $u_{pq}(t)$ *: number of phytomers* m_{pq} *in a growth unit of PA* p*, appearing at growth cycle* t
– $v_{pq}(t)$ *: number of active axillary buds of PA* q *in a growth unit of PA* p*, appearing at growth cycle* t
– $\tau(p)$ *: number of GC after which a bud of PA* p *changes to PA* $\mu(p)$*. Vectors* τ *and* μ *characterize meristem differentiation.*
In the following, we will denote the F0L-system of GreenLab organogenesis as Γ.

2.4 Structural Factorization

A fundamental result (or rather hint) given by Smith (1984) was surprisingly never used in the context of models of plant development : it is the ability to factorize the L-system productions. It showed particularly adapted to the concept of physiological age, *cf.* de Reffye et al. (2003), Cournède et al. (2006). We are thus able to factorize plant structure into smaller parts that may repeat themselves a large number of times.

2.4.1 Substructures

Definition 2.5 (Substructure) *At growth cycle* $t \geq 0$*, a substructure of physiological age* $p,\ 1 \leq p \leq P$ *and chronological age* $n,\ 0 \leq n \leq t$ *is a word in* $\mathcal{G}^*$ *defined as the complete plant structure that is generated after* n *cycles by a bud of physiological age* p*. It is also characterized by the ontogenic age* k *of the bud generating it. It is denoted by* $S_p^t(n, k)$ *(corresponding to the structure generated by* $b_p^{t-n}(k)$ *after* n *growth cycles).*

We deduce straightforwardly the following results.

Remark 2.1 *Substructures of chronological age 0 are buds.*

Remark 2.2 *In the deterministic case, at a given growth cycle, all the substructures with the same physiological, chronological and ontogenic ages are identical.*

Remark 2.3 *The set of all substructures is the language generated by* $\Gamma = \langle \mathcal{G}, \mathcal{B}, P_r \rangle$*. It will be denoted by* $\mathcal{S}^*$.

Remark 2.4 *The set of substructures* $\left(S_1^t(t,0)\right)_{t\geq 0}$ *is the language generated by* $\Gamma\left[b_1^0(0)\right]$. *It represents the evolution of the structure of the whole plant across time.*

Figure 2.2 illustrates the way substructures are organized. The total number of different substructures in a plant of chronological age t is very small even if the total number of organs is high. It is comparable to Pt if there is no differentiation, and to Pt^2 if there is.

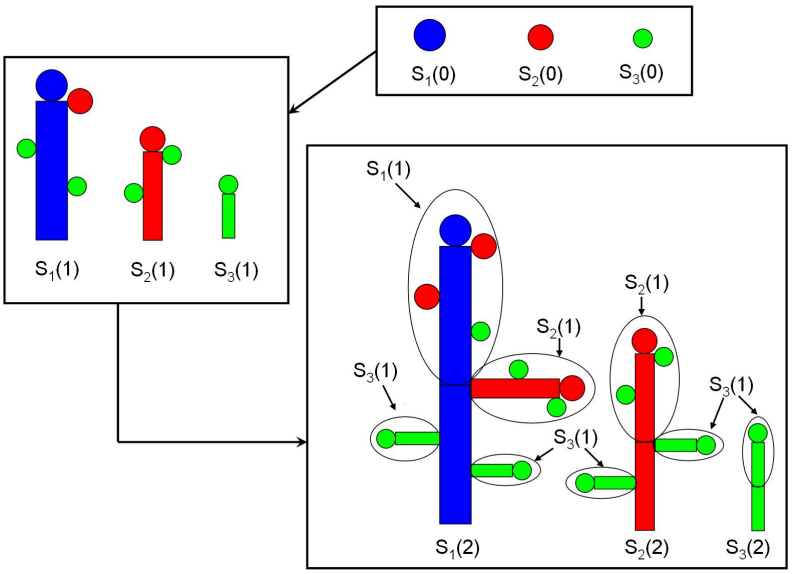


FIG. 2.2 – *Substructures of physiological ages 1, 2, 3 at chronological ages 0, 1, 2 and their organization :* $S_1(0)$, $S_2(0)$, $S_3(0)$ *are buds of physiological age* $1, 2, 3$ *respectively. In this example, a growth unit of physiological age 1 is composed of 2 metamers of type* m_{13} *and 1 metamer of type* m_{12} *; 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_{33} *(without axillary bud).*

Remark 2.5 (Notations) *If there is no differentiation phenomenon, all buds of the same physiological age are identical and they will simply be denoted by* b_p^t. *Likewise, the substructure it generates after* n *growth cycles will be denoted* $S_p^{t+n}(n)$.

2.4.2 Dynamic Equation of Development

Inspired by Rozenberg and Salomaa (1980a), we find convenient to introduce the mathematical concept of production function associated to the production rules.

Definition 2.6 (Production Function) *Let* $G = \langle V, W, P_r\rangle$ *be a F0L-system and* $(V^*, ., 1)$ *the monoid generated by* V. *The* $production\ function$ Ψ *derivating from* G *is the endomorphism on* $(V^*, ., 1)$ *generated by the production rules* P_r *: if* $(s, \chi) \in P_r$, $\Psi(s) = \chi$. *Note that it clearly defines an endomorphism on* V^* *since the production rules define it on a generating set.*

In the context of plant development, the production function Ψ derivating from Γ (*cf.* Definition 2.4) will be often referred to as the $growth\ endomorphism$, Cournède et al. (2006).

Remark 2.6 (Notations) Ψ^n *will denote the* n*-th iterate of* Ψ. *From the definition of substructures, we have :* $S_p^t(n,k) = \Psi^n\left(b_p^{t-n}(k)\right)$. *We also have* $\Psi^n\left(m_{pq}^t(i)\right) = m_{pq}^{t+n}(i+n)$.

The F0L-system $G = \langle V, W, P_r \rangle$ will also be denoted $\langle V, W, \Psi \rangle$ when Ψ is the production function associated to P_r.

We deduce the following important result.

Theorem 2.1 (Dynamic Equation of Plant Development) *For all $t \geq 1$, $n \in [1, t]$, $p \in [1, P]$, $q \in [p, P]$, $k \in [0, \min(\tau(p), t)]$, we have :*

If $0 \leq k < \tau(p)$:

$$S_p^t(n,k) = \left[\prod_{p \leq q \leq P} \left(m_{pq}^t(n)\right)^{u_{pq}(t+1-n)} \left(S_q^t(n-1,0)\right)^{v_{pq}(t+1-n)} \right] S_p^t(n-1,k+1)$$

$else$:

$$S_p^t(n,\tau(p)) = \left[\prod_{p \leq q \leq P} \left(m_{pq}^t(n)\right)^{u_{pq}(t+1-n)} \left(S_q^t(n-1,0)\right)^{v_{pq}(t+1-n)} \right] S_{\mu(p)}^t(n-1,0) \tag{2.4}$$

Proof: We first consider $k < \tau(p)$. Using the morphism property of Ψ, we get successively :

$$S_p^t(n,k) = \Psi^n\left(b_p^{t-n}(k)\right) = \Psi^{n-1}\left(\Psi\left(b_p^{t-n}(k)\right)\right)$$

$$= \Psi^{n-1}\left(\left(\prod_{q=p}^{P} \left(m_{pq}^{t-n+1}(1)\right)^{u_{pq}(t-n+1)} \left(b_q^{t-n+1}(0)\right)^{v_{pq}(t-n+1)}\right) b_p^{t-n+1}(k+1)\right)$$

$$= \left(\prod_{q=p}^{P} \Psi^{n-1}\left(m_{pq}^{t-n+1}(1)\right)^{u_{pq}(t-n+1)} \Psi^{n-1}\left(b_q^{t-n+1}(0)\right)^{v_{pq}(t-n+1)}\right) \Psi^{n-1}\left(b_p^{t-n+1}(k+1)\right)$$

$$= \left(\prod_{q=p}^{P} m_{pq}^t(n)^{u_{pq}(t-n+1)} S_q^t(n-1,0)^{v_{pq}(t-n+1)}\right) S_p^t(n-1,k+1)$$

If $k = \tau(p)$, the same principle applies except that the production rule $\Psi\left(b_p^{t-n}(k)\right)$ is different. □

In equation (2.4), substructure $S_p^t(n,k)$ is decomposed into :

- its oldest growth unit, called base growth unit :

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

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

$$\prod_{p \leq q \leq P} \left(S_q^t(n-1,0)\right)^{v_{pq}(t+1-n)}$$

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

$S_p^t(n-1,k+1)$ if $k < \tau(p)$,
or
$S_{\mu(p)}^t(n-1,0)$ if $k = \tau(p)$

This decomposition is illustrated on $S_1(2)$ in Figure 2.2.

2.4.3 Recursive Construction

If we suppose that all the elements of the alphabet (set of metamers $m_{pq}^t(n)$, set of buds $b_p^t(n)$) as well as the sequences $(u_{pq}(t))_t$ and $(v_{pq}(t))_t$ are known, Theorem 2.1 shows us how to build the topological structure of the plant at any growth cycle t recursively, as follows :

• Substructures of chronological age 0 are buds :

$$S_p^t(0,k) = b_p^t(k) \ , \tag{2.5}$$

• and if all the substructures of chronological age $n-1$ are built, we deduce the substructures of chronological age n from Equation 2.4 as functions of $m_{pq}^t(n)$, $u_{pq}(t-n+1)$ and $v_{pq}(t-n+1)$.

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.

If we only consider topology, we do not need to characterize the metamers by their chronological ages nor by the plant age. $m_{pq}^t(n)$ can thus be simply be denoted by m_{pq}. But if we consider the functional growth of a plant, metamers of different chronological ages have different masses and sizes, as well as metamers of the same chronological age, but at different plant ages.

(de Reffye et al., 2003) introduced how the dynamic development equation can be extended to build plant geometry with geometric operators replacing the concatenation operator. We will not detail this point here since our objective is to study plant functional growth which only relies on topology and not geometry in GreenLab.

2.4.4 Development Sequences

Plant toplogy can simply be seen as a function of :

– the sequences $(u_{pq}(t))_t$ and $(v_{pq}(t))_t$ for all (p,q) such that $1 \leq p \leq q \leq P$ (they will be called $development\ sequences$),
– the vectors $(\tau(p))_{1\leq p\leq P}$, $(\mu(p))_{1\leq p\leq P}$ in case of meristem differentation.

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 v_{pq} constant, see Yan et al. (2004). In such 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.
– GL2 corresponds to a stochastic model of organogenesis, u_{pq} and v_{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 $\mathcal{S}^*$. More details are given in section 2.7 and in Loi and Cournède (2008a).
– GL3 corresponds to a deterministic model with total retroaction between organogenesis and photosynthesis. $u_{pq}(t)$ and $v_{pq}(t)$ are functions of the biomass produced by the plant at growth cycle $t-1$ and t respectively, see Mathieu (2006), Mathieu et al. (2006), Mathieu et al. (2006).

Notations

U^t will denote the vector $(u_{pq}(t))_{1\leq p\leq q\leq P}$ and $\left(U^t\right)_t$ the associated sequence of vectors. Likewise, V^t will denote the vector $(v_{pq}(t))_{1\leq p\leq q\leq P}$ and $\left(V^t\right)_t$ the associated vector of sequences.

2.5 Structural Morphism

Definition 2.7 (Structural Morphism) *We define a structural morphism as a morphism from $(\mathcal{G}^*, \cdot)$ into $\left(\mathbb{R}^N, +\right)$ (for any $N \in \mathbb{N}$). It is completely characterized if defined on the generating set $\mathcal{G}$.*

Remark 2.7 *Let Φ denote a structural morphism. The dynamic equation of development (2.4) provides a powerful way to compute recursively Φ on the set of substructures $\mathcal{S}^*$. We apply Φ to Equation (2.4) and use the morphism property. Let $\Phi_p^t(n,k)$ denote $\Phi\left(S_p^t(n,k)\right)$. We have :*
For all $t \geq 1$, $n \in [1,t]$, $p \in [1,P]$, $q \in [p,P]$, $k \in [0, \min(\tau(p), t)]$, we have :

If $0 \leq k < \tau(p)$:

$$\Phi_p^t(n,k) = \left[\sum_{p \leq q \leq P} u_{pq}(t+1-n)\Phi\left(m_{pq}^t(n)\right) + v_{pq}(t+1-n)\Phi_q^t(n-1,0)\right] + \Phi_p^t(n-1,k+1)$$

$else$:

$$\Phi_p^t(n,\tau(p)) = \left[\sum_{p \leq q \leq P} u_{pq}(t+1-n)\Phi\left(m_{pq}^t(n)\right) + v_{pq}(t+1-n)\Phi_q^t(n-1,0)\right] + \Phi_{\mu(p)}^t(n-1,0) \tag{2.6}$$

Since all $\Phi\left(m_{pq}^t(n)\right)$ and $\Phi\left(S_p^t(0,k)\right) = \Phi\left(b_p^t(k)\right)$ are known (Φ defined on the generating set), we deduce all $\Phi\left(S_p^t(n)\right)$ recursively.

Example 2.1 (Leaf Surface Area)

As detailed in Section 3.4, it might be useful to determine leaf area in each substructure since a possibility to model the sink of an organ is to relate it to the leaf surface area above the organ in plant topology. In this case, it is necessary to determine leaf area in substructures of each type. It can be done by considering the structural morphism $\mathcal{A}$, from $(\mathcal{G}^*, \cdot)$ into $(\mathbb{R}, +)$. On the generating set, we simply have for all t, p, q, n :

$$\mathcal{A}\left(b_p^t(n)\right) = 0 \quad \text{and } \mathcal{A}\left(m_{pq}^t(n)\right) = A_{b,p}^t(n)$$

where $A_{b,p}^t(n)$ denotes the blade area of a leaf (or of the leaves) on a metamer of type $m_p^t(n)$.

2.6 Numbers of Metamers

The numbers of metamers (and thus the numbers of organs) can be computed as a result of organogenesis. These numbers will prove of crucial importance in section 3 to compute plant functioning. Without factorization of the plant structure, counting the numbers of metamers and of organs is a typical bottleneck ; the computing time can be tremendous for big trees and forests. Here, we show how the structural factorization helps to overcome this difficulty.

2.6.1 A structural morphism

The application Λ defined on $\mathcal{G}^*$ and giving the numbers of metamers of each physiological age is a structural morphism into $(\mathbb{R}^P, +)$. On the generating set $\mathcal{G}$:

$$\Lambda(m_{pq}^t(n)) = I_p \quad \text{and} \quad \Lambda(b_p^t(n)) = (0) \tag{2.7}$$

where :

$$I_p = \begin{pmatrix} 0 \\ \vdots \\ 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \leftarrow \text{rank p} \tag{2.8}$$

Then we deduce $\Lambda(S)$ for all $S \in \mathcal{S}^*$ as shown in Equation(2.6).

2.6.2 GL1 case

An interesting result can be derived for the GL1 model, that is to say when u_{pq} and v_{pq} are constant, and if there is no mutation (*cf.* Cournède et al. (2006)). In this case, $\Lambda(S_p^t(n))$ is independent of t. Thus, if we introduce the square matrix of order P :

$$\Lambda(n) = (\Lambda(S_1(n)), \cdots, \Lambda(S_P(n))) \quad , \tag{2.9}$$

we obtain the following system :

$$\Lambda(n) = \Omega + \Lambda(n-1)(I+N) \quad , \tag{2.10}$$

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

$$\Lambda(n) = \Omega \left(\sum_{i=0}^{n-1} (I+N)^i \right) \quad . \tag{2.11}$$

Without reiteration (if $v_{pp} = 0$ for all p), N is nilpotent and the equation simplifies. This result was already used by Lindenmayer and Jürgensen (1992).

2.6.3 Numbers of metamers and organs in the whole plant

A particular variable of interest is the total number of metamers of each physiological age at growth cycle t. ν^t will denote $\Lambda\left(S_1^t(t,0)\right)$. Likewise, to compute the functional growth, we will generally need the numbers of metamers (and more specifically the numbers of organs) of each physiological and chronological ages at growth cycle t. Les $N_p^t(n)$ (respectively $N_{o,p}^t(n)$) denote the number of metamers (respectively of organs of type o) of physiological age p, chronological age n in the whole plant at growth cycle t. Note that $N_p^t(0)$ corresponds to the numbers of metamers preformed in buds at growth cycle t, that is to say the numbers of metamers that will appear at growth cycle $t+1$. Instead of defining a special structural morphism to compute $N_p^t(n)$, it is interesting to note that

$$N_p^t(n) = N_p^{t-n}(0) = \nu_p^{t-n+1} - \nu_p^{t-n} \ . \tag{2.12}$$

However, computation based on structural morphism requires a detailed knowledge of all types of substructures. When only the total numbers of metamers and organs are needed, and not these numbers in each type of substructures, we can even bypass the structural reconstruction as in the GL1 case (*cf.* Mathieu (2006)).
Let β_p^t denote the number of active buds of physiological age p at growth cycle t. It can be deduced from the production rules (*cf.* Definition 2.4). If there is no meristem differentiation :

$$\beta_q^{t+1} = \sum_{1 \leq p \leq q} \beta_p^t v_{pq}(t+1) + \beta_q^t$$

We thus have :

$$\beta^{t+1} = \left(I + N^{t+1}\right)\beta^t \ . \tag{2.13}$$

with $I, N^{t+1} \in \mathcal{M}_P(\mathbb{R})$, I the identitiy matrix and $N_{qp}^{t+1} = v_{pq}(t+1)$ (inferior triangular matrix). The initial state is given by the seed, which can be considered as a bud of physiological age 1,

$$\beta^0 = I_1 = \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \ .$$

In case of meristem differentiation, we can also retrieve a recurrent formula, *cf.* Mathieu (2006) p68. The meristems of physiological age q that will differentiate at growth cycle $t+1$ were initiated at growth cycle $t+1-\tau(q)$.

$$\beta_q^{t+1} = \beta_q^t + \sum_{1\leq p\leq q} \beta_p^t v_{pq}(t+1) - \sum_{1\leq p\leq q} \beta_p^{t-\tau(q)} v_{pq}(t+1-\tau(q)) + \sum_{l|\mu(l)=q} \sum_{1\leq p\leq l} \beta_p^{t-\tau(l)} v_{pl}(t+1-\tau(l)) \ .$$

We deduce the number of metamers $N_p^t(0)$ preformed in buds at growth cycle t, for all $p \in [1; P]$:

$$N_p^t(0) = \beta_p^t \left(\sum_{p\leq q\leq P} u_{pq}(t+1) \right) \ , \tag{2.14}$$

and subsequently the numbers of organs of every type and every physiological age.

2.7 Stochastic Organogenesis

In the deterministic case, at a given growth cycle, all the substructures of the same type (same physiological, chronological, ontogenical ages) are identical. In many cases, especially for trees, this hypothesis is too restrictive. If it is not possible to model buds' behavior more precisely (according to their locations in the whole plant architecture, to exogeneous variables ...), we may consider the productions of buds as stochastic variables, *cf.* Kang et al. (2008a). In this section, we show how the L-system formalism naturally extends to model stochastic organogenesis. To compute functional growth, the numbers of organs are crucial variables. An interesting result in this prospect is the link between L-Systems and multitype branching processes. It allows the derivation of a recursive formula for the generating functions of the number of organs and hence the explicit computation of the distribution moments. The formalism introduced for the deterministic organogenesis based on the production function appears as a particular multitype branching process. This section is mostly extracted from Loi and Cournède (2008a).

2.7.1 Stochastic L-system

Let $\mathcal{V} = \{v_1, v_2, \cdots, v_m\}$ denote an alphabet and $\mathcal{V}^*$ the set of words over $\mathcal{V}$. The definitions of 0L-systems (Def. 2.2), F0L-systems (Def. 2.3) and production function (Def. 2.6) easily extend to the stochastic case :

Definition 2.8 (Stochastic 0L-system) *A stochastic 0L-system is a construct* $G = \langle \mathcal{V}, \omega_a, P_r, \pi \rangle$ *where :*

- $\omega_a \in \mathcal{V}^*$ *is a nonempty word (called the* axiom*) which represents the structure initiating the growth.*
- $P_r \subset \mathcal{V} \times \mathcal{V}^*$ *is a finite* set of productions.

– π *is called the* probability distribution. *It is a function from* P_r *into* $[0;1]$ *that maps the set of productions into the set of production probabilities. For instance,* $\pi(s,\chi)$ *gives the probability that* s *evolves into* χ. π *satisfies the following equality :*

$$\forall s \in \mathcal{V}, \qquad \sum_{(s,\chi)\in P_r} \pi(s,\chi) = 1.$$

Definition 2.9 (Stochastic F0L-system) *A stochastic F0L-system is a construct* $G = \langle \mathcal{V}, A, P_r, \pi \rangle$ *where :*
– A *is a finite nonempty subset of* $\mathcal{V}^*$ *(called the set of axioms of* G*).*
– *for every* $\omega_a \in A$, $G[\omega_a] = \langle V, \omega_a, P_r, \pi \rangle$ *is a stochastic 0L-system (called a component system of* G*).*

The concept of stochastic production functions sets up a probabilistic framework for stochastic F0L-systems.

Definition 2.10 (stochastic production function) *Let* $(\Omega, \mathcal{F}, \mathcal{P})$ *be a probability space associated to a stochastic F0L-system* $G = \langle \mathcal{V}, A, P_r, \pi \rangle$. *The mapping* F *from* $\Omega \times \mathcal{V}^*$ *into* $\mathcal{V}^*$ *is said to be the stochastic production function derivating from* G *if the following properties are true :*

1. $\forall \omega \in \Omega$, $F(\omega, \mathbf{1}) = \mathbf{1}$.
2. $\forall (\omega, x_1, x_2) \in \Omega \times \mathcal{V}^* \times \mathcal{V}^*$, $F(\omega, x_1x_2) = F(\omega, x_1)F(\omega, x_2)$
3. $\forall \omega \in \Omega$, $v \in \mathcal{V} \Rightarrow (v, F(\omega, v)) \in P_r$ *and* $P(F(.,v) = \chi) = \pi(v,\chi)$

An important result presented in Loi and Cournède (2008a) reconsiders the definition of stochastic L-systems in the context of Markov chain theory. The production function is shown to be equivalent to a Markov kernel on $\mathcal{V}^*$.

2.7.2 Markov kernel and stochastic L-system

We first recall some basic definitions and properties of the Markov chain theory (Stroock (2005)). Throughout the following all sets are finite or countable.

Definition 2.11 (Transition matrix) *A transition matrix (or kernel) from* A *to* B *is a map* $k : A \times B \to \mathbb{R}$ *such that* $k(a,b) \geq 0$ *and* $\sum_{b\in B} k(a,b) = 1$ *for all* $a \in A$. *A transition matrix or Markov kernel on* A *is a transition matrix from* A *to* A.

Definition 2.12 (Markov chain) *A Markov chain with state space* S *is a stochastic process* $(X_n)_{n\in\mathbb{N}}$ *on some probability space* $(\Omega, \mathcal{F}, P)$, *such that* $P(X_{n+1} = y \mid X_n = x) = p_{x,y}$ *for all* $x, y \in S$. $P = (p_{x,y})$ *is a Markov kernel on* S.

Theorem 2.2 *For every tupel* (P, μ) *of Markov kernel* P *on* S *and probability measure* μ *on* S *there exists a Markov Chain* $(X_n)_{n\in\mathbb{N}}$ *of kernel* P *and starting measure* $\mu = P(X_0 \in \cdot)$.

Remark 2.8 *The production rules of a stochastic 0L-system* $G = \langle \mathcal{V}, \omega_a, P_r, \pi \rangle$ *(or F0L-system) define a transition matrix* k *from* $\mathcal{V}$ *to* $\mathcal{V}^*$ *: if* $(x,y) \in P_r$, $k(x,y) = \pi(x,y)$, *else* $k_{x,y} = 0$. *For this reason, we may identify* k *with* π *without confusion.*

We deduce directly the following result :

Proposition 2.1 *Let* $P = (p_{x,y})$ *be the square matrix on* $\mathcal{V}^*$ *such that, for all* $x = x_1x_2\cdots x_n \in \mathcal{V}^*$ *with* $(x_1, x_2, \cdots, x_n) \in \mathcal{V}^n$ *and for all* $y \in \mathcal{V}^*$ *:*

$$p_{x,y} = \sum_{\substack{(y_1,y_2,\cdots,y_n)\in(\mathcal{V}^*)^n, \\ y_1y_2\cdots y_n = y}} \prod_{i=1}^{n} k(x_i, y_i)$$

Then, P *is a Markov kernel on* $\mathcal{V}^*$.

While the transition matrix k is defined by the production rules P_r and its distribution function π, the associated Markov kernel P on $\mathcal{V}^*$ is a convenient representation of the production function. The stochastic L-system (0L or F0L) will thus be defined indifferently by the production rules and probability distribution, its production function with probability distribution, a transition matrix from $\mathcal{V}$ to $\mathcal{V}^*$ or a Markov kernel on $\mathcal{V}^*$.

Remark 2.9 *To every component system $G[\omega_a] = \langle \omega_a, \pi \rangle$ we associate a Markov chain $(F_n[\omega_a])_{n\in\mathbb{N}}$ via the associated kernel P and the starting measure $\mu = \delta_{\omega_a}$ where δ_{ω_a} is the Dirac measure concentrated on ω_a.*

Beside plant topological structure directly deduced from L-systems, the numbers of organs are crucial variables in order to compute plant functioning (see for example de Reffye et al. (2008b)). To determine them, the order of symbols in words does not play any role, and we can consider the L-systems as commutative.
Let R be an equivalence relation on $\mathcal{V}^*$ defined as follows : $w_1 R w_2 \Leftrightarrow$ there exists Π, a permutation on the symbol ranks, such that $\Pi(w_1) = w_2$. Let us denote the quotient set $\mathcal{V}^*/R$ by $\tilde{\mathcal{V}}^*$. From now on, each word $w \in \mathcal{V}^*$ will be assimilated to the ordered representative $\tilde{w}$ of its equivalence class (*i.e.* $wR\tilde{w}$ and $\tilde{w} = \prod_{i=1}^{m} {v_i}^{\alpha_i}$ for $(\alpha_1, \cdots, \alpha_m) \in \mathbb{N}$). $\tilde{\mathcal{V}}^*$ is isomorphic to $\mathbb{N}^m$. Let Υ be the canonical isomorphism from $\tilde{\mathcal{V}}^*$ into $\mathbb{N}^m$. For a word $w = \prod_{i=1}^{m} {v_i}^{\alpha_i}$ in $\tilde{\mathcal{V}}^*$, we have $\Upsilon(w) = (\alpha_1, \cdots, \alpha_m)$. Let us denote by $\Upsilon_i(w)$ the i-th component of $\Upsilon(w)$ (*i.e.* $\Upsilon_i(w) = \alpha_i$).
In the following, the transition matrix k is thus considered as a map from $\mathcal{V} \times \tilde{\mathcal{V}}^*$ into $\mathbb{R}$.

2.7.3 Multitype branching processes and stochastic L-systems

Stochastic L-systems are closely related to multitype branching processes. Let us recall first the definition of a multitype branching process (Harris (1963), Athreya and Ney (2004)). Let us consider a population with m types of individuals. Assume a type i individual produces children of all types according to a probability distribution $\{P_i(j) : j = (j_1, \cdots, j_m), j_i \in \mathbb{N}, 1 \leq i \leq m\}$. Assume all individuals produce offspring independently of each other and of the past history of the process. Let $Z_{n,i}$ be the number of type i individuals in the n-th generation. Let $\{\xi_{n,i}^{(k)} : n \in \mathbb{N}^*, k \in \mathbb{N}, 1 \leq i \leq m\}$ be independent random vectors in $\mathbb{N}^m$ with $\xi_{n,i}^{(k)}$ having distribution $P_i(.)$.

Definition 2.13 (multitype Galton-Watson branching process)
If the vector $Z_n = (Z_{n,1}, \cdots, Z_{n,m})$ of population sizes in the n-th generation evolves by the recursive relation

$$Z_{n+1} = \sum_{i=1}^{m} \sum_{k=1}^{Z_{n,i}} \xi_{n,i}^{(k)}, \tag{2.15}$$

then $(Z_n)_{n\in\mathbb{N}}$ is a multitype Galton-Watson branching process.

The j-th component of $\xi_{n,i}^{(k)}$ represents the number of type j individuals produced by the k-th type i individual in the n-th generation. The set $\{P_i(.)\}_{i\in\{1,\cdots,m\}}$ is called the offspring distribution. Considering this definition, we have the following theorem :

Theorem 2.3 *Let $G = \langle \mathcal{V}, A, \pi \rangle$ be a stochastic F0L-system on $\mathcal{V} = \{v_1, \cdots, v_m\}$. Let $G[\omega_a] = \langle \mathcal{V}, \omega_a, \pi \rangle$ be a component system of G and $(F_n[\omega_a])_{n\in\mathbb{N}}$ the corresponding*

Markov chain. Then, $(\Upsilon(F_n[\omega_a]))_{n\in\mathbb{N}}$ *is a multitype Galton-Watson branching process whose offspring distributions* $\{P_i(.)\}_{i\in\{1,\cdots,m\}}$ *are given by :*

$$\forall i \in \{1,\cdots,m\}, \forall (j_1,\cdots,j_m) \in \mathbb{N}^m, \quad P_i(j_1,\cdots,j_m) = \pi\left(v_i, \prod_{k=1}^{m} v_k^{j_k}\right).$$

Proof: Let $\{P_i(.)\}_{i\in\{1,\cdots,m\}}$ be a set of probability distributions defined as follows :

$$\forall i \in \{1,\cdots,m\}, \forall (j_1,\cdots,j_m) \in \mathbb{N}^m, \quad P_i(j_1,\cdots,j_m) = \pi\left(v_i, \prod_{k=1}^{m} v_k^{j_k}\right).$$

Let $\{\xi_{n,i}^{(k)} : n \in \mathbb{N}^*, k \in \mathbb{N}, 1 \leq i \leq m\}$ be independent random vectors with $\xi_{n,i}^{(k)}$ having distribution $P_i(.)$. Thus, the q-th component of $\xi_{n,i}^{(k)}$ is the random variable that gives the number of type "v_q" symbols in a word generated by a type "v_i" symbol throughout G. Since there are $\Upsilon_i(F_n[\omega_a])$ type "v_i" symbols in $F_n[\omega_a]$, we have :

$$\Upsilon_q(F_{n+1}[\omega_a]) = \sum_{i=1}^{m} \sum_{k=1}^{\Upsilon_i(F_n[\omega_a])} \left(\xi_{n,i}^{(k)}\right)_q$$

and then,

$$\Upsilon(F_{n+1}[\omega_a]) = \sum_{i=1}^{m} \sum_{k=1}^{\Upsilon_i(F_n[\omega_a])} \xi_{n,i}^{(k)}$$

N.B. : the converse is true. Giving a multitype Galton-Watson branching process, there is a commutative stochastic L-system whose stochastic production function is entirely determined by the offspring distribution.

Let us now define the generating function associated to a stochastic 0L-system. Let $S = (s_1,\cdots,s_m) \in [0,1]^m$.

Definition 2.14 (generating function associated to a stochastic 0L-system) *Let* $G = \langle \mathcal{V}, A, \pi\rangle$ *be a stochastic F0L-system on* $\mathcal{V} = \{v_1,\cdots,v_m\}$. *Let* $G[\omega_a] = \langle \omega_a, \pi\rangle$ *be a component system of* G *and* $(F_n[\omega_a])_{n\in\mathbb{N}}$ *the corresponding Markov chain. For* $n \in \mathbb{N}$, *the generating function* $\psi_n[\omega_a]$ *of* $F_n[\omega_a]$ *is defined as follows :*

$$\psi_n[\omega_a](S) = \sum_{\chi\in\tilde{\mathcal{V}}^*} P(F_n[\omega_a] = \chi) \prod_{i=1}^{m} s_i^{\Upsilon_i(\chi)}.$$

By using the classical composition of generating functions (Harris (1963)) for a multitype Galton-Watson branching process, we deduce directly the following theorem :

Theorem 2.4 *Let* $G = \langle \mathcal{V}, \pi\rangle$ *be a stochastic F0L-system on* $\mathcal{V} = \{v_1,\cdots,v_m\}$. *For all* $w \in \mathcal{V}^*$, *let* $G[w] = \langle w, \pi\rangle$ *be a component system of* G *and and* $(F_n[w])_{n\in\mathbb{N}}$ *the corresponding Markov chain. For all* $n \in \mathbb{N}$, *let* $\psi_n[w]$ *be the generating function associated to* $F_n[w]$. *Then,*

$$\forall n \in \mathbb{N}, \forall w \in \mathcal{V}^*, \quad \psi_{n+1}[w](S) = \psi_1[w](\psi_n[v_1](S),\cdots,\psi_n[v_m](S))$$

We see that we can determine the generating function $\psi_{n+1}[w]$ at every cycle from the simple knowledge of $\psi_1[w]$ and $\psi_1[v_i]$, for all $1 \leq i \leq m$ (since then $\psi_n[v]$ can be determined recursively by taking $w = v_i$). And $\psi_1[v_i]$, for all $1 \leq i \leq m$, is directly given by the transition matrix of the L-system :

$$\psi_1[v_i](S) = \sum_{\chi\in\tilde{\mathcal{V}}^*} k(v_i,\chi) \prod_{i=1}^{m} s_i^{\Upsilon_i(\chi)} .$$

Moreover

$$\psi_1[\omega_a](S) = \prod_{i=1}^{m} \left(\psi_1[v_i](S)\right)^{\Upsilon_i(\omega_a)},$$

so that the transition matrix of the L-system is sufficient to determine recursively the generating functions.

2.7.4 Application to stochastic plant development

The productions of buds given in the definition of GreenLab organogenesis 2.4 may deeply vary according to their positions, environmental conditions, plant age... The deterministic modelling of such phenomena is not always easy. For this reason, the development processes may be considered as stochastic. The F0L-system $\Gamma = \langle \mathcal{G}, \mathcal{B}, P_r \rangle$ describing GreenLab organogenesis can be derived to the stochastic case by associating a probability distribution to the production rules, *cf.* Kang et al. (2007a), which amounts to give a transition matrix π from $\mathcal{G}$ (or more precisely from $\mathcal{B}$, since only bud productions are supposed stochastic) to $\mathcal{G}^*$. Experimental observations on plant populations and statistical analyses allow the estimation of the parameters of stochastic organogenesis models, cf. de Reffye (1979) for original works or Guédon et al. (2001) more recently.
For $n \in \mathbb{N}$, the generating function $\psi_n\left[b_1^0(0)\right]$ takes into account all the possible realisations of a plant starting with $b_1^0(0)$ (that is to say the seed) after n growth cycles, weighted by their probability of occurence. It can be determined recursively thanks to Theorem 2.4 from the knowledge of ψ_1 on $\mathcal{B}$, which is directly obtained provided the transition matrix π of the organogenesis L-system is known.
However, the development processes underlying GreenLab organogenesis are complex. The dynamics of the population of buds results from the combination of both branching and differentiation (see Section 2.1). As a consequence, we do not have an easy access to the distributions of the production rules (that is to say to $\psi_1(b)$, for all $b \in \mathcal{B}$). Thus, the idea is to break down the whole system into two simpler subsystems, one corresponding to development and branching without meristem differentiation, and the other simply corresponding to meristem differentiation. It proves easier to derive a stochastic model for each of this phenomenon and to compute accordingly the generating functions of each subprocess and then to compose them. We do not detail how each of this organogenesis subprocess are modelled. We refer to Kang et al. (2007a) and Loi and Cournède (2008a) for development without differentiation, and to Loi and Cournède (2008a) for meristem differentiation. In this last paper, the phase-type theory is introduced to derive the generating function.

Compound L-systems

The decomposition into simpler subsystems is made possible by defining the concept of compound L-system.

Definition 2.15 (compound L-system) *Let $G = \langle \mathcal{V}, \mathcal{V}^*, \pi \rangle$, $G_1 = \langle \mathcal{V}, \mathcal{V}^*, \pi_1 \rangle$ and $G_2 = \langle \mathcal{V}, \mathcal{V}^*, \pi_2 \rangle$ be stochastic L-systems on the same alphabet $\mathcal{V}$. Let P, P_1 and P_2 be the Markov kernels associated respectively to G, G_1 and G_2. G is said to be the composition of G_1 by G_2 if $P = P_1.P_2$. We write $G = G_1 \circ G_2$.*

Remark 2.10 *This notation is used to be coherent with the composition of the generating functions and the multiplication of the markov kernels. Nevertheless, from an algorithmic point of view, it corresponds to first applying the production rules of G_1, and then those of G_2.*

Remark 2.11 *Note that it is necessary to use F0L-systems, since the axiom of G_2 is given by the production of G_1.*

For GreenLab organogenesis, we can consider two elementary F0L-systems, one corresponding to development and branching without meristem differentiation, $\Gamma_{br} = \langle \mathcal{V}, \mathcal{V}^*, \pi_{br} \rangle$, and the other simply corresponding to meristem differentiation $\Gamma_{dif} = \langle \mathcal{V}, \mathcal{V}^*, \pi_{dif} \rangle$. We can show that (*cf.* Loi and Cournède (2008a)) :

Proposition 2.2 *the F0L-system corresponding to the complete organogenesis* Γ_{tot} *can be written as a compound L-system :* $\Gamma_{tot} = \Gamma_{br} \circ \Gamma_{diff}$.

We will let ψ^{tot}, ψ^{dif} and ψ^{br} denote respectively the generating functions of Γ_{tot}, Γ_{dif} and Γ_{br}. To the alphabet $\mathcal{B} \cup \mathcal{M}$ we let correspond the variables $S \in [0;1]^{Card(\mathcal{B})}$ and $M \in [0;1]^{P}$. Here we simply consider the number of metamers according to their physiological age , therefore $Card(\mathcal{M}) = P$. The composition result in Proposition 2.2 leads to the following theorem, whose proof is given in (Loi and Cournède, 2008a) :

Theorem 2.5

$$\boldsymbol{\psi}^{\boldsymbol{tot}}_{\boldsymbol{n+1}}(S, M) = \boldsymbol{\psi}^{\boldsymbol{br}}_{\boldsymbol{1}}(\boldsymbol{\psi}^{\boldsymbol{dif}}_{\boldsymbol{1}}(\boldsymbol{\psi}^{\boldsymbol{tot}}_{\boldsymbol{n}}(S, M), M), M)$$

with $\boldsymbol{\psi}^{\boldsymbol{tot}}_{\boldsymbol{n}}$, $\boldsymbol{\psi}^{\boldsymbol{dif}}_{\boldsymbol{1}}$, $\boldsymbol{\psi}^{\boldsymbol{br}}_{\boldsymbol{1}}$ *vectors of size* $Card(\mathcal{B})$ *:*

$$\boldsymbol{\psi}^{\boldsymbol{tot}}_{\boldsymbol{n}} = \left(\psi^{tot}_{n}[b]\right)_{b \in \mathcal{B}},$$

$$\boldsymbol{\psi}^{\boldsymbol{dif}}_{\boldsymbol{1}} = \left(\psi^{dif}_{1}[b]\right)_{b \in \mathcal{B}},$$

and

$$\boldsymbol{\psi}^{\boldsymbol{br}}_{\boldsymbol{1}} = \left(\psi^{br}_{1}[b]\right)_{b \in \mathcal{B}}.$$

Recurrent equations to compute the distribution moments of the numbers of metamers

Let $m_j[b, n]$ be the number of type j metamers in a structure initiated by a bud b after n growth cycles and let M_n be the matrix of size $Card(\mathcal{B})$ by P whose j-th column vector is $(E[m_j[b, n]])_{b \in \mathcal{B}}$, where $E[x]$ denotes the expectation of a random variable x. Let e_k be the vector of size k with all its components set to 1. We deduce the fundamental recursion equation for the expectations of the numbers of organs on each type of structures. It is important to note that the first line of the matrix corresponds to those of the whole plant.

Theorem 2.6 *for* $N > 0$,

$$M_{N+1} = \frac{\partial \boldsymbol{\psi}^{\boldsymbol{br}}}{\partial S}(e_{Card(\mathcal{B})}, e_P)\frac{\partial \boldsymbol{\psi}^{\boldsymbol{dif}}}{\partial S}(e_{Card(\mathcal{B})}, e_P)M_N + \frac{\partial \boldsymbol{\psi}^{\boldsymbol{br}}}{\partial M}(e_{Card(\mathcal{B})}, e_P)$$

Proof: Using the properties of generating functions, we have $M_n = \dfrac{\partial \boldsymbol{\psi}_{\boldsymbol{n}}}{\partial M}(e_{Card(\mathcal{B})}, e_P)$. Derivating the equation from Theorem 2.5 and taking $(S, M) = (e_{Card(\mathcal{B})}, e_P)$ ends the proof. □

In the same way, we can give recursion formulas for the variance. The detailed results, as well as examples of simulations and biological discussions, are given in Loi and Cournède (2008b).

Chapitre 3

The Equations of Plant Growth

Literature is already abundant on the functional concepts underlying the GreeLab model and its various versions (de Reffye and Hu (2003), Yan et al. (2004), Guo et al. (2006), Mathieu (2006), Letort (2008), de Reffye et al. (2008b), Cournède et al. (2008), Mathieu et al. (2009)). The fundamental principles common to all these versions are actually very general and are shared (to some extent) with various other models (for example LIGNUM (Perttunen et al., 1996), (Perttunen et al., 2005), TOMSIM (Heuvelink, 1999), GRAAL (Drouet and Pagès, 2003), SUNFLO (Lecoeur et al., 2008), ECO-Meristem (Luquet et al., 2007a)). The central equation of GreenLab describes the growth of an individual plant potentially in a population (field crops, forest stands...). Competition with other individuals for light, water, nutrients can thus be taken into account and affects the computation of biomass production, see Cournède and de Reffye (2007), Cournède et al. (2008) regarding competition for light or Le Chevalier et al. (2007a), Le Chevalier et al. (2007b) for competition for water.

GreenLab aims at describing the source-sink dynamics during plant growth. Sources correspond initially to the seed and then to biomass production and reserve remobilization. Sinks are demands for biomass of all living organs. So far, all sinks have access to all sources since we consider a common pool of biomass. However, this hypothesis is not fundamental and could be relaxed. Since structural development is described in GreenLab at the level of organs, the computation of demands is coupled with organogenesis. For this reason, a consistent time unit for architectural growth and photosynthetic production is defined in order to handle a constant structure. It allows the derivation of the discrete dynamic system of growth.

Our objective in this chapter is thus to present the general mathematical equations corresponding to these fundamental principles, and how the main versions of GreenLab can be derived in this framework.

3.1 Allocation

GreenLab is a source-sink model with a common pool of reserves for biomass allocation to all living organs. This remark leads us to consider the allocation equation as the central one in GreenLab, and not the production equation as generally done in the GreenLab context.

3.1.1 GreenLab fundamental equation

We recall that $m_p^t(n)$ denotes at growth cycle t a metamer of physiological age p and chronological age n (see chapter 2). It contains organs of type o (where $o = b,\ p,\ i,\ f$, for blades, petioles, internodes, flowers or fruits respectively) whose masses are denoted by $q_{o,p}^t(n)$ at growth cycle t. Let $T_{o,p}$ denote the maximal life span of organs of type o and

physiological age p. The allocation equation is thus given for all $t \geq 0$ by :

$$\begin{cases} q_{o,p}^{t+1}(n+1) = q_{o,p}^{t}(n) + \dfrac{p_{o,p}^{t}(n)}{D^t} u^t \text{ if } 0 \leq n \leq \min\left(T_{o,p} - 1, t\right) \\ \\ q_{o,p}^{t}(n) = 0 \text{ if } n = 0 \text{ or } n \geq \min\left(T_{o,p}, t+1\right) \end{cases} \tag{3.1}$$

where :
– $p_{o,p}^{t}(n)$ is the sink of an organ of type o in $m_p^t(n)$
– D^t is the total demand of the plant at growth cycle t (that is to say the sum of all sinks)

$$D^t = \sum_{o,p} \sum_{n=0}^{T_{o,p}} N_{o,p}^{t}(n) p_{o,p}^{t}(n) \tag{3.2}$$

– u^t is the biomass available for allocation (and can be considered as the system input).

Remark 3.1 *Equation (3.1) appears as the eulerian discretization of an ordinary differential equation of the form*

$$\frac{d\left[q_{o,p}(n+1)\right](t)}{dt} = f(t, u(t)) \ .$$

Notations

For any variable v ($v = N$, p, q, r, ...) we define the vectors

$$v_{o,p}^{t} = \left(v_{o,p}^{t}(n)\right)_{0 \leq n \leq T_{o,p}} \ ,$$

$$v_{o}^{t} = \begin{pmatrix} v_{o,1}^{t} \\ \vdots \\ v_{o,P}^{t} \end{pmatrix} \ ,$$

and (depending on the types of organs considered)

$$v^t = \begin{pmatrix} v_b^t \\ v_p^t \\ v_i^t \\ v_f^t \\ v_r^t \end{pmatrix} \ .$$

If v_1 and v_2 are two vectors of the same dimension, we will denote $v_1 \cdot v_2$ their scalar product.

Fundamental Equation

Finally, we can write :

$$q^{t+1} = Aq^t + Ap^t \frac{u^t}{N^t \cdot p^t} \ . \tag{3.3}$$

with A a square block-diagonal matrix.
For the sub-vectors $q_{o,p}^{t} = \left(q_{o,p}^{t}(n)\right)_{0 \leq n \leq T_{o,p}}$, we have

$$q_{o,p}^{t+1} = A_{o,p} q_{o,p}^{t} + A_{o,p} p_{o,p}^{t} \frac{u^t}{N^t \cdot p^t} \ . \tag{3.4}$$

with $A_{o,p} \in \mathcal{M}\left(T_{o,p}+1\right)$:

$$A_{o,p} = \begin{pmatrix} 0 & 0 & \dots & 0 & 0 \\ 1 & 0 & \dots & 0 & 0 \\ 0 & 1 & 0 & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}$$

Remark 3.2 *For plants with a growth cycle short compared to organ life spans ($e.g.$ sunflower, sugar beet), we may observe strong variations between the life spans of organs of the same type according to the growth cycle at which they appeared. In such cases, we let $T_{o,p}(t)$ denote the life span of an organ of type o and physiological age p initiated at growth cycle t (appeared at growth cycle $t+1$). We define $T_{o,p}$ as $T_{o,p} = \max_t\left(T_{o,p}(t)\right)$. And the subvectors remains unchanged : $q_{o,p}^t = \left(q_{o,p}^t(n)\right)_{0\leq n\leq T_{o,p}}$. The elementary allocation equation (3.1) becomes :*

$$\begin{cases} q_{o,p}^{t+1}(n+1) = q_{o,p}^t(n) + \dfrac{p_{o,p}^t(n)}{D^t}u^t \text{ if } 0 \leq n \leq \min\left(T_{o,p}(t-n)-1, t\right) \\ \\ q_{o,p}^t(n) = 0 \text{ if } n = 0 \text{ or } n \geq \min\left(T_{o,p}(t-n), t+1\right) \end{cases} .$$

3.1.2 Primary and secondary growthes

Two physiological phenomena may be differentiated for allocation : primary growth (corresponding to organ expansion) and secondary growth (also called structural growth, mostly for internodes but potentially for the other types of organs). To handle these two phenomena, several modelling strategies may be chosen.

– First, primary ($p_{o,p}^t(n)$) and secondary sinks (denoted $r_{o,p}^t(n)$) compete together. The allocation equation (3.1) simply rewrites :

$$q_{o,p}^{t+1}(n+1) = q_{o,p}^t(n) + \frac{p_{o,p}^t(n) + r_{o,p}^t(n)}{D^t}u^t \tag{3.5}$$

with

$$D^t = \sum_{o,p}\sum_{n=0}^{T_o} N_{o,p}^t(n)\left(p_{o,p}^t(n) + r_{o,p}^t(n)\right) \ . \tag{3.6}$$

– Second, there is first a global allocation of biomass dedicated to primary and secondary growthes respectively. It corresponds to the hypothesis that plants have a global allocation strategy which generally favours primary growth. Secondary growth is mostly used to stock biomass reserves. It is of course not always the case, for example in cases of mechanical stresses. We denote by α^t, the proportion of biomass allocated to primary growth. After this first step, $\alpha^t u^t$ is distributed among the primary sinks and $\left(1-\alpha^t\right)u^t$ among all the secondary sinks. Equation (3.1) becomes :

$$q_{o,p}^{t+1}(n+1) = q_{o,p}^t(n) + \left(\frac{\alpha^t p_{o,p}^t(n)}{D_1^t} + \frac{\left(1-\alpha^t\right)r_{o,p}^t(n)}{D_2^t}\right)u^t \tag{3.7}$$

with

$$D_1^t = \sum_{o,p}\sum_{n=0}^{T_o} N_{o,p}^t(n)p_{o,p}^t(n) \ , \tag{3.8}$$

and

$$D_2^t = \sum_{o,p}\sum_{n=0}^{T_o} N_{o,p}^t(n)r_{o,p}^t(n) \ . \tag{3.9}$$

3.1.3 Root System

In the current version of GreenLab, roots are considered as a compartment. No topological structure is considered, since detailed models of root system development are difficult to define and to identify on real data. Let q_r^t denote the root system biomass at growth cycle t.

As for the secondary growth, there are two possibilities to model allocation to the root system.

– First, the root system is considered as a normal sink p_r^t competing with all the organs of the above-ground part :

$$q_r^t = \frac{p_r^t}{D^t} u^t \tag{3.10}$$

with

$$D^t = \left(\sum_{o,p} \sum_{n=0}^{T_o} N_{o,p}^t(n) p_{o,p}^t(n) \right) \; + \; p_r^t = N^t \cdot p^t \; . \tag{3.11}$$

– Second, we consider a global step of allocation :

$$q_r^t = \beta^t u^t \tag{3.12}$$

where β^t is the proportion of biomass allocated to the root system.

3.1.4 Organ dimensions

If organ dimensions are needed to compute growth, we generally suppose that they are directly related to $q_{o,p}^t(n)$ through allometric relationships (potentially, by distinguishing biomass increment resulting from primary and secondary growthes). The allometric relationships may vary with the environmental conditions (quality and intensity of light for example, *cf.* Munier-Jolain et al. (2005), Dong et al. (2008)). Let $A_{b,p}^t(n)$ denote the blade area of a leaf (or of the leaves) on a metamer of type $m_p^t(n)$. In the simplest case, we assume a constant Specific Blade Mass (SBM), denoted by e, and we get :

$$A_{b,p}^t(n) = \frac{q_{b,p}^t(n)}{e} \; . \tag{3.13}$$

Let $l_{i,p}^t(n)$ denote the length of the internode of a metamer of type $m_p^t(n)$ and $A_{i,p}^t(n)$ its section area. We may choose (*cf.* Letort (2008)) :

$$\begin{cases} l_{i,p}^t(n) = \alpha \left(q_{b,p}^t(n) \right)^{\frac{1+\beta}{2}} \\ A_{i,p}^t(n) = \dfrac{1}{\alpha} \left(q_{b,p}^t(n) \right)^{\frac{1-\beta}{2}} \end{cases} \tag{3.14}$$

In the following, we will consider in details how the variables of the fundamental allocation equation (3.3) (p^t, N^t and finally the input variable u^t) are determined.

3.2 Sink variation function : $\mathbf{p}^t$

(a) Parametric function independent of the state variable

Organ sink variation functions are usually chosen independent of the system state variables and identical for all organs of the same type and same physiological age. We suppose the expansion of an organ of type o and physiological age p lasts from chronological age $T_{o,p}^i$ to chronological age $T_{o,p}^f$, $0 \leq T_{o,p}^i < T_{o,p}^f \leq T_{o,p}$ and organ sink variation functions are modelled with flexible functions able to fit typical biological kinetics. Discretized and

normalized beta-functions multiplied by the global sink value give interesting results.
We choose, for $n \in \left[T^i_{o,p}; T^f_{o,p} - 1\right]$:

$$p^t_{o,p}(n) = P_{o,p} \left(\frac{n+0.5}{T^f_{o,p} - T^i_{o,p}} \right)^{\alpha_o} \left(1 - \frac{n+0.5}{T^f_{o,p} - T^i_{o,p}} \right)^{\beta_o} \frac{1}{\max_{x \in [0,1]} x^{\alpha_o}(1-x)^{\beta_o}} \tag{3.15}$$

and $p^t_{o,p}(n) = 0$ if $n \notin \left[T^i_{o,p}; T^f_{o,p} - 1\right]$. $P_{o,p}$ is the global organ sink value, α_o and β_o are the parameters of the beta-function associated to organ sink variation.
Usually, $T^i_{o,p} = 0$. However, especially for herbaceous plants, we may observe a delay of expansion for axillary branches, (*e.g.* for *Arabidopsis Thaliana*, Letort (2008), Christophe et al. (2008)). Note that the global sink is usually chosen depending on organ's physiological age, but not the parameters of the beta-function.
With this formulation, $p^t_{o,p}(n)$ only depends on organ's chronological age and not on the growth cycle t. However, as for organ life spans, we may observe strong variations between the expansion times of organs of the same type according to the growth cycle at which they appeared. We therefore let $T^i_o(t)$ and $T^f_o(t)$ denote the initial and final chronological ages of expansion for organs of type o initiated at growth cycle t (appeared at growth cycle $t+1$). We then choose for $n \in \left[T^i_o(t-n); T^f_o(t-n) - 1\right]$ (Lemaire et al. (2008)).

$$p^t_{o,p}(n) = P_{o,p} \left(\frac{n+0.5}{T^f_{o,p}(t-n) - T^i_{o,p}(t-n)} \right)^{\alpha_o} \left(1 - \frac{n+0.5}{T^f_{o,p}(t-n) - T^i_{o,p}(t-n)} \right)^{\beta_o} \frac{1}{\max_{x \in [0,1]} x^{\alpha_o}(1-x)^{\beta_o}} \tag{3.16}$$

(b) Function of the input variables

Mathieu (2006) proposed that for some plants, the global sink value $P_{o,p}$ of an organ may depend on the trophic competition when it is initiated. The trophic competition in a plant can be well represented by the ratio of available biomass to demand. If we denote $P_{o,p}(t)$ the global sink value of organs of type o and physiological age p initiated at growth cycle t, we thus choose a function $f_{o,p}$ (linear for example) :

$$P_{o,p}(t) = f_{o,p}\left(\frac{u^t}{D^t}\right) = f_{o,p}\left(\frac{u^t}{N^t \cdot p^t}\right) \ . \tag{3.17}$$

Since $p^t_{o,p}(0)$ depends on $P_{o,p}(t)$, we see that we need to solve an implicit equation. We can thus write implicitly :

$$P_{o,p}(t) = \psi\left(\frac{N^t}{u^t}, P_{o,p}(t-1), P_{o,p}(t-2), \ldots, P_{o,p}(t-T_o)\right) \tag{3.18}$$

By substituting recursively, we finaly see that p^t depends on $\left(\frac{N^t}{u^t}, \frac{N^{t-1}}{u^{t-1}}, \ldots, \frac{N^0}{u^0}\right)$.
If we do not want to keep all the past inputs in the state equation, we shall consider p^t as a state variable.

(c) Function of the state variables and of topology

We may also consider that the sink of an organ belonging to a metamer $m^t_{pq}(n)$ is proportional to the leaf area borne by the substructure it belongs to, that is to say to the total leaf area in $S^t_p(n)$. It corresponds to the hypothesis that allocation to an organ depends on the fluxes that it sees (with matter fluxes from leaves to roots). This hypothesis is particularly relevant for secondary growth (*cf.* Shinozaki et al. (1964) for the pipe-model theory, in coherence with the Pressler Law).

For the sake of clarity, we consider no reiteration and no differenciation. Leaf area can be seen as structural morphism (*cf.* Definition 2.7) from $(\mathcal{G}^*, \cdot)$ into $(\mathbb{R}, +)$. Let Γ denote this morphism. By applying the morphism property (*cf.* Equation (2.6)) at growth cycle t, we deduce $\Gamma\left(S_p^t(n)\right) = \Gamma_p^t(n)$ for all p, n recursively on n.

$$\Gamma_p^t(n) = \sum_{p \leq q \leq P} u_{pq}(t+1-n)\ A_p^t(n) + \sum_{p \leq q \leq P} v_{pq}(t+1-n)\Gamma_q^t(n-1) + \Gamma_p^t(n-1) \quad (3.19)$$

and we may choose

$$p_{o,p}^t(n) = {p'}_{o,p}^t(n)\Gamma_p^t(n) \quad (3.20)$$

where ${p'}_{o,p}^t(n)$ is usually an independent function (as in (a)), but potentially a more complex function of the input variables (as in (b)). $\Gamma_p^t(n)$ is a function of q^t and of plant topology through the sequences $\left(U^k\right)_{1 \leq k \leq t+1}$ and $\left(V^k\right)_{1 \leq k \leq t}$. We need U^k on $[1; t+1]$ since $N^t(0)$ represents organs that will only appear at growth cycle $t+1$.

3.3 Number of organs N^t and topology variables U^t, V^t

N^t is a direct function of the sequences $\left(U^k\right)_{1 \leq k \leq t+1}$, $\left(B^k\right)_{1 \leq k \leq t}$. However, we distinguish it as a specific variable to differentiate the cases when only the numbers of organs play a role and those when the way they are connected has to be taken into account through the whole sequences $\left(U^t\right)_t$, $\left(B^t\right)_t$.
In Section 2.6.3, we showed that there exists a function f_1 such that :

$$N^t(0) = f_1\left(\beta^t, U^{t+1}\right),$$

thus, there exists a function f_2 such that

$$N^t = f_2\left(N^{t-1}, \beta^t, U^{t+1}\right) \quad (3.21)$$

where β^t is the vector of active buds of every physiological age at growth cycle t and

$$\beta^t = f_3\left(\beta^{t-1}, V^t\right)\ . \quad (3.22)$$

(a) Parametric functions independent of the state variables

In the simplest case, $\left(U^t\right)_t$, $\left(B^t\right)_t$ are given sequences. For example in the GL1 model (Zhan et al. (2003)), the sequences are constant. As shown in section 2.5, we can even compute ν^t and thus N^t explicitly. Another important case is when the sequences are not constant but can be fully observed empirically.

(b) Functions of the input variables

As for the sink variation function, Mathieu (2006) modelled a retroaction of plant functioning on plant topology controlled by the state of trophic competition (GL3 model). U^t and V^t are chosen as functions of the ratio of available biomass to demand. Integer part of linear functions are usually chosen. In Figure 3.1, we see the difference between the GL1 model (with constant sequences $\left(U^t\right)_t$, $\left(V^t\right)_t$) and the GL3 model, (with sequences $\left(U^t\right)_t$, $\left(V^t\right)_t$ functions of u^t).
There is a slight complexity induced by the fact that U^t is a function of $\frac{u^{t-1}}{N^{t-1} \cdot p^{t-1}}$:

$$U^t = g_1\left(\frac{u^{t-1}}{N^{t-1} \cdot p^{t-1}}\right) \quad (3.23)$$

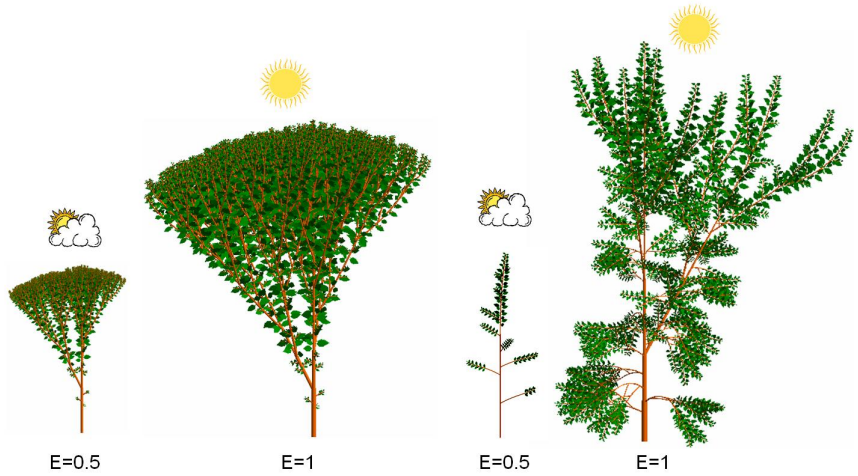


FIG. 3.1 – *Effects of climatic variations on the GL1 and GL3 organogenesis models : on the left side (GL1), topology is fixed, organs are simply smaller to adjust to the environment ; on the right side (GL3), toplogy is variable and the plant adapts its development to the environmental conditions.*

while V^t is a function of $\frac{u^t}{N^t \cdot p^t}$:

$$V^t = g_2\left(\frac{u^t}{N^t \cdot p^t}\right) \ . \tag{3.24}$$

Finally, combining Equations (3.21), (3.22), (3.23), (3.24), there exists h_1 such that :

$$N^t = h_1\left(N^{t-1}, \beta^{t-1}, \frac{u^t}{N^t \cdot p^t}\right) \tag{3.25}$$

We have to solve an implicit equation to determine N^t.
Note that in practice, we usually write and solve the implicit equations in $\left(U^{t+1}, V^t\right)$. Combining differently Equations (3.21), (3.22), (3.23), (3.24), we also show that there exists h_2 such that :

$$\left(U^{t+1}, V^t\right) = h_2\left(N^{t-1}, \beta^{t-1}, \frac{p^t}{u^t}, U^{t+1}, V^t\right) \ .$$

We thus determine $\left(U^{t+1}, V^t\right)$ as an implicit function of $\left(N^{t-1}, \beta^{t-1}, \frac{p^t}{u^t}\right)$ and finally deduce N^t.

Remark 3.3 *We choose not to change the notations that were used in previous works, even though it appears that it would have been more appropriate to define $u_{pq}(t)$ as the number of metamers m_{pq} initiated at growth cycle t and appearing at growth cycle $t+1$ instead of the number of metamers appearing at growth cycle t.*

3.4 Control variable for biomass allocation : u^t

In Equation (3.3), there are several ways to choose the input variable u^t, that is to say the available biomass for allocation. Different strategies were used and are currently used in GreenLab according to plants. There are mostly 3 sources to provide biomass for allocation : seed, resource acquisition by photosynthesis and reserve remobilization (from roots, layers or senescent organs). We will not speak about the seed, which usually gives u^0 and may

be considered as emptying on several growth cycles, that is to say also contributing to u^1, u^2 ...
From a mathematical point of view, we distinguish 3 cases : without feedback, with feedback but independent of plant topology, with feedback and dependent on plant topology.

(a) Open-loop system (without feedback)

This approach is quite frequent in crop models and usually considers the Leaf Area Index (LAI) as a parametric function of the thermal time, water stresses or other environmental conditions, *cf.* models like PILOTE (Mailhol et al. (1997)) or SUNFLO (Lecoeur et al. (2008)). Biomass production calculation is based on radiation interception and biological efficiencies. u_t is thus only a function of external variables (related to the environmental conditions). Such an approach is perfectly compatible with the GreenLab model, even though not generally used.

(b) Closed-loop system (with feedback) independent from plant topology

Instead of considering leaf area and thus biomass production as a direct function of the environmental conditions, it is also possible to deduce it from the biomass allocated to blades. This approach is very typical of the Wageningen modelling school (see for example Heuvelink (1999) or Marcelis et al. (1998) for a review). Let us denote $\mathcal{A}^t$ the total green leaf area at growth cycle t. In the simplest case, we assume a constant Specific Blade Mass (SBM), denoted by e, and we get :

$$\mathcal{A}^t = \frac{1}{e} \sum_p \sum_{n=1}^{T_o} N_{b,p}^t(n)\, q_{b,p}^t(n) \tag{3.26}$$

The biomass production at growth cycle t Q^t is taken as a parametric function of $\mathcal{A}^t$ and of the environmental conditions. To keep it simple, we sum up all the environmental effects in E^t, which may be chosen as a function of the Photosynthetically Active Radiation (PAR), hydric stress (Fraction of Transpirable Soil Water), temperature (for biologic efficiency)... Only simple functions have been tested so far in GreenLab. A possible production function inspired by classical crop models and Beer-Lambert's law (*cf.* Guo et al. (2006) and Cournède et al. (2008) for details) can be written :

$$Q^t = E^t \beta \left(1 - \exp\left(-\gamma \mathcal{A}^t\right)\right) \tag{3.27}$$

where β and γ are empirical parameters.
We, the thus have :

$$q^{t+1} = Aq^t + Ap^t \frac{u^t}{N^t \cdot p^t} \tag{3.28}$$

with the feedback :

$$u^t = E^t \beta \left(1 - \exp\left(-\eta N_b^t \cdot q_b^t\right)\right) \ . \tag{3.29}$$

Again, we see that plant structural development is only taken into account through the numbers of organs.
Another type of feedback can be introduced by modelling reserve remobilization. For example, Letort (2008) proposes that, after a time $\tau_{o,p}$, a proportion of organ biomass is given back to the common pool and reallocated to new organs in expansion. It implies

a change in matrices $A_{o,p}$ of Equation (3.4), as follows :

$$A_{o,p} = \begin{pmatrix} 0 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ 1 & 0 & \dots & \dots & \dots & \dots & \dots & 0 \\ 0 & 1 & 0 & \dots & \dots & \dots & \dots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots & \vdots & \vdots & \vdots \\ 0 & \dots & 0 & 1 & 0 & \dots & \dots & 0 \\ 0 & \dots & \dots & 0 & \theta(1) & 0 & \dots & 0 \\ \vdots & \dots & \dots & \dots & \ddots & \ddots & \ddots & \vdots \\ 0 & \dots & \dots & \dots & \dots & 0 & \theta(T_{o,p} - \tau_{o,p}) & 0 \end{pmatrix} \leftarrow \text{ line } \tau_{o,p} + 2$$

where for all $k \in [1; T_{o,p} - \tau_{o,p}]$:

$$\theta(k) = \frac{1 - F_{o,p}(1 - \alpha_{o,p}^{k})}{1 - F_{o,p}(1 - \alpha_{o,p}^{k-1})}$$

and $F_{o,p}, \alpha_{o,p}$ are parameters denoting respectively the proportion of biomass that can be remobilized and the speed of remobilization ($\alpha_{o,p} \in (0; 1)$).
Moreover, the feedback is now given by :

$$u^t = Q^t + \sum_{o,p} \sum_{n=\tau_{o,p}}^{T_{o,p}} N_{o,p}^t(n) q_{o,p}^t(n) \frac{F_{o,p} \alpha_{o,p}^{n-\tau_o} (1 - \alpha_{o,p})}{1 - F_{o,p} \left(1 - \alpha_{o,p}^{n-\tau_{o,p}}\right)}$$

(c) Closed-loop system (with feedback), with dependence on plant topology

Another strategy to compute biomass production consists in calculating the radiation intercepted by each leaf in plant architecture by modelling the radiative exchanges in the whole structure (see Soler et al. (2003) for an example of the radiosity method, or Chelle and Andrieu (2007) for a review). Models like LIGNUM Perttunen et al. (1996) or L-PEACH Allen et al. (2005) use this strategy. Such models thus introduce a strong feedback of architecture on functioning. However, it may seem a bit artificial since exact geometric representation of real canopies remains attainable only through the very laborious work of plant digitalization (e.g. Sinoquet et al. (2005)). Moreover, digitalization only allows represention of the particular plants studied, and most results are difficult to generalize. Thus, light models are usually used on virtual canopies generated by 3-D geometric simulations of crops or stands (Chelle and Andrieu (1998), Soler et al. (2003)). These virtual canopies may quite satisfactorily mimic some global geometric properties of real canopies, such as the gap fraction (see Fournier and Andrieu, 1998), but their detailed geometries necessarily differ, which may entail a significant bias in the computation of radiative fluxes.
If exact plant geometry is unattainable (especially for trees), a faithful representation of plant topology remains possible, by studying precisely the development of topological structures Godin and Caraglio (1998) and branching processes Guédon et al. (2001).
Some models of biomass acquisition are based on plant topology, for example transport-resistance models (de Reffye et al. (1997a), Allen et al. (1998)). Dry matter production is related to plant water transpiration Jones (1992), which is limited by water uptake from roots and by water transport in plant architecture. In de Reffye et al. (1997a) or Mathieu (2006), the resistance to water flow circulation is defined for leaves and internodes :

– A leaf of physiological age p and chronological age n, at growth cycle t has a resistance $R_{l,p}^t(n)$ given by :

$$R_{l,p}^t(n) = \frac{R_1}{A_{b,p}^t(n)} + R_2 = \frac{R_1'}{q_{b,p}^t(n)} + R_2$$

where $A_{b,p}^t(n)$ is its blade surface area, R_1, R_2, R_1' parameters ($R_1' = R_1/e$ and R_2 is petiole resistance supposed constant).

– An internode of physiological age p and chronological age n, at growth cycle t has a resistance $R^t_{i,p}(n)$ given by :

$$R^t_{l,p}(n) = \frac{\rho_p l^t_{i,p}(n)}{A^t_{i,p}(n)} = \rho'_p \left(q^t_{i,p}(n)\right)^{\beta_p}$$

where $l^t_{i,p}(n)$ is its length and $A^t_{i,p}(n)$ its section area. Both are related to $q^t_{i,p}(n)$ through allometric relationships (Letort (2008)). More exactly, instead of its total section area, we should only consider sapwood area, that is to say only the conductive part of the section. We can do it without difficulty by considering $A^t_{i,p}(n)$ as a function of $\left(q^t_{i,p}(n) - q^{t-k}_{i,p}(n-k)\right)$ instead of $q^t_{i,p}(n)$ (k would be the number of conductive rings).

These organs are connected in parallel or in series to form the hydraulic network into which water circulates from the roots to every leaf under the control of evapotranspiration potential (mm.day-1) as defined in Allen et al. (1998). Under the assumption that every leaf is subjected to the same potential of evapotranspiration E^t, plant biomass production can be calculated as :

$$Q^t = \frac{E^t}{\mathcal{R}^t} \tag{3.30}$$

where $\mathcal{R}^t$, the total resistance of plant structure, is computed recursively thanks to the structural factorization given in Equation (2.4). Let Φ be defined on the set of structures $\mathcal{A}^*$ and give the equivalent resistance. For the sake of clarity, we suppose that there is no differentiation. We can take for example (Mathieu (2006)) :

$$\Phi\left(S^t_p(n)\right) = \sum_{q=p}^{P} u_{pq} R^t_{i,p}(n) + \frac{1}{\sum_{q=p}^{P} u_{pq} \frac{1}{R^t_{l,p}(n)} + b_{pq} \frac{1}{\Phi\left(S^t_q(n-1)\right)} + \frac{1}{\Phi\left(S^t_p(n-1)\right)}} \tag{3.31}$$

Finally

$$\mathcal{R}^t = \Phi\left(S^t_1(t)\right) \tag{3.32}$$

and the feedback can be written

$$u^t = E^t \mathcal{R}^t \left(q^t, \left(U^k\right)_{1\leq k\leq t}, \left(B^k\right)_{1\leq k\leq t-1})\right) \ . \tag{3.33}$$

It depends on the state variables q^t, but also on the topology through the sequences $\left(U^t\right)_t$ and $\left(B^t\right)_t$.

Remark 3.4 *Note that if the hydraulic resistance in the trunk and branches can be neglected (which is the case for young trees, Cruiziat et al. (2002)), we have :*

$$Q^t = E^t \sum_{p,n} N^t_{b,p}(n) \frac{1}{R^t_{l,p}(n)} = E^t \sum_{p,n} N^t_{b,p}(n) \frac{1}{\frac{R'_1}{q^t_{b,p}(n)} + R_2} \tag{3.34}$$

and we see that again Q^t does not depend on topology. We are in the same situation as in paragraph (b). Such a model was studied in details by Mathieu (2006).

3.5 Discussion and examples

We have explored a lot of possibilities to model plant functional growth in the GreenLab mathematical framework, mostly in terms of feedback and interactions with plant topology. Through this approach, we see that the levels of retroaction can become tremendously intricate. It raises some difficult issues about the mathematical properties of the dynamic

system thus obtained, like stability or identifiability. A comprehensive study of the dynamic systems of growth has not been achieved yet. Likewise, the parametric sensitivity analysis has not been done systematically. It will prove of crucial importance for complex systems for which the biophysical meaning of parameters is blurred by interactions of several kinds. We present here two versions of the model which have been particularly studied and used.

3.5.1 First example : given numbers of organs, no influence of plant topology, biomass production as a feedback function

In this first case, the sink functions are independent of the functioning (case (a) of section 3.2), usually without secondary growth (or with sinks for the secondary growth independent of topology), the number of organs is also independent of plant functioning (usually of type GreenLab 1) and biomass available for allocation is determined with feedback (case (b) of section 3.4). The function for biomass production may be chosen by analogy with Beer-lambert's law. N^t and p^t are given parametric functions. This type of model corresponds to the one used for example for maize (Guo et al. (2006), Ma et al. (2008)) or sugar beet (Lemaire et al. (2008)).
We thus have :

$$q^{t+1} = Aq^t + u^t b^t \tag{3.35}$$

with b^t a parametric function only depending on time :

$$b^t = \frac{Ap^t}{N^t \cdot p^t}$$

and the feedback :

$$u^t = f\left(E^t, N_b^t \cdot q_b^t\right) \ . \tag{3.36}$$

where E^t is the environmental input.
We recall that a classical choice in GreenLab is given by Equation (3.27) :

$$u^t = E^t \beta \left(1 - \exp\left(-\eta N_b^t \cdot q_b^t\right)\right) \ . \tag{3.37}$$

β and η are empirical parameters.

Remark 3.5 *In some situations, the variable of interest may simply be the production at every groth cycle. We show below that (in the case with feedback) a recurrent equation can be deduced for u^t.*

We have $u^t = f\left(N_b^t \cdot q_b^t\right)$. From equation (3.35), we deduce :

$$N_b^t \cdot q_b^t = Aq_b^{t-1} \cdot N_b^t + \frac{Ap_b^{t-1} \cdot N_b^t}{N^{t-1} \cdot p^{t-1}} u^{t-1}$$

And it is easy to see that

$$Ap_b^t \cdot N_b^t = p_b^{t-1} \cdot N_b^{t-1}$$

and

$$Aq_b^{t-1} \cdot N_b^t = q_b^{t-1} \cdot N_b^{t-1} - q_b^{t-1}\left(T_b\right) N_b^{t-1}\left(T_b\right) \ .$$

Therefore,

$$N_b^t \cdot q_b^t = N_b^{t-1} \cdot q_b^{t-1} - q_b^{t-1}\left(T_b\right) N_b^{t-1}\left(T_b\right) + \frac{N_b^{t-1} \cdot p_b^{t-1}}{N^{t-1} \cdot p^{t-1}} u^{t-1} \ ,$$

and recursively

$$N_b^t \cdot q_b^t = -\sum_{k=0}^{t-1} q_b^k\left(T_b\right) N_b^k\left(T_b\right) + \sum_{k=0}^{t-1} \frac{N_b^k \cdot p_b^k}{N^k \cdot p^k} u^k \ .$$

After some simplifications, we can finally rewrite :

$$N_b^t \cdot q_b^t = \sum_{k=\max(0,t-T_b)}^{t-1} \frac{u^k}{N^k \cdot p^k} \left(\sum_{j=0}^{\min(k+T_b-t,k)} N_b^k(j) p_b^k(j) \right) . \tag{3.38}$$

If we adopt the following conventions $u^t = 0$ if $t < 0$ and $N^t(j) = 0$ for $j > t$, we write more simply :

$$N_b^t \cdot q_b^t = \sum_{k=t-T_b}^{t-1} \frac{u^k}{N^k \cdot p^k} \left(\sum_{j=0}^{k+T_b-t} N_b^k(j) p_b^k(j) \right) . \tag{3.39}$$

Substituting into Equation (3.37), we see that u^t is simply written as a function of $\left(u^{t-1}, u^{t-2}, \ldots, u^{t-T_b}\right)$. :

$$u^t = f\left(E^t, \sum_{k=t-T_b}^{t-1} \frac{u^k}{N^k \cdot p^k} \left(\sum_{j=0}^{k+T_b-t} N_b^k(j) p_b^k(j) \right) \right) . \tag{3.40}$$

Such a formulation is interesting to study system stability since u^t is characteristic of plant global performance, especially when it is restricted to the available biomass produced by photosynthesis.
We may be interested in determining the limit production of the system. A general approach in front of iterative systems like (3.40) is to look for the fixed points of the function $g : \mathbb{R}^{T_b} \to \mathbb{R}^{T_b}$, where :

$$g\begin{pmatrix} u_{t-T_0} \\ \vdots \\ u_{t-2} \\ u_{t-1} \end{pmatrix} = \begin{pmatrix} u_{t-T_0+1} \\ \vdots \\ u_{t-1} \\ f\left(E^t, \sum_{k=t-T_b}^{t-1} \frac{u^k}{N^k \cdot p^k} \left(\sum_{j=0}^{k+T_b-t} N_b^k(j) p_b^k(j) \right) \right) \end{pmatrix} . \tag{3.41}$$

It is easy to see that, if it exists, a fixed point of g is necessarily of the form

$$\begin{pmatrix} \alpha \\ \vdots \\ \alpha \end{pmatrix}$$

where α is solution of

$$\alpha = f\left(E^t, \sum_{k=t-T_b}^{t-1} \frac{\alpha}{N^k \cdot p^k} \left(\sum_{j=0}^{k+T_b-t} N_b^k(j) p_b^k(j) \right) \right) . \tag{3.42}$$

In the following, we will call α the fixed point of the sequence $\left(u^t\right)_t$ when

$$\begin{pmatrix} \alpha \\ \vdots \\ \alpha \end{pmatrix}$$

is the fixed point of g.
Some studies were achieved by Grange (2006) and Letort (2008) to determine the fixed points of the sequence $\left(u^t\right)_t$, when the feedback equation is given by (3.27). 0 is of course always a fixed point. We recall an interesting result.

Lemma 3.1 *Let us consider the sequence defined by (3.37,3.40) and suppose that :*

1. $\sum_{k=t-T_b}^{t-1} \frac{1}{N^k \cdot p^k} \left(\sum_{j=0}^{k+T_b-t} N_b^k(j) p_b^k(j) \right)$ *has a finite limit* C_∞ *(which is achieved for example when the numbers of organs have finite limits and when* p^t *is time-independent),*
2. E^t *has a finite limit* E *.*

Let $C_1 = E\beta$ *and* $C_2 = \eta C_\infty$.
Then there exists a non-zero fixed point u_∞ *for* $\left(u^t\right)_t$ *if and only if* $C_1 C_2 > 1$. u_∞ *satisfies :*

$$u_\infty = C_1 + \frac{1}{C_2} W\left(-C_1 C_2 \exp\left(-C_1 C_2\right)\right) \tag{3.43}$$

where W *is the Lambert function (cf. for example Corless et al. (1996),* $W(x)$ *is defined as the solution of the equation* $W(x)e^{W(x)} = x$*).*

We refer to Grange (2006) for the simple proof of this lemma.
However, proving that fixed points are limits of the production function is not trivial. Some interesting preliminary results are given by Chauveheid (2008) when $p_o^t(j) = 0$ for $j \geq 1$, which corresponds to an immediate expansion of organs.
The study of the fixed points and of the convergence of the sequence of biomass productions was also studied in the hydraulic case, that is to say when resource acquisition is given by Equation (3.34), by Mathieu (2006), Mathieu et al. (2007a).

3.5.2 Second example : retroaction of biomass production on the numbers of organs (GL3)

The source-sink model considered here is the same as the one presented above. Sink functions are given, that is to say independent of plant topology and of plant functioning. Biomass available for allocation is a function of leaf biomass (case with feedback) and independent of plant topology as well. But here, we consider that there is a retroaction of functioning on the development (case (b) of section 3.3).
We thus have :

$$q^{t+1} = Aq^t + u^t \frac{Ap^t}{N^t \cdot p^t} \tag{3.44}$$

with the feedback for available biomass :

$$u^t = f\left(E^t, N_b^t \cdot q_b^t\right) \ . \tag{3.45}$$

and the feedback of biomass production on plant topology modelled as follows :

$$U^t = g_1\left(\frac{u^{t-1}}{N^{t-1} \cdot p^{t-1}}\right)$$

and

$$V^t = g_2\left(\frac{u^t}{N^t \cdot p^t}\right) \ .$$

In Mathieu (2006), integer part of linear functions were chosen, so that :

$$U^t = E\left(A_u\left(\frac{u^{t-1}}{N^{t-1} \cdot p^{t-1}}\right) + B_u\right)$$

and

$$V^t = E\left(A_v\left(\frac{u^t}{N^t \cdot p^t}\right) + B_v\right)$$

where A_u, B_u, A_v and B_v are real-vectors of proper dimensions and $E(x)$ denotes for a real-vector x the vector of the integer parts of all components of x. These functions are usually truncated, since botanical constraints exist on the number of metamers per growth units or on the number of axillary buds on a metamer.
There exist h_1, h_2 (*cf.* Equations 2.13, 2.14) such that :

$$\beta^t = h_1\left(V^t, \beta^{t-1}\right) \ . \tag{3.46}$$

and

$$N_p^t(0) = h_2\left(\beta^t, U^{t+1}\right) \ . \tag{3.47}$$

Substituting the values of U^t and V^t into these expressions we get that there exist h_3, h_4 such that :

$$\beta^t = h_3\left(u^t, N^t, \beta^{t-1}\right) \ . \tag{3.48}$$

and

$$N_p^t(0) = h_4\left(\beta^t, u^t, N^t\right) \ . \tag{3.49}$$

We see that we have to solve an implicit equation to determine N^t.
For such systems, it is convenient to consider as state variables $X^t = \left(q^t, N^t, \beta^t\right)$ so that X^t can be expressed as a function of X^{t-1} only. However, other choices of variables can be made. Mathematical studies of this type of systems were achieved in Mathieu (2006), Mathieu et al. (2007a) or Chauveheid (2008). For example, it is possible to determine at which cycle a plant will produce its first branch or what is the condition for a plant to never branch.
To illustrate such type of retroaction, we can consider a simple type of plant with two physiological ages, with branches of physiological age 2 potentially appearing on the main axis of physiological age 1. Topology is described by :

$$\begin{aligned}
&U_{11}^t = 0\\
&U_{12}^t = 1\\
&U_{22}^t = 1\\
&V_{12}^t = \min\left(1, E\left(a\tfrac{u^t}{N^t.p^t} + b\right)\right)\\
&V^t 22 = 0
\end{aligned}$$

and a classical choice of biomass production given by Beer's law (*cf.* equation 3.27) :

$$u^t = E^t\beta\left(1 - \exp\left(-\eta \sum_{k=t-T_b}^{t-1} \frac{u^k}{N^k \cdot p^k}\left(\sum_{j=0}^{k+T_b-t} N_b^k(j) p_b^k(j)\right)\right)\right) \ . \tag{3.50}$$

In Figure 3.2, we show some examples of branching patterns and of evolutions of the ratio of available biomass to demand for different environmental factors.
The cases $E = 0.9$ and $E = 1.0$ in Figure 3.2 show oscillations of this ratio around the threshold value. It illutsrates an important result underlined by Mathieu et al. (2008), Mathieu et al. (2007a). In some conditions, plant's behaviour can become rhythmic. When the ratio of available biomass to demand reaches a given threshold at growth cycle t, a branch (or a fruit) appears, which may induce a decrease of the ratio under the threshold at growth cycle $t+1$ caused by an increase of plant demand. It may take some time for the ratio to increase again over the threshold level and for a new branch to appear. It results in a succession of branched and unbranched growth units, potentially on a periodic (or pseudo-periodic) basis, see Figure 3.3.

In real situations, the dynamic system of plant growth results from very complex interactions. Feedbacks can be seen as expressions of plant plasticity. In Figure 3.4, are presented

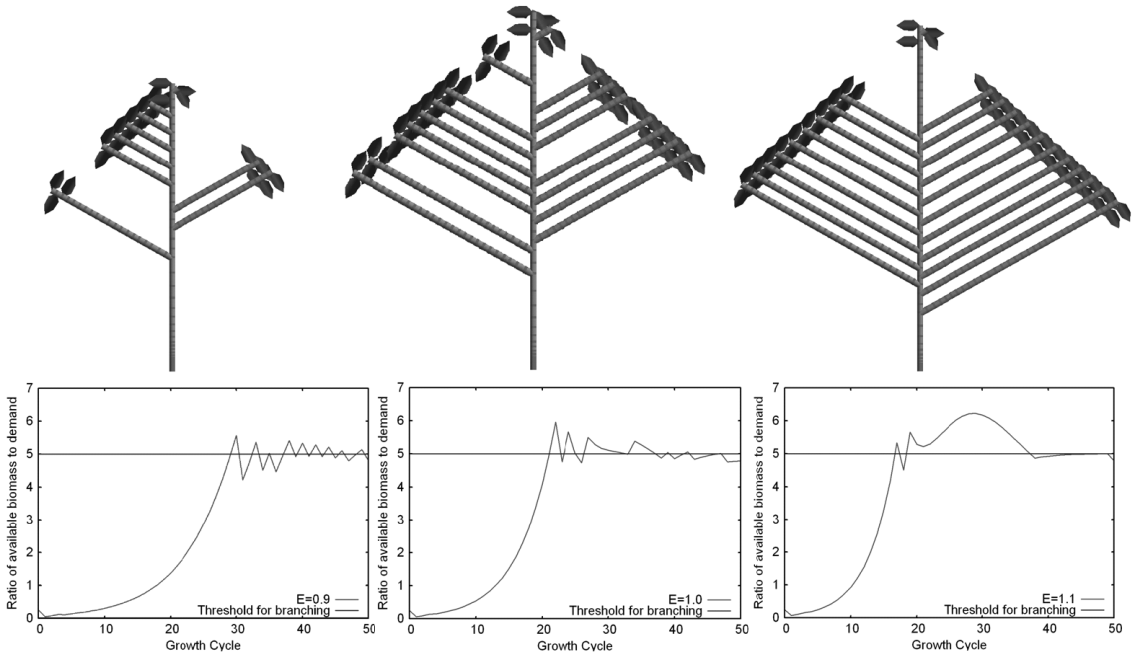


FIG. 3.2 – *Different branching patterns according to the environmental factor. From left to right :* $E = 0.9$, $E = 1.0$, $E = 1.1$ *: schematic representation of the plant and evolution of the ratio of available biomass to demand (immediate expansion,* $p_b = 1$, $p_i = 3$, $T_b = 3$, $\beta = 0.06$, $\eta = 0.0033$, $a = 0.2$, $b = 0$*).*

the simulation results for trees of the same type and same age growing in different competition conditions. The growth model is extended at stand scale to take into account competition for light (*cf.* Cournède et al. (2008) Cournède and de Reffye (2007)). It involves several types of retroaction (retroaction of growth on development, secondary growth depending on the ratio of biomass to demand and on topology, ...).
Studying mathematically such type of system is difficult regarding all the phenomena involved and their complex interactions. A method based on principal component analysis was proposed by Chauveheid (2008). A Proper Orthogonal Decomposition (POD - Galerkin) is applied to the dynamic system to determine its approximation on a simple basis. This method seems promising and will be tested more deeply.

3.6 Stochasticity

There are two specific cases for which we need to handle stochasticity in plant growth models :

– The first case corresponds to the study of the growth of an individual plant, for which some phenomena during plant growth are modelled stochastically. It is typically the case for stochastic models of development as described in Section 2.7 or in (de Reffye et al., 1991b), (Costes et al., 1992) for example. Besides, stochasticity could also result from model uncertainties. For example, in Equation 3.27, the environmental factor E^t could be modeled as a stochastic variable centered on the photosynthetically active radiation PAR^t, when the effects of stresses are not precisely known and when a variation of biomass production around the modelled value is observed.
– The second case occurs when studying variability in a population of plants through an individual-based growth model. The population of plants is described by a stochastic distribution of parameters and exogeneous variables, and each individual corresponds to

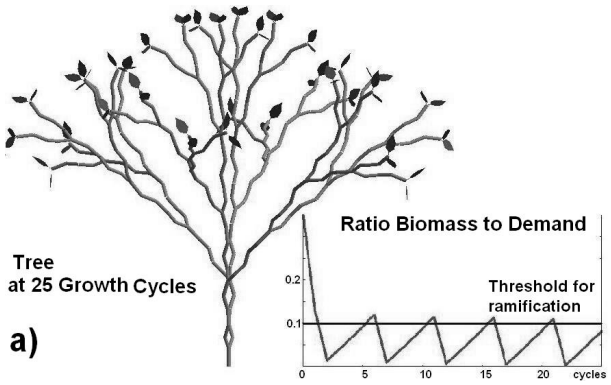


FIG. 3.3 – Simulation of periodic branching patterns, *cf.* Mathieu et al. (2007a). Plant representation at growth cycle 25 and evolution of the ratio of available biomass to demand wrt to growth cycle.

a realization of a stochastic variable of this distribution.
Of course both cases can be mixed when studying population of plants for which the individual-based model is itself stochastic.
Contrary to the stochastic model of plant development, for which we manage to write an explicit relation to determine the distribution moments of the numbers of organs, it is usually not possible to determine explicitly the moments of the state vector (corresponding to biomass accumulation) for stochastic growth models. We have to face the problem of computing the statistics of a random variable which undergoes a nonlinear transformation. This problem is quite classical in the theory of error analysis when studying error propagation, *cf.* for example Bevington and Robinson (2003). We recall in the following section some useful concepts.

3.6.1 Propagation of uncertainty

Let x be a random vector in $\mathbb{R}^d$ with mean $\bar{x}$ and covaraince matrix P. Let $g : \mathbb{R}^d \to \mathbb{R}$ a transformation of x, $y = g(x)$. Statistics on y can be obtained by Monte-Carlo simulations. However, the evaluation cost of g can be prohibitive for such simulations. A classical approximation of the moments of $g(x)$ can be obtained by considering the multivariate Taylor series expansion of $g(x)$ about $\bar{x}$. If we consider x as the sum of $\bar{x}$ and a zero-mean disturbance Δx, we have for all $n > 0$:

$$g(x) = g(\bar{x}) + D_{\Delta x} g(\bar{x}) + \frac{1}{2!} D^2_{\Delta x} g(\bar{x}) + \cdots + \frac{1}{n!} D^n_{\Delta x} g(\bar{x}) + o(\|\Delta x\|^n) \qquad (3.51)$$

where $D_{\Delta x}$ is the differential operator corresponding to the perturbation Δx

$$D_{\Delta x} = \sum_{i=1}^{d} \Delta x_i \frac{\partial}{\partial x_i}$$

such that

$$D_{\Delta x} g(\bar{x}) = \sum_{i=1}^{d} \Delta x_i \frac{\partial g}{\partial x_i}(\bar{x})$$

and

$$D^i_{\Delta x} g(\bar{x}) = \left[\left(\sum_{i=1}^{d} \Delta x_i \frac{\partial}{\partial x_i} \right)^i g \right](\bar{x}) \ .$$

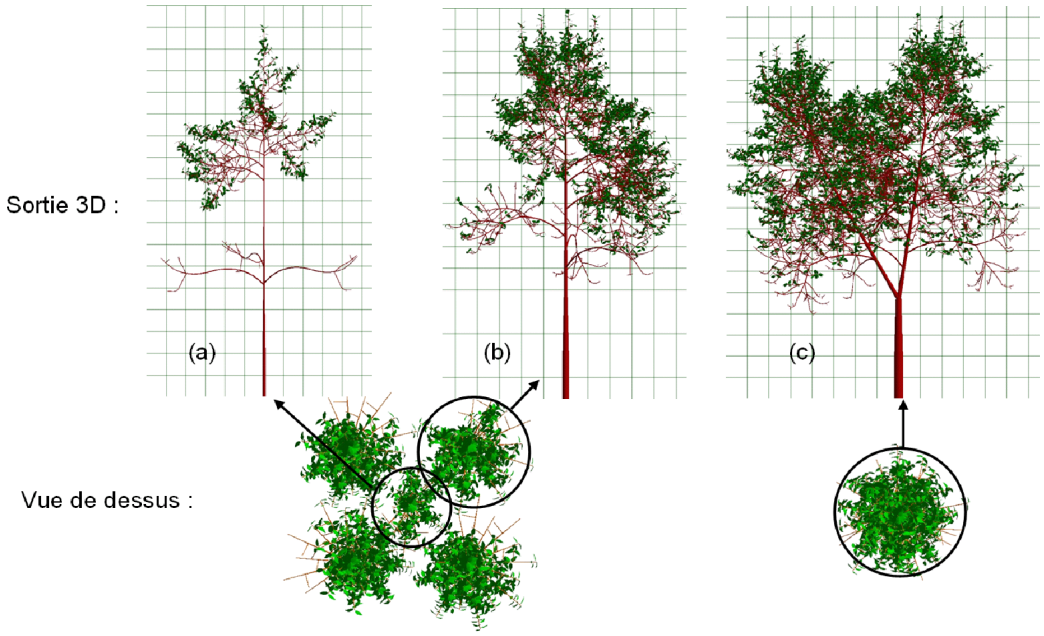


FIG. 3.4 – *Simulation of tree growth in heterogeneous conditions : view from above and detailed architectures of individuals. (a) shows a tree surrounded by four neighbours in a clump severely suffering from competition ; in (b) is shwon a tree on the edge of the clump ; in (c), a tree growing in open-field like conditions reiterates.*

We can approximate the expectation and variance of $g(x)$, respectively $E[g(x)]$ and $V[g(x)]$ thanks to the Taylor series (3.51). Since Δx is a zero-mean variable, all its odd moments are zero, and we have :

$$E[g(x)] = g(\bar{x}) + \frac{1}{2}\sum_{i,j}\frac{\partial^2 g}{\partial x_i \partial x_j}(\bar{x})\, E[\Delta x_i \Delta x_j] + E\left[\frac{1}{4!}D^4_{\Delta x} g(\bar{x}) + \cdots\right] ,$$

and we finally have the approximation, with an error at the fourth order :

$$E[g(x)] \approx g(\bar{x}) + \frac{1}{2}\sum_{i,j}\frac{\partial^2 g}{\partial x_i \partial x_j}(\bar{x})\, P_{ij} . \tag{3.52}$$

For the variance, we have :

$$Var[g(x)] = E\left[(g(x) - E[g(x)])^2\right] ,$$

with

$$\begin{aligned} g(x) - E[g(x)] = \; & D_{\Delta x} g(\bar{x}) + \frac{1}{2!}D^2_{\Delta x} g(\bar{x}) + \left(\frac{1}{3!}D^3_{\Delta x} g(\bar{x}) + \cdots\right) \\ & -\frac{1}{2}\sum_{i,j}\frac{\partial g}{\partial x_i}(\bar{x})\frac{\partial g}{\partial x_i}(\bar{x})\, P_{ij} - E\left[\frac{1}{4!}D^4_{\Delta x} g(\bar{x}) + \cdots\right] \end{aligned}$$

and since all odd moments are zero, we have the approximation with an error at the fourth order given by :

$$Var[g(x)] \approx E\left[(D_{\Delta x} g(\bar{x}))^2\right]$$

and thus

$$Var[g(x)] \approx \sum_{i,j}\frac{\partial g}{\partial x_i}(\bar{x})\frac{\partial g}{\partial x_j}(\bar{x})\, P_{ij} . \tag{3.53}$$

Remark 3.6 *Since the standard deviation is the square root of the variance and is comparable to $\bar{x}$ (same unit), the order of the approximation for the variance in 3.53 may prove unsufficient. We shall consider increasing the order of the approximation, but it requires the knowledge of the moments of higer orders for the distribution of x.*
If x is a gaussian vector $x = \mathcal{N}(\bar{x}, M)$, we shall take :

$$Var[g(x)] \approx \sum_{i,j} \frac{\partial g}{\partial x_i}(\bar{x}) \frac{\partial g}{\partial x_j}(\bar{x}) P_{ij} + C^4_{\Delta x}(\bar{x}) \tag{3.54}$$

with $C^4_{\Delta x}(\bar{x})$ a corrective term of order 4 (ensuring an error of order 3 for the corresponding approximation of the standard deviation) :

$$\begin{aligned} C^4_{\Delta x}(\bar{x}) = \quad & \frac{1}{4} \sum_{i,j,k,l} \frac{\partial^2 g}{\partial x_i \partial x_j}(\bar{x}) \frac{\partial^2 g}{\partial x_k \partial x_l}(\bar{x}) (P_{ik}P_{jl} + P_{il}P_{jk}) \\ & + \frac{1}{3} \sum_{i,j,k,l} \frac{\partial^3 g}{\partial x_i \partial x_j \partial x_k}(\bar{x}) \frac{\partial g}{\partial x_l}(\bar{x}) (P_{ij}P_{kl} + P_{ik}P_{jl} + P_{il}P_{jk}) \ . \end{aligned}$$

For example, if we take $x = \mathcal{N}\left(\begin{pmatrix} 1 \\ 2 \end{pmatrix}, \begin{pmatrix} 0.01 & 0 \\ 0 & 0.04 \end{pmatrix}\right)$ and $f(x) = x_1^2 - 3x_1x_2^3$, we get :

Monte-Carlo(10^6)	$g(\bar{x})$	Mean given by Eq. (3.52)
-23.71	-23.0	-23.71

TAB. 3.1 – Approximation of different orders for the mean

Monte-Carlo(10^6)	Variance given by Eq. (3.53)(without $C^4_{\Delta x}(\bar{x})$)	Variance given by Eq. (3.54)(with $C^4_{\Delta x}(\bar{x})$)
59.61	56.68	59.59

TAB. 3.2 – Approximation of different orders for the variance

Remark 3.7 *If g is not scalar and takes its value in $\mathbb{R}^m$, the same principle applies and we can also determine an approximation of $Cov(g_k(x), g_l(x))$:*

$$Cov(g_k(x), g_l(x)) \approx \sum_{i,j} \frac{\partial g_k}{\partial x_i}(\bar{x}) \frac{\partial g_l}{\partial x_j}(\bar{x}) P_{ij} \ .$$

3.6.2 Application to the GreenLab equation of growth

In this section, we show how to characterize the production distribution in a plant population on a simple example. We suppose that the main source of variability in the population is the seed mass, which is given by a gaussian law $\mathcal{N}(u_0, \sigma_0)$, and we want to deduce the distribution of the total biomass produced by each individual in the population after T growth cycles. For this purpose, we choose the classical production equation (3.50) :

$$u^t = E^t \beta \left(1 - \exp\left(-\eta \sum_{k=t-T_b}^{t-1} \frac{u^k}{N^k \cdot p^k} \left(\sum_{j=0}^{k+T_b-t} N_b^k(j) p_b^k(j)\right)\right)\right) \ . \tag{3.55}$$

We define $X^t \in \mathbb{R}^{T+1}$ as :

$$X^t = \begin{pmatrix} u^0 \\ u^1 \\ \vdots \\ u^t \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \tag{3.56}$$

The system evolution equation is thus written for all $n \leq T-1$:

$$X^{t+1} = F^t(X^t) \tag{3.57}$$

with $\left(F^t\right)_i$, $1 \leq i \leq T+1$, the components of F^t given by :

$$F_i^t(X) = (X)_i \quad \text{for} \quad 1 \leq i \leq t+1 \quad , \tag{3.58}$$

$$F_{t+2}^t(X) = E^{t+1}\beta\left(1 - \exp\left(-\eta \sum_{k=t-T_b+1}^{t} \frac{X_{k+1}}{N^k \cdot p^k}\left(\sum_{j=0}^{k+T_b-t-1} N_b^k(j) p_b^k(j)\right)\right)\right) \ . \tag{3.59}$$

and

$$\left(F^t\right)_i(X) = 0 \quad \text{for} \quad t+2 < i \leq T+1 \ . \tag{3.60}$$

We can apply the methodology introduced in section 3.6.1 to determine recursively an approximation of the distribution moments of X^T. We know :

$$\overline{X^0} = \begin{pmatrix} u^0 \\ 0 \\ \vdots \\ 0 \end{pmatrix} \text{ and } Cov\left(X^0\right) = \begin{pmatrix} \sigma_0^2 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & 0 \end{pmatrix}$$

and if we suppose known $\overline{X^t}$ and $Cov\left(X^t\right)$, we deduce $\overline{X^{t+1}}$ and $Cov\left(X^{t+1}\right)$ as follows.

- For the mean :
  - if $1 \leq i \leq T+1,\ i \neq t+2$:
$$\overline{X_i^{t+1}} = \overline{X_i^t} \ ,$$
  - and
$$\overline{X_{t+2}^{t+1}} = F_{t+2}^t\left(\overline{X^t}\right) + \frac{1}{2}\sum_{i,j} \frac{\partial^2 F_{t+2}^t}{\partial X_i \partial X_j}\left(\overline{X^t}\right) Cov\left(X^t\right)_{ij} \ .$$
- For the covariance :
  - for $1 \leq i \leq t+1,\ 1 \leq j \leq t+1$, $Cov\left(X^{t+1}\right)_{ij} = Cov\left(X_i^{t+1}, X_j^{t+1}\right)_{ij} = Cov\left(X^t\right)_{ij}$
  - for $i > t+2$ ou $j > t+2$, $Cov\left(X^{t+1}\right)_{ij} = 0$
  - finally, for $j < t+2$:

$$Cov\left(X_{t+2}^{t+1}, X_j^{t+1}\right) = \sum_{k,l} \frac{\partial F_{t+2}^t}{\partial X_k}\left(\overline{X^t}\right) \frac{\partial F_j^t}{\partial X_l}\left(\overline{X^t}\right) Cov\left(X^t\right)_{kl} \ ,$$

and

$$Cov\left(X_j^{t+1}, X_{t+2}^{t+1}\right) = Cov\left(X_{t+2}^{t+1}, X_j^{t+1}\right) \ .$$

It is also possible to choose an approximation of higer order as in Equation (3.54). $Q(T)$, the plant total biomass at growth cycle T is given by :

$$Q(T) = \sum_{i=1}^{T+1} X_i^T \ ,$$

and we finally get

$$\overline{Q(T)} = \sum_{i=1}^{T+1} \overline{X_i^T} \ ,$$

with variance

$$Var\left(Q(T)\right) = \sum_{k,l} Cov\left(X^T\right)_{kl} \ .$$

Chapitre 4

System Identification and Control

The first interest of models is their ability to describe and reproduce (simulate) physical or biological processes and phenomena for a better understanding of these processes, Varenne (2007). When studying complex systems (like plants), modelling and simulation can also be tools to reveal emerging properties of systems. For example, Mathieu et al. (2008) showed that modelling the retroaction of growth on development induced the possible appearance of periodic patterns for branching or fructification.
When the purpose is restricted to simulation, models may be validated from a qualitative point of view only. However, for applicative objectives in agriculture or forestry, it is fundamental to validate models quantitatively by confronting them to experimental data. In order to be able to use models as predictive tools and to implement control techniques, it is necessary to determine the endogeneous parameters of plant species and the parameters ruling the interactions of plants with their environment. In other words, the genotype-environment interaction has to be fully characterized. If the ambition is quite clear, the scientific task is of great difficulty.
The original approach used in quantitative genetics to link phenotypic traits to genotypes was based on statistical analysis. Regarding plants, there is now a general agreement on the potential interest of mechanistic models of plant growth in interaction with the environement as an intermediate step between genetics and phenotypic traits, *cf.* Hammer et al. (2002, 2006); Reymond et al. (2003); Tardieu (2003); Dingkuhn et al. (2005). If parameters of a plant growth model are shown to be stable in a proper range of environmental conditions, they can be considered as species-specific and characteristic of a genotype. In (Letort et al., 2008b), a virtual Quantitative Trait Loci detection was simulated in the case of phenotypic traits (such as crop yield) and in the case of GreenLab parameters, and was shown to be more accurate in the latter case.
In this chapter, we will first focus on the determination of the species-specific model parameters. They can be estimated from experimental data and some preliminary results were obtained on a wide variety of plants. If the same principles would apply to estimate the parameters ruling the interactions of plants with the environment, applicative studies are only at their early stages, and will thus not be fully detailed here. Finally, we will introduce the mathematical framework for optimal control of plant growth.

4.1 Parameter Estimation

Estimation of the parameters within models of dynamic system generally involves adopting a probalistic framework for model equations by taking into account process and measurement errors, *cf.* for example Goodwin and Payne (1977), Walter and Pronzato (2006), Ljung (1999). When system observations are regular, very efficient methods based on Kalman filtering have been devised. The principle is to treat observed data successively and to improve estimation at each step according to error prediction, see for example extended

Kalman filters in Ljung (1979) or recursive least squares in Ljung and Söderström (1983). However, some systems do not allow regular data acquisition, particularly living systems, as measurements are too complex. For functional plant growth models, the variables of interest are biomasses (dried or fresh matter). The system output is multivariate and generally corresponds to the masses of some elementary units of the plant (for example all organs, see Guo et al. (2006)). Measurements are thus destructive, they can be done only once for a given plant.

4.1.1 Estimation of Discrete Dynamic Systems

In this section, we recall some useful results on the estimation of discrete dynamic systems and introduce the notations for our study. We refer for example to Goodwin and Payne (1977) for a comprehensive presentation of the subject.
We consider a dynamic system described by a discrete time model. Let $(t_n)_{n\in[0;N]}$ be the finite sequence of succesive times corresponding to the evolution steps. We denote by $X_n \in \mathbb{R}^x$ the set of characteristic system variables at t_n, $U_n \in \mathbb{R}^u$, the set of exogeneous variables (entries, controls...) at t_n, and $P \in \mathbb{R}^p$, the vector of model parameters. The type of models addressed in this paper can be written in the following form :

$$X_{n+1} = F_n(X_n, U_n, P) \tag{4.1}$$

with X_0 given.
Let $Y \in \mathbb{R}^y$ be a vector of experimental observations made on the system at a given time t_N. These observations correspond to model outputs $\tilde{Y}$ deduced from X_N :

$$\tilde{Y} = G(X_N, P) \ . \tag{4.2}$$

If the initial state X_0 and $(U_n)_n$ are known, $\tilde{Y}$ is (implicitly) a function of P. Model errors are represented by the random vector ϵ :

$$\epsilon(P) = Y - \tilde{Y}(P) \ . \tag{4.3}$$

We choose to use the maximum likelihood estimator since it often has some optimum properties, like asymptotic efficiency (see for example Cramer (1946)). The likelihood of Y as a function of P, $\Pi(Y|P)$, is the probability density of Y when P is the vector of model parameters. Identification of the system parameters can be performed by finding $\hat{P}$ that maximizes $\Pi(Y|P)$.
We suppose that ϵ is a gaussian vector of covariance matrix Σ. Since the model is deterministic, $\Pi(Y|P) = \Pi\left(\epsilon(P)\right)$ and we have :

$$\Pi(Y|P) = \left[(2\pi)^y \mathrm{det}(\Sigma)\right]^{-1/2} \exp\left[-\frac{1}{2}\left(Y-\tilde{Y}(P)\right)^T \Sigma^{-1}\left(Y-\tilde{Y}(P)\right)\right] \ . \tag{4.4}$$

If Σ is known, $\hat{P}$ is given by the Gauss-Markov estimator, that is to say by minimizing the quadratic criteria :

$$\hat{P} = Argmin_P\left(\left(Y-\tilde{Y}(P)\right)^T \Sigma^{-1}\left(Y-\tilde{Y}(P)\right)\right) \ . \tag{4.5}$$

However, the covariance matrix Σ is generally unknown. In such cases, estimators of Σ and $\hat{P}$ can be derived if we have a large number of repetitions of the experimental observations. It is not the type of problem addressed here, since for plants, repetitions cause other problems due to the strong variability between individuals. This issue will be tackled in further studies.
The classical way to overcome this problem is thus to model the error covariance. The simplest possibility is to suppose that output errors are independent and homoscedastic, that is to say $\Sigma = \sigma^2 I$, where I is the identity matrix. The maximum likelihood estimator is thus equivalent to the ordinary least-square estimator. Other types of model for the error covariance matrix are possible.

4.1.2 Application to the GreenLab model

Houllier (1999) first proposed a methology for the estimation of GreenLab parameters. It was also briefly presented in Zhan et al. (2003). We present in this section an adapations of this method.
The fundamental GreenLab equation (3.3) is in the form of a classical discrete dynamic system (4.1). Depending on the particular model chosen, the set of parameters may vary. Some parameters can be directly measured on the field (organ expansion and life spans, specific blade mass ...) and the others are to estimate from experimental data. Since organ masses can be measured, the observation vector Y correspond to the state variables (or part of them). When available, the ring profiles for some specific growth units are also interesting data to estimate the secondary growth parameters for trees, *cf.* Letort et al. (2008a).

Heteroscedasticity

The error covariance matrix is not known and the homoscedastic hypothesis is not a very good option for the GreenLab model since the masses in Y corresponds to different types of organs with potentially very different size orders (*e.g.* petioles of a few grammes in sugar beet versus the whole root system of 250 g, *cf.* Lemaire et al. (2008)).
We thus consider a more general heteroscedastic hypothesis, corresponding to an error model studied in details by Taylor (1977) : the y observation are classified into q groups, with each of the y_i error terms in group i has common unknown variance θ_i ($1 \leq i \leq q$) and errors supposed mutually independent. For the observation vector on plants, each group corresponds to a type of organ. We suppose :

$$\epsilon(P) = Y - \tilde{Y}(P) = \mathcal{N}(0; \Sigma) \quad .$$

with Σ a diagonal matrix of rank y :

$$\Sigma = \begin{pmatrix} \theta_1 I_{y_1} & 0 & 0 & \cdots & 0 \\ 0 & \theta_2 I_{y_2} & 0 & \cdots & 0 \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ 0 & \cdots & 0 & \theta_{q-1} I_{y_{q-1}} & 0 \\ 0 & \cdots & 0 & 0 & \theta_q I_{y_q} \end{pmatrix} \tag{4.6}$$

with I_k, the identity matrix of order k.

Two stage Aitken estimator

For such type of sytems, a two stage Aitken estimator can be used (*cf.* Taylor (1977)). It will be denoted $\hat{P}_{2SA}$. We first find estimates $\hat{\theta}_i$ for all θ_i and then use the Gauss-Markov estimator for P (*cf.* (4.5)) with an estimated covariance matrix $\hat{\Sigma}$:

$$\hat{P}_{2SA} = Argmin_P \left(\left(Y - \tilde{Y}(P) \right)^T \hat{\Sigma}^{-1} \left(Y - \tilde{Y}(P) \right) \right) \quad . \tag{4.7}$$

If $y_i - p \geq 2$, a usual choice for the estimation of θ_i is given by :

$$\hat{\theta}_i = \frac{1}{y_i - p} \left(Y_i - \tilde{Y}_i \left(\hat{P}_{LS,i} \right) \right)^T \left(Y_i - \tilde{Y}_i \left(\hat{P}_{LS,i} \right) \right) \tag{4.8}$$

with $\hat{P}_{LS,i}$ the least-square estimator of P on the i-th sub-sample (the i-th group), *i.e.* :

$$\hat{P}_{LS,i} = Argmin_P \left(\left(Y_i - \tilde{Y}_i(P) \right)^T \left(Y_i - \tilde{Y}_i(P) \right) \right)$$

In the linear case, the properties of this estimator are studied by Taylor (1977).

Remark 4.1 *The negative log-likelihood denoted by $L(P,\theta)$ is :*

$$L(P,\theta) = \frac{y}{2}2\pi + \sum_{i=1}^{q} \frac{y_i}{2} ln(\theta_i) + \frac{1}{2}\sum_{i=1}^{q} \frac{1}{\theta_i}\left(\left(Y_i - \tilde{Y}_i(P)\right)^T \Sigma^{-1}\left(Y_i - \tilde{Y}_i(P)\right)\right)$$

so that the maximum-likelihood estimators for P and θ (denoted by $\hat{P}$ and $\hat{\theta}$) are :

$$\hat{\theta}_i = \frac{1}{y_i}\left(Y_i - \tilde{Y}_i(P)\right)^T \left(Y_i - \tilde{Y}_i(P)\right), \text{ for } 1 \leq i \leq q, \tag{4.9}$$

and

$$\hat{P} = Argmin_P \left(\sum_{i=1}^{q} \frac{1}{\theta_i}\left(Y_i - \tilde{Y}_i(P)\right)^T \left(Y_i - \tilde{Y}_i(P)\right)\right) . \tag{4.10}$$

$\hat{P}$ is thus solution of the implicit p-dimensional system of equations :

$$\sum_{i=1}^{q} \frac{y_i}{\left(Y_i - \tilde{Y}_i(\hat{P})\right)^T \left(Y_i - \tilde{Y}_i(\hat{P})\right)} \left(\frac{\partial \tilde{Y}}{\partial P}\left(\hat{P}\right)\right)^T \left(Y_i - \tilde{Y}_i(\hat{P})\right) = 0 .$$

Such equation is generally intractable, and we'd rather solve iteratively Equations (4.9, 4.10) (as suggested by Houllier (1999)), starting with

$$\hat{\theta}_i = \frac{1}{y_i}\left(Y_i - \tilde{Y}_i\left(\hat{P}_{LS,i}\right)\right)^T \left(Y_i - \tilde{Y}_i\left(\hat{P}_{LS,i}\right)\right)$$

then deducing $\hat{P}$ and again correcting $\hat{\theta}$ with this new version of $\hat{P}$, until this procedure stabilizes. Numerical tests however showed that there is no significant difference compared with the $2SA$ estimator given by (4.7,4.8).

Remark 4.2 *In the linear case, that is to say if there exist matrices $A_i \in \mathcal{M}_{y_i p}(\mathbb{R})$ for all $1 \leq i \leq q$ such that :*

$$\begin{pmatrix} Y_1 \\ \vdots \\ Y_i \\ \vdots \\ Y_q \end{pmatrix} = \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_q \end{pmatrix} P + \begin{pmatrix} \epsilon_1 \\ \vdots \\ \epsilon_i \\ \vdots \\ \epsilon_q \end{pmatrix} , \tag{4.11}$$

Equation (4.7) can be solved explicitly :

$$\hat{P}_{2SA} = \left(A^T \hat{\Sigma}^{-1} A\right)^{-1} A \hat{\Sigma}^{-1}\left(Y - \tilde{Y}(P)\right) ,$$

with

$$A = \begin{pmatrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_q \end{pmatrix} .$$

Such formulation is interesting in order to solve numerically Equation (4.7) in the non-linear case. An iterative procedure can be derived. Given $\hat{P}_n$, we deduce $\hat{P}_{n+1}$ by a linear approximation near $\hat{P}_n$. We have :

$$J(P) = \left(Y - \tilde{Y}(P)\right)^T \hat{\Sigma}^{-1}\left(Y - \tilde{Y}(P)\right)$$

and we define $G_n : \mathbb{R}^p \to \mathbb{R}$ *:*

$$G_n(\beta) = J\left(\hat{P}_n + \beta\right) \approx \left(Y - \tilde{Y}\left(\hat{P}_n\right) - \frac{\partial \tilde{Y}}{\partial P}\left(\hat{P}_n\right)\beta\right)^T \hat{\Sigma}^{-1}\left(Y - \tilde{Y}\left(\hat{P}_n\right) - \frac{\partial \tilde{Y}}{\partial P}\left(\hat{P}_n\right)\beta\right)$$

and

$$\left(\frac{\partial G_n}{\partial \beta}\right)(\beta) = -2\left(Y - \tilde{Y}\left(\hat{P}_n\right) - \frac{\partial \tilde{Y}}{\partial P}\left(\hat{P}_n\right)\beta\right)^T \hat{\Sigma}^{-1}\frac{\partial \tilde{Y}}{\partial P}\left(\hat{P}_n\right)$$

We choose $\hat{P}_{n+1} = \hat{P}_n + \beta_n$ *where* β_n *satisfies* $\left(\frac{\partial G_n}{\partial \beta}\right)(\beta_n) = 0$ *that is to say*

$$\beta_n = \left(\left(\frac{\partial \tilde{Y}}{\partial P}\left(\hat{P}_n\right)\right)^T \hat{\Sigma}^{-1}\left(\frac{\partial \tilde{Y}}{\partial P}\left(\hat{P}_n\right)\right)\right)^{-1}\left(\frac{\partial \tilde{Y}}{\partial P}\left(\hat{P}_n\right)\right)^T \hat{\Sigma}^{-1}\left(Y - \tilde{Y}\left(\hat{P}_n\right)\right) \ .$$

This iterative method to solve the minimization problem given in Equation (4.7) is called Gauss-Newton method, cf. Walter and Pronzato (2006). Numerical tests for the estimation of the GreenLab parameters seem to show that this method is more efficient than quasi-Newton minimization methods.

Estimator variance

In the linear case (*cf.* Eq. (4.11)), the variance of the Gauss-Markov estimator for P is :

$$Var\left(\hat{P}\right) = \left(A^T \Sigma^{-1} A\right)^{-1} \ , \tag{4.12}$$

it corresponds to the Cramer-Rao bound since the Gauss-Markov estimator provides the best estimate for P. When Σ is diagonanl and in the case of the 2-stage Aitken estimator, that is to say when Σ is estimated for using the Gauss-Markov estimator, it is proved by Taylor (1977) that to order $\left(\frac{1}{y}\right)$, the variance of $\hat{P}_{2SA}$ is that of the Gauss-Markov estimator in Eq. (4.12). A correction of the variance to order $\left(\frac{1}{y^2}\right)$ is given by Taylor (1977).
For the non-linear GreenLab model, Equation (4.12) was shown to be a very good approximation for $Var\left(\hat{P}_{2SA}\right)$ by Monte-Carlo simulations.

Multi-Fitting

In order to take into account the dynamics of plant growth, it is interesting to consider an observation vector composed of several intermediate observations at different growth stages $(\tau_i)_{1 \le i \le k}$, that is to say :

$$Y = \left(Y^{\tau_1}, \cdots, Y^{\tau_k}\right)^T$$

and corresponding model outputs

$$\tilde{Y} = \left(G\left(X^{\tau_1}, P\right), \cdots, G\left(X^{\tau_k}, P\right)\right)^T$$

with the observation function G defined in Eq. (4.2). The same heteroscedasticity hypothesis can be done, considering a constant error variance for each group of organs across time or for each group of organs at each observation time.
In GreenLab vocabulary, this method is called *multi-fitting*. However, since measurements are destructive, observations are done on different plants, that are supposed representative of the same dynamic system, with the same parameters. This hypothesis is practically

never true due to the high variability in plant populations at field level. Since taking into account this variability in an appropriate population growth model is only at its early stage, a practical method is first to define an average plant or an average of plants at each observation time (*cf.* Lemaire et al. (2008)) and to use multi-fitting on this average.

This methodology for parametric identification was implemented in $Digiplant$ software, and used on several plants : Maize (Ma et al., 2008), Arabidopsis, (Letort et al., 2007a), (Christophe et al., 2008), Pine saplings Guo et al. (2007), Cucumber (Mathieu et al., 2007b), Sugar-Beet (Lemaire et al., 2008), Cotton (Zhan et al., 2007), Rapeseed (Jullien et al., 2007), Beech (Letort et al., 2007b), Chrysanthemum (Kang and de Reffye, 2007), Sunflower (unpublished data, (de Reffye et al., 2006)), Grapevine (Pallas et al., 2008a)...

4.1.3 An example on experimental data of sugar-beet

In this section, we report the results of the parametric identification of the GreenLab model applied to sugar beet ($Radar$ variety). This study was conducted at the French research institute for sugar beet (ITB) and the results were presented by Lemaire et al. (2008).

Model specificities for sugar beet

We recall the fundamental equation (3.3) :

$$q^{t+1} = Aq^t + Ap^t \frac{u^t}{N^t \cdot p^t} \ . \tag{4.13}$$

- We have only one physiological age, and a mono-stem plant. The number of organs N^t is given, according to the observed phyllochron (with a decay in the leaf rhythm of appearance at full-cover)
- The organs concerned are blades, petioles, and the root system considered as an organ (*cf.* equations (3.10), (3.11)).
- A constant specific blade mass e is assumed, *cf.* Eq. (3.13)
- The input variable u^t is a feedback function of the state variables, as in Eq. (3.27) :

$$Q^t = PAR^t S_p \mu \left(1 - \exp\left(-k\mathcal{A}^t/S_p\right)\right) \tag{4.14}$$

  with
  - PAR^t the photosynthetically active radiation received during growth cycle t,
  - S_p an empirical coefficient corresponding to a characteristic surface (related to the two-dimensional projection of space potentially occupied by the plant onto the x-y plane)
  - μ is an empirical coefficient related to the Radiation Use Efficiency
  - k is the Beer-Lambert extinction coefficient
  - $\mathcal{A}^t$ is the total green leaf area
- In sugar beet, leaf life spans and expansion times vary from one leaf to the other. We therefore choose p_o^t as in Eq. (3.16). For all n, $T_b^i(n) = T_p^i(n) = 0$ (no expansion delay), but $T_b^f(n)$ and $T_b(n)$ vary with n (we observe experimentally that $T_b^f(n) = T_p^f(n)$ and $T_b(n) = T_p(n)$).
- Moreover, we choose the sink of petioles as a function of the state variables ($P_p(t)$ is the global sink value of a petiole initiated at growth cycle t) :

$$P_p(t) \ = \ P'_p \ + \ P''_p \, I_c(t)$$

  where $I_c(t)$ denotes a competition index at growth cycle t given by the ratio of leaf interception surface area to total leaf area :

$$I_c(t) = \frac{S_p\left(1 - e^{-k\mathcal{A}^t/S_p}\right)}{\mathcal{A}^t}$$

Parametric identification

Some parameters can be directly observed experimentally : e, $T_b(n)$, $T_p(n)$, $T_b^f(n)$, $T_p^f(n)$, T_r, k and some parameters are fixed $P_b = 1$ (since sinks are relative to each other, one of them has to be fixed for unicity) and $P_r = 400$ (since it appears that the fitting quality is not influenced by the root sink provided this one is big enough). The vector of hidden parameters is thus :

$$P = \left(\mu, S_p, P'_p, P''_p, \alpha_b, \beta_b, \alpha_p, \beta_p, \alpha_r, \beta_r\right) \ .$$

Concerning the experimental data from which the model parameters are estimated, biomass measurements were carried out at seven different stages. At each date, seven individuals plants were selected and the dry mass of every individual organ (blades, petioles and root storage) was measured. The average of the seven plants was constructed.

The numerical values of the parameters as well as their variances are given in (Lemaire et al., 2008) and are not recalled here. As an illustration, Figure 4.1 shows the fitting curves.

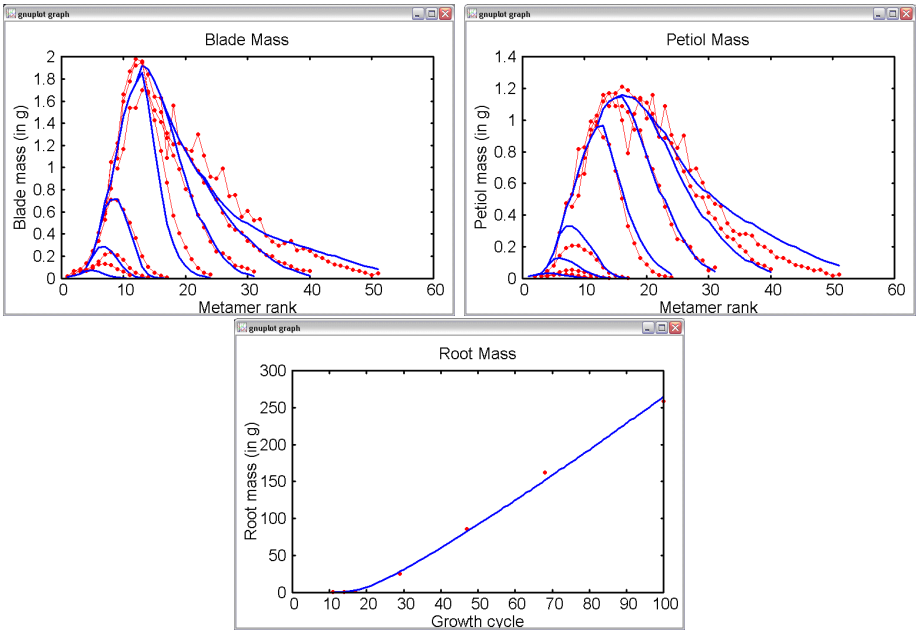


FIG. 4.1 – *Fitting results : Average plants at organ level : experimental data (in red) and simulated data (lines) (a) Blade dry masses and (b) petiole dry masses according to phytomer ranks, at 7 different growth stages ; (c) root dry mass as a function of the growth cycle (with experimental measures at the same 7 stages)*

If the goodness of fit is excellent, the statistical analysis of the estimation results is not properly done yet. The variances of the estimates are given under the assumption that the error model chosen is the good one. But the hypothesis of Σ diagonal, with masses of the same organ type having common variance θ_i $(1 \leq i \leq q)$ and errors supposed mutually independentand, *cf.* (4.6), is quite gross. It corresponds to using only the deterministic part of the dynamical model, which is really restrictive since in dynamic models, errors propagate (see an example on plant growth models in Schabenberger and Pierce (2001)). Current studies aim at improving estimation by modelling the dynamics of the covariance matrix of the model noise.

4.2 Optimization and control

The fundamental GreenLab equations can be written in the general form of a dynamic system :

$$q^{t+1} = f^t(q^t, P, c^t) \tag{4.15}$$

with the initial condition

$$q^0 = 0 \ .$$

q^t is the vector of state variables at growth cycle t (corresponding to the masses of organs), P is the vector of parameters, supposed species-specific, and c^t is the vector of exogeneous variables at growth cycle t. c^t can be decomposed into c_1^t, the exogeneous variables that can not be controlled (like temperature or radiation for field crops), and c_2^t, the control variables (that we can impose on the system, like water or nitrogen supplies, pruning). For applications to agronomy or forestry, we may want to optimize a function of the state variables and control variables $J(q, c_2)$. Let T denote the yield growth cycle. J may be the sum of an integral criteria related to the control cost and a final criteria related to the yield. We are mainly confronted to two types of optimization problems, direct optimization for the design of ideotypes and optimal control to design optimal cultivation modes.

Ideotype design:

In given environmental conditions (or in a given range of environmental conditions), that is to say for given c, what is the best set of parameters P to optimize $J(q, c_2)$. Since c_2 is supposed given, we can write the state variables as functions of P, and we are confronted to a direct optimization problem that can be solved classically :

$$\text{Find } P^* = Argmin\left(\tilde{J}(P)\right).$$

The main difficulties are the potential non-convexity of the criteria and the dimension of the search space. In such cases, we may resort to heuristic optimization techniques (such as evolutionary algorithms or simulated annealing, *cf.* for example Ratschek and Rokne (1988)). Letort (2008) gives an example for the determination of an ideotype for wheat, and show how to theoretically design this ideotype through a simple genetic model.

Optimal control:

For a given set of parameters P, and imposed exogeneous variables $\left(c_1^t\right)_t$, we aim at determining the best cultivation modes (among the acceptable ones), that is to say the best $\left(c_2^t\right)_{1\leq t\leq T-1} \in \Omega_c$ to minimize $J(q, c_2)$ (where Ω_c corresponds to the set of acceptable controls).
In the following, in order to simplify the notations, we will no longer distinguish c_1 and c_2 and consider the exogeneous variables $\left(c^t\right)_t$ as the control vector. Since P is given, we consider that the initial condition is now given at growth cycle 1, $q^1 = q_1$, with q_1 depending on the seed biomass and on P. With the initial condition q_1 and state equation (4.15), $\left(q^t\right)_{1\leq t\leq T}$ is fully determined as a function of the input $\left(c^t\right)_{1\leq t\leq T-1}$. There exists ϕ such that $q = \phi\left(q_1, c\right)$ Therefore, the optimal control problem consists in determining $\left(c^{*t}\right)_{1\leq t\leq T-1} \in \Omega_c$ minimizing $K(c)$ where $K(c) = J\left(\phi\left(q_1, c\right), c\right)$:

$$c^* = Argmin\left(K(c)\right) \quad (\mathcal{P}) \, .$$

Note that several adaptations of this formalism are possible to take into account other optimal control problems ($e.g.$ the optimal yield date) without inducing major difficulties.

A hint for resolution

To solve the optimal control problem $(\mathcal{P})$, we may apply the variational approach and the Lagrange theory to the discrete-time case, *cf.* Lewis and Syrmos (1995). For all t, $1 \leq t \leq N-1$, Eq. (4.15) :

$$q^{t+1} = f^t(q^t, c^t) \tag{4.16}$$

is a constraint to which we associate a Lagrange multiplier λ^{t+1}. We thus define the augmented criteria $\tilde{J}$ taking into account the constraints :

$$\tilde{J}(q, c, \lambda) = h\left(q^N\right) + \sum_{t=1}^{N-1} \left[g^t\left(q^t, c^t\right) + \left(\lambda^{t+1}\right)^T \left(f^t(q^t, c^t) - q^{t+1} \right) \right] . \tag{4.17}$$

We define the sequence of Hamiltonian functions for all $i \in [1; N-1]$ as :

$$H^t\left(x_q, x_c, x_\lambda\right) = g^t\left(x_q, x_c\right) + x_\lambda^T f^t(x_q, x_c)$$

and we get :

$$\tilde{J}(q, c, \lambda) = h\left(q^N\right) + \sum_{t=1}^{N-1} \left[H^t\left(q^t, c^t, \lambda^{t+1}\right) - \left(\lambda^{t+1}\right)^T q^{t+1} \right] .$$

If Ω_c is an open set, we know from the Lagrange-multiplier theory that all increments in J' due to increments in all the variables q^t, λ^t and c^t should be zero at a constrained minimum. We have

$$\begin{aligned} d\tilde{J} = \; & \left(\nabla h\left(q^N\right)\right)^T dq^N \\ & + \sum_{t=1}^{N-1} \left[\left(\frac{\partial H^t}{\partial x_q}\left(q^t, c^t, \lambda^{t+1}\right) \right)^T dq^t + \left(\frac{\partial H^t}{\partial x_c}\left(q^t, c^t, \lambda^{t+1}\right) \right)^T dc^t + \left(\frac{\partial H^t}{\partial x_\lambda}\left(q^t, c^t, \lambda^{t+1}\right) \right)^T d\lambda^{t+1} \right] \\ & - \sum_{t=1}^{N-1} \left[\left(q^{t+1}\right)^T d\lambda^{t+1} + \left(\lambda^{t+1}\right)^T dq^{t+1} \right] . \end{aligned}$$

After regrouping with respect to the increments, we get :

$$\begin{aligned} d\tilde{J} = \; & \left(\nabla h\left(q^N\right) - \lambda^N\right)^T dq^N + \left(\frac{\partial H^1}{\partial x_q}\left(q^1, c^1, \lambda^2\right) \right)^T dq^1 \\ & + \sum_{t=2}^{N-1} \left[\left(\frac{\partial H^t}{\partial x_q}\left(q^t, c^t, \lambda^{t+1}\right) - \lambda^t \right)^T dq^t \right] \\ & + \sum_{t=1}^{N-1} \left[\left(\frac{\partial H^t}{\partial x_c}\left(q^t, c^t, \lambda^{t+1}\right) \right)^T dc^t + \left(\frac{\partial H^t}{\partial x_\lambda}\left(q^t, c^t, \lambda^{t+1}\right) - q^{t+1} \right)^T d\lambda^{t+1} \right] \end{aligned}$$

Since q^1 is given, $dq^1 = 0$. The necessary conditions for a constrained minimum are thus given by the following equations :

$$\begin{aligned} q^{t+1} &= \frac{\partial H^t}{\partial x_\lambda}\left(q^t, c^t, \lambda^{t+1}\right) \quad \text{for all } t,\ 1 \leq t \leq N-1 \\ 0 &= \frac{\partial H^t}{\partial x_c}\left(q^t, c^t, \lambda^{t+1}\right) \quad \text{for all } t,\ 1 \leq t \leq N-1 \\ \lambda^t &= \frac{\partial H^t}{\partial x_q}\left(q^t, c^t, \lambda^{t+1}\right) \quad \text{for all } t,\ 2 \leq t \leq N-1 \end{aligned}$$

and the boundary conditions

$$\begin{aligned} q^1 &= q_1 \\ \lambda^N &= \nabla h\left(q^N\right) \end{aligned}$$

Finally,

$$q^{t+1} = f^t\left(q^t, c^t\right) \quad \text{for all } t,\ 1 \leq t \leq N-1 \tag{4.18}$$

$$0 = \frac{\partial g^t}{\partial x_c}\left(q^t, c^t\right) + \lambda^{t+1}\frac{\partial f^t}{\partial x_c}\left(q^t, c^t\right) \quad \text{for all } t,\ 1 \leq t \leq N-1 \tag{4.19}$$

$$\lambda^t = \frac{\partial g^t}{\partial x_q}\left(q^t, c^t\right) + \lambda^{t+1}\frac{\partial f^t}{\partial x_q}\left(q^t, c^t\right) \quad \text{for all } t,\ 2 \leq t \leq N-1 \tag{4.20}$$

$$q^1 = q_1 \tag{4.21}$$

$$\lambda^N = \nabla h\left(q^N\right) \tag{4.22}$$

From this formulation, we also deduce $\nabla K\left(c\right)$. For all t, $1 \leq t \leq T-1$:

$$\frac{\partial K}{\partial c^t}\left(c\right) = \frac{\partial g^t}{\partial x_c}\left(q^t, c^t\right) + \left(\lambda^{t+1}\right)^T \frac{\partial f^t}{\partial x_c}\left(q^t, c^t\right) \ . \tag{4.23}$$

A numerical iterative resolution algorithm can be derived to solve the optimal control problem $(\mathcal{P})$, from an initial control $c_0 \in \Omega_c$:

1. For $c_n \in \Omega_c$, we compute $q_n = \left(q_n^t\right)_{1 \leq t \leq N}$, from the initial condition (4.21) and the the state equation (4.18).
2. Knowing q_n, we can compute the adjoint state $\lambda_n = \left(\lambda_n^t\right)_{2 \leq t \leq N}$ backwards from the boundary condition (4.22) and from the adjoint equation (4.20).
3. We compute the gradient $\nabla K\left(c_n\right)$
4. We apply a gradient-based descent method :

   $$c_{n+1} = c_n + \rho_n \nabla K\left(c_n\right)$$

   with ρ_n a well-chosen step of descent.
5. The procedure is repeated until a convergence test is satisfied.

Some adaptations have to be made in order to take into account constraints on the state vector or on the control (for example if Ω_c is not open).

A simple example on optimal pruning

We consider here a theoretical plant, with immediate expansion and a leaf life span of one growth cycle. The available biomass for allocation is given by the seed q_0 at the first growth cycle and then by leaf photosynthesis, as in Eq. (3.27). We take a constant environmental factor $E^t = 1$. The proportion of biomass allocated to leaves at each growth cycle is a constant α. In such cases, Letort (2008) showed that there is no influence of topology. Therefore, we choose as state variable q^t, the total leaf biomass at growth cycle t. The fundamental equation can be written :

$$q^{t+1} = \alpha\beta\left(1 - \exp\left(-\gamma q^t / e\right)\right)$$

that we rewrite

$$q^{t+1} = A\left(1 - \exp\left(-Bq^t\right)\right) \ .$$

with q^1 given, $q^1 = \alpha q_0$. Convergence results regardig this particular type of sequence were proved by Chauveheid (2008).

We suppose that at each growth cycle, a proportion c^t, $c^t \in [0;1]$, of leaf biomass can be pruned, and the objective is to maximize the total yield of pruned leaves, $i.e.$ we want to solve the optimal control problem :

$$\text{Find } c* = \left(c^{*t}\right)_{1\leq t\leq N-1} \in [0;1]^{N-1} \text{ minimizing } K(c) = -\sum_{t=1}^{N-1} c^t q^t - q^N \ , \qquad (4.24)$$

where q satisfies $q^1 = \alpha q_0$ and for all $t \in [1; N-1]$:

$$q^{t+1} = A\left(1 - \exp\left(-B\left(1-c^t\right)q^t\right)\right) \ .$$

The main difference with the general case is that we have included some constraints on the control variables. They can be simply taken into account in the Lagrangian (4.17) by adding the terms

$$\sum_{t=1}^{N-1} \mu^t\left(c^t - 1\right) - \nu^t c^t$$

where $\left(\mu^t\right)_{1\leq t\leq N-1}$ and $\left(\nu^t\right)_{1\leq t\leq N-1}$ in $\mathbb{R}^{N-1}$ are Karush-Kuhn-Tucker multipliers. The optimality conditions are now given by :

$$q^{t+1} = f^t\left(q^t, c^t\right) \quad \text{for all } t,\ 1 \leq t \leq N-1$$

$$\mu^t \geq 0 \text{ and } \mu^t\left(c^t - 1\right) = 0 \quad \text{for all } t,\ 1 \leq t \leq N-1$$

$$\nu^t \geq 0 \text{ and } \nu^t c^t = 0 \quad \text{for all } t,\ 1 \leq t \leq N-1$$

$$\frac{\partial g^t}{\partial x_c}\left(q^t, c^t\right) + \lambda^{t+1}\frac{\partial f^t}{\partial x_c}\left(q^t, c^t\right) + \mu^t - \nu^t = 0 \quad \text{for all } t,\ 1 \leq t \leq N-1$$

$$\lambda^t = \frac{\partial g^t}{\partial x_q}\left(q^t, c^t\right) + \lambda^{t+1}\frac{\partial f^t}{\partial x_q}\left(q^t, c^t\right) \quad \text{for all } t,\ 2 \leq t \leq N-1$$

$$q^1 = q_1$$

$$\lambda^N = \nabla h\left(q^N\right)$$

Applied to our example, these optimality conditions become :

$$q^{t+1} = A\left(1 - \exp\left(-B\left(1-c^t\right)q^t\right)\right) \quad \text{for all } t,\ 1 \leq t \leq N-1,$$

$$\mu^t \geq 0 \text{ and } \mu^t\left(c^t - 1\right) = 0 \quad \text{for all } t,\ 1 \leq t \leq N-1$$

$$\nu^t \geq 0 \text{ and } \nu^t c^t = 0 \quad \text{for all } t,\ 1 \leq t \leq N-1$$

$$-q^t - AB\lambda^{t+1}q^t\exp\left(-B\left(1-c^t\right)q^t\right) + \mu^t - \nu^t = 0 \quad \text{for all } t,\ 1 \leq t \leq N-1,$$

$$\lambda^t = -c^t + AB\lambda^{t+1}\left(1-c^t\right)\exp\left(-B\left(1-c^t\right)q^t\right) \quad \text{for all } t,\ 2 \leq t \leq N-1,$$

and the boundary conditions

$$q^1 = q_1,$$

$$\lambda^N = -1.$$

We can deduce by considering the different cases that the optimal control is given by

$$c^{*t} = \min\left(\max\left(1 - \frac{ln(AB)}{Bq^{*t}}, 0\right), 1\right),$$

where q^* is the state vector associated to c^*.
More precisely, the following situations occur :

– if $AB \leq 1$, we know (*cf.* Chauveheid (2008)) that, without control, the sequence decreases to 0. In such a case, the best strategy is to prune all the leaves straight away at growth cycle 1, even though it makes the plant die. We have :

$$c^{*1} = 1,\ c^{*t} = q^{*t} = 0,\ \forall t \geq 2$$

– if $AB > 1$, we have $c^{*t} = 0$ while $\dfrac{\ln(AB)}{Bq^{*t}} \geq 1$. Therefore :

  – If $\dfrac{\ln(AB)}{Bq_1} < 1$, we have $c^{*1} = \dfrac{\ln(AB)}{Bq^1}$, $0 < c^{*1} < 1$. We deduce

$$q^{*t} = \frac{AB-1}{B}, \forall t \geq 2$$

  and

$$c^{*t} = 1 - \frac{\ln(AB)}{AB-1}, \forall t \geq 2$$

  – If $\dfrac{\ln(AB)}{Bq_1} \geq 1$, we can show that necessarily $q_1 < q_\infty$, with q_∞ the limit of the sequence $\left(q^t\right)_t$, when $\left(c^t\right)_t = 0$ (limit that exists when $AB > 1$, *cf.* Chauveheid (2008)). In such a case, and while $c^t = 0$, $\left(q^t\right)$ is a stricly increasing sequence.

  If for all $t < N$, $q^t \leq \dfrac{\ln(AB)}{B}$, then $c^{*t} = 0$ for all $t \in [1; N-1]$.

  If there exists τ such that $q^\tau > \dfrac{\ln(AB)}{B}$ and $q^t \leq \dfrac{\ln(AB)}{B}$ for all $t < \tau$, then $c^{*t} = 0$ for all $t < \tau$, $c^{*\tau} = \dfrac{\ln(AB)}{Bq^{*\tau}}$, and we deduce

$$q^{*t} = \frac{AB-1}{B},\ \forall t > \tau$$

  and

$$c^{*t} = 1 - \frac{\ln(AB)}{AB-1},\ \forall t > \tau.$$

  This last case is illustrated in Figure 4.2.

Another example of optimal control for plant growth concerns water supply. Preliminary results can be found in (Wu, 2005), (Wu et al., 2005).

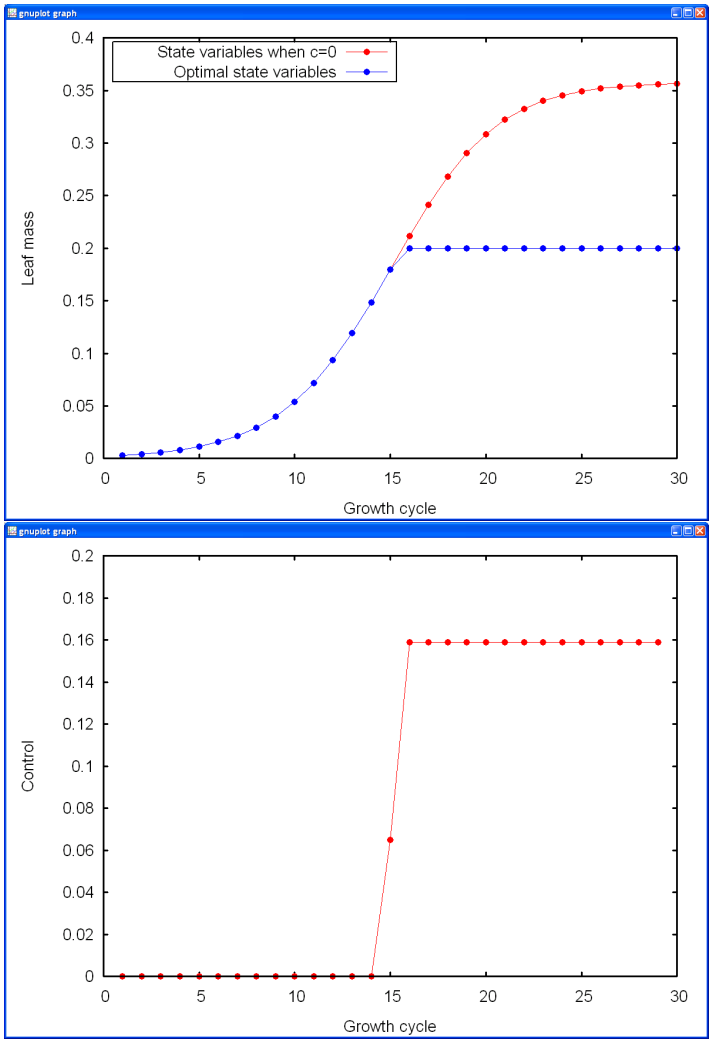


FIG. 4.2 – *Results of the optimal control problem for* $a = 0.7$, $b = 2$, $q_1 = 0.003$, $N = 30$ *: on the top graph are shown the evolutions of* $\left(q^t\right)_t$, *when* $c = 0$ *(final yield* $J(0) = 0.356$*) and when* $c = c^*$ *(final yield* $J(c^*) = 0.656$*). The below graph shows the optimal control* $\left(c^{*t}\right)_t$.

Chapitre 5

Perspectives

Dans la démarche de confrontation du modèle GreenLab aux données expérimentales sur plantes réelles, première étape vers de possibles applications, l'aspect qui a de loin été privilégié est l'étude de la généricité du modèle. Une des raisons est sans doute l'intérêt manifesté par différentes équipes de recherche en agronomie ou foresterie, la plupart travaillant sur des plantes différentes. Ainsi, les paramètres du modèle ont été estimés sur : cotonnier (mono-tige) (Zhan et al., 2007), maïs (Guo et al., 2006), tournesol (sur une variété chinoise (Guo et al., 2003) et une variété montpelliéraine (de Reffye et al., 2006)), la tomate (dans trois équipes différentes de la CAU, voir en particulier (Dong et al., 2008)), betterave (Lemaire et al., 2008), blé (Kang et al., 2008a), chrysanthème (Kang et al., 2007b), Arabidopsis (Christophe et al., 2008), concombre (Mathieu et al., 2007a), poivron, vigne (Pallas et al., 2008a), pin (Guo et al., 2007), hêtre (Letort et al., 2008a), cécropia, caféier (Mathieu, 2006), riz (également sur une variété montpelliéraine et une variété chinoise), et quelques autres ...

Dans la plupart des cas, on peut dire que l'objectif poursuivi a été l'ajustement de courbes expérimentales, c'est à dire trouver un jeu de paramètres qui, bien calés, permettent de reproduire de façon approchée des jeux de données. Il convient de préciser que ceci n'est pas une mince affaire, une dizaine de paramètres permettent généralement d'approcher des jeux de plusieurs centaines de données, correspondant à plusieurs stades de croissance. La principale conclusion de ces résultats est que les concepts mis en oeuvre dans GreenLab permettent de décrire de façon tout à fait satisfaisante les processus de production et d'allocations aux organes, tout au long de la croissance, ($i.e.$ la dynamique des interactions sources-puits) pour une large gamme de plantes.

Dans de rares cas, une étude approfondie a permis de tester la robustesse de l'identification, en étudiant une même variété dans différentes conditions environnementales, et plusieurs saisons. Dong et al. (2008) a montré à partir d'expérimentations sur tomate, comment les paramètres du modèle GreenLab calibré variaient de façon très significative dans des conditions de densité différentes. Ma et al. (2008) arrive à la même conclusion pour le maïs, même si les variations ne concernent qu'un petit nombre de paramètres. Les données maïs sont les plus riches car 6 campagnes expérimentales ont été réalisées depuis 2000 par l'équipe de Guo Yan à la China Agricultural University. La capacité prédictive du modèle a donc pu être testée, $cf.$ (Ma et al., 2007b), pour des résultats mitigés.

Ainsi, si l'approche "horizontale" des études GreenLab (c'est à dire la confrontation à un grand nombre de plantes) est très prolifique, et plutôt riche de succès, l'approche "verticale" (c'est à dire l'étude approfondie d'une variété) laisse encore des perspectives intéressantes. C'est cette deuxième voie, verticale, plus en phase avec les enjeux de la biologie intégrative, mais également porteuse du progrès nécessaire pour un jour passer aux applications, que nous souhaitons approfondir. Elle nous semble comporter quatre volets sur lesquels des progrès importants peuvent être réalisés : modélisation des interactions avec l'environnement, analyse statistique des résultats de l'estimation, analyse de la variabilité inter-plantes au

sein de la population, optimisation et contrôle optimal des cultures pour les applications.

Modélisation des interactions avec l'environnement et validation:

Comme souvent, les avantages du modèle GreenLab vont de pair avec ses points faibles. Ainsi, pour parvenir à la formulation mathématique, nous sommes passés par une simplification extrême des processus biologiques, en particulier pour l'interaction avec l'environnement : c'est le fameux "E=1", que même nos meilleurs amis agronomes ont du mal à avaler. Il paraît évident que les progrès en terme de robustesse des paramètres entre différentes saisons, et de prédiction, ne pourront être réalisés que par une prise en compte plus fine des facteurs environnementaux, en particulier des stress.
La méthodologie à employer consiste à réaliser les ajustements de données expérimentales en conditions environnementales contrastées. D'une part, un défaut dans la capacité du modèle à décrire les dynamiques de production - allocation doit conduire à l'introduction de facteurs de stress, par exemple un facteur de stress hydrique dans la fonction de production (Mailhol et al., 1997), (Casadebaig et al., 2008). Une des difficultés, commune en agronomie, est la gestion des stress couplés. La démarche classique consistant à utiliser des facteurs multiplicatifs n'est pas totalement satisfaisante. D'autre part, des différences significatives dans les paramètres estimés sur des expériences différentes peuvent conduire à construire ces paramètres comme des fonctions de facteurs de stress (qui pourraient être constantes dans une certaine gamme).
Cette démarche doit nécessairement être menée avec des biologistes et agronomes, en particulier le Laboratoire d'Ecophysiologie des Plantes sous Stress Environnementaux (LEPSE) à Montpellier. La thèse de Marlène Puech qui débute doit étudier la capacité de GreenLab a prendre en compte le stress hydrique, à partir du couplage de GreenLab avec le modèle de bilan hydrique de Pilote ($cf.$ Mailhol et al. (1997), Mailhol and Merrot (2008)). Des expérimentations sont également en cours sur l'effet du stress azoté à l'Institut Technique de la Betterave (sur betterave sucrière) et à l'UMR EGC de Grignon sur colza. Plutôt que le développement d'un modèle d'absorption-allocation pour l'azote comme pour le carbone ($cf.$ une tentative en ce sens par Drouet and Pagès (2007)), il est d'abord privilégié d'étudier les variations des paramètres fonctionnels du bilan carbone en fonction de l'indice d'azote, de façon à ne pas multiplier les paramètres intervenant dans le modèle. En ce qui concerne l'effet de la compétition pour la lumière et en particulier les effets de la densité de culture, des travaux de modélisation préliminaires ont été réalisés, (Cournède and de Reffye, 2007), (Cournède et al., 2008) et sont en cours de validation sur maïs et betterave.
L'objectif est donc de caractériser une espèce de façon aussi découplée que possible des conditions environnementales, de façon à pouvoir réellement explorer le potentiel phénotypique d'un génotype, étape préalable à l'exploration du lien entre modèle génétique et modèle de croissance. Une collaboration débutant avec l'équipe MAPI du LEPSE doit mener à l'analyse d'un grand nombre de génotypes de tournesol, en utilisant un couplage du modèle GreenLab (comme solveur sources-puits) et le modèle SUNFLO (comme modèle écophysiologique).
L'étape de mise en place de ces nouvelles potentialités du modèle doit en particulier s'accompagner d'une étude plus approfondie que ce qui a été fait jusqu'à présent du système dynamique, en particulier en ce qui concerne l'analyse de sensibilité paramétrique. En effet, l'augmentation du nombre de paramètres ne doit pas se traduire par une perte de significativité du modèle. Des méthodes globales d'analyse de sensibilité paramétrique ($cf.$ Cariboni et al. (2007)) et les techniques de réduction de modèle (Galerkin proper orthogonal decomposition, $cf.$ Volkwein (2008)) sont particulièrement adaptées aux systèmes avec un fort niveau d'interaction et doivent permettre d'améliorer la paramétrisation du modèle.
Dans la même optique de validation du modèle, il est important d'approfondir la comparaison de GreenLab avec d'autres modèles de type similaire, testant leurs comportements, leurs sensibilités, les confrontant à un même jeu de données. Nous ne rappelons pas ici l'ensemble des modèles structure-fonction avec lesquels GreenLab partage quelques simi-

larités, et nous renvoyons par exemple à Sievänen et al. (2000), Prusinkiewicz (2004), Fourcaud et al. (2008) pour des revues de littérature assez générales. Malgré tout, en ce qui concerne les arbres, nous pouvons citer le modèle LIGNUM (Perttunen et al. (1996), Perttunen et al. (2005)), assez proche de GreenLab dans son calcul de la production et de l'allocation, avec une description simplifiée de l'architecture. Dans cet objectif de comparaison, il nous semble intéressant de bénéficier de méthodes unificatrices comme celles découlant du language XL basé sur les "Relational Growth Grammars" (grammaires de croissance relationnelles) introduites par Kniemeyer et al. (2004), Kurth et al. (2005) et implémenté dans la plateforme GroIMP (Kniemeyer et al., 2007) : le modèle GreenLab pourrait être implémenté dans cette plateforme, comme LIGNUM l'a déjà été, de façon à analyser les différences de comportement des deux modèles, et voir dans quelle mesure ils peuvent s'enrichir l'un l'autre.

Analyse statistique de l'estimation:

Au-delà de l'étape d'ajustement, c'est à dire de reproduction de données expérimentales, nous devons parvenir à la véritable identification paramétrique du système. Des efforts ont déjà été effectués en ce sens, (*cf.* Chapitre 4) et la mise en place de l'estimateur de Aitken. Cependant, l'hypothèse d'indépendance des erreurs qu'il utilise est grossière. Cela correspond à ne pas prendre en compte la dynamique du modèle, ce qui est très restrictif, puisque dans les modèles dynamique, les erreurs se propagent, voir par exemple Schabenberger and Pierce (2001) pour un exemple en écologie. L'objectif est donc d'améliorer l'estimateur du maximum de vraisemblance en modélisant la dynamique de la matrice de variance des bruits de modélisation. Ceci doit conduire à un meilleur calcul des intervalles de confiance pour l'estimation. La question se pose de façon encore plus complexe dans le cas d'un modèle d'organogenèse stochastique dont on souhaite identifier les paramètres à partir des sorties fonctionnelles de la plante (les biomasses).

D'autre part, Quach et al. (2007) propose de nouvelles méthodes d'estimation développées dans le cadre d'applications à la biologie, en particulier pour les réseaux de gène. Elles sont spécifiquement adaptées à de faibles jeux de données, à faible fréquence, faible répétitions. La méthode est bayesienne, avec l'utilisation de l'"Unscented Kalman Filter". Elle semble pouvoir s'adapter au cadre du multi-fitting présenté au Chapitre 4), et sera donc testée et comparée avec les méthodes dérivées de l'estimation du maximum de vraisemblance.

Après ce premier point permettant d'améliorer la modélisation de l'interaction avec l'environnement, cette deuxième étape permettant de caractériser statistiquement les résultats de l'estimation, est bien sûr fondamentale pour l'élaboration d'un modèle robuste et prédictif : les résultats de la prédiction s'accompagnent également d'indications sur leurs précisions, par l'application de la théorie des incertitudes.

Une fois la qualité de l'estimation finement caractérisée, il devient également possible d'étudier la simplification des protocoles expérimentaux pour l'acquisition des données. Celle-ci est fastidieuse puisque elle implique un important suivi pour la détermination du schéma phénologique de la plante et les mesures de biomasse au niveau des organes. L'intérêt de la simplification du protocole est particulièrement important pour deux objectifs.

D'une part, dans une optique de sélection ou caractérisation variétale, car on peut être amené à tester un grand nombre de variétés (par exemple dans l'étude en cours sur tournesol dans l'équipe MAPI à Montpellier), et car l'objectif de la modélisation est justement par le biais de l'expérimentation virtuelle de faire gagner du temps d'expérimentation. On voit bien que si le protocole pour calibrer le modèle est trop lourd, cet objectif ne pourra être atteint. Des essais ont pu être réalisés sur les données du maïs par Ma Yun Tao (Ma et al., 2007a) sur la combinaison de mesures intermédiaires par compartiments et des mesures échantillonnées pour les biomasses des organes.

L'autre domaine dans lequel la simplification du protocole est cruciale est l'application de GreenLab aux arbres, pour lesquels le nombre d'organes rend inenvisageable (en conditions

normales) des mesures exhaustives. Ce protocole a déjà pu être simplifié par Véronique Letort pour les arbres (*cf.* (Letort, 2008), (Letort et al., 2008a)) en combinant un échantillonnage et des mesures cumulées pour les branches.
Si les méthodes de simplification déjà testées donnent de bons résultats pour les ajustements, l'impact sur la perte de précision quant à l'estimation des paramètres n'a pu être réellement quantifié. Cet objectif sera donc poursuivi dans le cadre de la théorie de la planification d'expériences optimales (*cf.* par exemple Walter and Pronzato (2006)).

Analyse de la variabilité inter-plantes:

Si l'étude de la plante individuelle joue un rôle fondamental dans la compréhension des phénomènes biologiques de la croissance, dans une optique applicative, il convient d'étudier le passage au peuplement. Ceci implique d'une part de développer la modélisation de la compétition (*cf.* Cournède and de Reffye (2007), Cournède et al. (2008)), mais également de caractériser au sein du peuplement la variabilité inter-plantes à partir du modèle de croissance individu-centré. C'est ce point qui nous intéresse particulièrement. Comme décrit dans la Section 3.6, la population de plantes est décrite par un vecteur aléatoire de paramètres endogènes (pour la variabilité génétique, *cf.* Fig. 5.1) et de variables exogènes (pour la variabilité environnementale). Chaque plante de la population est une réalisation de ce vecteur aléatoire. Comme présenté dans le paragraphe 3.6.2, la théorie des incertitudes

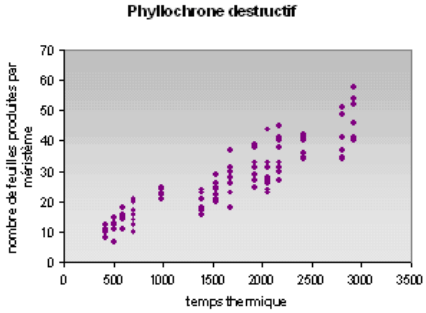


FIG. 5.1 – Dispersion du nombre de feuilles produites par les méristèmes d'un peuplement de betteraves en fonction du temps thermique (données Institut Technique de la Betterave, Sébastien Lemaire et Fabienne Maupas) : le phyllochrone présente une forte variabilité inter-plantes.

nous permet de déterminer la dispersion des sorties du modèle au niveau du peuplement.
Par ailleurs, il convient de développer les méthodes d'identification permettant de caractériser les facteurs de variabilité, d'estimer moyenne et écart-type pour chacun de ces facteurs, avec leurs intervalles de confiance. Une première tentative a été ébauchée par Cournède and de Reffye (2009) pour un peuplement de betteraves afin de caractériser la variabilité du phyllochrone et des densités, avec des résultats encourageants. Encore plus que pour l'estimation simple, la caractérisation de la variabilité s'appuie sur des données expérimentales lourdes, avec un certain nombre d'individus, la mise au point de protocoles optimaux est donc primordiale.
Une des premières plantes cibles pour l'étude de la variabilité au niveau du peuplement est la betterave sucrière, en partenariat avec l'ITB, car la population ne correspond pas à un hybride pur, et la variabilité génétique est forte. D'autre part, autour d'Amélie Mathieu à l'UMR EGC-Grignon, nous visons également une application au colza. Une difficulté supplémentaire est que les variations locales d'environnement se traduisent par des modifications de l'architecture du colza (les ramifications se développent en fonction des conditions de croissance). Il conviendra alors de prendre en compte la rétroaction du fonctionnement sur l'architecture.

Optimisation et contrôle optimal des cultures:

Dans la section 4.2, nous avons présenté comment le modèle de croissance de plantes pouvait servir à la résolution de problèmes d'optimisation pour la sélection végétale, ou de contrôle des itinéraires culturaux. Cette démarche mérite d'être poursuivie. La thèse de Qi Rui s'intéresse en particulier à la modélisation de problèmes d'optimisation dans divers contextes de la croissance des plantes : par exemple, le couplage du modèle de croissance GreenLab avec un modèle de développement d'une colonie d'insectes parasites, et le traitement insecticide optimal. De son côté, Thomas Guillon (sous la direction de Thierry Fourcaud) s'intéresse à la stratégie d'allocation optimale de la plante face à des contraintes mécaniques. Zhongping Li, qui travaille au couplage de GreenLab à un modèle de bilan hydrique, doit reprendre en détails le sujet de l'optimisation de l'irrigation sur lequel des résultats préliminaires ont été obtenus par Wu et al. (2005).
Si la modélisation des problèmes d'optimisation dans le contexte de la croissance des plantes est déjà un problème en soi, il convient bien sûr de développer les méthodes numériques de résolution de ces problèmes. Si les méthodes de contrôle optimal sont déjà bien rôdées dans d'autres domaines, il reste à les adapter aux cas rencontrées, en particulier les méthodes d'optimisation dans l'incertain, quand on ne connait que des scénarios climatiques moyens et leurs déviations.

Ces voies de recherche doivent nous permettre pour une variété donnée, de mettre en place un protocole expérimental optimisé, afin d'estimer les paramètres endogènes de la plante caractérisant sa dynamique sources-puits, son interaction avec l'environnement et sa réponse aux effets du stress, ainsi que l'incertitude sur cette estimation, pour une plante individuelle ou au niveau du peuplement, et de répondre à des questions précises quant à cette variété, questions biologiques sur le fonctionnement de la plante, ou agronomiques sur des prévisions de rendement en fonction de scénarios climatiques ou sur le contrôle optimal d'itinéraires culturaux.
Le développement d'outils de gestion des parcelles pour les agriculteurs (comme Visioplaine, *cf.* www.visioplaine.com) est un support intéressant pour la mise en application des modèles de cultures (ou plus généralement de croissance de plantes), car ils rassemblent les données d'entrées nécessaires au modèle (données climatiques, sol, culturales au niveau de la parcelle). Ils peuvent également, en lien avec l'imagerie satellitaire, ou par des relevés terrains en cours de culture, permettre de faire du recalage de modèle grâce aux techniques d'assimilation de données, dont l'application aux cultures est également un sujet intéressant.
En conclusion, cette thématique de recherche dans le cadre général de la modélisation du système plante bénéficie aujourd'hui d'une convergence intéressante entre la biologie intégrative et les mathématiques des systèmes complexes. Ce que nous proposons de mettre en oeuvre, c'est une consolidation de la démarche de modélisation, en approfondissant l'analyse de la robustesse et de la validité du modèle, dans un esprit de synthèse et d'intégration des connaissances biologiques et avec un espoir d'ingénierie potentielle pour des applications aux cultures.

Références

M. Allen, P. Prusinkiewicz, and T. Dejong. Using L-systems for modeling source-sink interactions, architecture and physiology of growing trees, the L-peach model. *New Phytologist*, 166 :869–880, 2005.

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

K. Athreya and P. Ney. *Branching Processes*. Dover Publications, 2004.

J.-F. Barczi, H. Rey, Y. Caraglio, P. de Reffye, D. Barthélémy, Q. Dong, and T. Fourcaud. Amapsim : a structural whole-plant simulator based on botanical knowledge and designed to host external functional models. *Annals of Botany*, 101(8), 2008.

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

D. Barthélémy and Y. Caraglio. Plant architecture : a dynamic, multilevel and comprehensive approach to plant form, structure and ontogeny. *Annals of Botany*, 99(3) :375–407, 2007.

P. R. Bevington and D. K. Robinson. *Data reduction and error analysis for the physical sciences*. Boston, MA : McGraw-Hill, 2003.

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

P. Casadebaig, P. Debaeke, and J. Lecoeur. Thresholds for leaf expansion and transpiration response to soil water deficit in a range of sunflower genotypes. *European Journal of Agronomy*, 28(4) :646–654, 2008.

D. Chauveheid. Formalisation et étude d'un système complexe dans la théorie des systèmes dynamiques. Master's thesis, Ecole Centrale Paris, France, 2008. Rapport de projet scientifique.

M. Chelle and B. Andrieu. Modelling the light environment of virtual crop canopies. In J. Vos, L. Marcelis, P. de Visser, P. Struik, and J. Evers, editors, *Functional-Structural Plant Modelling in Crop Production*. Springer, 2007.

M. Chelle and B. Andrieu. The nested radiosity model for the distribution of light within plant canopies. *Ecological Modelling*, 111 :75–91, 1998.

A. Christophe, V. Letort, I. Hummel, P.-H. Cournède, P. de Reffye, and J. Lecoeur. A model-based analysis of the dynamics of carbon balance at the whole plant level in *Arabidopsis Thaliana*. *Functional Plant Biology*, 35(11) :1147–1162, 2008.

R. M. Corless, G. H. Gonnet, D. E. G. Hare, D. J. Jeffrey, and D. E. Knuth. On the lambert w function. *Advances in Computational Mathematics*, 5 :329–359, 1996.

E. Costes, P. de Reffye, J. Lichou, Y. Guédon, A. Audubert, and M. Jay. Stochastic modelling of apricot growth units and branching. *Acta Horticulturae*, 313 :89–98, 1992.

P.-H. Cournède and P. de Reffye. A generalized poisson model to estimate inter-plant competition for light. In T. Fourcaud and X. Zhang, editors, *Plant growth Modeling, simulation, visualization and their Applications.* IEEE Computer Society (Los Alamitos, California), 2007.

P.-H. Cournède and P. de Reffye. Parametric identification of individual variability in plant population models. In M. Jaeger and B. Li, editors, *Plant growth Modeling, simulation, visualization and their Applications. Proceedings of PMA09*. IEEE Computer Society (Los Alamitos, California), 2009. (Submitted).

P.-H. Cournède, M.-Z. Kang, A. Mathieu, J.-F. Barczi, H.-P. Yan, B.-G. Hu, and P. de Reffye. Structural Factorization of Plants to Compute their Functional and Architectural Growth. *Simulation*, 82(7) :427–438, 2006.

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

H. Cramer. *Mathematical Methods of Statistics*. Princeton : Princeton University Press, 1946.

P. Cruiziat, H. Cochard, and T. Amégio. Hydraulic architecture of trees : main concepts and results. *Annals of Forest Science*, (59) :723–752, 2002.

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*. PhD thesis, Université Paris-Sud, Centre d'Orsay, 1979.

P. de Reffye and B. Hu. Relevant choices in botany and mathematics for building efficient dynamic plant growth models : the greenlab case. In B. Hu and M. Jaeger, editors, *Plant Growth Models and Applications*, pages 87–107. Tsinghua University Press and Springer, 2003.

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

P. de Reffye, P. Dinouard, and M. Jaeger. Basic concepts of computer plants growth simulation. In *Proceedings of NICOGRAPH'90*, pages 219–234, Tokyo, 1990.

P. de Reffye, P. Dinouard, and M. Jaeger. Modélisation et simulation de l'architecture de l'orme du japon zelkova serrata (thunb.) makino (ulmaceae) : la notion d'axe de référence. In C. Edelin, editor, *2ème Colloque International sur l'Arbre*, volume A7, pages 251–266, Montpellier, 1991a. Naturalia Monspeliensa.

P. de Reffye, E. Elguero, and E. Costes. Growth units construction in trees : a stochastic approach. *Acta Biotheoretica*, 39 :325–342, 1991b.

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

P. de Reffye, F. Houllier, F. Blaise, and T. Fourcaud. Essai sur les relations entre l'architecture d'un arbre et la grosseur de ses axes végétatifs. In *Modélisation et simulation de l'architecture des végétaux*, Sciences Update, pages 205–254. Bouchon, J., inra edition, 1997b.

P. de Reffye, F. Blaise, S. Chemouny, T. Fourcaud, and F. Houllier. Calibration of hydraulic growth model on the architecture of cotton plants. *Agronomie*, 19 :265–280, 1999.

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

P. de Reffye, H. Rey, P.-H. Cournède, M. Dingkuhn, and J. Lecoeur. Solving the inverse problem from plant architecture towards the plant functioning with greenlab model. application on sunflower. Personal Communication., 2006.

P. de Reffye, D. Barthélémy, P.-H. Cournède, and M. Jaeger. Modélisation et simulation de l'architecture et de la production végétales. In F. Hallé, editor, *Aux origines des plantes*. Fayard, 2008a.

P. de Reffye, E. Heuvelink, D. Barthélémy, and P.-H. Cournède. Plant growth models. In S. Jorgensen and B. Fath, editors, *Ecological Models. Vol. 4 of Encyclopedia of Ecology (5 volumes)*, pages 2824–2837. Elsevier (Oxford), 2008b.

M. Dingkuhn, D. Luquet, B. Quilot, and P. de Reffye. Environmental and genetic control of morphogenesis in crops : Towards models simulating phenotypic plasticity. *Australian Journal of Agricultural Research*, 56(11) :1289–1302, 2005.

Q. Dong, G. Louarn, Y. Wang, J. Barczi, and P. de Reffye. Does the structure-fonction model greenlab deal with crop phenotypic plasticity induced by plant spacing ? a case study on tomato. *Annals of Botany*, 101(8), 2008.

J.-L. Drouet and L. Pagès. GRAAL : a model of GRowth Architecture and carbon ALlocation during the vegetative phase of the whole maize plant, model description and parameterisation. *Ecological Modelling*, 165 :147–173, 2003.

J.-L. Drouet and L. Pagès. GRAAL-CN : a model of GRowth Architecture and ALlocation for Carbon and Nitrogen dynamics within whole plants formalized at the organ level. *Ecological Modelling*, 206 :231–249, 2007.

T. Fourcaud. Modélisation des relations biomécanique - morphogenèse à l'échelle de l'arbre et applications, 2004. Habilitation à diriger des recherches, Université de Bordeaux 1.

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

J. Françon. Sur la modélisation informatique de l'architecture et du développement des végétaux. In *2ème Colloque International : L'Arbre. Institut de Botanique, Montpellier, France*, 1990.

C. Godin and Y. Caraglio. A multiscale model of plant topological structures. *Journal of Theoretical Biology*, 191 :1–46, 1998.

G. Goodwin and R. Payne. *Dynamic System Identification : Experiment Design and Data Analysis*. Academic Press, 1977.

V. Grange. Greenlab mathematical model behavior analysis on roux architectural model. Master's thesis, Ecole Polytechnique, Palaiseau, France, 2006. Rapport de stage d'option scientifique.

Y. Guédon, D. Barthélémy, Y. Caraglio, and E. Costes. Pattern analysis in branching and axillary flowering sequences. *Journal of Theoretical Biology*, 212 :481–520, 2001.

H. Guo, V. Letort, L. Hong, T. Fourcaud, P.-H. Cournède, Y. Lu, and P. de Reffye. Adaptation of the greenlab model for analyzing sink-source relationships in chinese pine saplings. In T. Fourcaud and X. Zhang, editors, *Plant growth Modeling, simulation, visualization and their Applications (PMA06).* IEEE Computer Society (Los Alamitos, California), 2007.

Y. Guo, P. de Reffye, Y. Song, Z. Zhan, B. Li, and M. Dingkuhn. Modeling of Biomass Acquisition and Partitioning in the Architecture of Sunflower. In M. Jeager and B. Hu, editors, *Plant Growth Models and Applications*, pages 271–284. Tsinghua University Press and Springer, 2003.

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

F. Hallé, R. Oldeman, and P. Tomlinson. *Tropical trees and forests, An architectural analysis*. Springer-Verlag, New-York, 1978.

G. Hammer, M. Kropff, T. Sinclair, and J. Porter. Future contributions of crop modelling - from heuristics and supporting decision making to understanding genetic regulation and aiding crop improvement. *European Journal of Agronomy*, 18 :15–31, 2002.

G. Hammer, M. Cooper, F. Tardieu, S. Welch, B. Walsh, F. Van Eeuwijk, S. Chapman, and D. Podlich. Models for navigating biological complexity in breeding improved crop plants. *Trends in Plant Science*, 11(12) :587–593, 2006.

T. Harris. *The theory of branching processes*. Springer, Berlin, 1963.

E. Heuvelink. Evaluation of a dynamic simulation model for tomato crop growth and development. *Annals of Botany*, 83 :413–422, 1999.

F. Houllier. Generalized least squares for fitting amapagro model. Personal comm., LIAMA, 1999.

B. Hu and M. Jaeger. *Plant Growth Modeling and Applications. Proceedings of PMA03*. Tsinghua University Press and Springer, Beijing, 2003.

M. Jaeger. *Représentation et simulation de croissance des végétaux*. PhD thesis, Université Louis Pasteur, Strasbourg, France, 1987.

M. Jaeger and P. de Reffye. Basic concepts of computer simulation of plant growth. *Journal of Biosciences*, 17(3) :275–291, September 1992.

H. Jones. *Plants and Microclimate*. Cambridge University Press, 1992.

A. Jullien, J.-M. Allirand, P.-H. Cournède, A. Mathieu, P. de Reffye, and B. Ney. Is it possible to simulate rapeseed organ mass in relation to n nutrition ? calibration of the functional-structural model greenlab for the oilseed rape brassica napus l. during the vegetative phase for two nitrogen nutrition levels. In *Proceedings of the 12th International Rapeseed Congress, Wuhan, China (2007).*, 2007.

M. Kang and P. de Reffye. A mathematical approach estimating source and sink functioning of competing organs. In J. Vos, L. Marcelis, P. de Visser, P. Struik, and J. Evers, editors, *Functional-structural plant modelling in crop production, Wageningen*, volume Chapter 06, pages 70–80. Springer, 2007.

M. Kang, P.-H. Cournède, J.-P. Quadrat, and P. de Reffye. A stochastic language for plant topology. In T. Fourcaud and X. Zhang, editors, *Plant growth Modeling, simulation, visualization and their Applications*. IEEE Computer Society (Los Alamitos, California), 2007a.

M. Kang, E. Heuvelink, V. Letort, P.-H. Cournède, S. Carvalho, and P. de Reffye. Visualizing reaction of chrysanthemum to temperature and light : model calibration and validation. In P. Prusinkiewicz, J. Hanan, and B. Lane, editors, *Proceedings of the 5^{th} international workshop on Functional-Structural Plant Models (FSPM07)*. HortResearch (Napier, New Zealand), 2007b.

M. Kang, P.-H. Cournède, P. de Reffye, D. Auclair, and B. Hu. Analytical study of a stochastic plant growth model : application to the greenlab model. *Mathematics and Computers in Simulation*, 78(1) :57–75, 2008a.

M.-Z. Kang, P. de Reffye, J.-F. Barczi, B.-G. Hu, and F. Houllier. Stochastic 3d tree simulation using substructure instancing. In B. Hu and M. Jaeger, editors, *Plant Growth Models and Applications : PMA03*, pages 154–168. Tsinghua University Press and Springer (Beijing, China), 2003.

M.-Z. Kang, P.-H. Cournède, J. Le Roux, P. de Reffye, and B.-G. Hu. Theoretical study and numerical simulation of a stochastic model for plant growth. In *CARI04, Tunisia*, 2004.

O. Kniemeyer, G. Buck-Sorlin, and W. Kurth. A graph grammar approach to artificial life. *Artificial Life*, 10(4) :413–431, 2004.

O. Kniemeyer, G. Buck-Sorlin, and W. Kurth. GroIMP as a platform for functional-structural modelling for plants. In J. Vos, L. Marcelis, P. de Visser, P. Struik, and J. Evers, editors, *Functional-structural plant modelling in crop production, Wageningen*, volume Chapter 04, pages 50–60. Springer, 2007.

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

W. Kurth, O. Kniemeyer, and G. Buck-Sorlin. Relational growth grammars - a graph rewriting approach to dynamical systems with a dynamical structure. *Lecture Notes in Computer Science*, 3566 :56–72, 2005.

V. Le Chevalier, M. Jaeger, X. Mei, and P.-H. Cournède. Simulation and visualisation of functional landscapes : effects of the water resource competition between plants. *Journal of computer science and technology*, 22(6) :835–845, Nov. 2007a.

V. Le Chevalier, M. Jaeger, X. Mei, A. Lesluye, and P.-H. Cournède. A functional landscape prototype to simulate water resource competition between plants. In T. Fourcaud and X. Zhang, editors, *Plant growth Modeling, simulation, visualization and their Applications*. IEEE Computer Society (Los Alamitos, California), 2007b.

J. Lecoeur, R. Poiré-Lassus, A. Christophe, and L. Guilioni. Sunflo : A joint phenotyping and modelling approach to analyse and predict the differences in yield potential of sunflower genotypes. In *17^{th} International Sunflower Conference, June 8-12, 2008, Cordoba, Spain.*, 2008.

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

V. Letort. *Multi-scale analysis of source-sink relationships in plant growth models for parameter identification. Case of the GreenLab model.* PhD thesis, Ecole Centrale Paris, 2008.

V. Letort, P.-H. Cournède, J. Lecoeur, I. Hummel, P. De Reffye, and A. Christophe. Effect of topological and phenological changes on biomass partitioning in *Arabidopsis thaliana* inflorescence : a preliminary model-based study. In T. Fourcaud and X. Zhang, editors, *Plant growth Modeling, simulation, visualization and their Applications.* IEEE Computer Society (Los Alamitos, California), 2007a.

V. Letort, P.-H. Cournède, A. Mathieu, P. de Reffye, and T. Constant. Parameter identification of a functional-structural tree growth model and application to beech trees (*Fagus sylvatica*). In P. Prusinkiewicz, J. Hanan, and B. Lane, editors, *Proceedings of the* 5^{th} *international workshop on Functional-Structural Plant Models (FSPM07).* HortResearch (Napier, New Zealand), 2007b.

V. Letort, P.-H. Cournède, A. Mathieu, P. de Reffye, and T. Constant. Parametric identification of a functional-structural tree growth model and application to beech trees (*Fagus sylvatica*). *Functional Plant Biology*, 35(10) :951–963, 2008a.

V. Letort, P. Mahe, P.-H. Cournède, P. de Reffye, and B. Courtois. Quantitative genetics and functional-structural plant growth models : Simulation of quantitative trait loci detection for model parameters and application to potential yield optimization. *Annals of Botany*, 101(8), 2008b.

F. Lewis and V. Syrmos. *Optimal Control*. Wiley-IEEE, 1995.

A. Lindenmayer. Mathematical models for cellular interactions in development. i. filaments with one-sided inputs. *Journal of Theoretical Biology*, 18 :280–289, 1968.

A. Lindenmayer and H. Jürgensen. Grammars of development : Discrete-space models for growth, differentation and gene expression in modular organisms. In G. Rozenberg and A. Salomaa, editors, *Lindenmayer Systems : Impacts on Theoretical Computer Science, Computer Graphics and Developmental Biology*, pages 3–21. Springer-Verlag, Berlin, 1992.

L. Ljung. Asymptotic behavior of the extended kalman filter as a parameter estimator for linear systems. *IEEE Trans. Autom. Control*, 20(5) :643–652, 1979.

L. Ljung. *System identification : Theory for the user.* New Jersey : Prentice-Hall, 1999.

L. Ljung and T. Söderström. *Theory and Practice of Recursive Identification*. MIT Press, Cambridge, MA, 1983.

C. Loi and P.-H. Cournède. Generating functions of stochastic L-systems and application to models of plant development. *Discrete Mathematics and Theoretical Computer Science Proceedings*, AI :325–338, 2008a.

C. Loi and P.-H. Cournède. A markovian framework to formalize stochastic L-systems and application to models of plant development. Technical Report 6830, INRIA, 2008b.

D. Luquet, M. Dingkuhn, H. Kim, L. Tambour, and A. Clément-Vidal. Ecomeristem, a model of morphogenesis and competition among sinks in rice. 1. concept, validation and sensitivity analysis. *Functional Plant Biology*, 33(4) :309–323, 2007a.

Y. Ma, M. Wen, B. Li, P.-H. Cournède, and P. de Reffye. Calibration of greenlab model for maize with sparse experimental data. In T. Fourcaud and X. Zhang, editors, *The Second International Symposium on Plant Growth Modeling, Simulation, Visualization and Applications (PMA06).* IEEE (Los Alamitos, California), 2007a.

Y. Ma, Z. Zhan, Y. Guo, D. Luquet, P. de Reffye, and M. Dingkuhn. Parameter stability of the structural-functional plant model greenlab as affected by variation within populations, among seasons and among growth stages. *Annals of Botany*, 99 :61–73, 2007b.

Y. Ma, M. Wen, Y. Guo, B. Li, P.-H. Cournède, and P. de Reffye. Parameter optimization and field validation of the functional-structural model greenlab for maize at different population densities. *Annals of Botany*, 101(8), 2008.

J. Mailhol and A. Merrot. Spfc : a tool to improve water management and hay production in the crau region. *Irrigation Science*, 26(4) :289–302, 2008.

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

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

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

A. Mathieu, P.-H. Cournède, and P. de Reffye. A dynamical model of plant growth with full retroaction between organogenesis and photosynthesis. *ARIMA Journal*, 4 :101–107, 2006.

A. Mathieu, P.-H. Cournède, D. Barthélémy, and P. de Reffye. Conditions for the generation of rhythms in a discrete dynamic system. case of a functional structural plant growth model. In T. Fourcaud and X. Zhang, editors, *PMA06 - Plant growth Modeling, simulation, visualization and their Applications*. IEEE Computer Society (Los Alamitos, California), 2007a.

A. Mathieu, B. Zhang, E. Heuvelink, S. Liu, P.-H. Cournède, and P. de Reffye. Calibration of fruit cyclic patterns in cucumber plants as a function of source-sink ratio with the greenlab model. In H. J. Prusinkiewicz P, editor, *Proceedings of the 5th international workshop on FSPM*, November 2007b.

A. Mathieu, P.-H. Cournède, D. Barthélémy, and P. de Reffye. Rhythms and alternating patterns in plants as emergent properties of a model of interaction between development and functioning. *Annals of Botany*, 101(8), 2008.

A. Mathieu, P.-H. Cournède, V. Letort, D. Barthélémy, and P. de Reffye. A dynamic model of plant growth with interactions between development and functional mechanisms to study plant structural plasticity related to trophic competition. *Annals of Botany*, 2009.

N. Munier-Jolain, V. Biarnès, I. Chaillet, J. Lecoeur, and M. Jeuffroy. *Agrophysiologie du pois protéagineux*. INRA, 2005.

B. Pallas, A. Christophe, P.-H. Cournède, and J. Lecoeur. Trophic and non-trophic determinism of the end of axis development in grapevine (*Vitis vinifera* l.). *submitted*, 2008a.

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

J. Perttunen, R. Sievänen, E. Nikinmaa, H. Salminen, H. Saarenmaa, and J. Väkevä. Incorporating lindenmayer systems for architectural development in a functional-structural tree model. *Ecological Modelling*, 181 :479–491, 2005.

P. Prusinkiewicz. Modeling plant growth and development. *Current opinion in plant biology*, 7(1) :79–84, 2004.

P. Prusinkiewicz and A. Lindenmayer. *The Algorithmic Beauty of Plants*. Springer-Verlag, New-York, 1990.

P. Prusinkiewicz, A. Lindenmayer, and J. Hanan. Developmental models of herbaceous plants for computer imagery purposes. *Computer Graphics*, 22(4) :141–150, 1988.

M. Quach, N. Brunel, and F. d'Alché Buc. Estimating parameters and hidden variables in non-linear state-space models based on odes for biological networks inference. *Bioinformatics*, 23(23) :3209–3216, 2007.

H. Ratschek and J. Rokne. *New Computer Methods for Global Optimization*. Horwood, Chichester, 1988.

M. Reymond, B. Muller, A. Leonardi, A. Charcosset, and F. Tardieu. Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. *Plant Physiology*, 131 :664–675, 2003.

G. Rozenberg and A. Salomaa. *The Mathematical Theory of L-systems*. Academic Press, New York, 1980a.

O. Schabenberger and F. Pierce. *Contemporary Statistical Models for the Plant and Soil Sciences*. Addison-Wesley, 2001.

K. Shinozaki, K. Yoda, K. Hozumi, and T. Kira. A quantitative analysis of plant form - the pipe model theory i. basic analysis. *Japanese Journal of Ecology.*, 14 :97–105, 1964.

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

H. Sinoquet, G. Sonohat, J. Phattaralerphong, and C. Godin. Foliage randomness and light interception in 3-d digitized trees : an analysis from multiscale discretization of the canopy. *Plant, Cell and environment*, 28 :1158–1170, 2005.

A. Smith. Plants, fractals and formal languages. *Computer Graphics (SIGGRAPH 84 Conference Proceedings)*, 18(3) :1–10, 1984.

C. Soler, F. Sillion, F. Blaise, and P. de Reffye. An efficient instantiation algorithm for simulating radiant energy transfer in plant models. *ACM Transactions on Graphics*, 22 (2) :204–233, 2003.

D. Stroock. *An introduction to Markov processes*. Springer, 2005.

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

W. Taylor. Small sample properties of a class of two-stage aitken estimator. *Econometrica*, 45(2) :497–508, 1977.

F. Varenne. La simulation informatique face à la méthode des modèles. le cas de la croissance des plantes. *Nature Sciences Sociétés*, 11 :16–28, 2003.

F. Varenne. *Du Modèle à la Simulation Informatique*. Mathesis. Vrin, Paris, 2007.

S. Volkwein. *Model reduction using proper orthogonal decomposition*. Lecture Notes, University of Graz, Austria, 2008. URL `http ://www.uni-graz.at/imawww/volkwein/POD.pdf`.

D. Wallach, D. Makowski, and J. Jones. *Working with Dynamic Crop Models*. Elsevier, 2006.

E. Walter and L. Pronzato. *Identification de modèles paramétriques*. Masson, 2006.

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

L. Wu, P. De Reffye, B. Hu, F.-X. Le Dimet, and P.-H. Cournède. A water supply optimization problem for plant growth based on greenlab model. *ARIMA*, 3 :194–207, 2005.

H. Yan, J.-F. Barczi, P. de Reffye, and B. Hu. Fast Algorithms of Plant Computation Based on Substructure Instances. *Journal of WSCG*, 10 :145–152, 2002a.

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

Z. Zhan, P. de Reffye, F. Houllier, and B. Hu. Fitting a structural-functional model with plant architectural data. In B. Hu and M. Jaeger, editors, *Plant Growth Models and Applications*, pages 236–249. Tsinghua University Press and Springer (Beijing, China), 2003.

Z. Zhan, H. Rey, D. Li, Y. Guo, P.-H. Cournède, and P. de Reffye. Study on the effects of defoliation on the growth of cotton plant using the functional structural model greenlab. In T. Fourcaud and X. Zhang, editors, *Plant growth Modeling, simulation, visualization and their Applications*. IEEE Computer Society (Los Alamitos, California), 2007.

X. Zhao, P. de Reffye, F. Xiong, B. Hu, and Z. Zhan. Dual-scale automaton model of virtual plant growth. *Chinese Journal of Computers*, 24(6) :608–615, 2001.